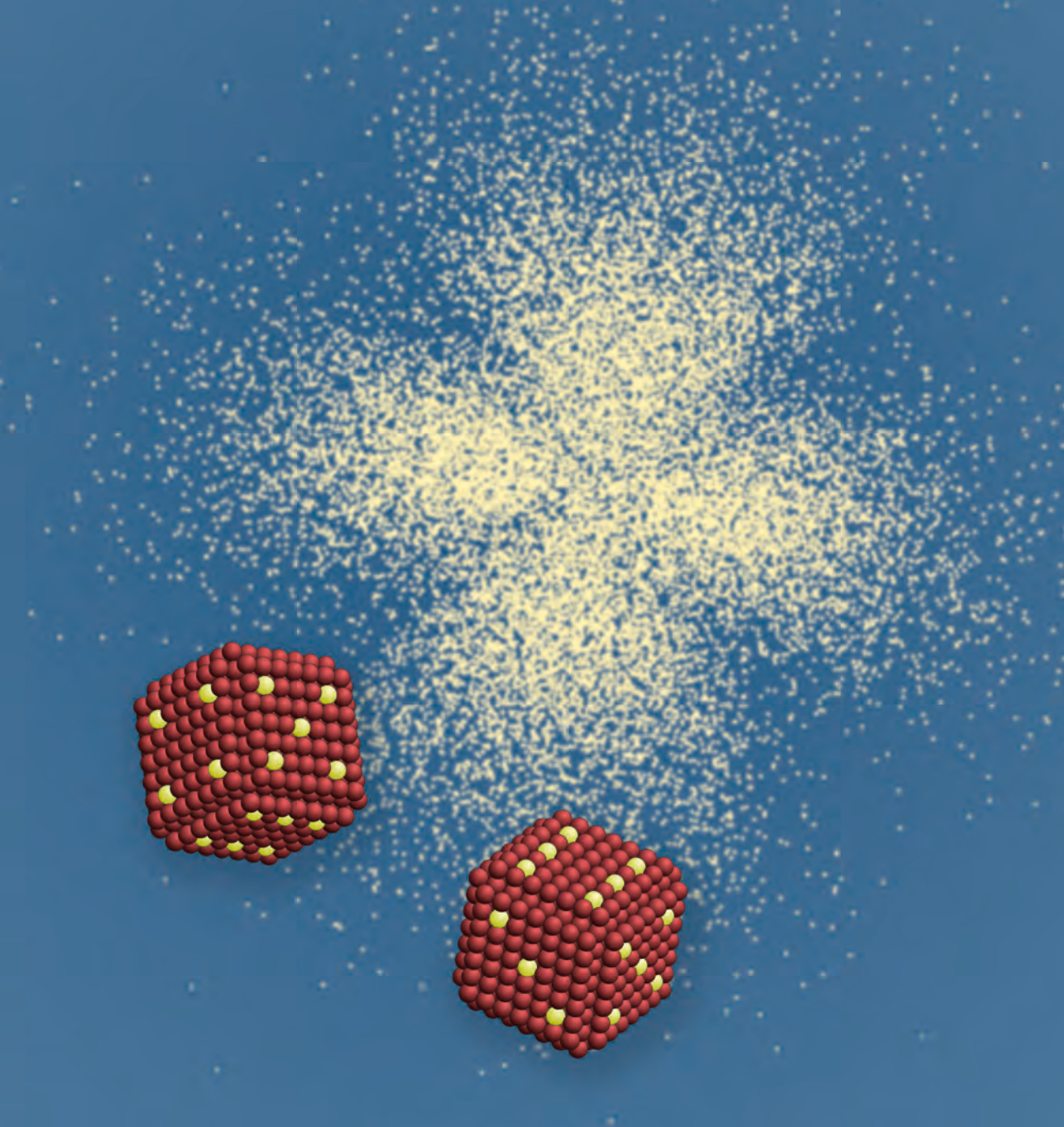




Many-Body Methods for Real Materials

Eva Pavarini, Erik Koch, and Shiwei Zhang (Eds.)

Forschungszentrum Jülich GmbH
Institute for Advanced Simulation

**Lecture Notes of the Autumn School on
Correlated Electrons 2019**

Eva Pavarini, Erik Koch, and Shiwei Zhang (Eds.)

Many-Body Methods for Real Materials

Autumn School organized by
the Institute for Advanced Simulation
at Forschungszentrum Jülich
16 – 20 September 2019

Schriften des Forschungszentrums Jülich
Modeling and Simulation

Band / Volume 9

ISSN 2192-8525

ISBN 978-3-95806-400-3

Bibliographic information published by the Deutsche Nationalbibliothek.
Die Deutsche Nationalbibliothek lists this publication in the Deutsche
Nationalbibliografie; detailed bibliographic data are available in the Internet
at <http://dnb.d-nb.de>.

Publisher: Forschungszentrum Jülich GmbH
Institute of Advanced Simulation

Cover Design: Grafische Medien, Forschungszentrum Jülich GmbH

Printer: Schloemer & Partner GmbH, Düren

Copyright: Forschungszentrum Jülich 2019

Distributor: Forschungszentrum Jülich
Zentralbibliothek, Verlag
52425 Jülich
Phone +49 (0)246 1 61-5368 · Fax +49 (0)246 1 61-6103
e-mail: zb-publikation@fz-juelich.de
www.fz-juelich.de/zb

Schriften des Forschungszentrums Jülich
Reihe Modeling and Simulation, Band / Volume 9

ISSN 2192-8525
ISBN 978-3-95806-400-3

Vollständig frei verfügbar über das Publikationsportal des Forschungszentrums Jülich (JuSER) unter
www.fz-juelich.de/zb/openaccess



This is an Open Access publication distributed under the terms of the [Creative Commons Attribution License 4.0](https://creativecommons.org/licenses/by/4.0/),
which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Contents

Preface

1. Introduction to Density Functional Theory
Xavier Blase
2. The Random Phase Approximation and its Applications to Real Materials
Xinguo Ren
3. Introduction to Variational and Projector Monte Carlo
Cyrus Umrigar
4. Optimized Quantum Monte Carlo Wave Functions
Arne Lüchow
5. Variational Wave Functions for Strongly Correlated Fermionic Systems
Federico Becca
6. Auxiliary-Field Quantum Monte Carlo at Zero- and Finite-Temperature
Shiwei Zhang
7. Exact Diagonalization and Lanczos Method
Erik Koch
8. Quantum Chemistry DMRG in a Local Basis
Miles Stoudenmire
9. Density Matrix Renormalization
Karen Hallberg
10. Dynamical Mean-Field Theory and the Mott Transition
Marcelo Rozenberg
11. Dynamical Mean-Field Theory for Materials
Eva Pavarini
12. Analytical Properties of Self-Energy and Luttinger-Ward Functional
Robert Eder
13. Introduction to Many-Body Green Functions In and Out Of Equilibrium
James Freericks
14. Electronic Transport in Correlated Single Molecule Junctions
Andrea Donarini
15. Diagrammatic Monte Carlo
Nikolay Prokof'ev
16. Stochastic Series Expansion Methods
Anders Sandvik
17. Algebraic Methods in Many-Body Physics
Gerardo Ortiz

Index

Preface

Emergent many-body phenomena are at the core of the exciting properties of strongly-correlated materials. Understanding them requires confronting the many-body problem. While, at first, this appears to be an impossible task, substantial progress has been made by combining physical insights with modern numerical approaches. A successful strategy is to devise methods that use the understanding gained from simple models for the construction of physically motivated wave-functions. Results for the ground state of real materials can then be obtained by optimizing them via deterministic or stochastic algorithms. The methods of choice for determining spectra are instead based on Green functions. The key idea is to map the complex realistic many-body Hamiltonian to a simpler auxiliary model that can be solved numerically.

This year's school will provide an overview of the state-of-the art of these techniques, their successes and their limitations. After introducing fundamental models and key concepts, lectures will focus on quantum Monte Carlo for optimizing correlated wave-functions, stochastically sampling series expansions for obtaining Green functions, and renormalization group techniques. Advanced lectures will address approaches to Mott physics, transport phenomena, and out-of-equilibrium dynamics. Applications will cover correlated systems ranging from transition-metal compounds and frustrated spin systems to correlated molecules.

The goal of the school is to introduce advanced graduate students and up to these modern approaches for the realistic modeling of strongly-correlated materials.

A school of this size and scope requires support and help from many sources. We are very grateful for all the financial and practical support we have received. The Institute for Advanced Simulation at the Forschungszentrum Jülich and the Jülich Supercomputer Centre provided the major part of the funding and were vital for the organization of the school and the production of this book. The Institute for Complex Adaptive Matter (ICAM) supported selected international speakers and participants.

The nature of a school makes it desirable to have the lecture notes available when the lectures are given. This way students get the chance to work through the lectures thoroughly while their memory is still fresh. We are therefore extremely grateful to the lecturers that, despite tight deadlines, provided their manuscripts in time for the production of this book. We are confident that the lecture notes collected here will not only serve the participants of the school but will also be useful for other students entering the exciting field of strongly correlated materials.

We are grateful to Mrs. H. Lexis of the Verlag des Forschungszentrum Jülich and to Mrs. D. Mans of the Grafische Betriebe for providing their expert support in producing the present volume on a tight schedule. We heartily thank our students and postdocs who helped with proofreading the manuscripts, often on quite short notice: Julian Mußhoff, Neda Samani, Qian Zhang, and Xue-Jing Zhang.

Finally, our special thanks go to Dipl.-Ing. R. Hölzle for his invaluable advice on the innumerable questions concerning the organization of such an endeavor, and to Mrs. L. Snyders for expertly handling all practical issues.

Eva Pavarini, Erik Koch, and Shiwei Zhang

August 2019

1 Introduction to Density Functional Theory

Xavier Blase

Institut Néel

CNRS, Grenoble, France

Contents

1	Overview	2
2	The many-body problem (selected considerations)	3
2.1	One and two-body operators act on the charge and pair density	4
2.2	Exchange-correlation hole and its sum-rule	4
3	Density functional theory	5
3.1	Hohenberg and Kohn theorems	5
3.2	The Euler-Lagrange equation: a density-only formulation	7
3.3	Kinetic-energy functionals: from hydrogen to the HEG	8
3.4	Kohn-Sham formulation: introducing auxiliary 1-body orbitals	9
4	Exchange-correlation functionals	10
4.1	The local density approximation: Ceperley and Alder QMC data	10
4.2	Structural properties within LDA: average impressive results	11
4.3	LDA satisfies the exchange-correlation sum-rule	13
4.4	Jacob's ladder of functionals: towards accuracy heaven?	14
5	On the meaning of Kohn-Sham auxiliary 1-body eigenstates	17
5.1	The band gap problem with DFT Kohn-Sham eigenvalues	18
5.2	Level ordering and self-interaction problems	19
5.3	What is the meaning of Kohn-Sham eigenvalues?	20
5.4	Janak theorem and fractional occupations in ensemble DFT	21

1 Overview

We present in this Chapter a brief overview of Density Functional Theory (DFT), an exact mean-field formalism for calculating ground-state total energies and charge densities. Several excellent books are devoted to DFT and two are listed in the bibliography [1, 2]. At the heart of DFT lies the idea that there is no need to know the details of the many-body wavefunction to calculate ground-state properties: the knowledge of the electronic density $\rho(\mathbf{r})$, a simple 3D scalar function, is enough to obtain the total energy of the system and all quantities that result (atomic structures, binding or atomization energies, elastic constants, phonon energies, activation barriers, the forces needed for molecular dynamics simulations, etc.) As an exact theorem, DFT applies to simple Fermi-liquid-like systems but also to strongly correlated materials.

Unfortunately, the DFT does not say how exactly the ground-state total energy depends on the ground-state electronic density. In practice, approximations are needed to express the kinetic energy and the many-body electron-electron interaction as a “functional” of $\rho(\mathbf{r})$. One thus leaves exact DFT and enters the difficult world of approximations with their specific range of validity. Here comes a second remarkable feature of DFT, namely that very simple approximations for such functionals, including the local density approximation (LDA), deliver excellent results, such as interatomic bond lengths within 1% of experimental data for a very large number of systems. Such an accuracy, combined with the simplicity of DFT that allows to study systems comprising several hundred atoms, can explain the formidable success of DFT in terms of the number of users and systems studied, with the development of very efficient and easy-to-use codes. This success can certainly contribute to explain that the 1998 Chemistry Nobel prize was awarded to Walter Kohn (and John Pople) for the development of DFT.

It remains, however, that DFT in its original formulation is limited to ground-state properties. As such, the field of electronic excitations, and in particular charged excitations as measured in a photoemission experiment for establishing “band-structures” does not formally lie within the reach of DFT. Fortunately, and unfortunately, a specific implementation of DFT, the Kohn-Sham formalism, introduces auxiliary one-body eigenstates and eigenvalues that are very widely used to calculate electronic energy levels. The rationale for doing so is not firmly established, but very valuable information about band dispersions, orbital shapes, etc. are usually obtained for “not-strongly-correlated” systems. Specific limitations are well established (too small band gaps, underbinding of localized states, etc.) that may find a partial cure by considering “generalized DFT” namely a mean-field approach combining DFT and Hartree-Fock: this is the field of hybrid functionals.

Clearly, there is the need for more solid foundations allowing, to build a formal link between DFT and excited states. This is where we stand nowadays, with a healthy competition between the world of DFT, attempting to bridge the gap with excited states properties, and other approaches that abandon mean-field techniques to tackle the explicit many-body problem, but at a cost that needs to be improved to compete in the study of large systems. This will be the subject of most of the other chapters in this book.

2 The many-body problem (selected considerations)

We start this chapter with a short reminder of quantum mechanics for many-particle (electrons) systems. The quantum states of an N -electron system are described by a wavefunction: $\psi(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2, \dots, \mathbf{r}_N\sigma_N)$ with $\mathbf{r}_i$ and σ_i space and spin variables. The probability to find N electrons with spins $(\sigma_1, \dots, \sigma_N)$ in the infinitesimal volume $d\mathbf{r}_1 d\mathbf{r}_2 \dots d\mathbf{r}_N$ centered at $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N$ is given by $d^{3N}\mathcal{P} = |\psi(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2, \dots, \mathbf{r}_N\sigma_N)|^2 d\mathbf{r}_1 d\mathbf{r}_2 \dots d\mathbf{r}_N$. By definition, the electron density can be found by integrating over $N-1$ spacial-variables and summing over all spins

$$\rho(\mathbf{r}) = N \int d\sigma dx_2 \dots dx_N |\psi(\mathbf{r}\sigma, x_2, \dots, x_N)|^2 \quad \text{where } x_i = (\mathbf{r}_i, \sigma_i)$$

with $\rho(\mathbf{r})d\mathbf{r}$ the number of electrons in the infinitesimal volume $d\mathbf{r}$ centered at $\mathbf{r}$. From the normalization of ψ one obtains $\int d\mathbf{r} \rho(\mathbf{r}) = N$. The charge density can be obtained by multiplying by the elementary charge $(-e)$. Note that it is customary to use the wording charge density for electron density.

The electronic Hamiltonian is actually known (atomic units):

$$\hat{H} = -\frac{1}{2} \sum_{i=1}^N \nabla_i^2 + \sum_{i=1}^N v^{\text{ion}}(\mathbf{r}_i) + \sum_{i < j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \quad \text{with } v^{\text{ion}}(\mathbf{r}) = - \sum_I \frac{Z_I}{|\mathbf{R}_I - \mathbf{r}|} \quad (1)$$

where we did not include the kinetic energy of the ions. v^{ion} is the ionic potential acting on electrons with $\{\mathbf{R}_I, Z_I\}$ the nuclear positions and charges. The energy of the system is given by the “expectation value” of the Hamiltonian (let’s forget spin)

$$\langle \psi | \hat{H} | \psi \rangle = \int \dots \int d\mathbf{r}_1 d\mathbf{r}_2 \dots d\mathbf{r}_N \psi^*(\{\mathbf{r}_i\}) \hat{H}(\{\mathbf{r}_i\}, \{\nabla_{\mathbf{r}_i}\}) \psi(\{\mathbf{r}_i\}) .$$

As such, it may seem that quantum mechanics is easy, with first-principles equations and formalisms developed in the first half of the 20th century. Let’s consider, however, the energy of a very small system, the water molecule with its 10 electrons. Let’s assume that we want to calculate its total energy for some wavefunction ψ (let’s not even ask how we obtained it ...) Applying naively some quadrature (e.g. trapezoidal rule) to calculate such an integral by paving the space around the molecule with a coarse $10 \times 10 \times 10$ grid, one obtains for $N=10$ electrons a sum of $10^{3N} = 10^{30}$ terms to calculate and add. Modern computers are “petaflop”: they perform 10^{15} floating point operations per second. We would therefore need of the order of 10^{15} seconds, namely 31 710 millennia for this simple evaluation! Clearly, the way we calculated this integral was very dumb, and clever sampling of phase space can be done much more efficiently using, e.g., Metropolis sampling. It remains that the exact many-body problem becomes dramatically expensive as soon as the number of electrons increases. Computers are handy, but the brains of the physicists and chemists to come up with nice approximations are luckily still required.

We now start by introducing simple considerations demonstrating that one does not always need all the details of the complex many-body wavefunction to calculate a physical observable. Unless stated otherwise, we will not display spin variables in the following for sake of brevity of the equations.

2.1 One and two-body operators act on the charge and pair density

Let us consider one-body operators, namely operators of the form $\hat{O} = \sum_{i=1}^N O(\mathbf{r}_i)$, where the $\{\mathbf{r}_i\}$ are the electronic positions. In particular, the electron-ion interaction energy is given by $E^{\text{ion-e}} = \sum_i \langle \psi | v^{\text{ion}}(\mathbf{r}_i) | \psi \rangle$ with

$$\begin{aligned} \langle \psi | v^{\text{ion}}(\mathbf{r}_i) | \psi \rangle &= \int d\mathbf{r}_1 d\mathbf{r}_2 \dots d\mathbf{r}_N v^{\text{ion}}(\mathbf{r}_i) |\psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_i, \dots, \mathbf{r}_N)|^2 \\ &= \int d\mathbf{r} d\mathbf{r}_2 \dots d\mathbf{r}_N v^{\text{ion}}(\mathbf{r}) |\psi(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N)|^2 \end{aligned}$$

where we have renamed all variables, in particular $\mathbf{r}_i \Rightarrow \mathbf{r}$, and reshuffled all space positions thanks to the symmetry properties of $|\psi|^2$. As a result, all N terms are identical, yielding

$$E^{\text{ion-e}} = N \int d\mathbf{r} v^{\text{ion}}(\mathbf{r}) \int d\mathbf{r}_2 \dots d\mathbf{r}_N |\psi(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N)|^2 = \int d\mathbf{r} v^{\text{ion}}(\mathbf{r}) \rho(\mathbf{r}).$$

As such, v^{ion} acts only on the electronic (or charge) density: there is no need for the full many-body wavefunction and its related $3N$ -integrals to get the $E^{\text{ion-e}}$ energy!

To conclude this paragraph, one can introduce another one-body operator, the electron density operator, $\hat{\rho}(\mathbf{r}) = \sum_{i=1}^N \delta(\mathbf{r} - \mathbf{r}_i)$, that “counts” the number of electrons at $\mathbf{r}$. The same demonstration as above (exercise!) allows to recover the expression for the electronic density: $\rho(\mathbf{r}) = \langle \psi | \hat{\rho}(\mathbf{r}) | \psi \rangle = N \int d\mathbf{r}_2 \dots d\mathbf{r}_N |\psi(\mathbf{r}, \mathbf{r}_2, \dots, \mathbf{r}_N)|^2$.

Let’s now consider the crucial case of two-body operators. The electron-electron interaction energy $E^{ee} = \langle \psi | \hat{V}^{ee} | \psi \rangle$ with $\hat{V}^{ee} = \sum_{i < j}^N \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}$ can also be simplified by renaming and reshuffling the integration variables:

$$\langle \psi | \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} | \psi \rangle = \int d\mathbf{r} d\mathbf{r}' d\mathbf{r}_3 \dots d\mathbf{r}_N \frac{|\psi(\mathbf{r}, \mathbf{r}', \mathbf{r}_3, \dots, \mathbf{r}_N)|^2}{|\mathbf{r} - \mathbf{r}'|}$$

independent of the specific (i, j) indices, yielding $N(N-1)/2$ identical terms so that $E^{ee} = \int d\mathbf{r} d\mathbf{r}' \rho_2(\mathbf{r}, \mathbf{r}') / |\mathbf{r} - \mathbf{r}'|$ with

$$\rho_2(\mathbf{r}, \mathbf{r}') = \frac{N(N-1)}{2} \int d\mathbf{r}_3 \dots d\mathbf{r}_N |\psi(\mathbf{r}, \mathbf{r}', \mathbf{r}_3, \dots, \mathbf{r}_N)|^2$$

which is the density of pairs satisfying $\int d\mathbf{r} d\mathbf{r}' \rho_2(\mathbf{r}, \mathbf{r}') = N(N-1)/2$.

For calculating the complex electron-electron energy responsible for electronic correlations, there is no need, in principle, for all the details of the many-body wavefunctions and the related $3N$ -integrals: we only need averaged (mean-field) quantities such as the 2-body pair-density! BUT we do not know at this stage how to build $\rho_2(\mathbf{r}, \mathbf{r}')$ without the knowledge of ψ .

2.2 Exchange-correlation hole and its sum-rule

Rewriting the pair density as $2\rho_2(\mathbf{r}, \mathbf{r}') = \rho(\mathbf{r})\rho(\mathbf{r}') [1 + h(\mathbf{r}, \mathbf{r}')]]$, with h called the pair-correlation function, one can express E^{ee} in terms of the (charge-charge) Hartree energy J

$$J = \frac{1}{2} \int d\mathbf{r} d\mathbf{r}' \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}$$

and the exchange-correlation (XC) hole density $\rho_{XC}(\mathbf{r}, \mathbf{r}') := \rho(\mathbf{r}')h(\mathbf{r}, \mathbf{r}')$ as

$$E^{ee} = J + \frac{1}{2} \int d\mathbf{r} d\mathbf{r}' \frac{\rho(\mathbf{r})\rho_{XC}(\mathbf{r}, \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}.$$

It is easy to demonstrate (exercise!) that

$$\int d\mathbf{r}' \rho_2(\mathbf{r}, \mathbf{r}') = \frac{N-1}{2} \rho(\mathbf{r}) \quad \text{and} \quad \int d\mathbf{r}' \rho_{XC}(\mathbf{r}, \mathbf{r}') = -1$$

yielding the XC-hole sum rule. The XC energy is the Coulomb energy between electrons and their respective XC-hole, namely the depletion of the charge density by one electron (through exchange and Coulomb repulsion) dynamically created around each electron. The 1/2 in the XC energy is an adiabatic factor: the XC hole grows with the electron density and would not exist without it.

The XC energy beyond the classical Hartree term can be written in a classical form as the Coulomb interaction between an electron and its XC hole, namely the dynamical depletion of exactly one charge created “locally” by Fermi and Coulomb repulsion. Due to the XC sum rule, the composite object (a quasiparticle) made out of the electron dressed by its XC hole is a neutral object weakly interacting with its surrounding. Independent-like particle theories, and related 1-body eigenvalue equations, are therefore more likely to be successful when applied to such quasiparticles.

3 Density functional theory

We have established that we do not need to know all the details of the many-body wavefunction to obtain in particular the electron-electron interaction energy. The pair density is sufficient. It can be shown further that the kinetic energy can be obtained from the knowledge of the 1st-order density matrix: $\gamma_1(\mathbf{r}, \mathbf{r}') = N \int \cdots \int d\mathbf{r}_2 \cdots d\mathbf{r}_N \psi(\mathbf{r}, \mathbf{r}_2, \cdots, \mathbf{r}_N) \psi^*(\mathbf{r}', \mathbf{r}_2, \cdots, \mathbf{r}_N)$ which is also a 2-body function averaging out most of the many-body wavefunction degrees of freedom. Density Functional Theory (DFT) goes one step beyond, demonstrating that the ground-state (GS) total energy only requires the knowledge of the electronic density, very much as for the action of the one-body ionic potential.

3.1 Hohenberg and Kohn theorems

Preliminaries: Room temperature is of the order of 25 meV, much smaller than typical electronic energy gaps or band dispersions, so most unperturbed (no strong light, etc.) solids or molecules are close to their lowest energy state with wavefunction ψ_{GS} and energy E_{GS} .

The variational principle provides a way to find ψ_{GS} and energy E_{GS} :

$$E_{GS} = \min_{\psi} E[\psi] \quad \text{with} \quad E[\psi] = \langle \psi | \hat{H} | \psi \rangle \quad \text{and} \quad \langle \psi | \psi \rangle = 1$$

This is the standard approach, where the energy is a *functional* of the many-body wavefunction ψ . The dramatic result from Hohenberg and Kohn [3] is that the ground-state energy can be written as a functional of the charge density:

$$E_{GS} = \min_{\psi} E[\psi] \xrightarrow{DFT} E_{GS} = \min_{\rho} E[\rho] \quad \text{with} \quad \int d\mathbf{r} \rho(\mathbf{r}) = N.$$

This is an *exact* result, namely there is an exact mean-field theory for the problem of the ground-state energy in N -electron systems!

Demonstration for non-degenerate ground-states: The ionic potential acts on the charge density and the electronic Hamiltonian (without ion-ion interaction) reads

$$\hat{H} = \hat{T} + \hat{V}^{ee} + \int d\mathbf{r} v^{ext}(\mathbf{r})\rho(\mathbf{r})$$

with $\hat{T}$ the kinetic energy and $v^{ext}=v^{\text{ion}}$ (the ions are “external” to the N -electron system).

Theorem: given a charge density $\rho(\mathbf{r})$, then there exist only one external potential $v^{ext}(\mathbf{r})$ (within a constant) such that the corresponding ground-state electronic density is equal to $\rho(\mathbf{r})$.

Reductio ad absurdum: Assume there exist 2 external potentials $v_1^{ext}(\mathbf{r})$ and $v_2^{ext}(\mathbf{r})$ that lead to the same ground-state charge density:

$$\begin{aligned} v_1^{ext}(\mathbf{r}) &\implies \hat{H}_1 \implies \psi_1^{GS} \implies \rho(\mathbf{r}) \\ v_2^{ext}(\mathbf{r}) &\implies \hat{H}_2 \implies \psi_2^{GS} \implies \rho(\mathbf{r}) \end{aligned}$$

Using the variational principle

$$\begin{aligned} E_1^{GS} = \langle \psi_1 | \hat{H}_1 | \psi_1 \rangle &< \langle \psi_2 | \hat{H}_1 | \psi_2 \rangle = \langle \psi_2 | \hat{H}_2 | \psi_2 \rangle + \langle \psi_2 | \hat{H}_1 - \hat{H}_2 | \psi_2 \rangle \\ &= E_2^{GS} + \int d\mathbf{r} (v_1^{ext} - v_2^{ext})(\mathbf{r})\rho(\mathbf{r}) \end{aligned}$$

Starting now from $\langle \psi_2 | \hat{H}_2 | \psi_2 \rangle$ (switching indices 1 and 2) one obtains

$$E_2^{GS} < E_1^{GS} + \int d\mathbf{r} (v_2^{ext} - v_1^{ext})(\mathbf{r})\rho(\mathbf{r})$$

and by adding the two inequalities

$$E_1^{GS} + E_2^{GS} < E_2^{GS} + E_1^{GS} \quad \text{IMPOSSIBLE!}$$

The demonstration hinges here on strict inequalities, namely assuming non-degenerate ground-states. This is the celebrated 1964 theorem by Hohenberg and Kohn [3].

Ground-state energy as a functional of the charge density: It follows that the charge density completely determines the external potential and thus the Hamiltonian (just add the universal kinetic $\hat{T}$ and V^{ee} operators) and thus the ground-state wavefunction ψ_{GS} :

$$\begin{array}{ccccccc} v^{ext} & \longrightarrow & \hat{H} & \longrightarrow & \psi_{GS} & \longrightarrow & \rho(\mathbf{r}) \\ \uparrow & & & & & & \downarrow \\ & \longleftarrow & & & & & \end{array}$$

Since $\rho(\mathbf{r})$ determines ψ_{GS} , it determines unequivocally the total energy of the ground-state, $E_{GS} = \langle \psi_{GS} | \hat{H} | \psi_{GS} \rangle$. It can be shown further as a corollary (Exercise) that the *variational principle* can be now used for E_{GS} as a functional of the charge density:

$$E_{GS} = \min_{\psi} E[\psi] \xrightarrow{DFT} E_{GS} = \min_{\rho} E[\rho] \quad \text{under the constraint that } \int d\mathbf{r} \rho(\mathbf{r}) = N.$$

Since the electron-ion interaction energy can be written explicitly as a functional of the charge density, one concludes that the sum of the kinetic and electron-electron interaction energy is also a functional of $\rho(\mathbf{r})$ that is labeled the universal Hohenberg and Kohn functional $F_{HK}[\rho]$ with $F_{HK}[\rho] = E_{GS}[\rho] - \int d\mathbf{r} v^{\text{ion}}(\mathbf{r})\rho(\mathbf{r})$. It is called universal since, contrary to the ionic potential that depends on the system of interest via the ionic charges and positions, the kinetic energy and electron-electron operators do not.

The basic ideas discussed in this section can be generalized without imposing the non-degeneracy of the ground state, bypassing further the problem of the v -representability of a given density $\rho(\mathbf{r})$: can we always find some v^{ext} potential that leads to a given $\rho(\mathbf{r})$? This can be done within the framework of Levy constrained-search formulation [4] that leads to the following definition for the universal Hohenberg and Kohn functional

$$F_{HK}[\rho] = \min_{\psi \rightarrow \rho} \langle \psi | \hat{T} + \hat{V}^{ee} | \psi \rangle$$

This is formally the standard search over many-body wavefunctions, but with the constraint that for a given density ρ , the search is restricted to many-body wavefunctions with $\langle \psi | \hat{\rho}(\mathbf{r}) | \psi \rangle = \rho(\mathbf{r})$. This further allows to define a kinetic energy and electron-electron interaction energy as independent functionals of the charge density: $T[\rho] = \min_{\psi \rightarrow \rho} \langle \psi | \hat{T} | \psi \rangle$ and $E^{ee}[\rho] = \min_{\psi \rightarrow \rho} \langle \psi | \hat{V}^{ee} | \psi \rangle$.

3.2 The Euler-Lagrange equation: a density-only formulation

The existence of $E[\rho]$ with an associated variational principle allows performing energy minimization with respect to the density under the constraint that the densities we consider integrate to the total number of electrons N . We thus introduce the Lagrangian

$$\Omega[\rho, \mu] = E[\rho] + \mu \left(N - \int d\mathbf{r} \rho(\mathbf{r}) \right)$$

where μ is a Lagrange parameter ensuring the conservation of the correct electron number. This leads by differentiation to the stationary equation

$$\frac{\partial \Omega[\rho, \mu]}{\partial \rho(\mathbf{r})} = 0 \quad \implies \quad \frac{\partial F_{HK}[\rho]}{\partial \rho(\mathbf{r})} + v^{\text{ion}}(\mathbf{r}) = \mu$$

which is a simple 3D differential equation. Its solution is the ground-state charge density from which the ground-state energy of the system can be calculated. This is dramatically simpler than the original many-body wave-function formulation. Unfortunately, we just do not know the universal functional $F_{HK}[\rho]$ expressing the kinetic energy and electron-electron interaction as a function of the density. While functionals of the density for the electron-electron interaction will be discussed and have met much success, one complicated issue remains, the kinetic energy for which we provide now two limiting expressions.

3.3 Kinetic-energy functionals: from hydrogen to the HEG

The first case is the well-known non-interacting homogeneous electron gas (the free-electron gas) that obeys the following relations (see e.g. Kittel)

$$T = \frac{3}{5} N \varepsilon_F \quad \text{with} \quad \varepsilon_F = \frac{\hbar^2 k_F^2}{2m_e} \quad \text{and} \quad k_F^3 = 3\pi^2 \frac{N}{V},$$

where V is the volume occupied by the N electrons, E_F the Fermi energy and T the total kinetic energy. In a homogeneous system, N/V is just the electronic density. It is traditional to rewrite

$$T = \frac{1.105}{r_s^2} \text{ (a.u.)} \quad \text{with} \quad \frac{4}{3}\pi r_s^3 = \frac{V}{N} = \frac{1}{\rho} \quad \text{the Wigner-Seitz radius.}$$

One may then “cook-up” some kinetic energy per electron

$$t_{TF}(\rho) = \frac{T}{N} = \frac{3}{5} \times \frac{\hbar^2}{2m_e} \times (3\pi^2 \rho)^{2/3}$$

yielding the historical Thomas-Fermi (TF) kinetic energy functional approximation for inhomogeneous systems with position dependent densities $\rho(\mathbf{r})$

$$T_{TF} = \int d\mathbf{r} \rho(\mathbf{r}) t^{TF}(\rho(\mathbf{r})) \simeq 2.871 \int d\mathbf{r} \rho(\mathbf{r})^{5/3} \text{ (a.u.)},$$

which is the first example of a local functional: the local kinetic energy density only depends on the density at the same space point. Numerical tests have shown that such an approximation is rather poor for real systems.

Another exact relation can be obtained for the hydrogen atom. In that case, the only occupied orbital is the 1s orbital: $\phi(\mathbf{r}) = A e^{-r}$ (r in a.u.) with A some normalizing constant taken to be positive. Since there is only one electron, the density reads $\rho(\mathbf{r}) = |\phi(\mathbf{r})|^2$. As such

$$T = -\frac{1}{2} \int d\mathbf{r} \phi(\mathbf{r}) \nabla^2 \phi(\mathbf{r}) = \frac{1}{2} \int d\mathbf{r} (\nabla \phi(\mathbf{r})) \cdot (\nabla \phi(\mathbf{r})),$$

where we used integration by part with the wavefunction cancelling at infinity. Using now $\phi(\mathbf{r}) = \sqrt{\rho(\mathbf{r})}$, with $\rho(\mathbf{r})$ positive, we obtain

$$T = \frac{1}{2} \int d\mathbf{r} (\nabla \sqrt{\rho(\mathbf{r})})^2 = \frac{1}{2} \int d\mathbf{r} \left(\frac{\nabla \rho(\mathbf{r})}{2\sqrt{\rho(\mathbf{r})}} \right)^2 = \frac{1}{8} \int d\mathbf{r} \frac{[\nabla \rho(\mathbf{r})]^2}{\rho(\mathbf{r})}.$$

This expression, called the von Weizsäcker functional, is yet another exact formula valid for a given specific system, but clearly very different from the one given above for the homogeneous gas. Anticipating the “gradient corrected functionals” we see that here the energy depends on the density and its gradient.

The search for a universal functional for the kinetic energy turns to be much more challenging than finding a decent functional for the electron-electron interaction. As such, “orbital-free” DFT, namely a DFT based exclusively on the density, has not yet met much success despite its remarkable simplicity.

3.4 Kohn-Sham formulation: introducing auxiliary 1-body orbitals

To bypass this problem of the expressing the kinetic energy as a functional of the density, Kohn and Sham introduced in 1965 their famous approach [5]. The idea goes as follows: in the case of a non-interacting electron system, electronic states can be described by one-body orbitals $\{\phi_n(\mathbf{r})\}$ arranged and populated by increasing energies $\{\varepsilon_n\}$. For such systems, the kinetic energy and electronic density are easily calculated

$$T_0 = -\frac{1}{2} \sum_{n=1}^N \int d\mathbf{r} \phi_n^*(\mathbf{r}) \nabla^2 \phi_n(\mathbf{r}) \quad \text{and} \quad \rho(\mathbf{r}) = \sum_{n=1}^N |\phi_n(\mathbf{r})|^2.$$

Can one imagine now a fictitious non-interacting electron gas submitted to an effective external potential $v^{\text{eff}}(\mathbf{r})$ such that its charge density is the same as that of the real interacting system? Since the ground-state electronic density fulfills the Euler-Lagrange equation, the real and fictitious systems should have the same Euler-Lagrange equation in each point, namely

$$\frac{\partial F_{HK}[\rho]}{\partial \rho(\mathbf{r})} + v^{\text{ion}}(\mathbf{r}) = \frac{\partial T_0}{\partial \rho(\mathbf{r})} + v^{\text{eff}}(\mathbf{r}),$$

yielding the definition of such an effective potential

$$v^{\text{eff}}(\mathbf{r}) = v^{\text{ion}}(\mathbf{r}) + \frac{\partial(F_{HK}[\rho] - T_0)}{\partial \rho(\mathbf{r})}.$$

Introducing the classical (Hartree) charge-charge interaction potential

$$v^H(\mathbf{r}) = \frac{\partial J[\rho]}{\partial \rho(\mathbf{r})} = \int d\mathbf{r}' \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}$$

with

$$J[\rho] = \frac{1}{2} \int d\mathbf{r} d\mathbf{r}' \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}$$

one can write

$$v^{\text{eff}}(\mathbf{r}) = v^{\text{ion}}(\mathbf{r}) + v^H(\mathbf{r}) + \frac{\partial E_{XC}[\rho]}{\partial \rho(\mathbf{r})}$$

with E_{XC} the DFT exchange-correlation energy

$$E_{XC} = T[\rho] - T_0[\rho] + E^{ee}[\rho] - J[\rho].$$

We observe that $T[\rho]$ and $E^{ee}[\rho]$ are still to be determined, but the standard argument is that $T_0[\rho]$ captures a significant fraction of $T[\rho]$ so that an approximation performed on E_{XC} is likely to have less detrimental effects as compared to directly approximating $T[\rho]$. We observe that within DFT, the exchange-correlation energy contains some correction to the kinetic energy, not solely the deviation $(E^{ee}[\rho] - J[\rho])$ from the electron-electron interaction to the classical Hartree term. Yet, what is the functional $E_{XC}[\rho]$?

4 Exchange-correlation functionals

The search for exchange-correlation functionals is the central on-going challenge within DFT with decades of failures and difficult successes. Below, we very briefly review some key considerations and terminology.

4.1 The local density approximation: Ceperley and Alder QMC data

Proposed in the seminal 1964 Hohenberg and Kohn paper, the local density approximation (LDA), reminiscent of the Thomas-Fermi kinetic energy functional described above, relies on the following approximation

$$E_{XC}[\rho] \simeq \int d\mathbf{r} \rho(\mathbf{r}) e_{XC}(\rho(\mathbf{r}))$$

where it is assumed that the exchange-correlation energy per electron, $e_{XC}(\mathbf{r})$, for an electron located in $\mathbf{r}$ only depends on the local value of the electronic density $\rho(\mathbf{r})$. Such an approximation is strictly valid only in the limit of a homogeneous electron gas. The next step was then taken by Ceperley and Alder [6] in 1986, who performed nearly exact (within numerical accuracy) quantum Monte Carlo (QMC) calculations for the interacting homogeneous electron gas (HEG) at various densities ρ^{hom} . For an homogeneous electron gas with homogeneous ionic positive background, the Hartree and electron-ion energies cancel exactly. The calculated QMC total energy $E^{QMC}[\rho^{hom}]$ contains thus the kinetic energy T and the $(E^{ee} - J)$ electron-electron interaction beyond the Hartree term. Subtracting now the kinetic energy $T^0[\rho^{hom}]$ for the non-interacting electron gas of same density, as given exactly by the Thomas-Fermi expression, one obtains straightforwardly

$$e_{XC}(\rho(\mathbf{r})) = \frac{E^{QMC}[\rho^{hom}] - T^0[\rho^{hom}]}{N} \quad \text{with} \quad \rho^{hom} = \rho(\mathbf{r})$$

with N the number of electrons for which the total energy is calculated in the QMC simulation and $\rho^{hom} = V/N$. One may further subtract the exact exchange energy, namely the total energy of the HEG in the Hartree-Fock approximation where the many-body wavefunction is described by a single Slater determinant. This is a famous calculation performed by Dirac in 1930 [7], yielding for the HEG exchange-energy

$$\frac{E_X}{N} = -\frac{3k_F}{4\pi} \simeq \frac{-0.458}{r_s} \text{ a.u.} \quad \text{with} \quad \frac{V}{N} = \frac{4}{3}\pi r_s^3,$$

with r_s the Wigner-Seitz radius, namely the radius of the sphere with volume V/N . Such an expression can alternatively be expressed as a function of the electronic density, yielding the Dirac exchange expression

$$E_X^{Dirac} = -C_X \int d\mathbf{r} \rho(\mathbf{r})^{4/3} \quad \text{with} \quad C_X = \frac{3}{4} \left(\frac{3}{\pi} \right)^{1/3}.$$

By subtraction, one obtains the correlation-only energy per electron $e_C^{hom}(\rho)$. The QMC data points have been fitted by clever functional forms (Vosko-Wilk-Nusair [8], Perdew-Zunger [9], etc.) that are used nowadays in standard DFT codes. We leave this section by noting that for the HEG with cancelling Hartree and electron-ion interactions, the kinetic and exchange energies add up to

$$E_{HEG}^{HF} = \frac{1.105}{r_s^2} - \frac{0.458}{r_s} \quad (\text{a.u.})$$

Without the exchange term, the HEG is unstable with no finite equilibrium r_s value.

Once the exchange-correlation energy is defined, one can proceed with solving the Kohn-Sham equations associated with the fictitious non-interacting electron gas

$$\left(-\frac{\nabla^2}{2} + v^{\text{eff}}(\mathbf{r})\right) \phi_n(\mathbf{r}) = \varepsilon_n \phi_n(\mathbf{r}) \quad \text{with} \quad v^{\text{eff}}(\mathbf{r}) = v^{\text{ions}}(\mathbf{r}) + V^H(\rho(\mathbf{r})) + V^{\text{XC}}(\rho(\mathbf{r})),$$

where we use the LDA exchange-correlation potential $V_{LDA}^{\text{XC}}(\rho(\mathbf{r})) = \partial E_{XC}^{\text{LDA}}(\rho(\mathbf{r}))/\partial \rho(\mathbf{r})$. Solving the Kohn-Sham equations yields the Kohn-Sham electronic energy levels $\{\varepsilon_n\}$ and the Kohn-Sham eigenstates $\{\phi_n\}$ from which the ground-state density and associated LDA total energy can be obtained

$$\begin{aligned} \rho(\mathbf{r}) &= \sum_{n=1}^N |\phi_n(\mathbf{r})|^2 \quad \text{and} \quad T^0 = \sum_{n=1}^N \langle \phi_n | -\nabla^2/2 | \phi_n \rangle \\ E_0^{\text{LDA}} &= T^0 + \int d\mathbf{r} v^{\text{ions}}(\mathbf{r}) \rho(\mathbf{r}) + E^H[\rho] + E_{LDA}^{\text{XC}}[\rho] \end{aligned}$$

where the sum extends over the N lowest energy levels (ground-state at zero temperature). An important aspect of such equations is that they are self-consistent: to obtain the $\{\phi_n\}$ one needs to solve the Kohn-Sham equations with a potential that depends *via* the density on ... the $\{\phi_n\}$! In practice, one takes some guess input density ρ_0 , such as a simple superposition of tabulated atomic densities, to build an input XC potential $V_{LDA}^{\text{XC}}(\rho_0(\mathbf{r}))$ that allows obtaining a first guess of Kohn-Sham orbitals. These orbitals allow to build an updated electronic density and a related updated XC potential and Kohn-Sham Hamiltonian, leading to a updated set of orbitals and density. When the input and output electronic densities and/or XC potentials are the same, the self-consistent cycle is stopped, leading to the ground-state self-consistent density and total energy. The technicalities of converging to the correct energy minimum (are there local minima?) and of the convergence rate are difficult issues not dealt with here.

4.2 Structural properties within LDA: average impressive results

We now know how to calculate in practice the ground-state energy within the LDA approximation for a given system characterized by ionic positions and nuclear charges (defining the v^{ion} potential) and the number N of electrons. We thus can, in particular, answer the important question: how good is the LDA approximation? We must remember here that what we are targeting with DFT are the ground-state total energy and charge density. Other observables, such as electronic energy levels, are not *in principle* within the scope of what DFT is designed for.

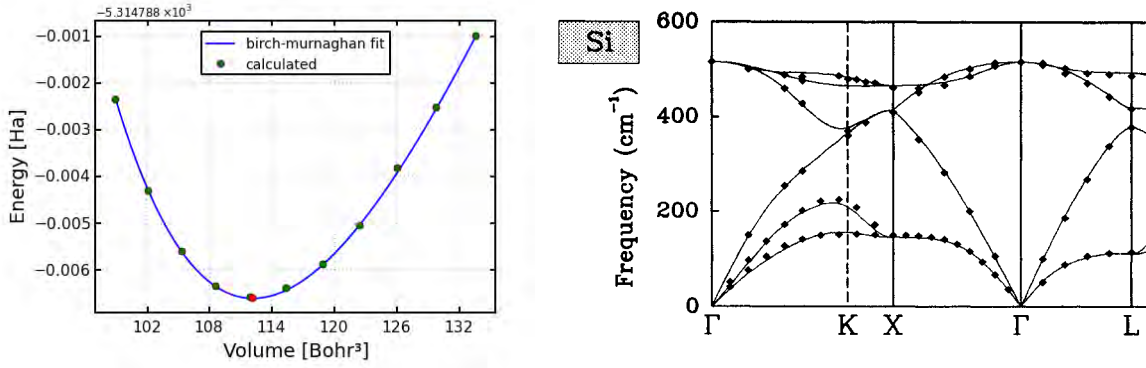


Fig. 1: (Left) Energy versus unit-cell volume for silver. DFT energy data points are fitted by some clever functional form (Birch-Murnaghan fit). The energy minimum gives the equilibrium volume and related lattice parameter as well as interatomic distances that can be compared to experiment (courtesy <http://exciting-code.org/beryllium-volume-optimization-for-cubic-systems>). (Right) Calculated LDA versus experimental phonon dispersion for silicon (from Ref. [13]).

We will come back to that point in the following Sections. In their original papers, Hohenberg, Kohn, and Sham were actually critical about the potential success of the LDA, concluding that for actual systems variations of the charge density were so strong that a model exact in the limit of homogeneous distribution of charges appeared to have little chance of being successful.

We now present in Fig. 1 (left) the results of a standard numerical exercise consisting in finding the equilibrium cell volume (or lattice parameter) for a solid using DFT (here silver in its FCC structure). With the space group known, one can calculate the DFT total energy for various unit cell volumes. The resulting calculated data points can be fitted by some polynomial law (or a more adapted functional form such as the Birch-Murnaghan law), yielding the equilibrium volume at zero temperature for a given approximation to the XC potential. The associated error for several functionals and several crystal families (metals, non-metals) are given in Table 1. Considering non-metallic structures, that include typical covalent systems such as silicon and diamond with very inhomogeneous density distributions, one finds that LDA predicts the lattice

	Metals(14)				Nonmetals (10)			
	LDA	PBEsol	PBE	TPSS	LDA	PBEsol	PBE	TPSS
ME (Å)	-0.136	-0.039	0.046	0.039	-0.042	0.026	0.085	0.066
MAE (Å)	0.136	0.042	0.060	0.060	0.042	0.026	0.085	0.066
MRE (%)	-2.71	-0.76	0.95	0.74	-0.86	0.56	1.76	1.35
MARE (%)	2.71	0.83	1.24	1.15	0.86	0.56	1.76	1.35

Table 1: Statistical data: mean error, mean absolute error, mean relative error MRE % and mean absolute relative error MARE %, for lattice constants Å of a selection of 14 metals and 10 nonmetals. Errors with respected to experimental data corrected for ZPAE (zero point anharmonic expansion) contribution. The PBEsol functional is a modification of PBE for solids [10] while TPSS is a “metaGGA” functional [11] with a dependence on the density but also the Laplacian $\nabla^2 \rho(\mathbf{r})$ proportional to the kinetic energy (adapted from Table IV from Ref. [12]).

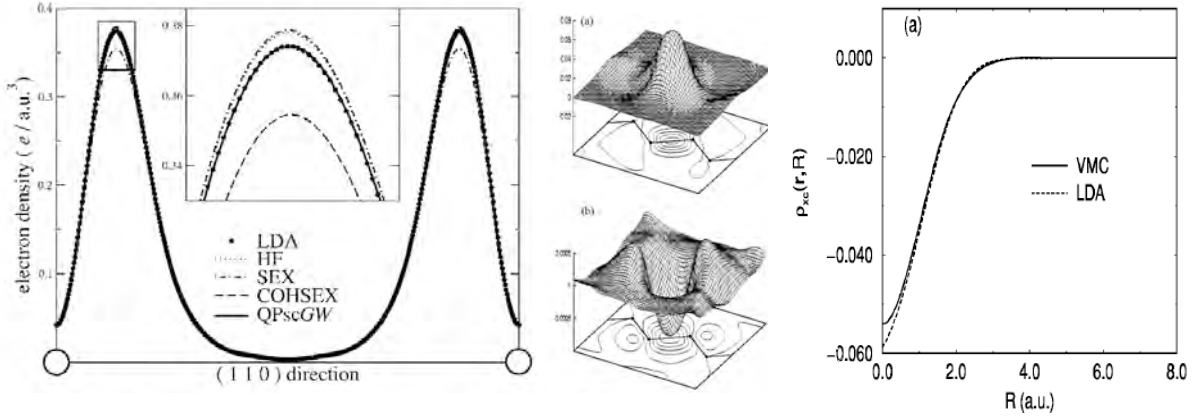


Fig. 2: (Left) Charge density for solid argon along an interatomic bond comparing LDA with the so-called quasiparticle self-consistent QPscGW approximation (from Ref. [14]). (Center and Right) LDA versus variational Monte Carlo (VMC) one-body density matrix and spherically averaged XC hole $\rho_{xc}(\mathbf{r}; R)$ for Silicon (from Ref. [15]).

parameters of solids with an average accuracy well within 1% (see boxed numbers). This is a remarkable results for an approach that is very simple and relies on a approximation assumed to be valid only for systems with homogeneous charge densities (e.g. alkali metals). The error is further systematic, with a tendency to overbind (too small lattice parameters). Such performance can certainly explain the success of DFT combined with the Kohn-Sham formalism and the local density approximation.

Beyond lattice parameters, we further plot in Fig. 1 (right) the LDA phonon band-structure for silicon as compared to experiment [13], with phonon energies standing as 2nd-order derivatives of total energies with respect to ionic positions via the force-constant matrix. Again, the agreement is very remarkable for a system far from displaying a homogeneous charge density. Since the DFT is designed to provide ground-state electronic densities, we now reproduce in Fig. 2 (left) the density of solid argon along a bond direction, comparing the LDA electronic density with results of a self-consistent many-body Green's function approach labeled the QPscGW formalism [14]. Clearly, the agreement is excellent, even in the present case of a very inhomogeneous density profile. Similar results are obtained for the one-body density matrix in Fig. 2 (center) in the case of silicon, comparing now the LDA density matrix built from the Kohn-Sham eigenstates: $\gamma_1(\mathbf{r}, \mathbf{r}') = \sum_{n=1}^N \phi_n(\mathbf{r})\phi_n(\mathbf{r}')$ with a variational Monte Carlo (VMC) reference: $\gamma_1(\mathbf{r}, \mathbf{r}') = N \int d\mathbf{r}_2 \cdots d\mathbf{r}_N \psi^*(\mathbf{r}, \mathbf{r}_2 \cdots \mathbf{r}_N) \psi(\mathbf{r}', \mathbf{r}_2 \cdots \mathbf{r}_N)$. The deviation between the two calculations is well within 1% again.

4.3 LDA satisfies the exchange-correlation sum-rule

This somehow unexpected success of the LDA for systems displaying strongly inhomogeneous charge density distributions relies in particular on the facts that (a) the exchange-correlation (XC) energy depends on a Coulomb-weighted spherical average over the XC-hole and (b) the LDA XC hole satisfies the exchange-correlation sum-rule. Taking the relation between the

electron-electron XC energy and the XC hole, one obtains indeed

$$E_{XC} = \frac{1}{2} \int d\mathbf{r} \rho(\mathbf{r}) \int_0^{+\infty} 4\pi R dR \rho_{XC}^{SA}(\mathbf{r}, R) \quad \text{and} \quad \rho_{XC}^{SA}(\mathbf{r}, R) = \frac{1}{4\pi} \int_{|\mathbf{r}-\mathbf{r}'|=R} d\mathbf{r}' \frac{\rho_{XC}(\mathbf{r}, \mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}$$

with $\rho_{XC}^{SA}(\mathbf{r}, R)$ the spherically averaged XC-hole around the $\mathbf{r}$ -point using the Coulomb norm. We plot in Fig. 2 (right) such a spherically-averaged XC-hole, comparing again LDA and VMC in the case of Silicon. While for a given $(\mathbf{r}, \mathbf{r}')$ pair of positions the LDA and VMC XC hole differ significantly, their spherically averaged values are remarkably close. Further, it can be shown [16] that the LDA XC hole satisfies an exact sum rule, namely

$$4\pi \int R^2 dR \rho_{XC}^{SA}(\mathbf{r}, R) = -1$$

indicating that if the LDA XC-hole is too small for a given R value, then it is too large in compensation for another R distance. Such properties are very strong desirable constraints that may explain the success of the LDA and ... the failure of subsequent approximations. The inclusion of some kinetic energy ($T - T_0$) in the definition of the DFT XC energy leads to caution when defining the DFT XC hole within what is called the “adiabatic connection formalism” that builds a connection between the non-interacting and interacting systems sharing the same density.

4.4 Jacob’s ladder of functionals: towards accuracy heaven?

Improving on the LDA approximation for a better description of observables related to the total energy (binding energy, atomization energy, structural phase diagrams, activation barriers, elastic constants, phonon spectra, etc.) remains the central issue in the field of DFT. Roughly speaking, two strategies can be followed: the first is to develop functionals that satisfy exact mathematical relations, such as satisfying the XC hole sum rule. The second strategy is more pragmatic and consists in fitting some general functional form with parameters on experimental data. An interesting article recently published by Perdew and coworkers in the journal *Science* and entitled “Density functional theory is straying from the path toward the exact functional” [17] provides a nice discussion on the two philosophies, illustrating further the “jungle” of functionals that exist nowadays. Clearly, functionals fitted to experimental data can be very accurate, but go away from the *ab initio* or first-principles character of DFT.

Concerning the strategy that consists in satisfying exact relations, an interesting illustration can be found in the early days of the so-called gradient corrected functionals that attempt to go away from the local density approximation by devising functionals that depend not only on the local density but also on its gradient to capture some information about charge inhomogeneities. Early functionals with low-order gradient corrections (LGC), such as the following one for the exchange (X) energy (σ the spin degree of freedom)

$$E_X^{LGC} = E_X^{LDA} - \beta \sum_{\sigma} \int d\mathbf{r} \rho_{\sigma}^{4/3} x_{\sigma}^2 \quad \text{with} \quad x_{\sigma} = \frac{|\nabla \rho_{\sigma}|}{\rho_{\sigma}^{4/3}} \quad \text{a-dimensional}$$

really failed in improving over the LDA. In fact, E_X^{LGC} does not satisfy the exchange-hole sum rule, leading to a divergency in the vacuum where $\rho(\mathbf{r})$ decays exponentially, etc. Further, the potential felt by an electron far away from an atom, molecule, surface, etc. should scale as $-1/r$. This term comes from the exchange potential, yielding for the exchange energy density at long distance in the vacuum

$$\lim_{r \rightarrow \infty} \epsilon_X(\mathbf{r}) \simeq -\frac{\rho(\mathbf{r})}{2r} \quad \text{with} \quad \rho(\mathbf{r}) = 2\rho_{\uparrow}(\mathbf{r}) = 2\rho_{\downarrow}(\mathbf{r}) \quad (\text{unpolarized systems}).$$

To cure such problems, Becke proposed in 1988 an exchange functional (B88) [18] that scales smoothly between the small and large $x_{\sigma} = |\nabla \rho_{\sigma}|/\rho_{\sigma}^{4/3}$

$$E_X^{B88} = E_X^{LDA} - \beta \sum_{\sigma} \int d\mathbf{r} \rho_{\sigma}^{4/3} \frac{x_{\sigma}^2}{1 + 6\beta \sinh^{-1}(x_{\sigma})} \quad (\beta \text{ parameter}).$$

The GGA correction vanishes for small gradients ($x_{\sigma} \rightarrow 0$). With $\rho(\mathbf{r}) \simeq e^{-\alpha r}$ for large r , $x_{\sigma} \rightarrow e^{\alpha r/3}$, and $\sinh^{-1}(x_{\sigma}) \rightarrow \alpha r/3$, the correct (vacuum) asymptotic behavior is recovered. This is the exchange functional used in the celebrated BLYP functional (B=Becke88). Here exact relations (asymptotic behavior, low or high density limit, sum rules, etc.) lead to a functional form that still contains one parameter fitted to experimental data, combining *de facto* the two above-mentioned strategies. Such generalized gradient approximations (GGA) constitute the “second rung” of the so-called Jacob’s ladder of functionals that provides a classification of functionals with increasing average accuracy [19]. Well known functionals of that family include the PW91 (Perdew-Wang 1991) [20], the PBE (Perdew-Becke-Ernzerhof) [21], or the BLYP (Becke-Lee-Yang-Parr) functionals [22]. As can be seen in Table 1, the PBE functional leads to better results for metals as compared to the LDA, with a tendency to underbind (too large lattice parameters), but no improvements for non-metallic systems (e.g. semiconductors). We witness here the fact that functionals developed for finite size systems, with, e.g., the proper treatment of long-range behavior in the vacuum, may not be relevant for solids where there is ... no vacuum. The PBEsol functional (see Table 1) is a modification of PBE for solids, yielding better results indeed for periodic extended systems.

We now introduce a key generalization of DFT, namely the merging of DFT and Hartree-Fock yielding “hybrid functionals” where density-dependent expressions are complemented by one-body-orbital dependent exact exchange. For the Kohn-Sham system, we indeed know how to calculate the “exact” exchange energy “associated with” the Kohn-Sham eigenstates $\{\phi_n\}$

$$E_X = -\frac{1}{2} \sum_{ij}^{occ} \iint d\mathbf{r} d\mathbf{r}' \frac{\phi_i^*(\mathbf{r}) \phi_j(\mathbf{r}) \phi_j^*(\mathbf{r}') \phi_i(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \delta_{\sigma_i \sigma_j},$$

where we have re-introduced the spin variables. This is more expensive than pure density functionals (Hartree-Fock (HF) scales as N^4 with system size) but helps in several directions:

- it offers clearly the correct asymptotic behavior for the electronic potential in the vacuum and satisfies the exchange-hole sum rule

- it helps curing *the self-interaction (SI) problem*: within DFT, since the charge density depends on the occupied orbitals, the action of $v^{\text{eff}}[\rho]$ on an occupied orbital amounts to having an electron interacting with itself (consider the H atom system!). This is a dramatic problem for localized orbitals. It does not exist within HF since the SI in the Hartree and exchange energies cancel out.

However, mixing 100% of exact exchange with a density dependent correlation functional leads (in general) to a failure. Density-dependent XC functionals are usually built together for reproducing the total XC potential properties. Namely, they benefit from large error cancellations. Considerations built on the “adiabatic connection” between the non-interacting and interacting electronic systems generated the historical Becke half-and-half functional: $E^{XC} = 0.5E^X(HF) + 0.5E^X(\text{Slater}) + E^C(\text{LYP=Lee-Yang-Parr})$. With a fitting strategy on 56 small molecules atomization energies, 8 proton affinities, and 10 first-row total atomic energies, the “Becke 3 parameters” (the B3 of B3LYP) exchange functional [23] mixes Slater LDA and B88 GGA exchange with 20% of exact exchange. Using perturbation theory, Perdew, Burke and Ernzerhof advocated 25% of exact exchange, leading to the 1996 PBE0 functional [24]. The B3LYP and PBE0 formula are probably the most popular functionals in quantum chemistry for finite size systems.

To combine the need for having 100% of exact exchange in the long-range in the case of finite size systems, while a much smaller amount of exchange in the short-range, range-separated hybrids were introduced [25]. The idea is to define a short-range (SR) and long-range (LR) Coulomb interaction thanks, e.g., to the error function erf, allowing the introduction of a long-range-only exchange potential

$$v_X^{LR}(\mathbf{r}, \mathbf{r}'; \omega) = - \sum_i^{\text{occ}} \phi_i(\mathbf{r}) \phi_i^*(\mathbf{r}') \frac{\text{erf}(\omega |\mathbf{r} - \mathbf{r}'|)}{|\mathbf{r} - \mathbf{r}'|}.$$

The use of the complementary error function allows to introduce a short-range-only exact exchange. As such, one can introduce different amounts of local (Dirac) and exact exchange in the short and long ranges. The ω parameter controls the (inverse) effective length that partition the interaction between short or long range. The very popular CAM-B3LYP functional [26] includes 65% of LR exact exchange with $\omega = 0.33$ while the LC- ω PBE includes 100% of LR exact exchange with $\omega = 0.4$.

We abandon here the hope to provide a thorough description of functionals. Let’s conclude on the fact that contrary to finite size molecules, long-range Coulomb interactions in solids are renormalized by the macroscopic dielectric constant ϵ_M that diverges in the case of metallic systems. As such, the long-range amount of exact exchange in solids should be qualitatively proportional to $(1/\epsilon_m)$, and even exponentially decaying in metals (Yukawa-type behavior). This is, e.g., the rationale behind the HSE functional, a range-separated hybrid for extended systems relying on the solid-state physics language of screened Coulomb potentials [27].

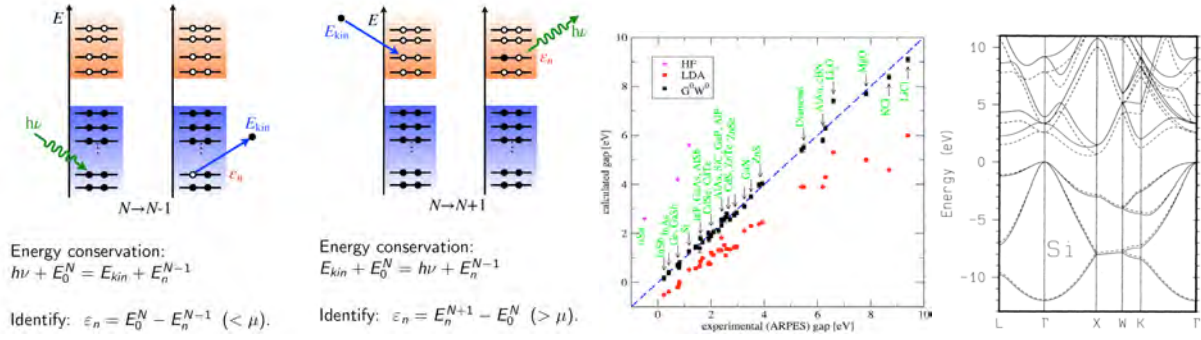


Fig. 3: (Left) Symbolic representation of direct and inverse photoemission experiments interpreted from the standard “one-electron energy levels” diagram. (Center) Compilation of LDA band gaps (red dots) in semiconductors and insulators as compared to experiments (first diagonal), Hartree-Fock data (cyan dots), and to a higher-level many-body Green’s function approach, the so-called $G_0W_0@LDA$ formalism (courtesy Valerio Olevano). (Right) LDA (dashed) versus GW (full lines) band-structure of Silicon. The zero of energy has been set to the valence bands top (from Ref. [28]).

5 On the meaning of Kohn-Sham auxiliary 1-body eigenstates

We tackle now a delicate problem within DFT, namely that of electronic properties. As discussed above, DFT is a ground state formalism designed to reproduce ground-state total energies and electronic density. As such, it is not designed to provide electronic energy levels. However, there is a very large literature exploiting Kohn-Sham eigenvalues for plotting the band-structure of realistic materials, with much success in many situations, but also well-documented limitations and failures. These are such aspects that we briefly explore now.

The meaning of what we call “electronic energy levels” when we plot a band-structure (or discrete energy levels in the case of a molecule) must be found in the experiment used to measure them, e.g., a photoemission experiment. In direct photoemission (see Fig. 3 (left)), a photon with energy $h\nu$ arrives on a piece of matter in its ground-state (with energy $E_0[N]$) and ejects an electron with some residual kinetic energy E_{kin} , leaving the system with $(N-1)$ electrons and an energy $E_n[N-1]$. The index n labels an eigenstate of the $(N-1)$ electron system. We define the energy of the electron in the system as $\varepsilon_n = E_0[N] - E_n[N-1]$, the energy of the level from which the electron was ejected. Using conservation of energy, $E_0[N] - E_n[N-1] = E_{kin} - h\nu$ measured experimentally. Similar considerations can be used to assimilate in inverse photoemission the differences of energy ($E_n[N+1] - E_0[N]$) between an excited state of the $(N+1)$ -electron system and the ground-state of the N -electron system as the unoccupied level energies.

Electronic energy levels as measured experimentally using photo-emission experiments are really differences of total energies between excited states of the $(N+1)$ or $(N-1)$ electron systems and the N -electron system in its ground state, namely $\varepsilon_n = E_n[N+1] - E_0[N]$ for unoccupied levels and $\varepsilon_n = E_0[N] - E_n[N-1]$ for occupied levels. What is the relation between the $\{\varepsilon_n^{KS}\}$ Kohn-Sham eigenvalues and such total energy differences between charged and neutral systems?

5.1 The band gap problem with DFT Kohn-Sham eigenvalues

Before trying to find a rationale for the Kohn-Sham energy levels, namely the energy of the fictitious independent electrons in the effective potential v^{eff} defined above, let's consider actual calculations, starting with the band gap of standard semiconductors and insulators. Data are compiled in Fig. 3 (center). The results of actual DFT calculations using the LDA XC potential (red dots) drive us to the conclusion that the band structure obtained with Kohn-Sham eigenvalues yields too small gaps. In the DFT Kohn-Sham world, the band gap of silicon turns to be about 0.5-0.6 eV, a factor two smaller than the 1.2 eV experimental value. Very similar results are obtained using “pure” DFT functionals, namely functionals not including some amount of exact exchange such as PBE. This is the band gap problem within DFT. A close inspection of Fig. 3 (center) for small gap systems further reveals that some systems, that are gaped semiconductors experimentally, turn out to be metallic (negative gap, namely an overlap of the occupied and empty bands) within LDA. This is the case, e.g., of the simple germanium system. It is fair to say that turning an insulator into a metal is a somehow dramatic failure. This tendency to underestimate band gaps can also be witnessed in the case of organic molecules for which the Kohn-Sham LDA HOMO-LUMO (highest occupied/lowest unoccupied molecular orbitals) energy gap can be underestimated by several eVs as compared to experiment or higher level techniques (see Fig. 4 (left)).

To better understand why DFT is still very widely used to study the electronic properties of a large variety of systems, we provide now in Fig. 3 (right) the LDA band structure for silicon (dashed line) that we compare to a much more accurate many-body perturbation theory approach, the *GW* formalism. We align the two band structures at the top of the valence band (zero of energy). The remarkable feature evidenced by this plot is that besides the band gap problem, the dispersion of bands in the valence and conduction manifolds are extremely close. Namely, the two band-structures could agree very well if a “scissor” operation, consisting in rigidly shifting the conduction bands by 0.5-0.6 eV higher in energy, is applied.

Contrary to Kohn-Sham DFT, Hartree-Fock (HF) yields too large gaps (see cyan dots in Fig. 3 (center) and HF data in Fig. 4 (left)). In the HF world, the band gap of Silicon is of the order of 6 eV, dramatically too large. This is due to the lack of correlations. Clearly, mixing some amount of exact exchange with DFT density-dependent XC potentials leads to much better gaps. This is exemplified in Fig. 4 (left) with the B3LYP data that are in better agreement with experiment. This is a clear incentive to mix exact exchange and density-dependent functionals, namely to use hybrid functionals for electronic properties, even though the criteria for selecting the proper amount of exact exchange for a given system is a difficult challenge if one wishes to preserve an *ab initio* (no fitting parameters) approach. The B3LYP functional includes 20% of exact exchange, an amount that does not seem sufficient to provide an accurate gap.

The band gap problem in DFT has been analyzed in depth and is related to the lack of a discontinuity of density-based XC potentials upon addition or removal of an electron [33, 34]. In a typical bulk system composed of the order of 10^{23} electrons per cm^3 , adding or removing an electron delocalized in some Bloch state hardly changes the charge density. As such, an

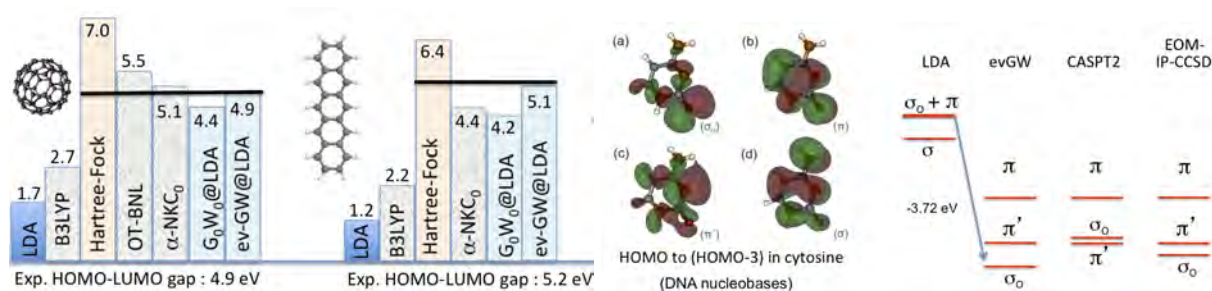


Fig. 4: (Left) HOMO-LUMO gap of gas phase (isolated) C_{60} fullerene and pentacene molecules. Various DFT and generalized DFT formalisms, including Kohn-Sham LDA, Hartree-Fock, B3LYP hybrid functionals, optimally-tuned range-separated hybrid (OT-BNL; Ref. [29]) or Koopmans' compliant (KNC; Ref. [30]) functionals are compared to many-body perturbation theories (G_0W_0 and evGW approaches) and to experiment (horizontal black line). (adapted from Ref. [31]) (Right) Top-most occupied levels for isolated cytosine DNA nucleobases as calculated using Kohn-Sham LDA approach and higher level many-body perturbation theories (evGW, CASPT2, EOM-IP-CCSD). Notice the reordering of levels from LDA to higher level approaches (adapted from Ref. [32]).

exchange-correlation potential built as a functional of the charge density will show no variations upon, e.g., adding an electron that will populate the bottom of the conduction bands across the gap. A simple analysis of the exact exchange operator clearly reveals that the XC potential should be discontinuous upon adding a charge to the neutral system.

We conclude this section with a warning: as discussed above, the DFT Kohn-Sham approach may turn a very standard band semiconductor, such as germanium, into a metal. This does not mean that the “band structure picture” fails and that strong correlations beyond mean-field (Mott transition) should be invoked to describe germanium. It just means that the DFT Kohn-Sham formalism with local functionals of the electronic density is not an accurate, not even a well-defined, formalism to capture band gaps.

5.2 Level ordering and self-interaction problems

This simple analysis in terms of band gap errors that can be cured by a rigid shift of, e.g., empty states needs however to be taken with care. While silicon and other simple sp -bonded semiconductors or insulators are characterized by Bloch states displaying equivalent spatial localization properties, systems mixing localized and extended states may suffer from a wrong ordering of levels at the Kohn-Sham DFT level with density-dependent functionals. Such systems include, e.g., transition metals displaying localized $3d$ levels together with itinerant (delocalized) sp bands, surfaces with extended bulk states versus localized surface states, or modern molecular electronic devices with extended states in the metallic electrodes but very localized molecular states in the junction. In the case of occupied levels, the self-interaction problem discussed above in Section 4.4 affects the localized states much more than the delocalized ones. As such, the ordering of levels within the occupied manifold can be wrong as well within DFT Kohn-Sham. This is exemplified in Fig. 4 (right) in the case of the gas phase (isolated) cytosine DNA

nucleobasis: while the highest occupied molecular orbital should be a delocalized π orbital, the Kohn-Sham LDA approach predicts erroneously at the top of the occupied manifold a very localized σ_O molecular orbital, localized on the oxygen. The self-interaction error is erroneously repelling high in energy this localized σ_O orbital. Hybrid functionals, with a portion of exact exchange, help in curing this problem since Hartree-Fock is self-interaction free.

5.3 What is the meaning of Kohn-Sham eigenvalues?

We now turn to a more formal analysis of the meaning of Kohn-Sham eigenvalues. We start by recovering the Kohn-Sham eigenvalue equation using the variational principle for the ground-state energy considered as a functional of the one-body Kohn-Sham auxiliary eigenstates $\{\phi_n\}$. We thus define a Lagrangian

$$\Omega[\{\phi_i\}, \{\lambda_{ij}\}] = E[\{\phi_i\}] + \sum_{i \leq j} \lambda_{ij} \left(\delta_{ij} - \langle \phi_i | \phi_j \rangle \right),$$

where the $\{\lambda_{ij}\}$ are the Lagrange parameters insuring that the $\{\phi_n\}$ are kept orthonormalized in the minimization process. The minimization of $\Omega[\{\phi_i, \lambda_{ij}\}]$ with respect to some $\phi_i^*(\mathbf{r})$ leads to

$$\left(-\frac{\nabla^2}{2} + v^{\text{eff}}(\mathbf{r}) \right) \phi_i(\mathbf{r}) = \sum_j \lambda_{ij} \phi_j(\mathbf{r}) \quad \text{with} \quad v^{\text{eff}}(\mathbf{r}) = v^{\text{ion}}(\mathbf{r}) + V^H(\rho(\mathbf{r})) + V^{\text{XC}}(\rho(\mathbf{r})),$$

using, e.g., the following chain rule for density-dependent potentials

$$\frac{\partial}{\partial \phi_i^*(\mathbf{r})} = \frac{\partial}{\partial \rho(\mathbf{r})} \times \frac{\partial \rho(\mathbf{r})}{\partial \phi_i^*(\mathbf{r})} = \frac{\partial}{\partial \rho(\mathbf{r})} \times \phi_i(\mathbf{r}).$$

A unitary rotation that diagonalizes the λ_{ij} matrix allows to recover the standard Kohn-Sham eigenvalue equation postulated in Section 3.4. However, we understand here that the Kohn-Sham eigenvalues are just Lagrange multipliers and their relation with addition/removal energies, as defined in a photoemission experiment, is far from clear! As another indication of the difficulty in identifying Kohn-Sham eigenvalues with total-energy differences between the neutral and charged systems, let's consider now the sum of Kohn-Sham eigenvalues over occupied states, namely

$$\sum_{i=1}^N \varepsilon_n = \sum_{i=1}^N \langle \phi_i | -\nabla^2/2 + v^{\text{eff}}(\mathbf{r}) | \phi_i \rangle = T_0 + \int d\mathbf{r} \left(v^{\text{ion}}(\mathbf{r}) + V^H(\mathbf{r}) + V^{\text{XC}}(\mathbf{r}) \right) \rho(\mathbf{r})$$

to be compared to the ground-state total energy

$$E[N] = T_0 + \int d\mathbf{r} v^{\text{ion}}(\mathbf{r}) \rho(\mathbf{r}) + J[\rho] + E^{\text{XC}}[\rho] \quad \text{with} \quad J[\rho] = \frac{1}{2} \int d\mathbf{r} V^H(\mathbf{r}) \rho(\mathbf{r}).$$

As a result, the ground-state total energy for, e.g., the N -electron system reads also

$$E[N] = \sum_{i=1}^N \varepsilon_n - J[\rho] + E^{\text{XC}}[\rho] - \int d\mathbf{r} V^{\text{XC}}(\mathbf{r}) \rho(\mathbf{r}) \quad \text{with} \quad V^{\text{XC}}(\mathbf{r}) = \frac{\partial E^{\text{XC}}[\rho]}{\partial \rho(\mathbf{r})},$$

where the sum of occupied level energies are completed by electron interaction terms. As such, differences of total energies such as $E[N+1] - E[N]$ cannot be identified simply to individual Kohn-Sham eigenstates and the meaning of Kohn-Sham eigenstates remains elusive.

5.4 Janak theorem and fractional occupations in ensemble DFT

Janak theorem [35] is an important relation that may pave the way to progress in the use of DFT to tackle not only ground-state total energies but also electronic excitations. The seminal idea is a generalization of the Kohn-Sham formalism to *fractional occupation numbers* of the electronic energy levels $\{\varepsilon_n\}$. In the standard Kohn-Sham approach, occupation numbers were set to unity below the Fermi level and zero above (zero temperature). In the generalized fractional approach, the kinetic energy and electronic density read

$$T_J[\rho] = \sum_i n_i \langle \phi_i | -\nabla^2/2 | \phi_i \rangle \quad \text{and} \quad \rho(\mathbf{r}) = \sum_i n_i |\phi_i(\mathbf{r})|^2,$$

where the index J stands for Janak. As such, the total energy is not only a functional of the one-body orbitals $\{\phi_n\}$, but also of the fractional occupations $\{n_i\}$ as additional variational parameters, namely

$$E_{GS} = \min_{\phi_i, n_i} E[\{n_i, \phi_i\}] \quad \text{with} \quad E = T_J[\{n_i, \phi_i\}] + \int d\mathbf{r} v^{ion}(\mathbf{r}) \rho(\mathbf{r}) + J[\rho] + E^{XC}[\rho],$$

where $\rho(\mathbf{r}) = \sum_i n_i |\phi_i(\mathbf{r})|^2$. The minimization with respect to the occupation factors leads to the Janak formula

$$\varepsilon_i = \frac{\partial E}{\partial n_i} \quad (\text{Janak formula})$$

telling us that one-body eigenvalues are related to the variation of the total energy with respect to an infinitesimal variation of the population. Since the Kohn-Sham formalism is a reduction of the Janak approach to $n_i = 1$ or $n_i = 0$ for occupied/empty levels, the Kohn-Sham eigenvalues can be interpreted as derivatives of the total energy at occupation numbers taken to be 0 or 1 namely

$$\varepsilon_i^{KS} = \left(\frac{\partial E}{\partial n_i} \right)_{n_i=0 \text{ or } 1}.$$

This differs from the experimental definition that electronic energy levels are variations of the total energy with respect to a unity (not infinitesimal) change of level population.

Apart from the technicalities of the derivation, a central question is related to the meaning of fractional occupations and, more generally, fractional number of electrons. The two concepts are not equivalent. One may consider a situation where fractional occupations are introduced while keeping the number of electrons fixed to N , namely $\sum_i n_i = N$. This is a very natural situation at finite temperature with the Fermi-Dirac distribution. Even at zero temperature, this is just reminiscent of Fermi liquid theory revealing that particle interactions lead to non-integer occupation number close to the Fermi level. On simpler grounds, it is also a useful exercise to consider the 1st-order density matrix

$$\gamma_1(\mathbf{r}, \mathbf{r}') = N \int \cdots \int d\mathbf{r}_2 \cdots d\mathbf{r}_N \psi(\mathbf{r}, \mathbf{r}_2, \cdots, \mathbf{r}_N) \psi^*(\mathbf{r}', \mathbf{r}_2, \cdots, \mathbf{r}_N)$$

for which it can be demonstrated that the expression of T and ρ as a function of fractional occupation is exact with the $\{n_i\}$ and $\{\phi_i\}$ the eigenvalues and eigenstates of γ_1 . These eigenstates are called “natural orbitals” and the eigenvalues fulfill $\sum_i n_i = N$.

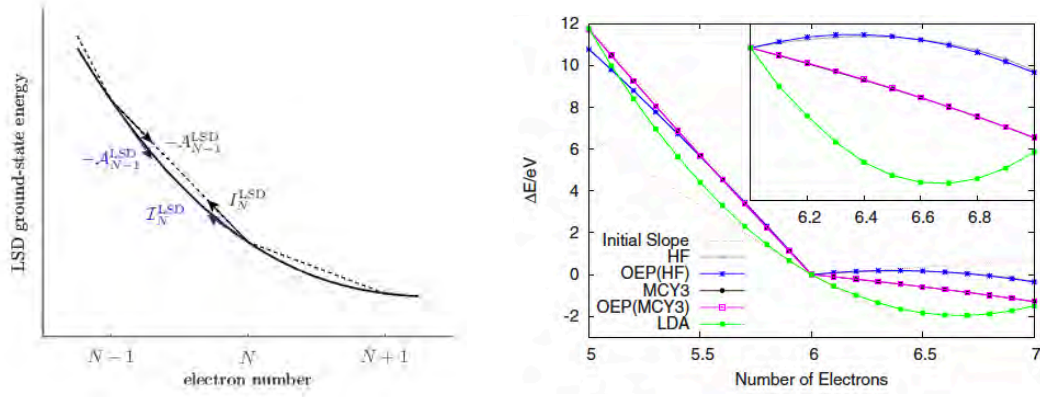


Fig. 5: (Left) Symbolic representation of the exact total energy (dotted lines) and DFT total energy (full line) as a function of the (continuous) electron number. The arrows indicate the right/left energy derivatives at integer values of the number of electrons in relation to the electronic affinity of the $N-1$ electron system (A_{N-1}^{LSD}) and the ionization potential (I_N^{LSD}) of the N -electron system. In the DFT Kohn-Sham approach, A_{N+1}^{LSD} and I_N^{LSD} are taken as the opposite of the LUMO energy of the $N-1$ -electron system and HOMO energy of the N -electron system (from Ref. [30]). (Right) Difference energy of the carbon atom $E = E(N_0 + \delta N) - E(N_0)$ ($N_0 = 6$) with several different functionals using OEP (optimized effective potentials) and GKS (generalized Kohn-Sham). Dotted line follows the initial slope for the non-straight functionals. The inset shows the range $6 < N < 7$ in more detail (from Ref. [39]).

Janak's theorem considers, however, the second situation where the total number of electrons is fractional, namely N becomes a continuous variable. As such, while photoemission measures how much the total energy changes upon the removal/addition of a full electron, Janak's theorem provides a relation for the removal/addition of an infinitesimal charge. Such a fractional number of electrons can be rationalized on the basis of the grand-canonical ensemble, namely when the system of interest can exchange electrons with a “bath.” The fractional charge can then be associated with a fractional probability of finding the charge in the (sub)system of interest. In quantum mechanics, this can be described by mixed states, that is a statistical ensemble of pure states. This is the basis for “ensemble DFT” [36, 37] that considers ensemble densities

$$\rho(\mathbf{r}) = (1 - \omega)\langle\psi_1|\hat{\rho}(\mathbf{r})|\psi_1\rangle + \omega\langle\psi_2|\hat{\rho}(\mathbf{r})|\psi_2\rangle,$$

where $|\psi_1\rangle$ and $|\psi_2\rangle$ are distinct many-body eigenstates, e.g., the ground-state and first-excited state of the N -electron system, or the ground-states of the N - and $N+1$ -electrons systems. In the first case, ensemble DFT strives to build a DFT approach to neutral excitations (e.g. optical excitations); in the second one, charged excitations (photoemission) are targeted.

We now summarize the main results associated with the variation of the total ground-state energy with respect to a continuous number of electrons. A first important result is that the total energy should be concave and piecewise linear between two integer values of N [38]. This is represented on the left of Fig. 5 as the dotted lines. We observe in particular that the derivative of the total energy with respect to the number of electrons is discontinuous across an integer value: the left and right derivatives are not identical. The piece-wise linearity on each side of an integer value means in particular that the fractional derivatives and the corresponding differences of total energy between integer N -values are identical.

Let's now turn to pure DFT with XC functionals of the density. It can be shown and confirmed by actual calculations that within DFT, the total energy is a concave and smooth function of the number of electrons (see Fig. 5 (left) full line and Fig. 5 (right) green line). In particular the left and right derivatives at integer values of N are identical and not equal to differences of total energy for integer variations of N . One recovers, in particular, the problem of the lack of discontinuity of the XC potential. As such, following the Janak theorem, the Kohn-Sham eigenvalues cannot be identified with electronic energy levels as measured by photoemission. On the contrary, it is found that the Hartree-Fock energy is a convex function of N with discontinuous derivatives at integer values (blue line Fig. 5 (right)). As such, mixing some DFT local functionals with some amount of exact exchange may result in a close-to-straight-line dependence of the total energy between integer values of N , which is just the condition we need for matching infinitesimal derivatives with total energy differences upon adding/removing an electron. This is yet another rationale for using hybrid functionals. However, which amount of mixing should be used for a given system, is again a difficult question to answer if we do not accept fitting strategies to known experimental data. See, e.g., Ref. [29,30] for a mathematically based strategy avoiding empirically adjusted parameters.

Here again, the hunt for a generalized DFT formalism able to tackle excited properties is an on-going boiling activity. In the case of weakly to moderately correlated systems, rather simple many-body perturbation theories such as the GW formalism, with a limited $\mathcal{O}(N^4)$ scaling with system size, is getting very popular in condensed matter physics and, more recently, quantum chemistry (see Figs. 3 and 4). We emphasize however that even if adopting alternative approaches to DFT, Kohn-Sham $\{\varepsilon_n^{KS}, \phi_n^{KS}\}$ eigenvectors remain very valuable zeroth-order one-body eigenstates to build higher-order correlation operators. In particular, GW calculations start generally with the knowledge of the time-ordered Green's function

$$G(\mathbf{r}, \mathbf{r}'; \omega) = \sum_n \frac{\phi_n^{KS}(\mathbf{r}) \phi_n^{KS,*}(\mathbf{r}')}{\omega - \varepsilon_n^{KS} + i \times \text{sign}(\varepsilon_n^{KS} - E_F) \times 0^+}$$

built with KS eigenstates and of the (RPA) screened Coulomb potential $W = V + V\chi_0 W$ and independent-electron susceptibility χ_0 relying again on Kohn-Sham eigenstates

$$\chi^0(\mathbf{r}, \mathbf{r}'; \omega) = \sum_{nm} (f_n - f_m) \frac{\phi_m^{KS}(\mathbf{r}) \phi_n^{KS,*}(\mathbf{r}) \phi_n^{KS}(\mathbf{r}') \phi_m^{KS,*}(\mathbf{r}')}{\omega - (\varepsilon_n^{KS} - \varepsilon_m^{KS}) + i0^+}$$

with $(f_{n/m})$ occupation factors. Even though DFT Kohn-Sham eigenstates do not represent here the final quantities that will be used to interpret the experimental data, they remain very valuable, representing affordable starting piece of information on the electronic properties.

References

- [1] R.M. Dreizler and E.K.U. Gross: *Density-Functional Theory* (Springer, Berlin, Heidelberg, 1990)
- [2] R.G. Parr and Y. Weitao: *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, 1994)
- [3] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964)
- [4] M. Levy, Proc. Natl. Acad. Sci. USA **76**, 6062 (1979)
- [5] W. Kohn and L.J. Sham, Phys. Rev. **140**, A1133 (1965)
- [6] D.M. Ceperley and B.J. Alder, Phys. Rev. Lett. **45**, 566 (1980)
- [7] P.A.M. Dirac, Proc. Camb. Phil. Soc. **26**, 376 (1930)
- [8] S.H. Vosko, L. Wilk, and M. Nusair, Can. J. Phys. **58**, 1200 (1980)
- [9] J.P. Perdew and Alex Zunger, Phys. Rev. B **23**, 5048 (1981)
- [10] J.P. Perdew, A. Ruzsinszky, G.I. Csonka, O.A. Vydrov, G.E. Scuseria, L.A. Constantin, X. Zhou, and K. Burke, Phys. Rev. Lett. **100**, 136406 (2008)
- [11] J. Tao, J.P. Perdew, V.N. Staroverov, and G.E. Scuseria, Phys. Rev. Lett. **91**, 146401 (2003)
- [12] G.I. Csonka *et al.*, Phys. Rev. B **79**, 155107 (2009)
- [13] P. Giannozzi, S. de Gironcoli, P. Pavone, and S. Baroni, Phys. Rev. B **43**, 7231 (1991)
- [14] F. Bruneval, N. Vast, and L. Reining, Phys. Rev. B **74**, 045102 (2006)
- [15] P.R.C. Kent, R.Q. Hood, M.D. Towler, R.J. Needs, and G. Rajagopal, Phys. Rev. B **57**, 15293 (1998)
- [16] O. Gunnarsson and B.I. Lundqvist, Phys. Rev. B **13**, 4274 (1976)
- [17] M.G. Medvedev, I.S. Bushmarinov, J. Sun, J.P. Perdew, and K.A. Lyssenko, Science **355**, 49 (2017)
- [18] A.D. Becke, Phys. Rev. A **38**, 3098 (1988)
- [19] J.P. Perdew, A. Ruzsinszky, and J. Tao, J. Chem. Phys. **123**, 062201 (2005)
- [20] J.P. Perdew *et al.*, Phys. Rev. B, **46**, 6671 (1992)
- [21] J.P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996)
- [22] C. Lee, W. Yang, and R.G. Parr, Phys. Rev. B **37**, 785 (1988)

- [23] A.D. Becke, J. Chem. Phys. **98**, 5648 (1993)
- [24] C. Adamo and V. Barone, J. Chem. Phys. **110**, 6158 (1999)
- [25] T. Leininger, H. Stoll, H.-J. Werner, and A. Savin, Chem. Phys. Lett. **275**, 151 (1997)
- [26] H. Iikura, T. Tsuneda, T. Yanai, and K. Hirao, J. Chem. Phys. **115**, 3540 (2001)
- [27] J. Heyd and Gustavo E. Scuseria, J. Chem. Phys. **118**, 8207 (2003)
- [28] M. Rohlfing, P. Krüger, and J. Pollmann, Phys. Rev. B **48**, 17791 (1993)
- [29] S. Refaely-Abramson, R. Baer, and L. Kronik, Phys. Rev. B **84**, 075144 (2011)
- [30] I. Dabo, A. Ferretti, N. Poilvert, Y. Li, N. Marzari, and M. Cococcioni, Phys. Rev. B **82**, 115121 (2010)
- [31] C. Faber, I. Duchemin, T. Deutsch, C. Attaccalite, V. Olevano, and X. Blase, J. Matter. Sci. **47**, 7472 (2012)
- [32] C. Faber, C. Attaccalite, V. Olevano, E. Runge, X. Blase, Phys. Rev. B **83**, 115123 (2011)
- [33] J.P. Perdew and M. Levy, Phys. Rev. Lett. **51**, 1884 (1983)
- [34] L.J. Sham and M. Schlüter, Phys. Rev. Lett. **51**, 1888 (1983)
- [35] J.F. Janak, Phys. Rev. B **18**, 7165 (1978)
- [36] A.K. Theophilou, J. Phys. C **12**, 5419 (1978)
- [37] E.K.U. Gross, L.N. Oliveira, and W. Kohn, Phys. Rev. A **37**, 2809 (1988)
- [38] J.P. Perdew, R.G. Parr, M. Levy, and J.L. Balduz, Phys. Rev. Lett. **49**, 1691 (1982)
- [39] A.J. Cohen, P. Mori-Sánchez, and W. Yang, Phys. Rev. B **77**, 115123 (2008)

2 The Random Phase Approximation and its Applications to Real Materials

Xinguo Ren

CAS Key Laboratory of Quantum Information

University of Science and Technology of China

Jinzhai Road 96, Hefei

Contents

1	Introduction	2
2	Theoretical foundation of RPA and beyond	3
2.1	Adiabatic connection approach to DFT	3
2.2	RPA derived from the ACFDT framework	6
2.3	Beyond-RPA computational schemes	9
3	Implementation of the RPA method	12
4	The performance of RPA-based methods and prototypical applications	15
4.1	Van der Waals interactions	15
4.2	Surface adsorption	18
4.3	Structural phase transitions	19
5	Recent developments	21
6	Summary	22

1 Introduction

The random phase approximation (RPA) is a fundamental concept that plays a central role in many-body physics. The concept was developed by Bohm and Pines [1,2] in the early 1950's in an endeavor to describe the cohesive properties of the so-called jellium—interacting electrons moving in a background of a uniform positive charge. Using a Hamiltonian formulation of interacting many-electron system, Bohm and Pines were able to decouple the collective motion of electrons—the plasma oscillations—from their individual motions, a procedure named as RPA. It was soon recognized [3] that the original RPA formulation is equivalent to the infinite summation of ring diagrams from the viewpoint of diagrammatic many-body perturbation theory [4]. Since then, the RPA concept has gone beyond the realm of condensed matter physics and significantly influenced all branches of physics.

Although the RPA concept can be applied to any interacting many-particle systems (for its applications in nuclear physics, see, e.g., Ref. [5]), in this review we shall restrict ourselves to electronic systems which are governed by Coulomb interactions. The next key step towards applying RPA to real materials was the incorporation of RPA into the Kohn-Sham (KS) density-functional theory (DFT) framework in 1970's [6–8]. This formulation turned RPA into a first-principles electronic-structure method, suitable for computing the ground-state energy of real materials. Within the KS-DFT framework, RPA can be viewed as a fifth-rung approximation to the exchange-correlation (XC) energy functional, according to a well-accepted classification scheme of the XC functionals, known as Jacob's ladder of DFT as proposed by Perdew [9]. However, the application of RPA to realistic systems was impeded by its high computational cost and the lack of efficient algorithms at the time. The first application of RPA to small molecules only appeared in early 2000's, carried out by Furche [10] and in [11]. Since then, RPA has been applied to a variety of systems including atoms [12,13], molecules [10,11,14–18], solids [19–26], surfaces [27,28], interfaces [29,30], layered materials [31], and defects [32,33]. The consensus arising from these studies is that RPA is capable of describing the delicate energy differences in complex chemical environments [27,28,24,26], the correct asymptotic behavior of van der Waals (vdW) complexes [34,35] and layered materials [31,36], and the correct dissociation limit of closed-shell molecules [10,11]. Evidence shows that RPA can provide unprecedented accuracy compared to lower-rung density-functional approximations at tractable computational cost. As such, RPA is expected to play an increasingly more important role in computational materials science, with the rapid development of more efficient algorithms and the availability of more powerful computing resources.

In a review paper [37] published in 2012, we discussed the history of the RPA concept, its formulation as a first-principles method, and its applications in quantum chemistry and computational materials science up to that time. These points were nicely summarized by David Pines in his recent review paper titled as “*Emergent behavior in strongly correlated electron systems*” [38]. In particular, Pines noted that,

Sixty-plus years later, the RPA continues to play a significant role in nuclear physics, bosonic field-theory, the quarkgluon plasma, many-fermion solvable models, and

especially in computational chemistry and materials science. A recent review by Ren et al., to which the interested reader is referred, describes the impact of the RPA in the theoretical chemistry and materials science community, cites some thirty articles that indicate the renewed and widespread interest in the RPA during the period 2001-2011, discusses how it enables one to derive the $1/r^6$ interaction between spatially separated closed shell electron systems, and, shows, in some detail, how the RPA enables one to go beyond density-functional theory in computing ground state energies.

This highlights the far-reaching impact of the RPA in a variety of fields, in particular in computational chemistry and materials science. It is worthwhile to mention that several other, complementary RPA review papers also appeared around that time, where the theoretical foundation and applications of RPA are discussed from different perspective. More recent account of RPA of review character can be found in Ref. [39].

In this lecture, we first give an account of the theoretical foundation of RPA, highlighting its unique role in electronic structure theory. Computational schemes beyond RPA are also briefly discussed in this section. This is followed by a sketching of the key algorithm of implementing RPA using the resolution-of-identity technique. We then present some prototypical applications illustrating the usefulness of RPA in computational materials science. Most recent efforts devoted to further developing the theoretical and computational aspects of RPA as well as extending its capabilities are briefly mentioned and commented, before we conclude.

2 Theoretical foundation of RPA and beyond

RPA as a first-principles method can be derived from several theoretical frameworks. One convenient framework to derive RPA is the adiabatic-connection fluctuation-dissipation theorem (ACFDT) which offers a powerful mathematical device to construct the exact XC energy functional via the density response functions of a series of partially interacting systems, connecting the KS system and the true physical system. In this formulation, an approximation to the density response function translates into a corresponding approximation to the XC energy functional. However, this is not the only approach to formulate RPA. In fact, RPA can also be derived from other theoretical perspectives, including coupled cluster theory [40, 41], the Green-function based many-body perturbation theory [42, 37, 43], and time-dependent DFT [10]. Below we focus on the ACFDT perspective.

2.1 Adiabatic connection approach to DFT

An interacting N -electron system is described by the following Hamiltonian

$$\hat{H} = \hat{T} + \hat{V}_{\text{ext}} + \hat{V}_{\text{ee}} = - \sum_i^N \frac{\nabla_i^2}{2} + \sum_i^N v^{\text{ext}}(\hat{\mathbf{r}}_i) + \frac{1}{2} \sum_{i \neq j}^N \frac{1}{|\hat{\mathbf{r}}_i - \hat{\mathbf{r}}_j|}, \quad (1)$$

where $\hat{T}$, $\hat{V}_{\text{ext}}$, and $\hat{V}_{\text{ee}}$ are, respectively, the kinetic energy, the external potential, and the electron-electron Coulomb interaction operators. Hartree atomic units ($\hbar = e = m_e = 1$) are used throughout this lecture note. Note that the operators $\hat{T}$ and $\hat{V}_{\text{ee}}$ are universal for any N -electron systems, and a system is completely specified by the external potential,

$$v^{\text{ext}}(\mathbf{r}) = - \sum_{\alpha} \frac{Z_{\alpha}}{|\mathbf{R}_{\alpha} - \mathbf{r}|}, \quad (2)$$

with Z_{α} and $\mathbf{R}_{\alpha}$ being, respectively, the nuclear charges and positions of the atoms in the system. The Hamiltonian in (1) cannot be solved for more than a few electrons. To deal with it, it is customary to separate $\hat{H}$ into two parts,

$$\hat{H} = \hat{H}_0 + \hat{H}' \quad (3)$$

where $\hat{H}_0$ is a *mean-field* (MF) Hamiltonian, describing a collection of noninteracting particles subjected to a self-consistently determined effective single-particle potential,

$$\hat{H}_0 = \sum_i^N \left(-\frac{\nabla_i^2}{2} + v^{\text{ext}}(\mathbf{r}_i) + v_i^{\text{MF}} \right). \quad (4)$$

Here the mean-field potential v^{MF} is an average potential that one electron experiences due to the presence of other electrons. Different underlying principles to determine v^{MF} lead to different self-consistency schemes. Suppose Φ_0 is the ground-state wave function (a Slater determinant) of $\hat{H}_0$, within Hartree-Fock theory, the effective single-particle potential v^{MF} is chosen such that the expectation value of $\langle \Phi_0 | \hat{H} | \Phi_0 \rangle$ is minimized. On the other hand, within KS-DFT, the effective v^{MF} (now the KS potential v^{KS}) is determined so that the electron density of the KS system $n(\mathbf{r}) = \langle \Phi_0 | \hat{n}(\mathbf{r}) | \Phi_0 \rangle$ reproduces the density $n(\mathbf{r})$ of the true physical system. Within the adiabatic connection (AC) approach to KS-DFT, one considers a continuous set of fictitious Hamiltonians

$$\hat{H}(\lambda) = - \sum_i^N \frac{\nabla_i^2}{2} + \sum_i^N v_{\lambda}^{\text{aux}}(\mathbf{r}_i) + \frac{\lambda}{2} \sum_{i \neq j}^N \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \quad (5)$$

which connects $\hat{H}_0$ at $\lambda = 0$ (with $v_{\lambda=0}^{\text{aux}} = v^{\text{ext}} + v^{\text{KS}}$) and $\hat{H}$ at $\lambda = 1$ (with $v_{\lambda=1}^{\text{aux}} = v^{\text{ext}}$). The Hamiltonian $\hat{H}(\lambda)$ for $0 < \lambda < 1$ describes a collection of particles moving under the auxiliary external potential $v_{\lambda}^{\text{aux}}(\mathbf{r})$ and interacting with a scaled Coulomb interaction $\lambda/|\mathbf{r} - \mathbf{r}'|$. The auxiliary potential $v_{\lambda}^{\text{aux}}(\mathbf{r})$ ($0 < \lambda < 1$) is chosen such that the density of λ -scaled systems is kept at the physical density, i.e., $n_{\lambda}(\mathbf{r}) = n(\mathbf{r})$, along the AC path. Denoting the ground-state wavefunction of $\hat{H}(\lambda)$ as $|\Psi_{\lambda}\rangle$,

$$\hat{H}(\lambda)|\Psi_{\lambda}\rangle = E_{\lambda}|\Psi_{\lambda}\rangle, \quad (6)$$

and adopting the normalization condition $\langle \Psi_{\lambda} | \Psi_{\lambda} \rangle = 1$, the Hellmann-Feynman theorem implies

that

$$\begin{aligned} \frac{dE_\lambda}{d\lambda} &= \langle \Psi_\lambda | \frac{\partial \hat{H}_\lambda}{\partial \lambda} | \Psi_\lambda \rangle = \left\langle \Psi_\lambda \left| \sum_{i=1}^N \frac{\partial v_\lambda^{\text{aux}}(\hat{\mathbf{r}}_i)}{\partial \lambda} + \frac{1}{2} \sum_{i \neq j}^N \frac{1}{|\hat{\mathbf{r}}_i - \hat{\mathbf{r}}_j|} \right| \Psi_\lambda \right\rangle \\ &= \int d\mathbf{r} \langle \Psi_\lambda | \hat{n}(\mathbf{r}) \frac{\partial v_\lambda^{\text{aux}}(\mathbf{r})}{\partial \lambda} | \Psi_\lambda \rangle + \frac{1}{2} \iint d\mathbf{r} d\mathbf{r}' \frac{\langle \Psi_\lambda | \hat{n}(\mathbf{r}) \hat{n}(\mathbf{r}') - \hat{n}(\mathbf{r}) \delta(\mathbf{r} - \mathbf{r}') | \Psi_\lambda \rangle}{|\mathbf{r} - \mathbf{r}'|} \end{aligned} \quad (7)$$

$$= \int d\mathbf{r} n(\mathbf{r}) \frac{\partial v_\lambda^{\text{aux}}(\mathbf{r})}{\partial \lambda} + \frac{1}{2} \iint d\mathbf{r} d\mathbf{r}' \frac{\langle \Psi_\lambda | \hat{n}(\mathbf{r}) \hat{n}(\mathbf{r}') | \Psi_\lambda \rangle - n(\mathbf{r}) \delta(\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \quad (8)$$

To derive (7) and (8), we have used the expression for the density operator of the N -electron systems

$$\hat{n}(\mathbf{r}) = \sum_{i=1}^N \delta(\mathbf{r} - \hat{\mathbf{r}}_i), \quad (9)$$

and the condition $\langle \Psi_\lambda | \hat{n}(\mathbf{r}) | \Psi_\lambda \rangle = n(\mathbf{r})$. The ground-state energy of the interacting system can then be obtained as

$$\begin{aligned} E &= E_0 + \int_0^1 \frac{dE_\lambda}{d\lambda} d\lambda \\ &= E_0 + \int d\mathbf{r} n(\mathbf{r}) (v_{\lambda=1}^{\text{aux}}(\mathbf{r}) - v_{\lambda=0}^{\text{aux}}(\mathbf{r})) + \int_0^1 d\lambda \frac{1}{2} \iint d\mathbf{r} d\mathbf{r}' \frac{\langle \Psi_\lambda | \hat{n}(\mathbf{r}) \hat{n}(\mathbf{r}') | \Psi_\lambda \rangle - n(\mathbf{r}) \delta(\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \\ &= E_0 - \int d\mathbf{r} n(\mathbf{r}) v^{\text{KS}}(\mathbf{r}) + \frac{1}{2} \iint d\mathbf{r} d\mathbf{r}' \frac{n(\mathbf{r}) n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \\ &\quad + \int_0^1 d\lambda \frac{1}{2} \iint d\mathbf{r} d\mathbf{r}' \frac{\langle \Psi_\lambda | \delta \hat{n}(\mathbf{r}) \delta \hat{n}(\mathbf{r}') | \Psi_\lambda \rangle - n(\mathbf{r}) \delta(\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}, \end{aligned} \quad (10)$$

where we have introduced the fluctuation of the density operator $\delta \hat{n}(\mathbf{r}) = \hat{n}(\mathbf{r}) - n(\mathbf{r})$ and used the fact that $\langle \Psi_\lambda | \delta \hat{n}(\mathbf{r}) | \Psi_\lambda \rangle = 0$. Now the ground-state energy of the reference KS state ($\lambda = 0$) E_0 is given by

$$E_0 = \langle \Psi_0 | \hat{H}_0 | \Psi_0 \rangle = \langle \Phi_0 | \hat{H}_0 | \Phi_0 \rangle = - \sum_{n=1}^N \langle \psi_n | \frac{\nabla^2}{2} | \psi_n \rangle + \int d\mathbf{r} v^{\text{ext}}(\mathbf{r}) n(\mathbf{r}) + \int d\mathbf{r} v^{\text{KS}}(\mathbf{r}) n(\mathbf{r}). \quad (11)$$

Combining (10) and (11), one arrives at the formal expression of the exact ground-state total energy of N -electron systems,

$$\begin{aligned} E &= - \sum_{n=1}^N \langle \psi_n | \frac{\nabla^2}{2} | \psi_n \rangle + \int d\mathbf{r} v^{\text{ext}}(\mathbf{r}) n(\mathbf{r}) + \frac{1}{2} \iint d\mathbf{r} d\mathbf{r}' \frac{n(\mathbf{r}) n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \\ &\quad + \int_0^1 d\lambda \frac{1}{2} \iint d\mathbf{r} d\mathbf{r}' \frac{\langle \Psi_\lambda | \delta \hat{n}(\mathbf{r}) \delta \hat{n}(\mathbf{r}') | \Psi_\lambda \rangle - n(\mathbf{r}) \delta(\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}. \end{aligned} \quad (12)$$

In KS-DFT, the ground-state total energy for an interacting N -electron systems is an (implicit) functional of the electron density $n(\mathbf{r})$ and can be conveniently separated into four terms:

$$E[n] = T_s[\psi_n] + E_{\text{ext}}[n] + E_{\text{H}}[n] + E_{\text{xc}}[n], \quad (13)$$

where

$$T_s = - \sum_{n=1}^N \langle \psi_n | \frac{\nabla^2}{2} | \psi_n \rangle \quad (14)$$

is the kinetic energy of the KS independent-particle system,

$$E_{\text{ext}}[n] = \int d\mathbf{r} v^{\text{ext}}(\mathbf{r}) n(\mathbf{r}) \quad (15)$$

is the external potential energy due to the nuclei,

$$E_{\text{H}}[n] = \frac{1}{2} \iint d\mathbf{r} d\mathbf{r}' \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \quad (16)$$

is the classic Hartree energy, and E_{xc} the exchange-correlation energy. Among the four terms in Eq. (13), only $E_{\text{ext}}[n]$ and $E_{\text{H}}[n]$ are explicit functionals of $n(\mathbf{r})$. The noninteracting kinetic energy T_s is treated exactly in KS-DFT in terms of the single-particle KS orbitals $\psi_n(\mathbf{r})$, which themselves are a functional of $n(\mathbf{r})$. All the many-body complexity is contained in the unknown XC energy term, whose exact form is not an explicit functional of the electron density $n(\mathbf{r})$, nor the KS orbitals $\psi_n(\mathbf{r})$. Comparing Eq. (13) to (12), one immediately obtains the formally exact AC expression for the XC energy, in terms of the density-density correlation function,

$$E_{\text{xc}}[n] = \int_0^1 d\lambda \frac{1}{2} \iint d\mathbf{r} d\mathbf{r}' \frac{\langle \Psi_\lambda | \delta \hat{n}(\mathbf{r}) \delta \hat{n}(\mathbf{r}') | \Psi_\lambda \rangle - n(\mathbf{r})\delta(\mathbf{r}-\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}. \quad (17)$$

A central task of the DFT community is to develop accurate and tractable approximations to $E_{\text{xc}}[n]$. The success of KS-DFT lies in the fact that usefully accurate approximations can be found, for which DFT calculations can be done at affordable cost. Widely used approximations to $E_{\text{xc}}[n]$ include local-density approximation (LDA) [44], generalized gradient approximations (GGAs) [45–47], meta-GGA [48,49], and hybrid functional approximations [50,51]. These approximations belong to the first four rungs of the Jacob’s ladder [9]. Despite their enormous success, these existing approximate functionals suffer from many-electron self-interaction errors [52], or delocalization errors [53] due to which the localized electronic states tend to delocalize over the system, leading to several severe consequences. Furthermore, the non-local correlation effects between widely separated subsystems are not captured within these approximations by constructions. These intrinsic deficiencies limit the possible accuracy that can be achieved in practical calculations. Qualitative and sometimes quantitative failures have been documented in a number of situations, among which the most prominent are van der Waals (vdW) bonded systems [54], materials with strong correlations [55], certain surface adsorption problems [56], and chemical reaction barrier heights [57].

2.2 RPA derived from the ACFDT framework

The construction of LDA, GGA, and meta-GGA functionals only incorporate local and semi-local quantities, such as the electron density, density gradients, and the kinetic energy density. For the construction of hybrid functionals, a non-local quantity—the reduced density matrix—is

required. To go beyond these approximations, it is necessary to include additional information in the functional construction. For this purpose, it is instrumental to invoke the zero-temperature fluctuation-dissipation theorem [58], which relates the density-density correlation (*fluctuations*) in Eq. (17) to the imaginary part of the density response function (*dissipation*) of the system

$$\langle \Psi_\lambda | \delta \hat{n}(\mathbf{r}) \delta \hat{n}(\mathbf{r}') | \Psi_\lambda \rangle = -\frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \operatorname{Im} \chi^\lambda(\mathbf{r}, \mathbf{r}', \omega). \quad (18)$$

Here the density response function $\chi^\lambda(\mathbf{r}, \mathbf{r}', \omega) = \delta n_\lambda(\mathbf{r}, \omega) / \delta v^{ext}(\mathbf{r}', \omega)$ describes the variation of the density of the partially interacting system at the spatial point $\mathbf{r}$, up to linear order, due to a change of the local external potential at $\mathbf{r}'$. From Eqs. (17) and (18), we arrive at the renowned ACFDT expression for the XC energy in DFT

$$\begin{aligned} E_{xc} &= \frac{1}{2} \int_0^1 d\lambda \iint d\mathbf{r} d\mathbf{r}' \frac{1}{|\mathbf{r}-\mathbf{r}'|} \left(-\frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \operatorname{Im} \chi^\lambda(\mathbf{r}, \mathbf{r}', \omega) - \delta(\mathbf{r}-\mathbf{r}') n(\mathbf{r}) \right) \\ &= \frac{1}{2} \int_0^1 d\lambda \iint d\mathbf{r} d\mathbf{r}' \frac{1}{|\mathbf{r}-\mathbf{r}'|} \left(-\frac{1}{\pi} \int_0^{\infty} d\omega \chi^\lambda(\mathbf{r}, \mathbf{r}', i\omega) - \delta(\mathbf{r}-\mathbf{r}') n(\mathbf{r}) \right). \end{aligned} \quad (19)$$

The fact that the above frequency integration can be performed along the imaginary axis originates from the analytical structure of $\chi^\lambda(\mathbf{r}, \mathbf{r}', \omega)$ and the fact that it becomes purely real on the imaginary axis. Such a property simplifies the ground-state energy calculation within the ACFDT framework considerably. The ACFDT expression in Eq. (19) transforms the problem of computing the XC energy to the computation of the response functions of a continuous set of fictitious systems along the AC path, which in practice has to be approximated as well.

The exact density response function $\chi^\lambda(\mathbf{r}, \mathbf{r}', i\omega)$ is not known either. However, according to linear-response time-dependent DFT (TDDFT), the interacting response function χ^λ for $\lambda > 0$ is linked to the noninteracting response function χ^0 via the Dyson-like equation

$$\chi^\lambda(\mathbf{r}, \mathbf{r}', i\omega) = \chi^0(\mathbf{r}, \mathbf{r}', i\omega) + \int d\mathbf{r}_1 d\mathbf{r}_2 \chi^0(\mathbf{r}, \mathbf{r}_1, i\omega) \left[\frac{\lambda}{|\mathbf{r}_1 - \mathbf{r}_2|} + f_{xc}^\lambda(\mathbf{r}_1, \mathbf{r}_2, i\omega) \right] \chi^\lambda(\mathbf{r}_2, \mathbf{r}', \omega), \quad (20)$$

where

$$f_{xc}(\mathbf{r}_1, \mathbf{r}_2, i\omega) = \frac{\delta v_{xc}(\mathbf{r}_1)}{\delta n(\mathbf{r}_2, \omega)} = \frac{\delta^2 E_{xc}[n]}{\delta n(\mathbf{r}_1) \delta n(\mathbf{r}_2, \omega)} \quad (21)$$

is the so-called f_{xc} kernel, given by the functional derivative of the XC potential with respect to the frequency-dependent density variation.

Obviously, the f_{xc} kernel is a very complex quantity to deal with, and there are considerable ongoing efforts to find better approximations for it [59]. In this context, the *random phase approximation* amounts to simply neglecting the f_{xc} kernel, and the resultant interacting response function is termed the RPA response function,

$$\chi_{\text{RPA}}^\lambda(\mathbf{r}, \mathbf{r}', i\omega) = \chi^0(\mathbf{r}, \mathbf{r}', i\omega) + \int d\mathbf{r}_1 d\mathbf{r}_2 \chi^0(\mathbf{r}, \mathbf{r}_1, i\omega) \frac{\lambda}{|\mathbf{r}-\mathbf{r}'|} \chi_{\text{RPA}}^\lambda(\mathbf{r}_2, \mathbf{r}', \omega). \quad (22)$$

In physical terms, the RPA here corresponds to the linearized time-dependent Hartree approximation, by which the variation of the exchange-correlation potential due to an external perturbation is neglected.

In Eqs. (20) and (22), $\chi^0(\mathbf{r}, \mathbf{r}_1, i\omega)$ is the independent-particle response function of the KS reference system ($\lambda = 0$) and is known explicitly in terms of the single-particle KS orbitals $\psi_n(\mathbf{r})$, orbital energies ε_n , and occupation factors f_n

$$\chi^0(\mathbf{r}, \mathbf{r}', i\omega) = \sum_{mn} \frac{(f_m - f_n) \psi_m^*(\mathbf{r}) \psi_n(\mathbf{r}) \psi_n^*(\mathbf{r}') \psi_m(\mathbf{r}')}{\varepsilon_m - \varepsilon_n - i\omega}. \quad (23)$$

Based on Eqs. (19) and (22), the XC energy in RPA can be further decomposed into an exact exchange (EX) term and a RPA correlation term,

$$E_{xc}^{\text{RPA}} = E_x^{\text{EX}} + E_c^{\text{RPA}}, \quad (24)$$

where

$$\begin{aligned} E_x^{\text{EX}} &= \frac{1}{2} \iint d\mathbf{r} d\mathbf{r}' \frac{1}{|\mathbf{r} - \mathbf{r}'|} \left(-\frac{1}{\pi} \int_0^\infty d\omega \chi^0(\mathbf{r}, \mathbf{r}', i\omega) - \delta(\mathbf{r} - \mathbf{r}') n(\mathbf{r}) \right) \\ &= -\sum_{mn} f_m f_n \iint d\mathbf{r} d\mathbf{r}' \frac{\psi_m^*(\mathbf{r}) \psi_n(\mathbf{r}) \psi_m^*(\mathbf{r}') \psi_n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \end{aligned} \quad (25)$$

and

$$\begin{aligned} E_c^{\text{RPA}} &= -\frac{1}{2\pi} \iint d\mathbf{r} d\mathbf{r}' \frac{1}{|\mathbf{r} - \mathbf{r}'|} \int_0^\infty d\omega \left(\int_0^1 d\lambda \chi_{\text{RPA}}^\lambda(\mathbf{r}, \mathbf{r}', i\omega) - \chi_0(\mathbf{r}, \mathbf{r}', i\omega) \right) \\ &= \frac{1}{2\pi} \int_0^\infty d\omega \text{Tr} \left(\ln(1 - \chi^0(i\omega)v) + \chi^0(i\omega)v \right), \end{aligned} \quad (26)$$

with $v(\mathbf{r}, \mathbf{r}') = 1/|\mathbf{r} - \mathbf{r}'|$. Equation (25) defines the exact-exchange energy in the KS-DFT context, which has the same expression as the Hartree-Fock exchange energy, but is extended here to fractional occupation numbers and to be evaluated with KS orbitals. Furthermore, in the second line of Eq. (26), for brevity the following convention

$$\text{Tr}[fg] = \iint d\mathbf{r} d\mathbf{r}' f(\mathbf{r}, \mathbf{r}') g(\mathbf{r}', \mathbf{r}) \quad (27)$$

has been adopted.

So far, we have described how the RPA method is defined as an approximate XC energy functional in the KS-DFT context, within the ACFDT framework. As mentioned above, RPA can also be derived from the perspective of coupled cluster theory and the Green-function based many-body perturbation theory. For instance, from the perspective of the ring coupled cluster doubles (rCCD) theory, the RPA correlation energy can be obtained as,

$$E_c^{\text{RPA}} = \frac{1}{2} \text{Tr}(BT^{\text{rCCD}}) = \frac{1}{2} \sum_{ij,ab} B_{ia,jb} T_{jb,ia}^{\text{rCCD}}, \quad (28)$$

where $T_{jb,ia}^{\text{rCCD}}$ is the rCCD amplitude, to be determined by solving the so-called Riccati equation [41], and $B_{ia,jb} = \langle ij|ab \rangle$ with $\langle ij|ab \rangle$ being two-electron Coulomb integrals,

$$\langle ij|ab \rangle = \iint d\mathbf{r} d\mathbf{r}' \frac{\psi_i^*(\mathbf{r}) \psi_a(\mathbf{r}) \psi_j^*(\mathbf{r}') \psi_b(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}. \quad (29)$$

Here and below, we adopt the convention that i, j, k refer to occupied KS orbitals, a, b, c to virtual (unoccupied) orbitals, and m, n to general ones. Due to limited space, here we will not elaborate on the alternative formulations of RPA any further, and interested readers are referred to Ref. [37] and the original references [41, 42] for more details.

2.3 Beyond-RPA computational schemes

Despite its appealing features, RPA does not go without shortcomings. The conventional wisdom is that RPA describes the long-range correlation very well, whereas it is not adequate for short-range correlations. This issue has been well known for the homogeneous electron gas. For real materials, RPA was found to underestimate the cohesive energies of both molecules [10, 60] and solids [21]. This stimulated much research interest and several beyond-RPA schemes have been developed to fix this deficiency. Among these, two approaches are particularly noteworthy: 1) improving RPA by restoring the contribution of the f_{xc} kernel within the ACFDT formalism; 2) improving RPA from the perspective of diagrammatic many-body perturbation theory. Successful beyond-RPA schemes of the first type include the truncated adiabatic LDA kernel correction [61], and the more systematic construction of the XC kernel with respect to a series expansion of the coupling constant λ within the ACFDT framework [62, 59].

The diagrammatic many-body expansion approach to correct RPA, on which we shall concentrate here, exploits the fact that RPA, in contrast with other density-functional approximations (DFAs), has a clear diagrammatic representation—an infinite summation of the ring diagrams. The question arises if there is a simple and systematic way to incorporate the missing diagrams to arrive at an improved theory. The fermionic nature of electrons requires the many-electron wavefunction to be antisymmetric, and the diagrammatic representation of the correlation energy contains graphs describing both *direct* processes and *exchange* processes. For example, in the Møller-Plesset perturbation theory, both types of processes are included at each order, which ensures the theory to be one-electron “self-correlation free”—the correlation energy for one-electron system being zero. The famous second-order Møller-Plesset perturbation theory (MP2), a widely used method in quantum chemistry, contains one *direct* term and one *exchange* term, represented respectively by the leading diagram in the first two rows of Fig. 1. These two terms cancel each other for one-electron systems. RPA is represented by the ring diagrams summed up to infinite order, as represented by the first-row diagrams in Fig. 1, where only “direct” processes are accounted for.¹ A simple way to include the exchange processes and eliminate the self-correlation error is to *antisymmetrize* the two-electron Coulomb integrals within the rCCD formulation of RPA [cf. Eq. (28)]. By doing so, a second-order screened exchange (SOSEX) term [63–65] is added, and the resultant RPA+SOSEX correlation energy is given by

$$E_c^{\text{RPA+SOSEX}} = \frac{1}{2} \sum_{ij,ab} (\langle ij|ab \rangle - \langle ij|ba \rangle) T_{ia,jb}^{\text{rCCD}}, \quad (30)$$

¹The RPA discussed here is often referred to as *direct* RPA in quantum chemistry literature. In some literature, RPA in fact corresponds to the time-dependent Hartree-Fock theory, where the exchange terms are also included. Nowadays, RPA with exchange included is usually referred to as RPAX or *full* RPA.

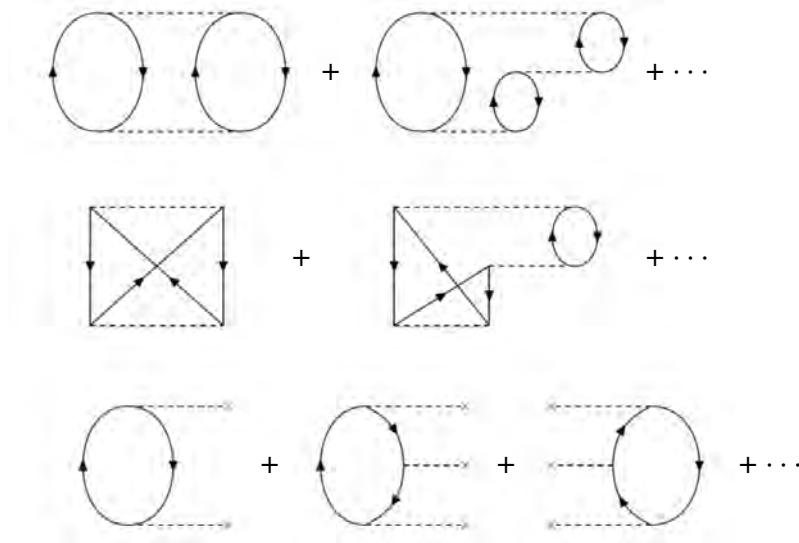


Fig. 1: Goldstone diagrams for $rPT2$. Diagrams in the three rows correspond to RPA, SOSEX, and rSE, respectively. Dashed lines ending with a cross in the third row denote the matrix element $\Delta v_{mn} = \langle \psi_m | \hat{v}^{HF} - \hat{v}^{MF} | \psi_n \rangle$. The rules to evaluate Goldstone diagrams can be found in Ref. [66].

with the two terms in Eq. (30) corresponding to RPA and SOSEX correlation energies respectively. The SOSEX contribution can be represented by the diagrams shown in the second row of Fig. 1, where the leading term corresponds to the exchange contribution in MP2. The RPA+SOSEX scheme has the interesting feature that it is one-electron self-correlation free, and improves substantially the total energy [64, 65]. The underbinding problem of RPA for chemically bonded molecules and solids is on average alleviated by the SOSEX correction.

As illustrated in Fig. 1, both RPA and SOSEX correlations can be interpreted as *infinite-order summations of selected types of diagrams*, with the MP2 terms as the leading order. This perspective is helpful for identifying important contributions that are still missing in the RPA+SOSEX scheme. In fact, at the second-order, in addition to the direct and exchange terms, there is yet another type of contribution, arising from the singles excitations (SE),

$$E_c^{SE} = \sum_i^{\text{occ}} \sum_a^{\text{unocc}} \frac{|\langle \Phi_0 | \hat{H}' | \Phi_i^a \rangle|^2}{E_0 - E_{i,a}^{(0)}} = \sum_{ia} \frac{|\langle \psi_i | \hat{v}^{HF} - \hat{v}^{MF} | \psi_a \rangle|^2}{\varepsilon_i - \varepsilon_a} = \sum_{ia} \frac{|\langle \psi_i | \hat{f} | \psi_a \rangle|^2}{\varepsilon_i - \varepsilon_a}, \quad (31)$$

where $\hat{v}^{HF}$ is the self-consistent Hartree-Fock potential, $\hat{v}^{MF}$ is the mean-field potential, based on which the reference state is generated, and $\hat{f} = -\nabla^2/2 + \hat{v}^{\text{ext}} + \hat{v}^{HF}$ is the single-particle Hartree-Fock Hamiltonian (also known as the Fock matrix in the quantum chemistry literature). Φ_0 and Φ_i^a are the KS Slater determinant and the singly-excited configuration generated by exciting one electron from an occupied state i to a virtual state a . A detailed derivation of Eq. (31) can be found in the supplementary material of Ref. [17]. Obviously for a Hartree-Fock reference where $\hat{v}^{MF} = \hat{v}^{HF}$, Eq. (31) becomes zero, a fact known as Brillouin theorem [66]. Therefore, this term is not present in standard MP2 theory based on the Hartree-Fock reference.

In Ref [17], it was shown that adding the SE term of Eq. (31) to RPA significantly improves the accuracy of vdW-bonded molecules, which the standard RPA scheme generally underbinds, and the SOSEX correction does not improve much. Similar to the RPA and SOSEX case, one can identify a sequence of single-excitation processes up to infinite order, as illustrated in the third row in Fig. 1. Summing these single-excitation diagrams up to infinite order represents a renormalization of the 2nd-order SE, and is termed as renormalized singles excitations (rSE) [37,35]. Remarkably, this infinite summation ends up with a closed expression similar to Eq. (31) ,

$$E_c^{\text{rSE}} = \sum_{ia} \frac{|\tilde{f}_{ia}|^2}{\tilde{\varepsilon}_i - \tilde{\varepsilon}_a}, \quad (32)$$

where $\tilde{\varepsilon}_i$ and $\tilde{\varepsilon}_a$ are the eigenvalues obtained by diagonalizing separately the occupied-occupied and virtual-virtual subblocks of the Fock matrix,

$$\begin{aligned} \sum_k f_{ik} \mathcal{O}_{kj} &= \tilde{\varepsilon}_i \mathcal{O}_{ij} \\ \sum_c f_{ac} \mathcal{U}_{cb} &= \tilde{\varepsilon}_a \mathcal{U}_{ab}, \end{aligned} \quad (33)$$

with $f_{mn} = \langle \psi_m | \hat{f} | \psi_n \rangle$. Note that the matrix f_{mn} is not diagonal since the ψ_n are the reference KS orbitals and hence not the eigenfunctions of $\hat{f}$, which is the single-particle Hartree-Fock Hamiltonian. Now $\tilde{f}_{ia}$ in the numerator of Eq. (32) are to the “transformed” off-diagonal block of the Fock matrix

$$\tilde{f}_{ia} = \sum_{jb} \mathcal{O}_{ij}^* f_{jb} \mathcal{U}_{ba}, \quad (34)$$

where the eigenvectors obtained in (33) are used here as the transformation coefficients. The physical origin of the SE corrections is that the commonly used KS references for RPA and beyond-RPA calculations are not the optimal starting point. The SE corrections accounts for the “orbital relaxation” effect, which leads to a lowering of the ground-state total energy.

Diagrammatically, RPA, SOSEX and rSE are three distinct infinite series of many-body terms, in which the three leading terms correspond to the three terms in second-order many-body perturbation theory. Summing them up, the resultant RPA+SOSEX+rSE scheme can be viewed a *renormalization* of the normal second-order many-body perturbation theory. Therefore the RPA+SOSEX+rSE scheme is also termed as “*renormalized second-order perturbation theory*” or rPT2 in the literature.

Independent of the rPT2 scheme described above, Bates, and Furche developed a beyond-RPA formalism termed as RPA renormalized many-body perturbation theory [67]. The essence of this formalism is to express the correlation energy in terms of an integration over the polarization propagator (closely related to the density response function) along the AC path. The correlation energy can be improved by improving the polarization propagator based on a series expansion in terms of the RPA polarization propagator multiplied with a four-point kernel. Benchmark calculations [68] show that the approximate exchange kernel (AXK) scheme within this formalism performs better than RPA+SOSEX discussed above. However, the rSE contribution is not included in the AXK scheme.

3 Implementation of the RPA method

In most practical RPA calculations, the evaluation of the RPA XC energy E_{xc}^{RPA} as given in Eqs. (24-26) is done employing the KS orbitals and orbital energies generated by a preceding KS-DFT calculation under certain lower-rung approximations. The GGA of Perdew, Burke, and Ernzerhof (PBE) is often used in the preceding KS-DFT calculation, and the RPA calculation based on a PBE reference is often denoted as “RPA@PBE.” In standard RPA@PBE calculations, the ground-state total energy is given by

$$\begin{aligned} E^{RPA@PBE} &= E^{PBE} - E_{xc}^{PBE} + E_{xc}^{RPA@PBE} \\ &= T_s^{PBE} + E_{ext}^{PBE} + E_H^{PBE} + E_x^{EX@PBE} + E_c^{RPA@PBE} \\ &= \langle \Phi_{PBE} | \hat{H} | \Phi_{PBE} \rangle + E_c^{RPA@PBE} \end{aligned} \quad (35)$$

where T_s^{PBE} , E_{ext}^{PBE} , E_H^{PBE} , and E_{xc}^{PBE} are, respectively, the noninteracting kinetic energy, the external potential energy, the Hartree energy, and the XC energy obtained from the self-consistent PBE calculation. The RPA@PBE total energy defined in (35) is obtained by subtracting the PBE XC energy from the PBE total energy, and then adding on top the RPA XC energy, evaluated using PBE orbitals and orbital energies. Equation (35) indicates that the RPA total energy can be seen as the sum of the Hartree-Fock energy evaluated with respect to the PBE reference and the RPA@PBE correlation energy.

Now it is clear that the key in RPA calculations is to evaluate the exact-exchange energy plus the RPA correlation energy. The algorithm for evaluating the exact-exchange energy in the present context is exactly the same as the Hartree-Fock exchange energy evaluation, which is routinely done in quantum chemistry calculations [69, 70]. The Hartree-Fock exchange is also a key component of hybrid density-functionals, which are available in increasingly more software that can deal with periodic systems [71, 72]. Here we shall not discuss the implementation of the Hartree-Fock exchange, but rather focus on the implementation of the RPA correlation energy part. To develop efficient algorithms to evaluate Eq. (26), the key is to realize that both χ^0 and v are non-local operators in space. Their real-space forms $\chi^0(\mathbf{r}, \mathbf{r}', i\omega)$ and $v(\mathbf{r}, \mathbf{r}')$ can be seen as basis representations of the corresponding operator χ^0 and v in terms of real-space grid points. Under such a discretization, χ^0 and v become matrices of dimension as large as the number of real space grid points. This is not an efficient representation since the number of grid points is rather large, especially in all-electron calculations. To deal with this problem, one can introduce an auxiliary basis set $\{P_\mu(\mathbf{r})\}$ to represent χ^0 and v , namely,

$$\chi^0(\mathbf{r}, \mathbf{r}', i\omega) = \sum_{\mu, \nu}^{N_{aux}} P_\mu(\mathbf{r}) \chi_{\mu\nu}^0(i\omega) P_\nu(\mathbf{r}'), \quad (36)$$

and

$$V_{\mu\nu} = \iint d\mathbf{r} d\mathbf{r}' \frac{P_\mu(\mathbf{r}) P_\nu(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \quad (37)$$

where N_{aux} is the number of auxiliary basis functions. To find the matrix form $\chi_{\mu, \nu}^0(i\omega)$, one

needs to expand the products of KS orbitals in terms of $P_\mu(\mathbf{r})$,

$$\psi_m^*(\mathbf{r})\psi_n(\mathbf{r}) = \sum_{\mu} C_{mn}^{\mu} P_{\mu}(\mathbf{r}), \quad (38)$$

where C_{mn}^{μ} is the expansion coefficients. Inserting the expansion (38) into (23), and comparing to (36), one arrives at

$$\chi_{\mu\nu}^0(i\omega) = \sum_{m,n} \frac{(f_m - f_n) C_{mn}^{\mu} C_{mn}^{\nu*}}{\varepsilon_m - \varepsilon_n - i\omega}. \quad (39)$$

Thus, within the auxiliary basis representation, χ^0 and v become $N_{\text{aux}} \times N_{\text{aux}}$ matrices. Since N_{aux} is typically much smaller than the number of real-space grid points, or the number of pair products of KS orbitals, the χ^0 matrix in (39) and the V matrix in (38) can be seen as compressed representation of their operators, with the trace of their product unchanged,

$$\text{Tr} \chi^0(i\omega) v = \iint \chi^0(\mathbf{r}, \mathbf{r}', i\omega) v(\mathbf{r}', \mathbf{r}) d\mathbf{r} d\mathbf{r}' = \sum_{\mu\nu} \chi_{\mu\nu}^0(i\omega) V_{\mu\nu}. \quad (40)$$

Based on this observation, one may conclude that the computed RPA correlation energy is independent of the basis representation of the χ^0 and v operators. Thus one may equivalently interpret Eq. (26) as matrix algebra with χ^0 and v represented within the auxiliary basis as given by (39) and (37). Using $\text{Tr} \ln A = \ln \det A$ (with $\det A$ being the determinant of the matrix A), one obtains the final working expression for evaluating the RPA correlation energy,

$$\begin{aligned} E_c^{\text{RPA}} &= \frac{1}{2\pi} \int_0^\infty d\omega \ln (\det(1 - \chi^0(i\omega) v) + \chi^0(i\omega) v) \\ &= \frac{1}{2\pi} \int_0^\infty d\omega \ln (\det(1 - \Pi(i\omega)) + \Pi(i\omega)). \end{aligned} \quad (41)$$

In Eq. (41), we have introduced an intermediate quantity $\Pi(i\omega) = V^{1/2} \chi^0(i\omega) V^{1/2}$ and used the property $\text{Tr}(V^{1/2} \chi^0 V^{1/2}) = \text{Tr}(\chi^0 V)$. The frequency integration in Eq. (41) can be done relatively easily, since the integrand is rather smooth and peaked at low frequency values. A modified Gauss-Legendre grid, which transforms a standard Gauss-Legendre grid in the range $[-1, 1]$ to $[0, \infty]$,

$$\tilde{x}_i = x_0 \frac{1 + x_i}{1 - x_i}, \quad \tilde{w}_i = w_i \frac{2x_0}{(1 - x_i)^2} \quad (42)$$

works rather well for most systems. In (42), (x_i, w_i) are the abscissas and weights of the grid points generated from the Gauss-Legendre quadrature formula in an integration range $[-1, 1]$, whereas $(\tilde{x}_i, \tilde{w}_i)$ are abscissas and weights of the transformed grid. Usually a few tens of frequency grid points are sufficient to get highly accurate results, except for systems with vanishing gaps, where considerably more grid points have to be used. Alternative frequency grids have been developed for evaluating the RPA correlation energies [73, 33], which have been shown to work well for small gap systems.

The implementation scheme described above is known as the resolution-of-identity (RI) approach to RPA [73, 74]. From the above discussion, one may see that once the matrix forms of χ^0 and v within the auxiliary basis are determined, the rest of the RPA calculation is rather straightforward. The key steps in RPA calculations are thus: 1) Generate the auxiliary basis functions $\{P_\mu(\mathbf{r})\}$; 2) Determine the triple coefficients $C_{m,n}^\mu$; and 3) Construct the $\chi_{\mu,\nu}^0$ matrix using (39). The choice of auxiliary basis functions $\{P_\mu(\mathbf{r})\}$ depends on the underlying one-electron basis functions used in the KS-DFT calculations. Our own RPA implementation was done in the FHI-aims code [75, 74] which employs atom-centered numeric orbitals (NAOs) as basis functions

$$\phi_i(\mathbf{r}) = u_k(r)Y_{lm}(\theta, \phi), \quad (43)$$

where $u_k(r)$ is a numerically tabulated radial function and $Y_{lm}(\theta, \phi)$ is a spherical harmonic. Our auxiliary basis functions are also constructed as a numerical radial function multiplied by spherical harmonics

$$P_\mu(\mathbf{r}) = \xi_s(r)Y_{lm}(\theta, \phi) \quad (44)$$

but the radial function $\xi_s(r)$ has a different shape from $u_k(r)$. In fact, they are generated in order to best represent the product of the one-electron orbitals $\phi_i\phi_j$. Details on how the auxiliary basis functions are constructed can be found in Refs. [74, 76].

Once the auxiliary basis functions $\{P_\mu(\mathbf{r})\}$ are constructed, we can start to determine the triple expansion coefficients $C_{m,n}^\mu$. For any finite auxiliary basis set, the expansion in (38) is an approximation, incurring an error $\delta\rho_{mn}(\mathbf{r}) = \sum_\mu C_{mn}^\mu P_\mu(\mathbf{r}) - \psi_m^*(\mathbf{r})\psi_n(\mathbf{r})$. The accuracy of this approximation will not only depend on the quality and size of the auxiliary basis, but also on the expansion coefficients $C_{m,n}^\mu$. In the RI approach with Coulomb metric [77], instead of minimizing the norm of the error ($\delta\rho_{mn}|\delta\rho_{mn}$), one minimizes the self Coulomb repulsion of this error ($\delta\rho_{mn}|v|\delta\rho_{mn}$), leading to the following expression for $C_{m,n}^\mu$,

$$C_{m,n}^\mu = \sum_\nu (mn|v|\nu) V_{\nu\mu}^{-1}, \quad (45)$$

where

$$(mn|v|\nu) = \iint d\mathbf{r} d\mathbf{r}' \frac{\psi_m^*(\mathbf{r})\psi_n(\mathbf{r})P_\nu(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}, \quad (46)$$

and V^{-1} is the inverted Coulomb matrix. This conventional (global) RI approach works extremely well for small molecules. For big molecules and periodic systems, one may switch to a localized variant of the RI approach [76]. With enhanced auxiliary basis functions, the localized RI approach has been shown to be sufficiently accurate in practical calculations [76], and is instrumental for periodic systems [72].

A final key step is to build the χ^0 matrix in (39). Since the number of auxiliary basis functions N_{aux} scales linearly with the number of one-electron basis functions, the computational cost of Eq. (39) scales as $O(N^4)$ with respect to the size of the system, and represents the computational bottleneck of RPA calculations. In recent years, lower-scaling algorithms have also been developed [33, 78], which holds promise for extending the applicability of RPA to large systems.

4 The performance of RPA-based methods and prototypical applications

The revived interest in RPA in recent years comes not only from its appealing concept, but also from its remarkable performance in practical applications. First of all, RPA can describe vdW interactions in a seamless way. The reason behind this capability has already been clearly explained by Dobson in the early days [79], and the performance of RPA has been demonstrated later for rare-gas crystals [19] and rare-gas dimers [17]. For lower-rung density-functional approximations (DFAs) ranging from LDA to hybrid functionals, the long-range vdW interactions are not captured, and *ad hoc* corrections have to be added in order to describe systems where vdW interactions play an important role. Furthermore, it seems that RPA is able to describe the delicate energy differences in complex bonding situations, including molecules adsorbed on surfaces [27, 28], the isostructural phase transition [24] and the relative stability of different polymorphs of crystals [26], the binding and ex-foliation energies of layered compounds [36], and the formation energy of defects [33]. Finally, RPA yields accurate chemical reaction barrier heights [37], which is crucial for reliably estimating the rate of chemical reactions.

4.1 Van der Waals interactions

Van der Waals (vdW) interactions arise from the coupling between spontaneous quantum charge fluctuations separated in space. For two well-separated, spherically-symmetric charge-neutral systems, the interaction energy goes like

$$\Delta E \sim \frac{C_6}{R^6} \quad \text{for } R \rightarrow \infty \quad (47)$$

where C_6 is the dispersion coefficient and R is the separation between the two systems. Dobson showed that the interaction energy between two subsystems A and B obtained by RPA exactly follows the asymptotic behavior given by (47), with the RPA C_6 coefficients given by

$$C_6^{\text{RPA}} = \frac{3}{\pi} \int d\omega \alpha_A^{\text{RPA}}(i\omega) \alpha_B^{\text{RPA}}(i\omega) \quad (48)$$

where $\alpha_A^{\text{RPA}}(i\omega)$ is the RPA polarizability of the subsystem A , which can be obtained by integrating over the microscopic RPA response function as,

$$\alpha_A^{\text{RPA}}(i\omega) = \frac{1}{3} \sum_{i=1}^3 \iint d\mathbf{r} d\mathbf{r}' r_i \chi^{\text{RPA}}(\mathbf{r}, \mathbf{r}', i\omega) r'_i \quad (49)$$

with $r_{1,2,3} = x, y, z$.

In Fig. 2, the binding energy curves of Ar_2 obtained using PBE and RPA-based methods are presented. MP2 is the simplest post-Hartree-Fock quantum chemistry approach capable of describing vdW interactions, and hence also included here for comparison. The reference curve is given by the Tang-Toennies potential model, with model parameters determined from experiment. PBE underestimates the bonding strength of Ar_2 considerably, and this is appreciably

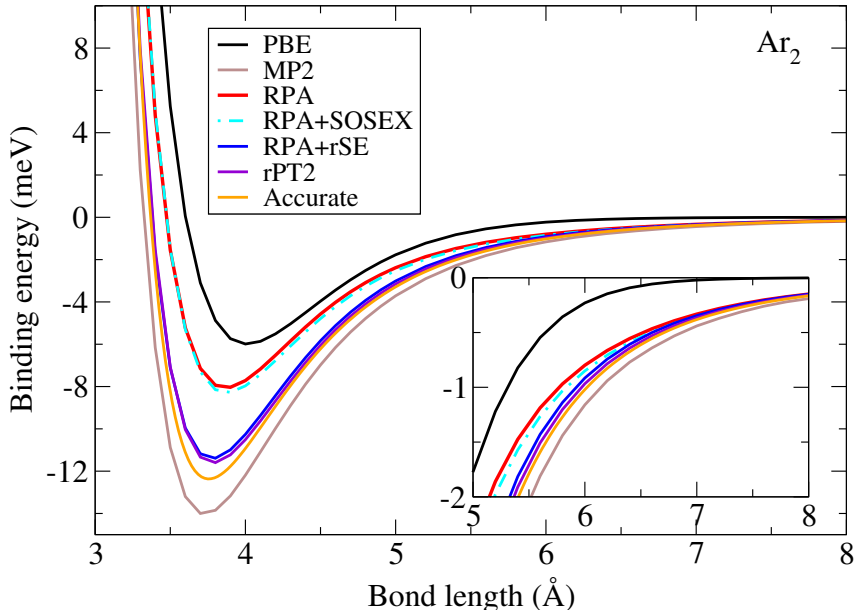


Fig. 2: Binding energy curves for Ar_2 obtained using PBE, MP2, and RPA-based methods. All RPA-based methods use PBE orbitals as input. Calculations are done using the FHI-aims code [75, 74]. The “Accurate” reference curve is given by the Tang-Toennies potential model for Ar_2 [80]. The inset is a zoom-in of the asymptotic behavior at large separations.

improved by RPA. More importantly, at large separations, the PBE binding energy decays rapidly (in fact exponentially) to the energy zero, but the RPA curve displays a correct C_6/R^6 asymptotic behavior. Compared to the reference result, however, the RPA still shows a substantial underbinding behavior, in contrast to the MP2 curve which shows too strong binding of Ar_2 . The underbinding issue of RPA for vdW dimers arises from the too strong Pauli repulsion, which is in turn due to the Hartree-Fock part of the RPA total energy, obtained with the semi-local GGA orbitals. This issue can be fixed by the single-excitations corrections [17], and in particular its renormalized version [35]. As shown in Fig. 2, the RPA+rSE scheme brings the binding energy curve of Ar_2 in close agreement with the reference one. The SOSEX correction, however, does not have a noticeable effect here. Thus the final rPT2 binding energy curve is almost on top of the RPA+rSE one. The performance of RPA-based methods for other rare-gas dimers can be found in Ref. [35].

A widely used benchmark set for weak interactions is the S22 test set designed by Jurečka et al., [82]. This test set collects 22 molecular dimers, among which 7 dimers are of hydrogen binding type, 8 of pure vdW (also called “dispersion”) bonding, and another 7 of mixed type. Figure 3(a) shows the structures of water dimer, adenine-thymine dimer, and water-benzene dimer, representing respectively the three bonding types. Because of its good representativity and the availability of accurate reference interaction energies obtained using the CCSD(T) method [81], S22 has been widely used for benchmarking the performance of or training the parameters for computational schemes that aim at describing weak interactions. The performance of RPA and some of the RPA-related methods has been benchmarked for this test set [15, 18, 17, 83].

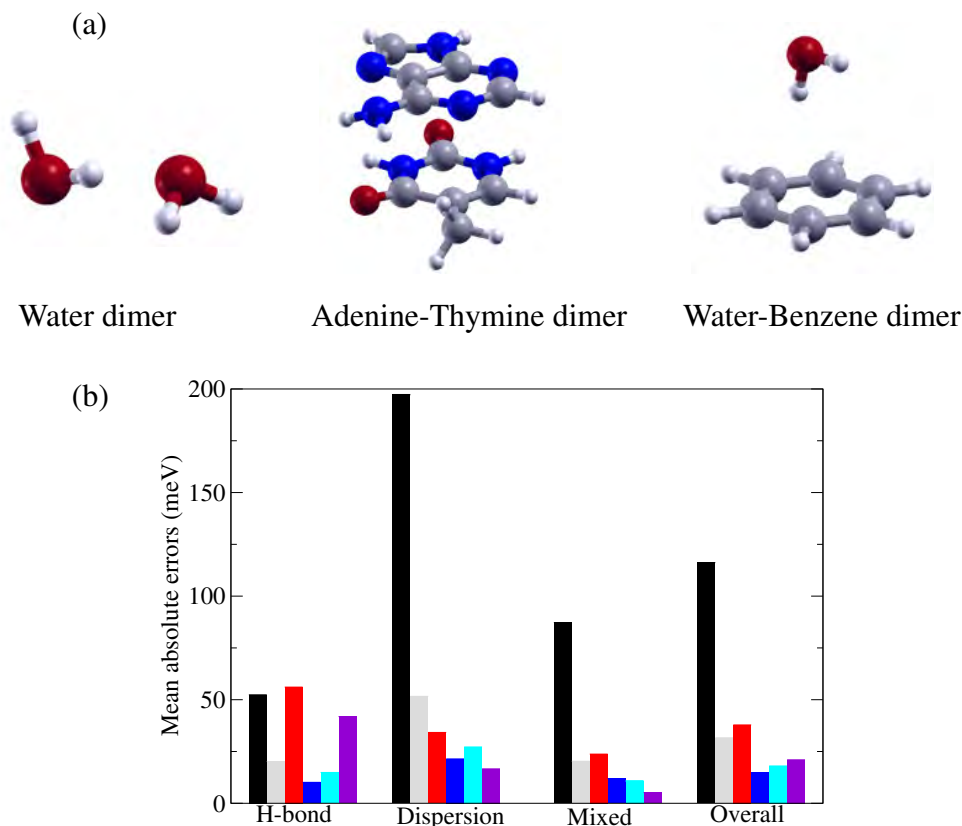


Fig. 3: (a) Structures of water dimer, adenine-thymine dimer, and water-benzene dimer. C, O, N, and H atoms are represented by grey, red, blue, and white balls. (b) MAEs (in meV) for the S22 test set given by PBE, MP2, RPA, RPA+rSE, RPA+SOSEX, and rPT2 methods. The CCSD(T) results of Takatani et al. [81] are used here as the reference.

In Fig. 3(b) the mean absolute errors (MAE) of PBE, MP2, and RPA-based methods are presented for the three subsets separately and for the entire S22 set. Both PBE and the correlated methods can describe the hydrogen bonding well, since this type of bonding is dominated by the electrostatic interactions, which has already been captured by the semi-local functionals to a large extent. The error of the standard RPA method is still appreciable ($\text{MAE} > 1 \text{ kcal/mol} \approx 43 \text{ meV}$) arising from its general underbinding behavior. This can again be corrected by rSE, or by SOSEX terms. However, adding the two types of corrections together, the rPT2 scheme tends to overbind the hydrogen-bonded molecules, and hence the MAE increases again. For dispersion and mixed bondings, RPA performs better, and the MAE can be further reduced by rSE and SOSEX corrections. The MP2 method, on the other hand, yields a relatively large MAE for dispersion-bonded molecules, owing to its well-known overbinding problem for this type of interaction. This benchmark test indicates that the RPA+rSE scheme is a suitable approach recommendable for describing weak interactions. The advantage of this scheme is that it does not noticeably increase the computational cost, compared to the standard RPA scheme. In addition to vdW complexes, RPA and its variants have also been applied to chemically bonded molecules and crystalline solids. Interested readers may look into the literature for further details [21, 84, 37, 35].

4.2 Surface adsorption

An accurate description of atoms and molecules interacting with surfaces is the key to understand important physical and chemical processes such as heterocatalysis, molecular electronics, and corrosion. Molecules adsorbed on metal surfaces represent a particularly difficult situation, since quantitatively accurate results can only be obtained if the approach is able to describe simultaneously the chemical bonding, vdW interactions, metallic screening, and charge transfer processes. The issue is best illustrated by the “CO adsorption puzzle,” where LDA and several popular GGAs predict the wrong adsorption site for the CO molecule adsorbed on several noble/transition metal surfaces at low coverages [56]. For CO adsorbed on the Cu(111) or Pt(111) surfaces, LDA and popular GGAs erroneously favor the highly-coordinated hollow site (see, e.g., Fig. 4(a)), whereas experiments clearly show that the singly-coordinated on-top site is the energetically most stable site [85, 86]. This posed a severe challenge to the first-principles modeling of molecular adsorption problems and represents a key test example for the RPA-based methods.

In Ref. [27], the performance of the RPA method for CO adsorbed on Cu(111) was investigated. The computed RPA adsorption energies for both the on-top and fcc (face centered cubic) hollow sites are presented in Fig. 4(b), together with the results from LDA, AM05 [87], PBE, and the hybrid PBE0 functional [88]. Figure 4 reveals what happens in the CO adsorption puzzle when climbing the Jacob’s ladder in DFT [9]—going from the first two rungs (LDA and GGAs) to the fourth (hybrid functionals), and finally to the 5th-rung RPA functional. It can be seen that, along the way, the adsorption energies on both sites are reduced, but the magnitude of the reduction is bigger for the fcc hollow site. The correct energy ordering is already restored at the PBE0 level, but the adsorption energy difference between the two sites is vanishingly small. RPA not only predicts the correct adsorption site, but also produces a reasonable adsorption energy difference of 0.22 eV, consistent with experiments. The effect of the starting reference state on the calculated RPA results has also been checked. In Ref. [27], in addition to the commonly used scheme RPA@PBE, RPA calculations were also done on top of the hybrid functional PBE0. Figure 4 indicates that the small difference between RPA@PBE and RPA@PBE0 results is insignificant for understanding the “CO adsorption puzzle.”

Schimka *et al.* extended the RPA benchmark studies of the CO adsorption problem to more surfaces [28], and found that RPA is the only approach that gives both good adsorption energies and surface energies. GGAs and hybrid functionals at most yield either good surface energies, or adsorption energies, but not both. Following these works, RPA has subsequently been applied to the adsorption of small alkanes in Na-exchanged chabazite [89], benzene on the Si(001) surface [90], graphene on the Ni(111) surface [29, 30] and the Cu(111) and Co(0001) surfaces [30]. In all these studies, RPA was demonstrated to be able to capture the delicate balance between chemical and dispersion interactions, and yields quantitatively reliable results.

We expect RPA to become an increasingly more important approach in surface science, with increasing computer power and more efficient implementations.

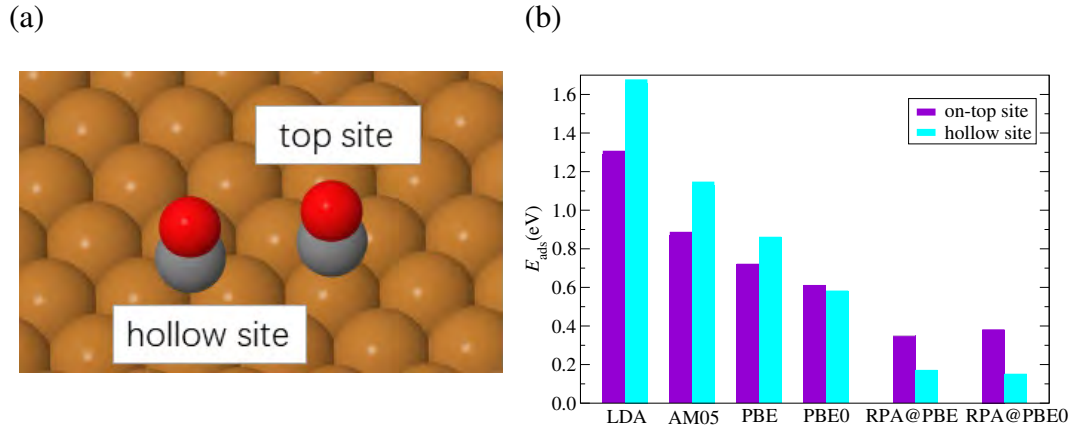


Fig. 4: (a): Schematic of CO adsorbed on the Cu(111) surface, with the on-top and hollow adsorption sites illustrated. (b): Adsorption energies for CO sitting on the on-top and fcc hollow sites as obtained using LDA, AM05, PBE, PBE0, and RPA. RPA results are presented for both PBE and PBE0 references. Calculations were done with the FHI-aims code. Adopted from Ref. [37].

4.3 Structural phase transitions

The f -electron materials, which contain rare-earth or actinide elements, pose a great challenge to first-principles approaches. A prominent example of an f -electron system is the Ce metal, which has an intriguing iso-structural α - γ phase transition, accompanied by a drastic volume collapse (as large as 15% at ambient pressure and zero temperature). The two phases are characterized by distinct spectroscopic and magnetic properties. Various theoretical approaches, including LDA plus dynamical mean-field theory (LDA+DMFT) have been employed to study this system [91, 92]. DFT within its local and semilocal approximations is unable to describe the two phases of Ce. In fact, due to the strong delocalization error in these functionals, the localized nature of the f -electrons in the γ phase cannot be properly described; only the α phase is described with some confidence within LDA/GGAs, although not at a quantitative level.

It was shown by Casadei *et al.* [24, 25] that, remarkably, hybrid functionals like PBE0 and HSE can yield two self-consistent solutions, with distinct lattice constants, cohesive energies, electronic band structures, local magnetic moments, and different degrees of f -electron localization. These two solutions can be reasonably associated with the phases of the Ce metal. However, the energetic ordering of the α -like and γ -like phases produced by PBE0 is not consistent with the experimental situation where the α phase is energetically more stable at low temperature and ambient pressure. Adding RPA corrections on top of the PBE cohesive energies for the two solutions, the energy ordering is reversed, and the α -like solution becomes energetically more stable. The transition pressure of the two phases given by the Gibbs construction is consistent with the experimental value.

The capability of RPA to capture the delicate energy difference between competing phases or polymorphs has also been demonstrated for Iron disulfide (FeS_2), a potentially interesting system for photovoltaic and photoelectrochemical applications. This material turns out to be rather challenging, since popular conventional density-functional approximations fail to produce the

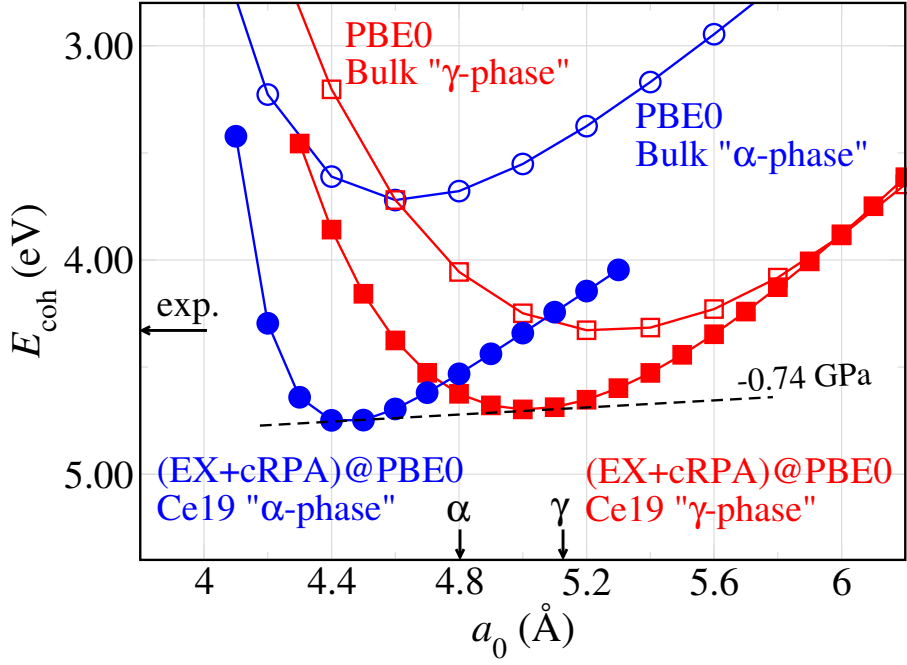


Fig. 5: Calculated PBE0 and RPA@PBE0 [RPA denoted here as exact-exchange plus RPA correlation (EX+cRPA)] cohesive energy (E_{coh}) as a function of the lattice constant a_0 for the two electronic states based on self-consistent PBE0 solutions. The correction of RPA with respect to the PBE0 cohesive energies was done for a 19-atom fcc-cerium cluster. The dashed line illustrates the Gibbs construction for the transition pressure from the α to the γ phase. The arrows on the energy axes indicates the experimental cohesive energy of the α phase. Adapted from Ref. [24].

relative stability of its two polymorphs: pyrite and marcasite, with the latter artificially stabilized. It was demonstrated by Zhang *et al.* [26] that RPA, regardless of its reference state, can correctly restore the energy ordering of the two polymorphs of FeS_2 . These authors further reveal that the fact that RPA tends to stabilize the pyrite polymorph is due to its smaller KS band gap, resulting in a large RPA correlation energy as compared to the marcasite polymorph. This observation is consistent with the case of the Ce metal [24], where the more metallic α -phase is stabilized within the RPA. Another successful application of this kind is that the tiny energy difference between the two allotropes of carbon, graphite and diamond, can be reliably described by RPA, with the correct prediction that graphite is energetically slightly lower than diamond [31]. More systematic benchmark studies of the performance of RPA and several beyond-RPA methods for predicting the transition pressure of structural phase transitions of a set of bulk materials have recently been reported in Ref. [93].

Other types of materials science problems to which RPA has been applied include layered compounds [94, 36] and defect systems [32, 33]. We shall not elaborate on these types of applications here due to limited space, and interested readers may look into the original literature for details. In summary, there is ample evidence that RPA-based methods perform well in capturing delicate energy differences in materials, and fix some of the qualitative failures of more conventional approaches.

5 Recent developments

The field of RPA-based methodology development and applications represents a rapidly evolving branch of computational electronic structure theory. Notable progress has been achieved in several directions within the last few years, which we would like to briefly recapitulate here.

1. RPA force calculations. Analytical gradients of the RPA total energy with respect to the atomic positions have been computed within both the atomic orbital basis [95–97, 78] and plane-wave basis [98] frameworks. This allows to compute interatomic forces and relax structures at the RPA level, which is a long-sought goal of the electronic-structure community. Moreover, it is now possible to calculate vibrational frequencies [96] and phonon spectra based on the RPA force constant [98], and even molecular dynamics simulations can be carried out based on the RPA forces [98]. These advancements greatly enhanced the capability of RPA in computational chemistry and materials science.
2. Low-scaling RPA implementations. Another noteworthy development is several low-scaling—ranging from linear scaling to cubic scaling—algorithms for RPA calculations have been designed and implemented [99, 100, 33, 101–103]. This paves the way for applying RPA to large-sized and complex materials that were inaccessible in the past.
3. Particle-particle RPA. The above-discussed RPA is represented by ring diagrams and called particle-hole RPA (phRPA) in the literature. In addition to phRPA, another type of RPA, consisting of an infinite summation of ladder diagrams, has also been discussed in the nuclear physics literature [5]. This type of RPA is referred to as particle-particle RPA and has recently been brought to the attention of electronic structure community [104–107]. Benchmark calculations show that ppRPA carries interesting features that are not present in phRPA. Attempts for combining the two types of RPA in one framework have been made both in a range-separated manner [108] and globally [109]. However, it seems that merging the two RPA channels into one united theory is a highly nontrivial task [109].
4. Self-consistent RPA in a generalized Kohn-Sham framework. As mentioned before, the majority of practical RPA calculations are done in a post-processing fashion, using orbitals and orbital energies generated from a preceding DFA calculation. The importance of the rSE contribution indicates that commonly used semi-local DFAs are not optimal starting points for RPA calculations. Thus running RPA calculations in a self-consistent way is highly desirable. However, for orbital-dependent functionals like RPA, the criterion for “self-consistency” is not uniquely defined. The optimized effective potential (OEP) RPA [110–112] is a well-defined procedure, but the obtained results for binding energies are not better than perturbative RPA. Most recently, self-consistent RPA schemes are developed by Jin *et al.* within a generalized OEP framework [113] and by Voora *et al.* within a generalized KS framework [114]. The two schemes differ in details, but the rSE contribution is captured in both schemes. Initial results obtained from these schemes look very promising and there is much to explore along this direction.

5. More beyond-RPA schemes to improve the accuracy. In addition to the beyond-RPA schemes already discussed in Sec. 2.3, one most recent development by Zhang and Xu [115] is to introduce a spin-pair distinctive algorithm in the ACFDT context, whereby the contributions of same-spin and opposite-spin pairs to the correlation energy are separated. By scaling the contributions for two types of spin pairs differently, one can achieve a simultaneous attenuation of both self-interaction errors and static correlation errors. Similar to the power series approximation of Görling and coauthors [62,59], the spin-pair distinctive formalism of Zhang and Xu [115] is particularly successful in dealing with systems with multi-reference characters.

6 Summary

In this lecture, we discussed the basic concept, the theoretical formulation, the implementation algorithm, the prototypical applications, and the recent development of RPA-based methods. We expect that RPA and its variants will have an ever-increasing impact on computational materials science, and become the mainstream methods in the near future. Since this is an actively developing field, and the number of papers is quickly growing, it is well possible some important developments are not covered in our discussion. The purpose of this manuscript is to inform the readers about the overall status of this field, and stimulate more work on the development and application of RPA-type methodology.

Acknowledgements

The author wishes to thank Patrick Rinke and Matthias Scheffler for the fruitful collaborations on the RPA. He also would like to thank Muhammad Naeem Tahir, Yanyong Wang, and Rong Shi for proofreading these lecture notes.

References

- [1] D. Bohm and D. Pines, Phys. Rev. **82**, 625 (1951)
- [2] D. Bohm and D. Pines, Phys. Rev. **92**, 609 (1953)
- [3] J. Hubbard, Proc. Roy. Soc. (London) **A240**, 539 (1957)
- [4] M. Gell-Mann and K.A. Brueckner, Phys. Rev. **106**, 364 (1957)
- [5] P. Ring and P. Schuck: *The Nuclear Many-Body Problem* (Springer, New York, 1980)
- [6] D.C. Langreth and J.P. Perdew, Solid State Commun. **17**, 1425 (1975)
- [7] D.C. Langreth and J.P. Perdew, Phys. Rev. B **15**, 2884 (1977)
- [8] O. Gunnarsson and B.I. Lundqvist, Phys. Rev. B **13**, 4274 (1976)
- [9] J.P. Perdew and K. Schmidt, in V. Van Doren, C. Van Alsenoy, and P. Geerlings (eds.): *Density Functional Theory and its Application to Materials* (AIP, Melville, NY, 2001)
- [10] F. Furche, Phys. Rev. B **64**, 195120 (2001)
- [11] M. Fuchs and X. Gonze, Phys. Rev. B **65**, 235109 (2002)
- [12] H. Jiang and E. Engel, J. Chem. Phys **127**, 184108 (2007)
- [13] M. Hellgren and U. von Barth, Phys. Rev. B **76**, 075107 (2007)
- [14] J. Toulouse, I.C. Gerber, G. Jansen, A. Savin, and J.G. Ángyán, Phys. Rev. Lett. **102**, 096404 (2009)
- [15] W. Zhu, J. Toulouse, A. Savin, and J.G. Ángyán, J. Chem. Phys. **132**, 244108 (2010)
- [16] A. Heßelmann and A. Görling, Phys. Rev. Lett. **106**, 093001 (2011)
- [17] X. Ren, A. Tkatchenko, P. Rinke, and M. Scheffler, Phys. Rev. Lett. **106**, 153003 (2011)
- [18] H. Eshuis and F. Furche, J. Phys. Chem. Lett. **2**, 983 (2011)
- [19] J. Harl and G. Kresse, Phys. Rev. B **77**, 045136 (2008)
- [20] J. Harl and G. Kresse, Phys. Rev. Lett. **103**, 056401 (2009)
- [21] J. Harl, L. Schimka, and G. Kresse, Phys. Rev. B **81**, 115126 (2010)
- [22] H.V. Nguyen and S. de Gironcoli, Phys. Rev. B **79**, 205114 (2009)
- [23] D. Lu, Y. Li, D. Rocca, and G. Galli, Phys. Rev. Lett. **102**, 206411 (2009)
- [24] M. Casadei, X. Ren, P. Rinke, A. Rubio, and M. Scheffler, Phys. Rev. Lett. **109**, 146402 (2012)

- [25] M. Casadei, X. Ren, P. Rinke, A. Rubio, and M. Scheffler, *Phys. Rev. B* **93**, 075153 (2016)
- [26] M.Y. Zhang, Z.H. Cui, and H. Jiang, *J. Mater. Chem. A* **6**, 6606 (2018)
- [27] X. Ren, P. Rinke, and M. Scheffler, *Phys. Rev. B* **80**, 045402 (2009)
- [28] L. Schimka, J. Harl, A. Stroppa, A. Grüneis, M. Marsman, F. Mittendorfer, and G. Kresse, *Nat. Mater.* **9**, 741 (2010)
- [29] F. Mittendorfer, A. Garhofer, J. Redinger, J. Klimeš, J. Harl, and G. Kresse, *Phys. Rev. B* **84**, 201401(R) (2011)
- [30] T. Olsen, J. Yan, J.J. Mortensen, and K.S. Thygesen, *Phys. Rev. Lett.* **107** (2011)
- [31] S. Lebègue, J. Harl, T. Gould, J.G. Ángyán, G. Kresse, and J.F. Dobson, *Phys. Rev. Lett.* **105**, 196401 (2010)
- [32] F. Bruneval, *Phys. Rev. Lett.* **108**, 256403 (2012)
- [33] M. Kaltak, J. Klimeš, and G. Kresse, *J. Chem. Theory Comput.* **10**, 2498 (2014)
- [34] J.F. Dobson, A. White, and A. Rubio, *Phys. Rev. Lett.* **96**, 073201 (2006)
- [35] X. Ren, P. Rinke, G.E. Scuseria, and M. Scheffler, *Phys. Rev. B* **88**, 035120 (2013)
- [36] T. Björkman, A. Gulans, A.V. Krasheninnikov, and R.M. Nieminen, *Phys. Rev. Lett.* **108**, 235502 (2012)
- [37] X. Ren, P. Rinke, C. Joas, and M. Scheffler, *J. Mater. Sci.* **47**, 7447 (2012)
- [38] D. Pines, *Rep. Prog. Phys.* **79**, 092501 (2016)
- [39] G.P. Chen, V.K. Voora, M.M. Agee, S.G. Balasubramani, and F. Furche, *Annu. Rev. Phys. Chem.* **68**, 421 (2017)
- [40] R.J. Bartlett and M. Musiał, *Rev. Mod. Phys.* **79**, 291 (2007)
- [41] G.E. Scuseria, T.M. Henderson, and D.C. Sorensen, *J. Chem. Phys.* **129**, 231101 (2008)
- [42] N.E. Dahlen, R. van Leeuwen, and U. von Barth, *Phys. Rev. A* **73**, 012511 (2006)
- [43] F. Caruso, D.R. Rohr, M. Hellgren, X. Ren, P. Rinke, A. Rubio, and M. Scheffler, *Phys. Rev. Lett.* **110**, 146403 (2013)
- [44] W. Kohn and L.J. Sham, *Phys. Rev.* **140**, A1133 (1965)
- [45] A. Becke, *J. Chem. Phys.* **88**, 2547 (1988)
- [46] C. Lee, W. Yang, and R.G. Parr, *Phys. Rev. B* **37**, 785 (1988)
- [47] J.P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996)

- [48] J. Tao, J.P. Perdew, V.N. Staroverov, and G.E. Scuseria, Phys. Rev. Lett. **91**, 146401 (2003)
- [49] J. Sun, A. Ruzsinszky, and J.P. Perdew, Phys. Rev. Lett. **115**, 036402 (2015)
- [50] A.D. Becke, J. Chem. Phys. **98**, 5648 (1993)
- [51] J. Heyd, G.E. Scuseria, and M. Ernzerhof, J. Chem. Phys. **118**, 8207 (2003)
- [52] J.P. Perdew, R.G. Parr, M. Levy, and J.L. Balduz, Jr., Phys. Rev. Lett. **49**, 1691 (1982)
- [53] A.J. Cohen, P. Mori-Sanchez, and W. Yang, Science **321**, 792 (2008)
- [54] S. Kristyán and P. Pulay, Chem. Phys. Lett. **229**, 175 (1994)
- [55] M. Imada, A. Fujimori, and Y. Tokura, Rev. Mol. Phys. **70**, 1039 (1998)
- [56] P.J. Feibelman, B. Hammer, J.K. Nørskov, F. Wagner, M. Scheffler, R. Stumpf, R. Watwe, and J. Dumestic, J. Phys. Chem. B **105**, 4018 (2001)
- [57] Y. Zhang and W. Yang, J. Chem. Phys. **109**, 2604 (1998)
- [58] P. Nozières and D. Pines: *The Theory of Quantum Liquids* (Benjamin, New York, 1966)
- [59] A. Görling, Phys. Rev. B **99**, 235120 (2019)
- [60] A. Ruzsinszky, J.P. Perdew, and G.I. Csonka, J. Chem. Theory Comput. **6**, 127 (2010)
- [61] T. Olsen and K.S. Thygesen, Phys. Rev. B **86**, 081103(R) (2012)
- [62] J. Erhard, P. Bleizier, and A. Görling, Phys. Rev. Lett. **117**, 143002 (2016)
- [63] D.L. Freeman, Phys. Rev. B **15**, 5512 (1977)
- [64] A. Grüneis, M. Marsman, J. Harl, L. Schimka, and G. Kresse, J. Chem. Phys. **131**, 154115 (2009)
- [65] J. Paier, B.G. Janesko, T.M. Henderson, G.E. Scuseria, A. Grüneis, and G. Kresse, J. Chem. Phys. **132**, 094103 (2010); Erratum: *ibid.* **133**, 179902 (2010)
- [66] A. Szabo and N.S. Ostlund: *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory* (McGraw-Hill, New York, 1989)
- [67] J.E. Bates and F. Furche, J. Chem. Phys. **139**, 171103 (2013)
- [68] G.P. Chen, M.M. Agee, and F. Furche, J. Chem. Theory Comput. **14**, 5701 (2018)
- [69] M.J. Frisch *et al.*: *Gaussian 09W Revision D.01*. Gaussian Inc. Wallingford CT 2009
- [70] *TURBOMOLE V7.3 2018, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from <http://www.turbomole.com>.*

- [71] J. Paier, M. Marsman, K. Hummer, G. Kresse, I.C. Gerber, and J.G. Ángyán, *J. Chem. Phys.* **124**, 154709 (2006)
- [72] S.V. Levchenko, X. Ren, J. Wieferink, R. Johanni, P. Rinke, V. Blum, and M. Scheffler, *Comp. Phys. Comm.* **192**, 60 (2015)
- [73] H. Eshuis, J. Yarkony, and F. Furche, *J. Chem. Phys.* **132**, 234114 (2010)
- [74] X. Ren, P. Rinke, V. Blum, J. Wieferink, A. Tkatchenko, A. Sanfilippo, K. Reuter, and M. Scheffler, *New J. Phys.* **14**, 053020 (2012)
- [75] V. Blum, F. Hanke, R. Gehrke, P. Havu, V. Havu, X. Ren, K. Reuter, and M. Scheffler, *Comp. Phys. Comm.* **180**, 2175 (2009)
- [76] A.C. Ihrig, J. Wieferink, I.Y. Zhang, M. Ropo, X. Ren, P. Rinke, M. Scheffler, and V. Blum, *New J. Phys.* **17**, 093020 (2015)
- [77] O. Vahtras, J. Almlöf, and M.W. Feyereisen, *Chem. Phys. Lett.* **213**, 514 (1993)
- [78] M. Beuerle and C. Ochsenfeld, *J. Chem. Phys.* **149**, 244111 (2018)
- [79] J.F. Dobson, in M.P. Das (ed.): *Topics in Condensed Matter Physics* (Nova, New York, 1994)
- [80] K.T. Tang and J.P. Toennies, *J. Chem. Phys.* **118**, 4976 (2003)
- [81] T. Takatani, E.G. Hohenstein, M. Malagoli, M.S. Marshall, and C.D. Sherrill, *J. Chem. Phys.* **132**, 144104 (2010)
- [82] P. Jurečka, J. Šponer, J. Černý, and P. Hobza, *Phys. Chem. Chem. Phys.* **8**, 1985 (2006)
- [83] H. Eshuis and F. Furche, *J. Chem. Phys.* **136**, 084105 (2012)
- [84] J. Paier, X. Ren, P. Rinke, G.E. Scuseria, A. Grüneis, G. Kresse, and M. Scheffler, *New J. Phys.* **14**, 043002 (2012)
- [85] H. Steininger, S. Lehwald, and H. Ibach, *Surf. Sci.* **123**, 264 (1982)
- [86] G.S. Blackman, M.L. Xu, D.F. Ogletree, M.A.V. Hove, and G.A. Somorjai, *Phys. Rev. Lett.* **61**, 2352 (1988)
- [87] R. Armiento and A.E. Mattsson, *Phys. Rev. B* **72**, 085108 (2005)
- [88] J.P. Perdew, M. Ernzerhof, and K. Burke, *J. Chem. Phys.* **105**, 9982 (1996)
- [89] F. Göttl and J. Hafner, *J. Chem. Phys.* **134**, 064102 (2011)
- [90] H.J. Kim, A. Tkatchenko, J.H. Cho, and M. Scheffler, *Phys. Rev. B* **85**, 041403 (2012)
- [91] K. Held, A. McMahan, and R. Scalettar, *Phys. Rev. Lett.* **87**, 1 (2001)
- [92] B. Amadon, S. Biermann, A. Georges, and F. Aryasetiawan, *Phys. Rev. Lett.* **96**, 1 (2006)

- [93] N. Sengupta, J.E. Bates, and A. Ruzsinszky, *Phys. Rev. B* **97**, 235136 (2018)
- [94] A. Marini, P. García-González, and A. Rubio, *Phys. Rev. Lett* **96**, 136404 (2006)
- [95] J. Rekkedal, S. Coriani, M.F. Iozzi, A.M. Teale, T. Helgaker, and T.B. Pedersen, *J. Chem. Phys.* **139**, 081101 (2013)
- [96] A.M. Burow, J.E. Bates, F. Furche, and H. Eshuis, *J. Chem. Theory Comput.* **10**, 180 (2014)
- [97] B. Mussard, P.G. Szalay, and J.G. Ángyán, *J. Chem. Theory. Comput.* **10**, 1968 (2014)
- [98] B. Ramberger, T. Schäfer, and G. Kresse, *Phys. Rev. Lett.* **118**, 106403 (2017)
- [99] D. Neuhauser, E. Rabani, and R. Baer, *J. Phys. Chem. Lett.* **4**, 1172 (2013)
- [100] J.E. Moussa, *J. Chem. Phys.* **140**, 014107 (2014)
- [101] M. Kállay, *J. Chem. Phys.* **142**, 204105 (2015)
- [102] J. Wilhelm, P. Seewald, M. Del Ben, and J. Hutter, *J. Chem. Theory Comput.* **12**, 5851 (2016)
- [103] D. Graf, M. Beuerle, H.F. Schurkus, A. Luenser, G. Savasci, and C. Ochsenfeld, *J. Chem. Theory Comput.* **14**, 2505 (2018)
- [104] H. van Aggelen, S.N. Steinmann, D. Peng, and W. Yang, *Phys. Rev. A* **88**, 030501(R) (2013)
- [105] D. Peng, S.N. Steinmann, H. van Aggelen, and W. Yang, *J. Chem. Phys.* **139**, 103112 (2013)
- [106] Y. Yang, H. van Aggelen, S.N. Steinmann, D. Peng, and W. Yang, *J. Chem. Phys.* **139**, 174110 (2013)
- [107] E. Scuseria, T.M. Henderson, and I.W. Bulik, *J. Chem. Phys.* **139**, 104113 (2013)
- [108] J.J. Shepherd, T.M. Henderson, and G.E. Scuseria, *Phys. Rev. Lett.* **112**, 133002 (2014)
- [109] M.N. Tahir and X. Ren, *Phys. Rev. B* **99**, 195149 (2019)
- [110] P. Verma and R.J. Bartlett, *The Journal of Chemical Physics* **136**, 044105 (2012)
- [111] P. Bleiziffer, A. Hesselmann, and A. Görling, *J. Chem. Phys.* **139**, 084113 (2013)
- [112] M. Hellgren, F. Caruso, D.R. Rohr, X. Ren, A. Rubio, M. Scheffler, and P. Rinke, *Phys. Rev. B* **91**, 165110 (2015)
- [113] Y. Jin, D. Zhang, Z. Chen, N.Q. Su, and W. Yang, *J. Phys. Chem. Lett.* **8**, 4746 (2017)
- [114] V.K. Voora, S.G. Balasubramani, and F. Furche, *Phys. Rev. A* **99**, 012518 (2019)
- [115] I.Y. Zhang and X. Xu, *J. Phys. Chem. Lett.* **10**, 2617 (2019)

3 Introduction to Variational and Projector Monte Carlo

Cyrus J. Umrigar

Laboratory of Atomic and Solid State Physics

Physics Department, Cornell University

Ithaca, NY 14853, USA

Contents

1	Introduction	2
2	QMC in a nutshell	2
2.1	Variational Monte Carlo	3
2.2	Projector Monte Carlo	4
3	Variational Monte Carlo	5
3.1	Metropolis-Hastings method	5
4	Projector Monte Carlo	10
4.1	Importance sampling	11
4.2	Branching random walks	12
4.3	Taxonomy of PMC methods	12
4.4	Diffusion Monte Carlo	13
4.5	Sign problem	15
5	Form of trial wavefunctions	19
5.1	Slater-Jastrow wavefunctions	19
5.2	Symmetry-projected broken-symmetry mean-field wavefunctions	20
5.3	Backflow wavefunctions	21
5.4	Wavefunctions in orbital-space QMC	21
6	Optimization of trial wavefunctions	21
6.1	Newton method	23
6.2	Linear method	28
6.3	Stochastic reconfiguration method	28
7	Outlook	29

1 Introduction

Quantum Monte Carlo (QMC) methods are a broad and versatile class of stochastic algorithms for solving the many-body Schrödinger equation. They have been applied to fermions and bosons at zero and finite temperatures. Zero temperature calculations include both ground and excited states. Application areas range from quantum chemistry to lattice models of correlated materials to nuclear physics. Although the various methods differ greatly in their details, there is a small core of underlying ideas that are common to all the methods, which we will emphasize in this lecture. This lecture will be limited to zero temperature methods for fermionic systems. The emphasis will be on providing a unified description of both variational and projector Monte Carlo methods, both in discrete and in continuous space. The wide range of QMC applications will not be discussed. The finite-temperature path-integral Monte Carlo method has been reviewed in Ref. [1]. Details of zero-temperature methods, which we do not have time for here, can be found in the original papers, in books and review articles [2–8], and in chapters by Becca, Lüchow, Prokof'ev, Sandvik, and Zhang in this volume.

The many-body Schrödinger equation can be solved straightforwardly by expanding the wavefunction in a linear combination of determinants of single-particle orbitals, a method that is known as full configuration interaction. The limitation is that the number of states scales combinatorially in the number of orbitals N_{orb} , and the number of electrons $N = N_{\uparrow} + N_{\downarrow}$, as $N_{\text{orb}} C_{N_{\uparrow}} \times N_{\text{orb}} C_{N_{\downarrow}}$, where $N_{\uparrow}$, $N_{\downarrow}$ are the number of up- and down-spin electrons respectively, so this brute-force method can be applied only to tiny systems. In contrast, some of the QMC methods scale as a low-order polynomial in N , provided that an approximate solution whose accuracy depends on the quality of a trial wavefunction is acceptable. Frequently high quality trial wavefunctions can be constructed, making QMC one of the few methods that provide accurate solutions for difficult problems.

2 QMC in a nutshell

I will distinguish between *QMC simulations* and *QMC calculations*, although almost all other researchers use these terms interchangeably. To my mind the stochasticity in *QMC simulations* mimics the stochasticity of the experimental system, e.g., the diffusion of neutrons in a nuclear reactor. On the other hand, *QMC calculations*, which is what we will be discussing here, introduce stochasticity into problems that are in fact deterministic. The solution of the many-body Schrödinger equation is perfectly deterministic (not to be confused with the probabilistic interpretation of the wavefunction amplitude) so it is in fact rather remarkable that introducing stochasticity makes the problem more tractable.

QMC methods are most useful when the dimension of the Hilbert space is very large and other many-body methods become impractical. They can be used both when the state of the system is described by discrete degrees of freedom, e.g., spin states or expansion coefficients of the wavefunction in a finite basis, as well as when the state is described by continuous degrees of freedom, e.g., the wavefunction amplitudes as a function of $3N$ –dimensional electron coordi-

nates, in which case the Hilbert space is in fact infinite! The basic ideas of QMC are the same in discrete and continuous systems, so I will provide a unified treatment of both cases and will use discrete and continuous notation (e.g., sums and integrals) interchangeably.

2.1 Variational Monte Carlo

Many QMC methods rely on having good approximate wavefunctions to improve their accuracy and efficiency. So, we consider three wavefunctions: the exact wavefunction $|\Psi_0\rangle$ and two approximate wavefunctions which we call the trial wavefunction $|\Psi_T\rangle$ and the guiding wavefunction $|\Psi_G\rangle$: Their expansions in a (complete or incomplete) basis of N_{st} states are

$$\text{Exact} \quad |\Psi_0\rangle = \sum_i^{N_{\text{st}}} e_i |\phi_i\rangle, \quad \text{where} \quad e_i = \langle \phi_i | \Psi_0 \rangle \quad (1)$$

$$\text{Trial} \quad |\Psi_T\rangle = \sum_i^{N_{\text{st}}} t_i |\phi_i\rangle, \quad \text{where} \quad t_i = \langle \phi_i | \Psi_T \rangle \quad (2)$$

$$\text{Guiding} \quad |\Psi_G\rangle = \sum_i^{N_{\text{st}}} g_i |\phi_i\rangle, \quad \text{where} \quad g_i = \langle \phi_i | \Psi_G \rangle \quad (3)$$

If the basis is incomplete then “exact” should be construed to mean “exact in that basis.” Ψ_T and Ψ_G are frequently chosen to be the same function, but they serve two different purposes and at times there are good reasons for choosing them to be different, as will become apparent shortly. The basis state indices may be discrete (e.g., determinants of single-particle orbitals) or continuous (e.g., the $3N$ coordinates of the N electrons). In either case one can use $|\Psi_T\rangle$ to define the “local energy” of that state

$$E_L(i) = \frac{\sum_j^{N_{\text{st}}} H_{ij} t_j}{t_i}. \quad (4)$$

The variational energy of $|\Psi_T\rangle$ can be evaluated as follows:

$$\begin{aligned} E_V &= \frac{\langle \Psi_T | \hat{H} | \Psi_T \rangle}{\langle \Psi_T | \Psi_T \rangle} = \frac{\sum_{ij}^{N_{\text{st}}} \langle \Psi_T | \phi_i \rangle \langle \phi_i | \hat{H} | \phi_j \rangle \langle \phi_j | \Psi_T \rangle}{\sum_i^{N_{\text{st}}} \langle \Psi_T | \phi_i \rangle \langle \phi_i | \Psi_T \rangle} \\ &= \frac{\sum_{ij}^{N_{\text{st}}} t_i H_{ij} t_j}{\sum_k^{N_{\text{st}}} t_k^2} = \sum_i^{N_{\text{st}}} \frac{t_i^2}{\sum_k^{N_{\text{st}}} t_k^2} \frac{\sum_j^{N_{\text{st}}} H_{ij} t_j}{t_i} = \sum_i^{N_{\text{st}}} \frac{t_i^2}{\sum_k^{N_{\text{st}}} t_k^2} E_L(i) \\ &\approx \frac{\left[\sum_i^{N_{\text{MC}}} E_L(i) \right]_{\Psi_T^2}}{N_{\text{MC}}} \xrightarrow{\Psi_G \neq \Psi_T} \frac{\left[\sum_i^{N_{\text{MC}}} \left(\frac{t_i}{g_i} \right)^2 E_L(i) \right]_{\Psi_G^2}}{\left[\sum_k^{N_{\text{MC}}} \left(\frac{t_k}{g_k} \right)^2 \right]_{\Psi_G^2}} \end{aligned} \quad (5)$$

In the last line, we switch from the sum over all states to a sum over states sampled with probability $g_i^2 / \sum_k^{N_{\text{st}}} g_k^2$, first for $\Psi_G = \Psi_T$ (the usual case) and then for $\Psi_G \neq \Psi_T$. In the limit

that $\Psi_T \rightarrow \Psi_0$, $E_{L,i} \rightarrow E_0$, the exact energy, independent of i . It is now apparent that the requirements on $|\Psi_T\rangle$ and $|\Psi_G\rangle$ are rather different. We wish to employ a $|\Psi_T\rangle$ which is not only a good approximation to $|\Psi_0\rangle$ (i.e., one that gives a low E_V and has small rms fluctuations in E_L), but also one for which E_L can be evaluated efficiently. For a discrete basis, this means that $\hat{H}$ should be very sparse in that basis, whereas for the continuous real space basis this requires that $\langle \phi_k | \Psi_T \rangle$ and its Laplacian (needed for calculating the kinetic energy) can be evaluated efficiently. Instead the requirements on $|\Psi_G\rangle$ are that $\langle \phi_i | \Psi_G \rangle$ can be evaluated efficiently and that it is nonzero whenever $\langle \phi_i | \Psi_0 \rangle$ is nonzero, since otherwise the expectation values would be biased relative to the actual variational value.

Hence the minimal ingredients for an accurate and efficient VMC algorithm are:

1. A method (Metropolis-Hastings) for sampling $\langle \phi_k | \Psi_T \rangle^2$.
2. A $|\Psi_T\rangle$ with variational parameters that is flexible enough to be a good approximation to $|\Psi_0\rangle$ for optimized parameters, and for which $\langle \phi_k | \Psi_T \rangle$ and E_L can be evaluated efficiently.
3. Robust and efficient methods for optimizing the variational parameters.

These ingredients will be discussed in Secs. 3, 5, and 6 respectively.

2.2 Projector Monte Carlo

Projector Monte Carlo (PMC) methods evaluate the true energy E_0 (in the absence of a “sign problem”) rather than the variational energy E_V using a “mixed estimator,” $\langle \Psi_0 | \hat{H} | \Psi_T \rangle / \langle \Psi_0 | \Psi_T \rangle$, for the energy. Following almost the same steps as in VMC

$$\begin{aligned}
 E_0 &= \frac{\langle \Psi_0 | \hat{H} | \Psi_0 \rangle}{\langle \Psi_0 | \Psi_0 \rangle} = \frac{\langle \Psi_0 | \hat{H} | \Psi_T \rangle}{\langle \Psi_0 | \Psi_T \rangle} = \frac{\sum_{ij}^{N_{st}} \langle \Psi_0 | \phi_i \rangle \langle \phi_i | \hat{H} | \phi_j \rangle \langle \phi_j | \Psi_T \rangle}{\sum_k^{N_{st}} \langle \Psi_0 | \phi_k \rangle \langle \phi_k | \Psi_T \rangle} \\
 &= \frac{\sum_{ij}^{N_{st}} e_i H_{ij} t_j}{\sum_k^{N_{st}} e_k t_k} = \sum_i^{N_{st}} \frac{e_i t_i}{\sum_k^{N_{st}} e_k t_k} \frac{\sum_j^{N_{st}} H_{ij} t_j}{t_i} = \sum_i^{N_{st}} \frac{e_i t_i}{\sum_k^{N_{st}} e_k t_k} E_L(i) \\
 &= \frac{\left[\sum_i^{N_{MC}} E_L(i) \right]_{\Psi_T \Psi_0}}{N_{MC}} \xrightarrow{\Psi_G \neq \Psi_T} \frac{\left[\sum_i^{N_{MC}} \left(\frac{t_i}{g_i} \right) E_L(i) \right]_{\Psi_G \Psi_0}}{\left[\sum_k^{N_{MC}} \left(\frac{t_k}{g_k} \right) \right]_{\Psi_G \Psi_0}} \quad (6)
 \end{aligned}$$

The only difference between Eqs. (6) and (5) is that the sampled distribution is $e_i g_i / \sum_k^{N_{st}} e_k g_k$ rather than $g_i^2 / \sum_k^{N_{st}} g_k^2$. At first sight this seems like an insurmountable obstacle since $|\Psi_G\rangle$ is known, but $|\Psi_0\rangle$ is not. In Sec. 4 we describe methods for sampling the mixed distribution needed for PMC. The properties required of $|\Psi_T\rangle$ and $|\Psi_G\rangle$ are exactly the same as for VMC. Note that although an unbiased energy is obtained in the absence of a sign problem regardless of $|\Psi_T\rangle$ and $|\Psi_G\rangle$, the statistical error depends on $|\Psi_T\rangle$ and $|\Psi_G\rangle$. When the sign problem is present, approximations will be needed and their accuracy and efficiency depend on $|\Psi_T\rangle$ and $|\Psi_G\rangle$.

3 Variational Monte Carlo

3.1 Metropolis-Hastings method

As discussed in Sec. 2.1 we need to sample with probability $g_i^2 / \sum_k^{N_{\text{st}}} g_k^2$. To familiarize the reader with both discrete and continuous notation, we now assume we are working in continuous real space, in which case we need to sample from the probability density function $\rho(\mathbf{R}) = \psi_{\text{T}}^2(\mathbf{R}) / \int d\mathbf{R} \psi_{\text{T}}^2(\mathbf{R})$, where we assume that the choice $|\Psi_{\text{G}}\rangle = |\Psi_{\text{T}}\rangle$ has been made since that is the usual practice. It is possible to sample a large number of probability density functions [9] directly (i.e., with no serial correlations) using the transformation or the rejection methods, but not such complicated probability density functions, so instead the Metropolis-Hastings method [10, 11] is used. It was originally developed to sample the thermal Boltzmann distribution, but is in fact a very powerful method for sampling any *known* discrete or continuous distribution. (In Sec. 4 we will see how to sample the unknown distribution $\psi_{\text{T}}\psi_0 / \int d\mathbf{R} \psi_{\text{T}}\psi_0$!)

3.1.1 Markov chains

The Metropolis-Hastings method employs a *Markov chain*. A Markov chain is specified by two ingredients:

- 1) an initial state
- 2) a transition matrix $M(\mathbf{R}_{\text{f}}|\mathbf{R}_{\text{i}})$ (probability of transition from $\mathbf{R}_{\text{i}} \rightarrow \mathbf{R}_{\text{f}}$.) with the properties

$$M(\mathbf{R}_{\text{f}}|\mathbf{R}_{\text{i}}) \geq 0, \quad \text{and} \quad \int d\mathbf{R}_{\text{f}} M(\mathbf{R}_{\text{f}}|\mathbf{R}_{\text{i}}) = 1. \quad (\text{Column-stochastic matrix}) \quad (7)$$

The first property expresses the fact that probabilities must be non negative. The second property expresses the fact that a point at $\mathbf{R}_{\text{i}}$ must go somewhere at the next step. The eigenvalues of a column-stochastic matrix are between 0 and 1, and there is at least one eigenvalue equal to 1 since the vector with all components equal to one is a left eigenvector with eigenvalue 1, and the left and right eigenvalues of any matrix are the same. If in addition M is a “primitive matrix,” i.e., there exists an integer n for which all the elements of M^n are strictly positive, then there is a *unique* eigenvector with eigenvalue 1 and the Markov chain is said to be “ergodic.”¹

We wish to choose an M such that repeated application of M results in sampling $\rho(\mathbf{R})$, so we choose an M that satisfies

$$\int d\mathbf{R}_{\text{f}} M(\mathbf{R}_{\text{i}}|\mathbf{R}_{\text{f}}) \rho(\mathbf{R}_{\text{f}}) = \rho(\mathbf{R}_{\text{i}}) = \int d\mathbf{R}_{\text{f}} M(\mathbf{R}_{\text{f}}|\mathbf{R}_{\text{i}}) \rho(\mathbf{R}_{\text{i}}) \quad \forall \mathbf{R}_{\text{i}} \quad (8)$$

The first equality expresses a *stationarity condition*, namely that the net flux in and out $\mathbf{R}_{\text{i}}$ is zero. Hence if one starts with the correct distribution $\rho(\mathbf{R})$, repeated application of M will continue to sample from $\rho(\mathbf{R})$. The second equality follows from the definition of a stochastic matrix. Eq. (8) shows that $\rho(\mathbf{R})$ is a right eigenvector with eigenvalue 1. Since all the other

¹Here we ignore some subtleties that arise when the space is infinite.

eigenvalues are between 0 and 1, but not equal to 1, $\rho(\mathbf{R})$ is the dominant right eigenvector of M and repeated application of M results eventually in sampling $\rho(\mathbf{R})$.

In practice, the length of Monte Carlo (MC) run should be long enough that there is a significant probability of the system making several transitions between the neighborhoods of any pair of representative states that make a significant contribution to the average. This ensures that states are visited with the correct probability with only small statistical fluctuations. For example in a double-well system many transitions between the two wells should occur, but we can choose our Markov matrix to achieve this even if the barrier between wells is high.

A drawback of the Metropolis-Hastings method is that the sampled states are serially correlated. The rate at which the initial density evolves to the desired density ρ and the autocorrelation time of estimates of various observables is governed by the subdominant eigenvalues. In the ideal situation all the other eigenvalues are zero and every sample is independent.

Construction of M : Detailed balance condition We have as yet not provided a prescription to construct M , such that ρ is its stationary state. To do this we impose the *detailed balance* condition

$$M(\mathbf{R}_f|\mathbf{R}_i) \rho(\mathbf{R}_i) = M(\mathbf{R}_i|\mathbf{R}_f) \rho(\mathbf{R}_f) \quad (9)$$

The detailed balance condition is more stringent than the stationarity condition. Instead of requiring that the net flux in and out of each state is zero, it requires that the net flux between *every pair* of states is zero. It is a sufficient, not a necessary condition, but it provides a practical way to construct M .

To go further, it is convenient employ a 2-step process. Moves from $\mathbf{R}_i$ to a provisional point $\mathbf{R}'_f$ are proposed with probability $T(\mathbf{R}'_f|\mathbf{R}_i)$ and then accepted with probability $A(\mathbf{R}'_f|\mathbf{R}_i)$. The corresponding Markov matrix is

$$M(\mathbf{R}_f|\mathbf{R}_i) = \begin{cases} A(\mathbf{R}_f|\mathbf{R}_i) T(\mathbf{R}_f|\mathbf{R}_i) & \text{if } \mathbf{R}_f \neq \mathbf{R}_i \\ 1 - \int d\mathbf{R}'_f A(\mathbf{R}'_f|\mathbf{R}_i) T(\mathbf{R}'_f|\mathbf{R}_i) & \text{if } \mathbf{R}_f = \mathbf{R}_i \end{cases} \quad (10)$$

$M(\mathbf{R}_f|\mathbf{R}_i)$ and $T(\mathbf{R}_f|\mathbf{R}_i)$ are stochastic matrices, but $A(\mathbf{R}_f|\mathbf{R}_i)$ is not. The detailed balance condition now becomes

$$\begin{aligned} A(\mathbf{R}_f|\mathbf{R}_i) T(\mathbf{R}_f|\mathbf{R}_i) \rho(\mathbf{R}_i) &= A(\mathbf{R}_i|\mathbf{R}_f) T(\mathbf{R}_i|\mathbf{R}_f) \rho(\mathbf{R}_f) \\ \text{i.e. } \frac{A(\mathbf{R}_f|\mathbf{R}_i)}{A(\mathbf{R}_i|\mathbf{R}_f)} &= \frac{T(\mathbf{R}_i|\mathbf{R}_f)}{T(\mathbf{R}_f|\mathbf{R}_i)} \frac{\rho(\mathbf{R}_f)}{\rho(\mathbf{R}_i)}. \end{aligned} \quad (11)$$

3.1.2 Choice of acceptance matrix

To satisfy Eq. (11) we can choose

$$A(\mathbf{R}_f|\mathbf{R}_i) = F \left(\frac{T(\mathbf{R}_i|\mathbf{R}_f)}{T(\mathbf{R}_f|\mathbf{R}_i)} \frac{\rho(\mathbf{R}_f)}{\rho(\mathbf{R}_i)} \right) \quad (12)$$

where F is any function for which $F(x)/F(1/x) = x$ and $0 \leq F(x) \leq 1$. Two possible choices are $F(x) = x/(1+x)$ and $F(x) = \min\{1, x\}$, corresponding to

$$A(\mathbf{R}_f|\mathbf{R}_i) = \frac{T(\mathbf{R}_i|\mathbf{R}_f) \rho(\mathbf{R}_f)}{T(\mathbf{R}_i|\mathbf{R}_f) \rho(\mathbf{R}_f) + T(\mathbf{R}_f|\mathbf{R}_i) \rho(\mathbf{R}_i)}, \quad (13)$$

$$\text{and } A(\mathbf{R}_f|\mathbf{R}_i) = \min \left\{ 1, \frac{T(\mathbf{R}_i|\mathbf{R}_f)}{T(\mathbf{R}_f|\mathbf{R}_i)} \frac{\rho(\mathbf{R}_f)}{\rho(\mathbf{R}_i)} \right\}. \quad (14)$$

The latter choice is the optimal choice since it maximizes the acceptance for given ρ and T , and is the choice made in the Metropolis et al. and the Hastings papers. Actually, Metropolis et al. assumed that $T(\mathbf{R}_i|\mathbf{R}_f) = T(\mathbf{R}_f|\mathbf{R}_i)$ in which case the factors of T drop out of Eq. (14), and Hastings made the generalization to $T(\mathbf{R}_i|\mathbf{R}_f) \neq T(\mathbf{R}_f|\mathbf{R}_i)$, which enables a more efficient algorithm.

3.1.3 Choice of proposal matrix T

The optimal choice for the acceptance matrix $A(\mathbf{R}_f|\mathbf{R}_i)$ is straightforward, Eq. (14), but there is considerable scope for using one's ingenuity to come up with good proposal matrices, $T(\mathbf{R}_f|\mathbf{R}_i)$, that give small serial correlations between samples. As mentioned before, the ideal choice of $M(\mathbf{R}_f|\mathbf{R}_i)$ has one eigenvalue equal to 1 and the rest zero. However, in practice the way to find efficient choices for $T(\mathbf{R}_f|\mathbf{R}_i)$ is not to think about eigenvalues, but instead to think about choosing a $T(\mathbf{R}_f|\mathbf{R}_i)$ that has large proposed moves and at the same time has high acceptance probabilities, i.e.,

$$\frac{T(\mathbf{R}_i|\mathbf{R}_f)}{T(\mathbf{R}_f|\mathbf{R}_i)} \frac{\rho(\mathbf{R}_f)}{\rho(\mathbf{R}_i)} \approx 1. \quad (15)$$

There is a great deal of freedom in the choice of $T(\mathbf{R}_f|\mathbf{R}_i)$, the only constraints being that it is a stochastic matrix leading to an ergodic Markov chain, and that it must be possible to efficiently sample $T(\mathbf{R}_f|\mathbf{R}_i)$ with a direct sampling method.

It may appear from Eq. (15) that our goal should be to make $T(\mathbf{R}_f|\mathbf{R}_i) \propto \rho(\mathbf{R}_f)$ since in that case the various factors cancel and the product becomes 1. This is in fact the case if it is possible to achieve that condition over all space, but it is not—if it were possible, we would not be using Metropolis-Hastings in the first place. So, we will discuss an alternative goal in a moment.

In order to prevent the acceptance from getting too small, it is common practice to restrict the moves to be in the neighborhood of $\mathbf{R}_i$ by choosing $T(\mathbf{R}_f|\mathbf{R}_i)$ to be non-zero, or at least not negligible, only within a domain $D(\mathbf{R}_i)$ of volume $\Omega(\mathbf{R}_i)$ around $\mathbf{R}_i$. For a given functional form of $T(\mathbf{R}_f|\mathbf{R}_i)$ the acceptance decreases as $\Omega(\mathbf{R}_i)$ increases, so, there exists an optimal $\Omega(\mathbf{R}_i)$ for which the system evolves the fastest.

To make further progress, we now make explicit that $T(\mathbf{R}_f|\mathbf{R}_i)$ is an ergodic matrix by writing

$$T(\mathbf{R}_f|\mathbf{R}_i) = \frac{S(\mathbf{R}_f|\mathbf{R}_i)}{\int d\mathbf{R}_f S(\mathbf{R}_f|\mathbf{R}_i)} \approx \frac{S(\mathbf{R}_f|\mathbf{R}_i)}{S(\mathbf{R}_i|\mathbf{R}_i)\Omega(\mathbf{R}_i)}, \quad (16)$$

where $S(\mathbf{R}_f|\mathbf{R}_i)$ is non-zero only in the domain $D(\mathbf{R}_i)$. Then,

$$\frac{A(\mathbf{R}_f, \mathbf{R}_i)}{A(\mathbf{R}_i, \mathbf{R}_f)} = \frac{T(\mathbf{R}_i|\mathbf{R}_f)}{T(\mathbf{R}_f|\mathbf{R}_i)} \frac{\rho(\mathbf{R}_f)}{\rho(\mathbf{R}_i)} \approx \frac{\Omega(\mathbf{R}_i)}{\Omega(\mathbf{R}_f)} \frac{S(\mathbf{R}_i|\mathbf{R}_i)}{S(\mathbf{R}_f|\mathbf{R}_f)} \frac{S(\mathbf{R}_i|\mathbf{R}_f)}{S(\mathbf{R}_f|\mathbf{R}_i)} \frac{\rho(\mathbf{R}_f)}{\rho(\mathbf{R}_i)}. \quad (17)$$

Noting that for the present purpose S should be viewed as a function of the left index only, it is apparent that the choice

$$S(\mathbf{R}_f|\mathbf{R}_i) \propto \sqrt{\rho(\mathbf{R}_f)/\Omega(\mathbf{R}_f)} \quad \text{yields} \quad A(\mathbf{R}_f, \mathbf{R}_i)/A(\mathbf{R}_i, \mathbf{R}_f) \approx 1. \quad (18)$$

To be more precise, if the log-derivatives of $S(\mathbf{R}_f|\mathbf{R}_i)$ equal those of $\sqrt{\rho(\mathbf{R}_f)/\Omega(\mathbf{R}_f)}$ at $\mathbf{R}_f = \mathbf{R}_i$, the average acceptance goes as $1 - \mathcal{O}(\Delta^3)$, where Δ is the linear dimension of $D(\mathbf{R}_i)$, provided that $D(\mathbf{R}_i)$ is inversion symmetric about $\mathbf{R}_i$.

Another good choice for $T(\mathbf{R}_f|\mathbf{R}_i)$, motivated by the diffusion Monte Carlo algorithm discussed in Sec. 4, is

$$T(\mathbf{R}_f|\mathbf{R}_i) = \frac{1}{(2\pi\tau)^{3N/2}} \exp \left[-\frac{(\mathbf{R}_f - \mathbf{R}_i - \mathbf{V}(\mathbf{R}_i)\tau)^2}{2\tau} \right], \quad (19)$$

where $\mathbf{V}(\mathbf{R}_i) = \nabla\Psi(\mathbf{R})/\Psi(\mathbf{R})|_{\mathbf{R}=\mathbf{R}_i}$ is called the *drift velocity* of the wave function and τ is the time step which can be adjusted so as to minimize the autocorrelation time of the local energy. This is sampled by choosing

$$\mathbf{R}_f = \mathbf{R}_i + \mathbf{V}(\mathbf{R}_i)\tau + \boldsymbol{\eta}, \quad (20)$$

where $\boldsymbol{\eta}$ is a vector of $3N$ random numbers drawn from the Gaussian distribution with average 0 and standard deviation $\sqrt{\tau}$.

We demonstrate the increase in the acceptance probability that can be achieved by using Eq. (18) or Eq. (19) with a simple one-dimensional example. We wish to sample $\Psi(\mathbf{R})^2$. The simplest choice for $T(\mathbf{R}_f|\mathbf{R}_i)$ is a uniform distribution in $\Omega(\mathbf{R}_i)$ specified by $|\mathbf{R}_f - \mathbf{R}_i| < \Delta$ and zero outside

$$T(\mathbf{R}_f|\mathbf{R}_i) = \begin{cases} \frac{1}{2\Delta} & \text{if } \mathbf{R}_f \in \Omega(\mathbf{R}_i), \\ 0 & \text{elsewhere.} \end{cases} \quad (21)$$

Instead, our recommended prescription from Eq. (18) is

$$T(\mathbf{R}_f|\mathbf{R}_i) = \begin{cases} \frac{1}{2\Delta} \left(1 + \frac{\nabla\Psi(\mathbf{R})}{\Psi(\mathbf{R})} \bigg|_{\mathbf{R}=\mathbf{R}_i} (\mathbf{R}_f - \mathbf{R}_i) \right) & \text{if } \mathbf{R}_f \in \Omega(\mathbf{R}_i), \\ 0 & \text{elsewhere} \end{cases} \quad (22)$$

and the prescription of Eq. (19) becomes

$$T(\mathbf{R}_f|\mathbf{R}_i) = \frac{1}{(2\pi\tau)^{1/2}} \exp \left[-\frac{(\mathbf{R}_f - \mathbf{R}_i - V(\mathbf{R}_i)\tau)^2}{2\tau} \right]. \quad (23)$$

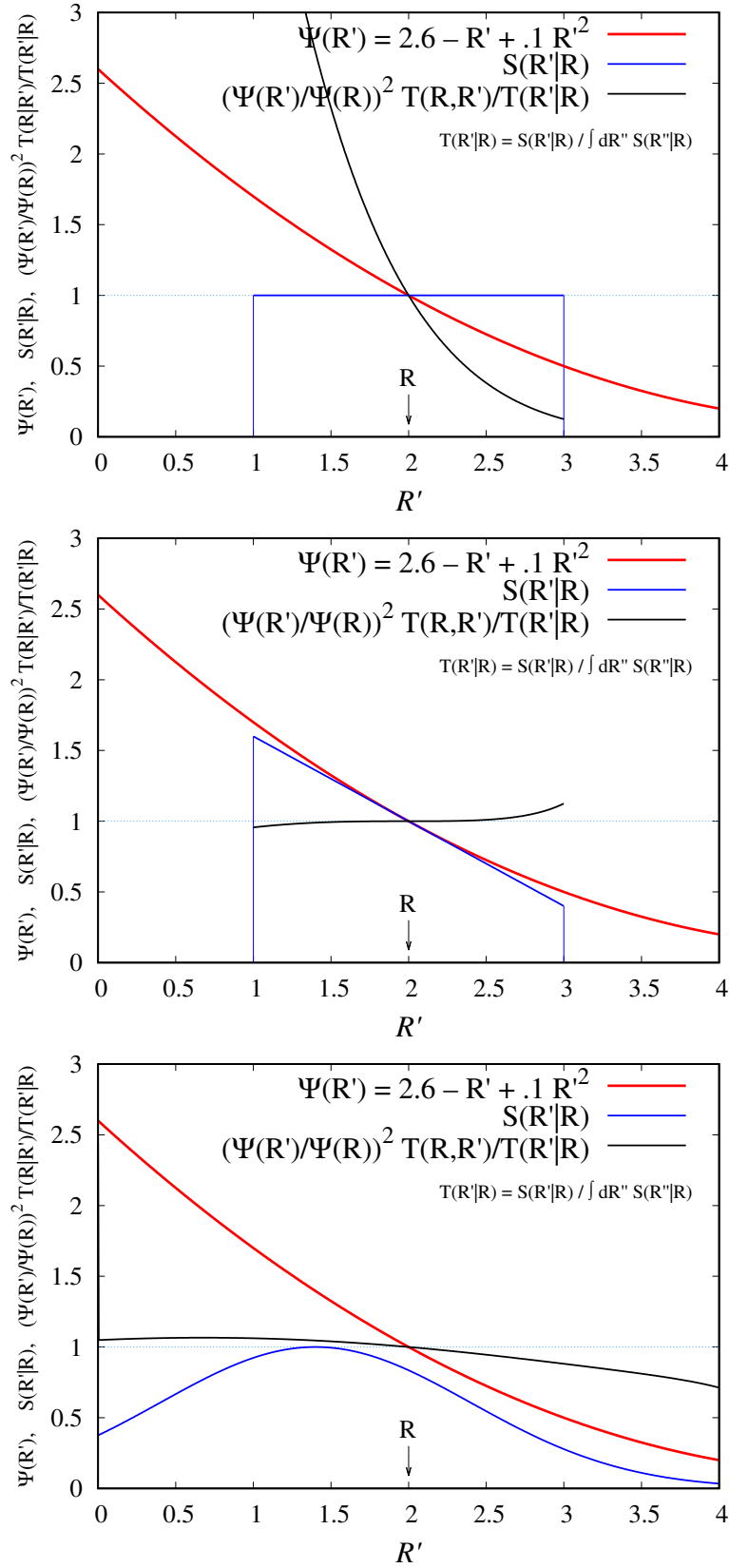


Fig. 1: Plots of $\Psi(R')$, $S(R'|R)$, $\left(\frac{\Psi(R')}{\Psi(R)}\right)^2 \frac{T(R, R')}{T(R', R)}$ for the three choices of $T(R', R)$ in Eqs. (21), (22), (23). The lower two plots have much larger average acceptances than the top.

In Fig. 1 we show plots of $\Psi(R')$, $S(R'|R)$, and $(\Psi(R')/\Psi(R))^2 T(R, R')/T(R', R)$ for the three different choices of $T(R', R)$ in Eqs. (21), (22), and (23). In the top plot $(\frac{\Psi(R')}{\Psi(R)})^2 \frac{T(R, R')}{T(R', R)}$ deviates linearly in $R' - R$ from 1 in the vicinity of $R' = R$ and deviates greatly from 1 over much of $\Omega(R)$, in the middle plot it deviates cubically in $R' - R$ from 1 in the vicinity of $R' = R$ and is close to 1 in all of $\Omega(R)$, and in the bottom plot it deviates linearly in $R' - R$ from 1 but stays close to 1 over a fairly wide range. Hence the choices of $T(\mathbf{R}_f|\mathbf{R}_i)$ in Eq. (18) and (19) result in much smaller autocorrelation times than the simple choice of a symmetric $T(\mathbf{R}_f|\mathbf{R}_i)$.

The analysis in Eqs. (16) to (19) and the examples in Eqs. (21) to (23) assume that $\rho(\mathbf{R})=\Psi(\mathbf{R})^2$ has continuous derivatives. In reality, $\Psi(\mathbf{R})^2$ has derivative discontinuities at electron-nucleus and electron-electron coincidence points. The former are much more important than the latter, since the wavefunction is large there and the velocity is directed towards the nucleus, so the electron tends to overshoot the nucleus. So, there is a high likelihood of having electrons near the nuclei and their acceptances are likely to be small. To tackle this problem, a highly efficient VMC algorithm with autocorrelation times of the local energy close to 1 is presented in Refs. [12, 13] wherein the electron moves are made in spherical coordinates centered on the nearest nucleus and the size of radial moves is proportional to the distance to the nearest nucleus. In addition, the size of the angular moves gets larger as one approaches a nucleus. This algorithm allows one to achieve, in many cases, an autocorrelation time of the local energy close to 1.

3.1.4 Moving the electrons all at once or one by one?

The accept-reject step can be performed after moving each electron, or after moving all the electrons. The former requires more computer time since the wavefunction and its gradient must be recalculated after each move. The increase in computer time is not a factor of N as one may naively expect, but more like a factor of 2. The reason is that it takes $\mathcal{O}(N^3)$ time to calculate a determinant from scratch, but only $\mathcal{O}(N^2)$ time to recalculate it, using the matrix determinant lemma and the Sherman-Morrison formula, after a single row or column has been changed. For systems with many electrons, moving the electrons one at a time leads to a more efficient algorithm because larger moves can be made for the same average acceptance, so the autocorrelation time of the local energy is smaller, more than compensating the increase of the calculation time per MC step.

4 Projector Monte Carlo

In Sec. 2 we wrote down the exact energy as a mixed estimator where the bra is the exact wavefunction and the ket the trial wavefunction, but we did not explain how one accesses the exact wavefunction, i.e., the ground state of the Hamiltonian. This is done using a projector. Projector Monte Carlo is a stochastic realization of the *power method* for finding the dominant eigenvector of a matrix. If one repeatedly multiplies an arbitrary eigenvector (not orthogonal to the dominant eigenvector) by the matrix, then one eventually gets the dominant eigenvector,

since at each step it gets multiplied by the largest in magnitude eigenvalue of the matrix. The power method is an example of an iterative eigensolver. Other iterative eigensolvers exist, such as the Lanczos method and the diagonally-preconditioned Davidson method, and they are much more efficient than the power method for deterministic calculations. So, if the dimension of the Hilbert space is not very large, say $< 10^{10}$ then one would just do a deterministic calculation using the Lanczos or Davidson methods. However, when the Hilbert space is so large (even infinite) that it is not possible to store even a few vectors of Hilbert space dimension, then PMC methods become the method of choice, since at any point in time they store only a sample of states.

The *projector* is any function of the Hamiltonian that maps the ground state eigenvalue of $\hat{H}$ to 1, and the higher eigenvalues of $\hat{H}$ to absolute values that are < 1 (preferably close to 0). We use the term “projector” somewhat loosely, since it is only repeated application of the projector that yields the desired state:

$$|\Psi_0\rangle = \lim_{n \rightarrow \infty} \hat{P}^n(\tau) |\Psi_T\rangle. \quad (24)$$

Possible choices for the projector are

$$\text{Exponential projector:} \quad \hat{P} = e^{\tau(E_T \hat{1} - \hat{H})} \quad (25)$$

$$\text{Linear projector:} \quad \hat{P} = \hat{1} + \tau(E_T \hat{1} - \hat{H}) \quad (\tau < 2/(E_{\max} - E_0)) \quad (26)$$

$$\text{Green function projector:} \quad \hat{P} = \frac{1}{\hat{1} - \tau(E_T \hat{1} - \hat{H})} \quad (27)$$

where E_T is an estimate of the ground state energy.

4.1 Importance sampling

The projectors above enable us to sample $e_i = \langle \phi_i | \Psi_0 \rangle$. However, according to Eq. (6) we want to sample from $g_i e_i = \langle \phi_i | \Psi_G \rangle \langle \phi_i | \Psi_0 \rangle$. Since

$$\sum_f P_{fi} e_i = e_f \quad (28)$$

the similarity transformed matrix with elements

$$\tilde{P}_{fi} = \frac{g_f P_{fi}}{g_i} \quad (29)$$

has an eigenstate with elements $g_i e_i$:

$$\sum_i \tilde{P}_{fi}(g_i e_i) = \sum_i \left(\frac{g_f P_{fi}}{g_i} \right) (g_i e_i) = g_f e_f. \quad (30)$$

$\hat{\tilde{P}}$ is called the *importance sampled* projector and it samples $\langle \phi_i | \Psi_G \rangle \langle \phi_i | \Psi_0 \rangle$.

4.2 Branching random walks

Note that unlike the Markov matrix in Eq. (10) used in the Metropolis-Hastings method, the columns of the projector are not normalized to 1. We can write $\tilde{P}_{fi}$ as

$$\tilde{P}_{fi} = \frac{\tilde{P}_{fi}}{\sum_f \tilde{P}_{fi}} \sum_f \tilde{P}_{fi} \equiv T_{fi} W_i \quad (31)$$

where now T_{fi} is a Markov matrix and W_i are multiplicative weights. So, instead of the unweighted Monte Carlo walk that we had when doing VMC, we now have *walkers* that are specified by not just their position and but also by their weight. At each Monte Carlo step, the weight gets multiplied by W_i . If we have a single weighted walker, then a few generations of the walk will dominate the averages and the computational effort expended on the rest of the walk would be largely wasted. It is possible to have a population of walkers of fluctuating population size, with each walker having unit weight, but this leads to unnecessary birth/death events. So, it is best to have a population of walkers, such that all walkers within a generation have roughly the same weight, say within a factor of 2, and birth/death events when the weights go outside that range. Even so, the weights of different generations will vary a lot in a sufficiently long run. So, efficiency demands that we exercise *population control* to make the weights of each generation approximately the same. The population control error is proportional to the inverse of the target population size N_{walk} . The error arises because of a negative correlation between the energy averaged over the generation and the weight of the generation. When the energy is low, the weight tends to be large and population control artificially reduces the weight and thereby creates a positive bias in the energy. Similarly, when the energy is high, the weight tends to be small and population control artificially increases the weight and this too creates a positive bias in the energy. Since the relative fluctuations in the energy and in the weight go as $1/\sqrt{N_{\text{walk}}}$, the relative fluctuations in their covariance goes as $1/N_{\text{walk}}$, resulting in a $\mathcal{O}(1/N_{\text{walk}})$ population control bias. So, one way to reduce the population control error is to simply use a large population, and this is what most people do. If one wishes to be sure that the error is sufficiently small, plot the energy versus $1/N_{\text{walk}}$ and take the limit $1/N_{\text{walk}} \rightarrow 0$. But there exists a better way that allows us to estimate and remove most of the population control error within a single run [14, 15].

4.3 Taxonomy of PMC methods

The various PMC methods can be characterized by a) the form of the projector, and, b) the space in which the walk is performed, i.e., the single-particle basis and the quantization used. By second quantization we mean that walkers are characterized by only the occupancy of the single-particle states, whereas by first quantization we mean that walkers are characterized by the occupancy and by which electrons are on which states. So, walkers that are related by electron permutations in a first quantized space map onto the same walker in a second quantized space. Table 1 is a taxonomy of some PMC methods. For example, DMC uses the exponential projector (imaginary-time propagator) with a 1st quantized continuous real space basis,

Method	Projector	SP Basis	Quantization
Diffusion Monte Carlo (DMC) [16, 17, 15]	$e^{\tau(E_T \hat{\mathbf{1}} - \hat{H})}$	$\mathbf{r}$	1^{st}
GFMC [18–20]	$e^{\tau(E_T \hat{\mathbf{1}} - \hat{H})}$ (samp. τ)	$\mathbf{r}$	1^{st}
Lattice-regularized DMC (Sorella, Casula)	$e^{\tau(E_T \hat{\mathbf{1}} - \hat{H})}$ (samp. τ)	$\mathbf{r}_i$	1^{st}
FCIQMC [21–23]	$\hat{\mathbf{1}} + \tau(E_T \hat{\mathbf{1}} - \hat{H})$	ϕ_i^{orthog}	2^{nd}
phaseless AFQMC [24, 5]	$e^{\tau(E_T \hat{\mathbf{1}} - \hat{H})}$	$\phi_i^{\text{nonorthog}}$	2^{nd}

Table 1: *Taxonomy of PMC methods. The methods are characterized by the projector, and the space in which the walk is performed, i.e., the single-particle basis and the quantization.*

AFQMC also uses the exponential projector but with a 2^{nd} quantized orbital basis, and FCIQMC uses a linear projector with a 2^{nd} quantized orbital basis. In AFQMC the orbitals are nonorthogonal and evolve continuously during the MC run, whereas in FCIQMC they are orthogonal and are fixed during the entire run. The linear projector has the advantage that if the Hamiltonian is known exactly in the chosen basis, so also is the projector. However, it can be used only when the spectrum of the Hamiltonian is bounded.

4.4 Diffusion Monte Carlo

We now discuss in some detail just one of the various PMC methods, the PMC method in real-space using first-quantized walkers. This method is more commonly known as diffusion Monte Carlo (DMC), and the projector is often referred to as the Green function since it is the Green function of the Fokker-Planck equation in the short-time approximation. We limit the following discussion to $\Psi_G(\mathbf{R}) = \Psi_T(\mathbf{R})$.

We now derive an approximate expression for

$$\langle \Psi_T | \mathbf{R}' \rangle \langle \mathbf{R}' | e^{\tau(E_T \hat{\mathbf{1}} - \hat{H})} | \mathbf{R} \rangle \frac{1}{\langle \mathbf{R} | \Psi_T \rangle} \equiv \Psi_T(\mathbf{R}') G(\mathbf{R}', \mathbf{R}, \tau) \frac{1}{\Psi_T(\mathbf{R})} \equiv \tilde{G}(\mathbf{R}', \mathbf{R}, \tau).$$

We multiply the imaginary-time the Schrödinger equation

$$-\frac{1}{2} \nabla^2 \Psi(\mathbf{R}, t) + (\mathcal{V}(\mathbf{R}) - E_T) \Psi(\mathbf{R}, t) = -\frac{\partial \Psi(\mathbf{R}, t)}{\partial t} \quad (32)$$

by $\Psi_T(\mathbf{R})$ and rearrange terms to obtain

$$-\frac{\nabla^2}{2} (\Psi \Psi_T) + \nabla \cdot \left(\frac{\nabla \Psi_T}{\Psi_T} \Psi \Psi_T \right) + \underbrace{\left(\frac{-\nabla^2 \Psi_T}{2 \Psi_T} + \mathcal{V} - E_T \right)}_{E_L(\mathbf{R})} (\Psi \Psi_T) = -\frac{\partial (\Psi \Psi_T)}{\partial t} \quad (33)$$

defining $f(\mathbf{R}, t) = \Psi(\mathbf{R}, t) \Psi_T(\mathbf{R})$, this is

$$\underbrace{-\frac{1}{2} \nabla^2 f}_{\text{diffusion}} + \underbrace{\nabla \cdot \left(\frac{\nabla \Psi_T}{\Psi_T} f \right)}_{\text{drift}} + \underbrace{(E_L(\mathbf{R}) - E_T) f}_{\text{growth/decay}} = -\frac{\partial f}{\partial t} \quad (34)$$

Since we know the solution for each individual term on the LHS, an approximation Green function is

$$\tilde{G}(\mathbf{R}', \mathbf{R}, \tau) = \frac{1}{(2\pi\tau)^{3N/2}} e^{-\left(\mathbf{R}' - \mathbf{R} - \mathbf{V}(\mathbf{R})\tau\right)^2/(2\tau) + \left(E_T - (E_L(\mathbf{R}') + E_L(\mathbf{R}))/2\right)\tau} + \mathcal{O}(\tau^2). \quad (35)$$

This is the same as the proposal matrix in Eq. (20) but there is now an additional growth/decay factor. So, the walker moves exactly as before, but the weight gets multiplied by this factor. Simply by turning on this multiplicative factor, one can switch between VMC and DMC.

This projector gives rather poor results—it has a very large time-step error and typically gives charge distributions for finite systems that are too diffuse. We describe next two improvements: a) take into account the singularities of the projector, and b) introduce an accept-reject step to ensure that the exact distribution is sampled in the limit that the $\Psi_T(\mathbf{R})$ is exact.

4.4.1 Singularities of $\tilde{G}(\mathbf{R}', \mathbf{R}, \tau)$

Following Ref. [15], Table 2 shows the singularities of $\tilde{G}(\mathbf{R}', \mathbf{R}, \tau)$ at the nodes of $\Psi_T(\mathbf{R})$ and the particle coincidences. $\tilde{G}(\mathbf{R}', \mathbf{R}, \tau)$ is accurate if $\mathbf{V}(\mathbf{R})$ is nearly constant over the time step, τ , and if E_L changes nearly linearly between $\mathbf{R}$ and $\mathbf{R}'$. Both assumptions fail dramatically at these singularities. For exact Ψ_T the local energy $E_L(\mathbf{R})$ is constant, but for the approximate Ψ_T used in practice, it diverges to $\pm\infty$ as the inverse of the distance to the nearest node. On the other hand the divergence of $E_L(\mathbf{R})$ at particle coincidences is easily removed simply by imposing the electron-nucleus and electron-electron cusps on Ψ_T . The velocity $\mathbf{V}(\mathbf{R})$ diverges at the nodes and has a discontinuity at particle coincidences even for exact Ψ_T , so this cannot be taken care of by improving Ψ_T . Instead, we must improve upon the implicit assumption in $\tilde{G}(\mathbf{R}', \mathbf{R}, \tau)$ that $\mathbf{V}(\mathbf{R})$ is constant during the time step τ . When an electron is near a node, the velocity diverges as the inverse distance to the node and is directed away from the node, so the large velocity does not persist over the entire time step τ . Taking this into account and integrating over the time step we find that the average velocity over the time-step τ is:

$$\bar{\mathbf{V}} = \frac{-1 + \sqrt{1 + 2V^2\tau}}{V^2\tau} \mathbf{V} \rightarrow \begin{cases} \mathbf{V} & \text{if } V^2\tau \ll 1 \\ \sqrt{2/\tau} \hat{\mathbf{V}} & \text{if } V^2\tau \gg 1 \end{cases} \quad (36)$$

Similar improvements to the average velocity in the vicinity of electron-nucleus coincidences, and the average of E_L can also be made.

Table 2: Singularities of $\tilde{G}(\mathbf{R}', \mathbf{R}, \tau)$.

Region	Local energy E_L	Velocity $\mathbf{V}$
nodes	$E_L \sim \pm 1/R_\perp$ for Ψ_T $E_L = E_0$ for Ψ_0	$V \sim 1/R_\perp$ for both Ψ_T and Ψ_0
e-n and e-e coincidences	$E_L \sim 1/x$ if cusps not imposed E_L finite if cusps are imposed $E_L = E_0$ for Ψ_0	V has a discontinuity for both Ψ_T and Ψ_0

4.4.2 Imposing detailed balance in DMC

Since $\tilde{G}(\mathbf{R}', \mathbf{R}, \tau)$ has a time-step error, it fails to sample the correct distribution even in the limit that $\Psi_T(\mathbf{R})$ is exact. So, for large values of τ , the version of DMC presented so far can actually be less accurate than VMC. Following Refs. [17, 15] this is easily remedied as follows: If we omit the third term on the left-hand side of Eq. (34) then it can be verified that $f(\mathbf{R}) = \Psi_T(\mathbf{R})^2$ is the solution since

$$-\frac{1}{2}\nabla^2 \Psi_T^2(\mathbf{R}) + \nabla \cdot \left(\frac{\nabla \Psi_T}{\Psi_T} \Psi_T^2(\mathbf{R}) \right) = 0. \quad (37)$$

However, we can sample $\Psi_T(\mathbf{R})^2$ using Metropolis-Hastings. So, we can view the drift-diffusion part of $\tilde{G}(\mathbf{R}', \mathbf{R}, \tau)$ as being the proposal matrix $T(\mathbf{R}', \mathbf{R})$ and introduce an accept-reject step after the drift and diffusion steps and before the reweighting step. With this simple modification, DMC is guaranteed to give an energy lower than the VMC energy, and in the limit of $\tau \rightarrow 0$ an energy that is variational, i.e., higher than the true energy. Finally, we account for the fact that the walker moves less far, since some of the moves are rejected, by using an effective time step τ_{eff} for reweighting

$$\tau_{\text{eff}} = \tau \frac{R_{\text{accep}}^2}{R_{\text{prop}}^2}, \quad (38)$$

where R_{prop}^2 is the sum of the squares of all the proposed one-electron move distances and R_{accep}^2 is the same sum, but including only the accepted moves.

4.5 Sign problem

PMC suffer from a sign problem except in a few special situations, e.g., 1-dimensional problems in real space. In all PMC methods, the sign problem occurs because an undesired state grows relative to the state of interest when the system is evolved by repeated stochastic applications of the projector. This results in a computational cost that grows exponentially with system size N . A reasonable definition [25] is that there is no sign problem if the computer time required to compute the value of an observable for an N -particle system with specified error, ϵ , scales as $T \propto N^\delta \epsilon^{-2}$, where δ is a small power (say, $\delta \leq 4$). (It is assumed that N is increased in some approximately homogeneous way, e.g., adding more atoms of the same species.) The details of how the sign problem manifests itself is different in the various PMC methods, and we discuss two examples.

4.5.1 Sign problem in DMC

In the absence of importance sampling (which has the important side effect of imposing the fixed-node constraint), the DMC projector of Eq. (35) becomes

$$G(\mathbf{R}', \mathbf{R}, \tau) \equiv \langle \mathbf{R}' | \hat{P}(\tau) | \mathbf{R} \rangle \approx \frac{e^{-\frac{(\mathbf{R}' - \mathbf{R})^2}{2\tau} + \left(E_T - \frac{V(\mathbf{R}') + V(\mathbf{R})}{2} \right) \tau}}{(2\pi\tau)^{3N/2}}. \quad (39)$$

It is nonnegative everywhere, so there is no sign problem if one were interested in the dominant state of this projector. However, the dominant state of this projector is the Bosonic ground state whereas the state of interest is the Fermionic ground state. If one started with a positive distribution and a negative distribution such that their sum is purely Fermionic as illustrated in Fig. 2, and applied the projector deterministically, both the positive and the negative distributions would tend to the Bosonic ground state, but, their sum would yield the Fermionic ground state, though with an amplitude that gets exponentially small relative to the amplitude of the individual components with increasing MC time. However, the projection is done stochastically and the probability of positive and negative walkers landing on the same state at the same MC time step and cancelling is very small if the portion of the state space that contributes significantly to the expectation values is finite and large, and it is zero if the state space is continuous (unless the dynamics of the walkers is modified to force opposite-sign walkers to land on the same spot). Hence it is not possible to sum the positive and negative contributions to extract the Fermionic ground state. This is demonstrated in Fig. 2. Furthermore, the problem cannot be solved by using an extremely large population of walkers. This enhances the probability of cancellations, but, because of fluctuations, eventually only positive or only negative walkers will survive and so the Fermionic state will completely disappear.

Fixed-node approximation: The importance-sampled Green function of Eq. (35) is not just a similarity transform, as in Eq. (29), of the Green function of Eq. (39). A *fixed-node* constraint has sneaked in as well. The velocity in Eq. (35) diverges at nodes and is directed away from them, so the number of node crossings per unit time goes to zero in the $\tau \rightarrow 0$ limit.² So, the solution obtained is the solution to the Schrödinger equation with the boundary condition that it goes to zero at the nodes of $\Psi_T(\mathbf{R})$. Since we are now adding in an artificial constraint, the resulting energy has a positive *fixed-node error*, which disappears in the limit that the nodal surface of $\Psi_T(\mathbf{R})$ is exact. The fixed-node approximation not only enables the calculation of the Fermionic ground state by eliminating the non-Fermionic states, it also enables the calculation of Fermionic excited states by preventing a collapse to the Fermionic ground state. For the excited states one loses the upper-bound property, but nevertheless the method has been used to calculate challenging excited states accurately [26–28].

4.5.2 Sign problem in FCIQMC

It may appear from the above discussion that the sign problem can be solved by performing the MC walk in a 2^{nd} -quantized, i.e., antisymmetrized, basis. Each 2^{nd} -quantized basis state consists of all the permutations of the corresponding 1^{st} -quantized basis states. Then there are no Bosonic states or states of any symmetry other than Fermionic, so there is no possibility of getting noise from non-Fermionic states. Of course, it is well known that this does not solve the sign problem. The problem is that different paths leading from a state to another can contribute with opposite sign. If the opposite sign contributions occur at the same MC step, then the

²The number of node crossings per MC step goes as $\tau^{3/2}$, so the number of crossings per unit time goes to zero as $\sqrt{\tau}$.

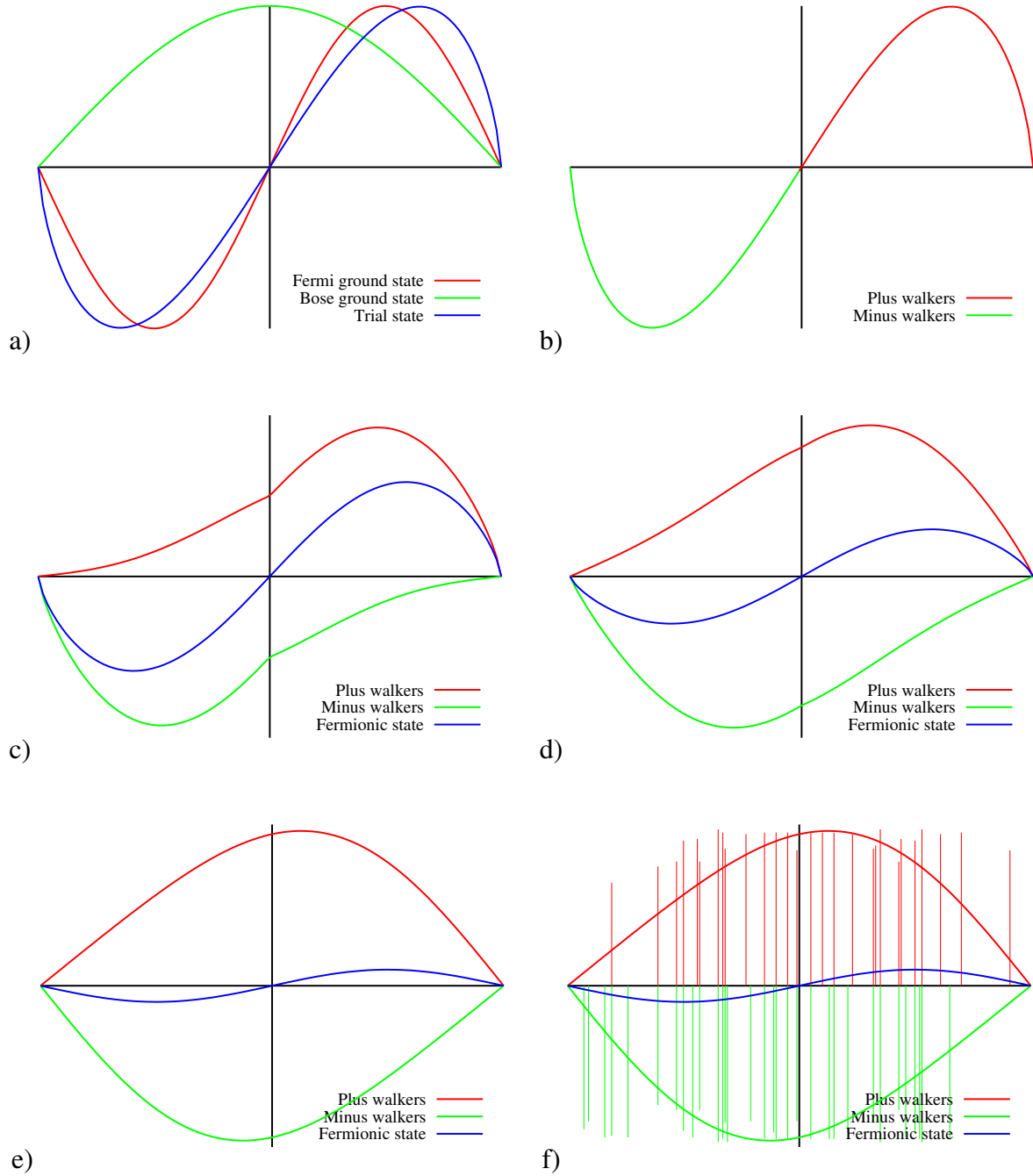


Fig. 2: Demonstration of the sign problem in DMC. a) The green curve schematically depicts the Bosonic ground state, the red curve the Fermionic ground state, and the blue curve an approximate Fermionic wavefunction. b) The starting positive distribution is shown in red and the starting negative distribution in green. Their sum is purely Fermionic. c-f) The red and the green curves show the evolution of the positive and negative distributions. Their sum, the blue curve, converges to the Fermionic ground state. f) For a finite population, the walkers are depicted as delta functions and in a continuous space they never meet and cancel (unless they are forced to in some way). Consequently there is an exponentially vanishing “signal to noise” ratio.

contributions cancel and yield a net contribution of smaller absolute magnitude, just as they would in a deterministic calculation. The problem occurs when opposite sign contributions occur at different MC steps. Further, since Ψ and $-\Psi$ are equally good, they are each sampled with equal probability in the course of a long MC run.

In a few special situations the sign problem is absent. The necessary and sufficient condition for there to be no sign problem is that all columns (or equivalently rows) of the projector have the same sign structure aside from an overall sign. Equivalently, there is no sign problem if it is possible to find a set of sign changes of the basis functions such that all elements of the projector are nonnegative. For example, the projector with the following sign structure

$$\begin{bmatrix} + & - & + & + \\ - & + & - & - \\ + & - & + & + \\ + & - & + & + \end{bmatrix} \quad (40)$$

does not have a sign problem, since changing the sign of the 2^{nd} basis state makes all the elements nonnegative. Note that it is not necessary to actually make these sign changes—the projectors before and after the sign changes are equally good.

Although a 2^{nd} -quantized basis does not solve the sign problem, it is advantageous to use a 2^{nd} -quantized basis when designing an exact Fermionic algorithm. First, an antisymmetrized basis is a factor of $N!$ smaller, and so the probability of opposite sign walkers meeting and cancelling is greater. Second, since each 2^{nd} -quantized basis state consists of a linear combination of 1^{st} -quantized basis states, 2^{nd} -quantized states that are connected by the projector may have multiple connections coming from several of the constituent 1^{st} -quantized states. Hence there is the possibility of internal cancellations in the 2^{nd} -quantized basis, which reduces the severity of the sign problem [29]. Third, since Bosonic and other symmetry states are eliminated it is clear that one can achieve a stable signal to noise for any large but finite basis by making the walker population very large. The limit of an infinite walker population is equivalent to doing a deterministic projection, which of course does not have a sign problem.

The FCIQMC method [21] does just this. It uses efficient algorithms for dealing with a large number of walkers and obtains essentially exact energies for very small molecules in small basis sets, using a large but manageable number of walkers. The MC walk, is done in an orbital occupation number (or equivalently determinantal) basis. For somewhat larger systems it is necessary to employ the *initiator approximation* [22] which greatly reduces the walker population needed to achieve a stable signal to noise ratio. Only states that have more walkers on them than some threshold value can spawn walkers on states that are not already occupied. The associated initiator error disappears of course in the limit of an infinite population and in practice it is possible to get accurate energies for interesting systems. However, the initiator error can be of either sign and it can be nonmonotonic, so extrapolation to the infinite population limit can be tricky. A large gain in efficiency can be gained by doing the projection on the most important states deterministically [23] since the sign-problem is present only for stochastic projection.

5 Form of trial wavefunctions

A major advantage of QMC methods is that since the integrals are done using Monte Carlo, one has a great deal of freedom in the choice of the form of the trial wavefunction $\Psi_T(\mathbf{R})$, allowing for compact yet accurate choices. As mentioned in Sec. 2.1, both the accuracy and the efficiency of both VMC and PMC (when there is a sign problem) depend on $\Psi_T(\mathbf{R})$. The only constraint on the choice of $\Psi_T(\mathbf{R})$ is that it should be possible to evaluate it and its local energy quickly (in low-order polynomial in N time).

5.1 Slater-Jastrow wavefunctions

The most commonly used choice for electronic systems is the so-called Jastrow-Slater form, that is, the product of a linear combination of determinants D_n of single-particle orbitals and a Jastrow correlation factor

$$\Psi_T = \left(\sum_n d_n D_n^\uparrow D_n^\downarrow \right) \times \mathcal{J}. \quad (41)$$

Note that in order to save computation time, we have replaced each determinant by a product of an up-spin and a down-spin determinant, which is not fully antisymmetric. This is legitimate because the expectation value is unchanged upon full antisymmetrization for any operator that does not have explicit spin dependence. The single-particle orbitals are usually expanded in spherical harmonics times Slater functions (monomial times an exponential in radial distance) for all-electron calculations in order to be able to impose the electron-nucleus cusps, and in spherical harmonics times gaussians or Gauss-Slater functions [30] for pseudopotential calculations. The minimal Jastrow function is a function of the electron-electron coordinates with the correct antiparallel- and parallel-spin cusps, but more typically it is written as a product of electron-nucleus, electron-electron and electron-electron-nucleus Jastrows:

$$\mathcal{J} = \prod_{\alpha i} \exp(A_{\alpha i}) \prod_{ij} \exp(B_{ij}) \prod_{\alpha ij} \exp(C_{\alpha ij}) \quad (42)$$

where α indexes the nuclei, and i and j index the electrons. Adding higher-order Jastrows, say 3 electrons and a nucleus, leads to minor gains relative to the increase in computational cost [31]. In all there are 4 kinds of parameters that can be optimized: a) the linear coefficients d_n multiplying the determinants, b) the orbital coefficients that specify the orbitals in terms of the basis functions, c) the exponents of the basis functions, and d) the Jastrow parameters. The number of basis exponents and the number of Jastrow parameters scale linearly in the number of atomic species, the number of orbital parameters scale as the number of electrons times the number of basis functions, and the number of d_n can be combinatorially large in the number of basis functions and electrons. However, in practice only a very tiny fraction of these are used. In fact one of the big advantages of QMC methods is that because of the effectiveness of the Jastrow in capturing some of the correlation, the number of determinants can be orders of magnitude smaller than in other methods for the same accuracy.

In real-space QMC, the Jastrow is particularly effective in introducing the so-called “dynamic correlation.” The multi-determinantal expansion is used mostly to include “near-degeneracy” or “static” correlation, which requires relatively few determinants. Consequently the number of determinants required to obtain a given energy is often orders of magnitude less in the presence of a flexible Jastrow than in its absence. Moreover, the size of the single-particle basis needed is reduced, particularly if the exponents of the basis functions are also optimized (though this is rarely done). Note that although the Jastrow does not directly change the nodes of $\Psi_T(\mathbf{R})$, when the wavefunction is optimized in the presence of the Jastrow it indirectly changes the nodes of $\Psi_T(\mathbf{R})$ and thereby enables accurate fixed-node DMC energies with compact wavefunctions.

The Jastrow plays another very important role in real-space QMC. The local energy, $E_L(\mathbf{R})$, diverges to $\pm\infty$ at electron-nucleus and electron-electron coincidences, unless cusp-conditions [32, 33] are imposed. The electron-nucleus cusps can be imposed by placing constraints on the orbitals (both intra-atom and inter-atom contributions need to be considered) but the electron-electron cusp requires the Jastrow function. Once the cusp conditions are imposed, $E_L(\mathbf{R})$ becomes finite at particle coincidences (though there is still a finite discontinuity in the limit that two electrons approach a nucleus [34]). This greatly reduces the fluctuations of the local energy and improves the efficiency of both VMC and DMC.

The multideterminant expansion is typically chosen by performing a small complete active space self consistent field (CASSCF) calculation and keeping the most important determinants. However, for challenging molecules there are several determinants outside the CAS space of affordable size that are more important than some of the determinants that are included from within the CAS space. Consequently, as the number of included determinants is increased, convergence of the energy is observed but to a spurious value. This problem is solved by selecting the determinants from a configuration interaction calculation [35, 36], which selects the most important determinants from the entire Hilbert space.

5.2 Symmetry-projected broken-symmetry mean-field wavefunctions

Recently there has been remarkable increase in the number of determinants that can be efficiently included in the multideterminant expansion [37–39]. Nevertheless, since the number of determinants grows combinatorially in the size of the single-particle basis and the number of electrons (though with a much reduced prefactor because of the Jastrow) there is considerable interest in using more flexible mean-field states than a single determinant, namely the antisymmetrical geminal power (AGP) [40] and Pfaffian [41, 42] as the antisymmetric part of QMC trial wavefunctions [43]. Recently these ideas have been extended to use, in QMC trial wavefunctions, symmetry-projected broken-symmetry mean-field states [43], first employed in traditional quantum chemistry calculations [44]. The symmetries that are broken and restored are combinations of particle-number, S^2 , S_z and complex conjugation. The most flexible of these breaks bonds correctly and yields remarkably good potential energy curves [43], considering that the computational cost is only marginally greater than that for a single determinant.

5.3 Backflow wavefunctions

Another direction (which can be combined with the above) is to use backflow wavefunctions, first introduced by Feynman to describe correlations in liquid Helium. The orbitals are evaluated at backflow-transformed coordinates that depend on the positions of all the electrons. These give considerably improved energies and fluctuations [45–48], but they incur a factor of N increase in the computational cost since the determinant lemma and Sherman-Morrison formula cannot be used to achieve an $\mathcal{O}(N^2)$ cost for updating determinants and their derivatives when a single electron is moved.

5.4 Wavefunctions in orbital-space QMC

Although most VMC and PMC calculations for electronic systems have been done in real space, recently there has been considerable interest in orbital-space QMC [49–51]. The orbital-space Jastrow plays a rather different role than its real-space counterpart—its most important contribution is to suppress double occupancy of orbitals and so it is effective in describing static correlations. The straightforward approach has a computational cost that scales as $\mathcal{O}(N^4)$ for constant error per electron, but ideas borrowed from the semistochastic heatbath configuration interaction (SHCI) method [52], reduce this cost to $\mathcal{O}(N^2)$ [51].

6 Optimization of trial wavefunctions

Accurate variational wavefunctions typically have a large number of linear and nonlinear parameters, that have to be optimized. As many as several hundred thousand have been used [53, 39]. One of the interesting features of QMC methods is that sometimes tiny changes in the algorithm, that may go unnoticed, can make a dramatic improvement to its efficiency. It is very helpful to think about ideal situations where the variance becomes zero; although this may never be achieved in practice, it is helpful for designing algorithms with low variance. This is particularly true in the case of wavefunction optimization algorithms. At the present time the 3 most used optimization algorithms are the stochastic reconfiguration method [54, 40], the Newton method [55], and the linear method [56–58]. We will present the Newton method in some detail to illustrate the sort of small algorithmic changes that can provide large efficiency improvements, but will mention the other two methods only cursorily.

Optimizing the wavefunctions is important for several reasons. We enumerate below the various errors in QMC calculations that are reduced by optimizing the wavefunction:

1. Statistical error (both the rms fluctuations of E_L and the autocorrelation time)
2. Variational error in E_{VMC}
3. Fixed-node error in E_{DMC}
4. Time-step error in DMC
5. Population control error in DMC
6. Pseudopotential locality error in DMC when using nonlocal pseudopotentials
7. Error of mixed estimates of observables that do not commute with the Hamiltonian in DMC

In fact all errors, aside from the finite-size errors present in calculations of periodic systems, benefit from wavefunction optimization.

The next question for optimizing wavefunctions is: precisely what quantity do we want to optimize? Possible choices are:

$$1) \text{ minimize } E_{\text{VMC}} = \frac{\langle \Psi_{\text{T}} | H | \Psi_{\text{T}} \rangle}{\langle \Psi_{\text{T}} | \Psi_{\text{T}} \rangle} = \langle E_{\text{L}} \rangle_{\Psi_{\text{T}}^2} \quad (43)$$

$$2) \text{ minimize } \sigma_{\text{VMC}}^2 = \frac{\langle \Psi_{\text{T}} | (H - E_{\text{T}})^2 | \Psi_{\text{T}} \rangle}{\langle \Psi_{\text{T}} | \Psi_{\text{T}} \rangle} = \langle E_{\text{L}}^2(\mathbf{R}_i) \rangle_{\Psi_{\text{T}}^2} - \langle E_{\text{L}}(\mathbf{R}_i) \rangle_{\Psi_{\text{T}}^2}^2 \quad (44)$$

$$3) \text{ maximize } \Omega^2 = \frac{|\langle \Psi_{\text{FN}} | \Psi_{\text{T}} \rangle|^2}{\langle \Psi_{\text{FN}} | \Psi_{\text{FN}} \rangle \langle \Psi_{\text{T}} | \Psi_{\text{T}} \rangle} = \frac{\left\langle \left| \frac{\Psi_{\text{FN}}}{\Psi_{\text{T}}} \right|^2 \right\rangle_{\Psi_{\text{T}}^2}}{\left\langle \left| \frac{\Psi_{\text{FN}}}{\Psi_{\text{T}}} \right|^2 \right\rangle_{\Psi_{\text{T}}^2}} \quad (45)$$

$$4) \text{ minimize } E_{\text{DMC}} = \frac{\langle \Psi_{\text{FN}} | H | \Psi_{\text{T}} \rangle}{\langle \Psi_{\text{FN}} | \Psi_{\text{T}} \rangle} = \langle E_{\text{L}} \rangle_{|\Psi_{\text{FN}} \Psi_{\text{T}}|} \quad (46)$$

In fact all of the above have been studied to some extent in published and unpublished work. For an infinitely flexible wave function all these optimization criteria will yield the exact wavefunction (except that minimizing σ could yield an excited state) but for the imperfect functional forms used in practice they differ. Since the final energy is obtained from E_{DMC} rather than E_{VMC} , the most desirable option is the last one. However, the very limited experience gained so far indicates that minimizing E_{VMC} with flexible wavefunctions results in approximately minimizing also E_{DMC} , so the additional effort of minimizing E_{DMC} is not worthwhile. Hence the common practice is to minimize E_{VMC} or a linear combination of E_{VMC} and σ_{VMC}^2 , with most of the weight on E_{VMC} .

Early work on optimizing wavefunctions used variance minimization [59, 60] because early attempts at minimizing E_{VMC} required many more Monte Carlo samples than for minimizing σ_{VMC}^2 . This is explained in Fig. 3. In green we schematically show the local energies for some sampled points. Since the wavefunction is not exact these energies have a spread. In red we show how these energies change as one attempts to minimize E_{VMC} . Most of the energies go down and some may go up, but the average on the sample goes down. As the wavefunction is made more flexible the average may go down well below the true energy. Now if one draws a fresh sample, then we find that the average on this sample has actually gone up rather than down! Once one realizes the problem with energy minimization it becomes clear that the way around it is to minimize the variance of the local energies as shown in blue. In that case local energies move closer to each other and the average energy also tends to go down, less so than for energy minimization, but this gain in energy is genuine—an independent sample shows the same effect. Of course energy minimization will work if one has a sufficiently large number of samples, but the point of the above thought experiment is that the number of samples needed is smaller for variance minimization than for naive energy minimization. We next discuss more clever ways to minimize the energy that overcome this problem by minimizing the expectation value of the energy without minimizing the energy of the chosen sample.

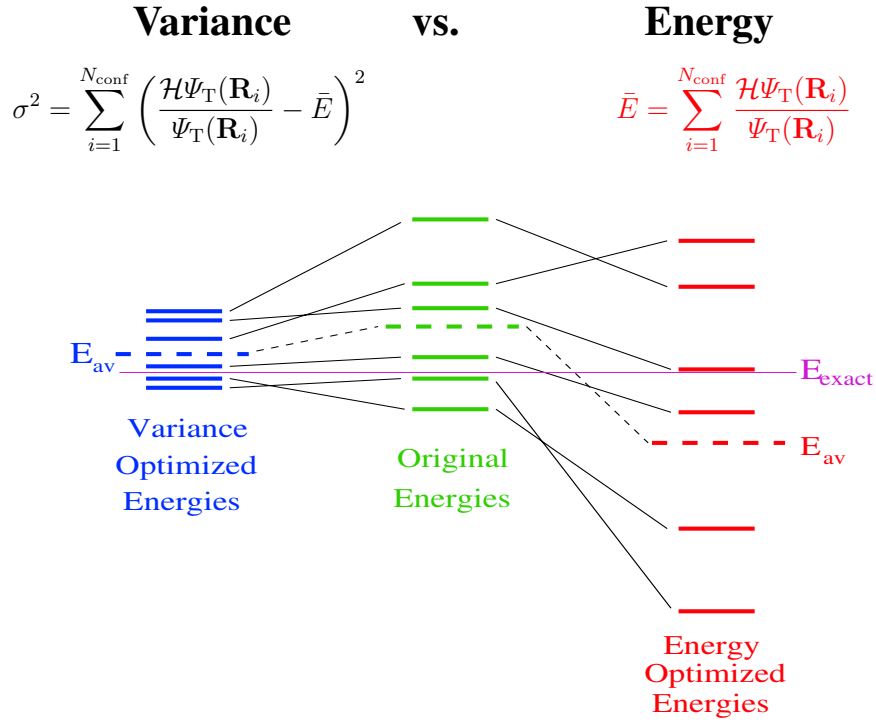


Fig. 3: Why variance minimization requires fewer samples than naive energy minimization.

6.1 Newton method

6.1.1 Minimization of E_{VMC}

In the Newton method, the parameter changes, $\delta \mathbf{p}$, are obtained by solving linear equations

$$\mathbf{h} \delta \mathbf{p} = -\mathbf{g}, \quad (47)$$

where $\mathbf{h}$ is the Hessian and $\mathbf{g}$ the gradient of E_{VMC} with respect to the variational parameters.

In the rest of this section Ψ_{T} and E_{VMC} are the only wavefunction and energy of relevance, so in the interest of notational brevity we replace these by Ψ and $\bar{E}$ respectively.

$$\bar{E} = \frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \langle E_{\text{L}} \rangle_{|\Psi|^2}; \quad E_{\text{L}}(\mathbf{R}) = \frac{H\Psi(\mathbf{R})}{\Psi(\mathbf{R})} \quad (48)$$

where the notation $\langle \cdots \rangle_{|\Psi|^2}$ denotes a Monte Carlo average over samples drawn from $|\Psi|^2$.

Following Ref. [55] the energy gradient wrt parameter p_i , denoted by $\bar{E}_i$, is

$$\bar{E}_i = \frac{\langle \Psi_i | H | \Psi \rangle + \langle \Psi | H | \Psi_i \rangle}{\langle \Psi | \Psi \rangle} - 2 \frac{\bar{E} \langle \Psi | \Psi_i \rangle}{\langle \Psi | \Psi \rangle} = 2 \frac{\langle \Psi_i | H | \Psi \rangle - \bar{E} \langle \Psi | \Psi_i \rangle}{\langle \Psi | \Psi \rangle} \quad (49)$$

$$\approx \left\langle \frac{\Psi_i}{\Psi} E_{\text{L}} + \frac{H\Psi_i}{\Psi} - 2\bar{E} \frac{\Psi_i}{\Psi} \right\rangle_{|\Psi|^2} \approx 2 \left\langle \frac{\Psi_i}{\Psi} (E_{\text{L}} - \bar{E}) \right\rangle_{|\Psi|^2} \quad (50)$$

In Eq. (49) we use the Hermiticity of the Hamiltonian to go from the expressions on the left to that on the right. The expressions in Eq. (50) are the MC estimates of the corresponding

expressions in Eq. (49). Note however, that the expressions on the left and right of Eq. (50) become precisely equal only in the limit of an infinite sample. For a finite sample, the expression on the right has much smaller fluctuations than the one on the left for sufficiently good trial wavefunctions³ because it has zero variance in the limit that Ψ is exact.

Rewriting the gradient as

$$\bar{E}_i = 2 \frac{\langle \Psi_i | H \Psi \rangle - \bar{E} \langle \Psi | \Psi_i \rangle}{\langle \Psi | \Psi \rangle} = 2 \frac{\langle \Psi_i \Psi (E_L - \bar{E}) \rangle}{\langle |\Psi|^2 \rangle} \quad (51)$$

and taking the derivative wrt p_j , we obtain the following expression for the hessian:

$$\bar{E}_{ij} = 2 \left[\frac{\langle (\Psi_{ij} \Psi + \Psi_i \Psi_j) (E_L - \bar{E}) \rangle + \langle \Psi_i \Psi (E_{L,j} - \bar{E}_j) \rangle - \bar{E}_i \langle \Psi \Psi_j \rangle}{\langle |\Psi|^2 \rangle} \right] \quad (52)$$

$$= 2 \left[\left\langle \left(\frac{\Psi_{ij}}{\Psi} + \frac{\Psi_i \Psi_j}{|\Psi|^2} \right) (E_L - \bar{E}) \right\rangle_{\psi^2} - \left\langle \frac{\Psi_i}{\Psi} \right\rangle_{\psi^2} \bar{E}_j - \left\langle \frac{\Psi_j}{\Psi} \right\rangle_{\psi^2} \bar{E}_i + \left\langle \frac{\Psi_i}{\Psi} E_{L,j} \right\rangle_{\psi^2} \right]. \quad (53)$$

What can be done to reduce the variance of this expression? The last term is not symmetric in i and j because we started from the right hand rather than the left hand expression in Eq. (50), so we can symmetrize it, but that does not affect the variance appreciably.

Next note that $\langle E_{L,i} \rangle_{\psi^2} = \frac{\langle |\Psi|^2 (\frac{H\Psi}{\Psi})_i \rangle}{\langle |\Psi|^2 \rangle} = \frac{\langle |\Psi|^2 (\frac{H\Psi_i}{\Psi} - \frac{\Psi_i}{|\Psi|^2} H\Psi) \rangle}{\langle |\Psi|^2 \rangle} = \frac{\langle \Psi H\Psi_i - \Psi_i H\Psi \rangle}{\langle |\Psi|^2 \rangle} = 0$, so we are free to add terms such as $\left\langle \frac{\Psi_j}{\Psi} \right\rangle_{\psi^2} \langle E_{L,i} \rangle_{\psi^2}$. Now, the fluctuations of the covariance $\langle ab \rangle - \langle a \rangle \langle b \rangle$ are smaller than those of the product $\langle ab \rangle$, when $\sqrt{\langle a^2 \rangle - \langle a \rangle^2} \ll |\langle a \rangle|$ and $\langle b \rangle$ is small. ($\langle E_{L,i} \rangle_{\psi^2}$ is 0 on an infinite sample and small on a finite sample.) Hence we make the replacement

$$\left\langle \frac{\Psi_i}{\Psi} E_{L,j} \right\rangle_{\psi^2} \rightarrow \frac{1}{2} \left(\left\langle \frac{\Psi_i}{\Psi} E_{L,j} \right\rangle_{\psi^2} - \left\langle \frac{\Psi_i}{\Psi} \right\rangle_{\psi^2} \langle E_{L,j} \rangle_{\psi^2} + \left\langle \frac{\Psi_j}{\Psi} E_{L,i} \right\rangle_{\psi^2} - \left\langle \frac{\Psi_j}{\Psi} \right\rangle_{\psi^2} \langle E_{L,i} \rangle_{\psi^2} \right).$$

The resulting expression is

$$\begin{aligned} \bar{E}_{ij} = & 2 \left[\left\langle \left(\frac{\Psi_{ij}}{\Psi} + \frac{\Psi_i \Psi_j}{|\Psi|^2} \right) (E_L - \bar{E}) \right\rangle_{\psi^2} - \left\langle \frac{\Psi_i}{\Psi} \right\rangle_{\psi^2} \bar{E}_j - \left\langle \frac{\Psi_j}{\Psi} \right\rangle_{\psi^2} \bar{E}_i \right] \\ & + \left\langle \frac{\Psi_i}{\Psi} E_{L,j} \right\rangle_{\psi^2} - \left\langle \frac{\Psi_i}{\Psi} \right\rangle_{\psi^2} \langle E_{L,j} \rangle_{\psi^2} + \left\langle \frac{\Psi_j}{\Psi} E_{L,i} \right\rangle_{\psi^2} - \left\langle \frac{\Psi_j}{\Psi} \right\rangle_{\psi^2} \langle E_{L,i} \rangle_{\psi^2} \end{aligned} \quad (54)$$

$$\begin{aligned} = & 2 \left[\left\langle \left(\frac{\Psi_{ij}}{\Psi} - \frac{\Psi_i \Psi_j}{|\Psi|^2} \right) (E_L - \bar{E}) \right\rangle_{\psi^2} \quad (0 \text{ for } p_i \text{ linear in exponent}) \right. \\ & + 2 \left\langle \left(\frac{\Psi_i}{\Psi} - \left\langle \frac{\Psi_i}{\Psi} \right\rangle_{\psi^2} \right) \left(\frac{\Psi_j}{\Psi} - \left\langle \frac{\Psi_j}{\Psi} \right\rangle_{\psi^2} \right) (E_L - \bar{E}) \right\rangle_{\psi^2} \\ & \left. + \left\langle \frac{\Psi_i}{\Psi} E_{L,j} \right\rangle_{\psi^2} - \left\langle \frac{\Psi_i}{\Psi} \right\rangle_{\psi^2} \langle E_{L,j} \rangle_{\psi^2} + \left\langle \frac{\Psi_j}{\Psi} E_{L,i} \right\rangle_{\psi^2} - \left\langle \frac{\Psi_j}{\Psi} \right\rangle_{\psi^2} \langle E_{L,i} \rangle_{\psi^2} \right]. \end{aligned} \quad (55)$$

³A subtle point is that the zero-variance estimator on the RHS of Eq. (50) is also an infinite-variance estimator when Ψ_T is not exact and the parameters being optimized can change the nodes of the wavefunction. So, for poor wavefunctions it may be preferable to use a known alternative expression that has finite variance for approximate Ψ_T .

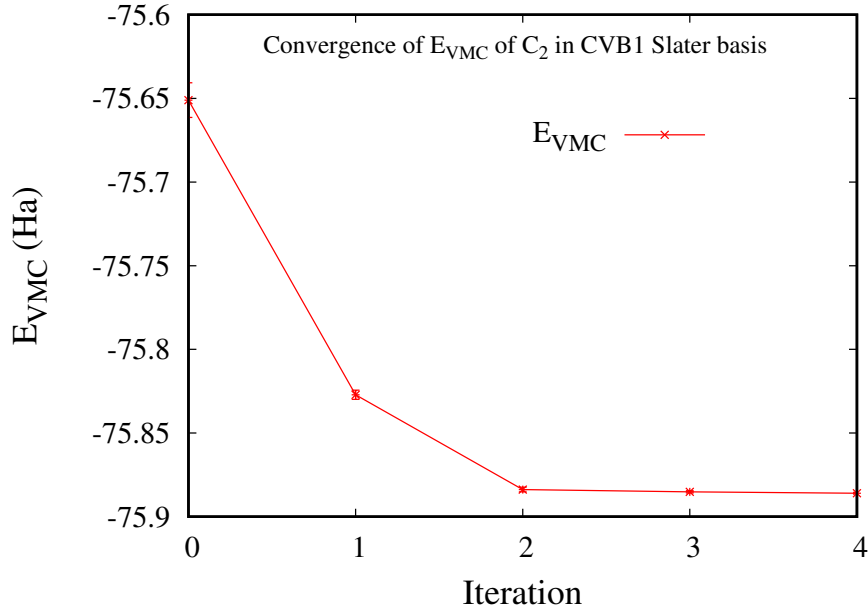


Fig. 4: Energy convergence of C_2 using a Slater CVB1 basis with 48 basis functions, optimizing 24 Jastrow, 164 CSF and 90 orbital parameters. If the basis exponents are optimized then with 44 basis functions, a considerably more compact wavefunction with just 13 CSFs gives an energy better than this by 1 mHa (not shown). However, converging the wavefunction takes many more optimization iterations.

The expressions in the first two lines have zero variance in the limit of an exact trial wavefunction and the last line has a much reduced variance compared to our starting point, Eq. (53). Finally, we note that the individual summands in Eq. (55) have a leading 3^{rd} -order divergence near the nodes of the trial wavefunction for parameters that change wavefunction nodes, but the leading divergences cancel each other, leaving a 2^{nd} -order divergence which gives a finite expectation value but infinite variance. Despite having infinite variance estimators for the gradient and the hessian, the method works remarkably well for small systems even for parameters that move the nodes of Ψ_T , as shown in Fig. 4. If finite variance estimators are needed they are obtained by using nodeless Ψ_G with $\Psi_G \approx |\Psi_T|$ except near the nodes of Ψ_T [61].

6.1.2 Minimization of σ_{VMC}

As mentioned earlier, another option is to minimize the variance of the energy

$$\sigma^2 = \frac{\int d^3R |\Psi|^2 (E_L - \bar{E})^2}{\int d^3R |\Psi|^2}. \quad (56)$$

The exact gradient and hessian have been derived and used but the following simpler option works just as well. When the parameters are changed, the variance changes both because E_L changes and because the distribution of sampled points changes. If we ignore the latter, which means that we are computing the variance on a fixed set of Monte Carlo configurations, we get

$$(\sigma^2)_i = 2 \langle E_{L,i} (E_L - \bar{E}) \rangle = 2 \langle (E_{L,i} - \bar{E}_i) (E_L - \bar{E}) \rangle. \quad (57)$$

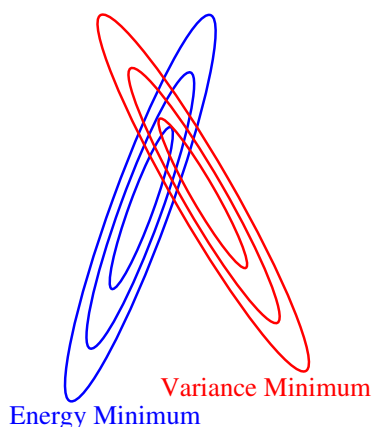


Fig. 5: *Schematic of energy and variance contours. Adding a small fraction of variance to the energy in the function to be minimized can reduce the variance while raising the energy only very slightly.*

In the case of energy minimization we added a term with zero expectation value to reduce the variance of the energy. Similarly, in the right hand expression above we have added a term to minimize the variance of the variance.

Then the (positive definite) Levenberg-Marquardt approximation to the Hessian matrix is

$$(\sigma^2)_{ij} = 2 \langle (E_{L,i} - \bar{E}_i)(E_{L,j} - \bar{E}_j) \rangle. \quad (58)$$

6.1.3 Mixed minimization

A linear combination of the energy and variance can be optimized simply by using the same linear combination for the gradient and the hessian. The reason for possibly wanting to use a linear combination is that by adding in small fraction of variance minimization (say 0.01-0.05) to energy minimization, it may be possible to reduce the variance without appreciably raising the energy. To explain why, we show in Fig. 5 schematic contours for the energy and variance. If only the energy is minimized then the parameters may lie anywhere close to the bottom of the energy well, but if a small fraction of the variance is added to the mix then the portion of the energy well closer to the bottom of the variance minimum is favored.

Fig. 6 has convergence plots for energy, variance, and mixed minimization for the NO₂ molecule using a pseudopotential. The optimization takes only a few iterations. We see that mixed minimization lowers the variance while raising the energy only slightly. However, energy minimization has the smallest autocorrelation time, T_{corr} , so the benefit from doing a mixed minimization rather than just an energy minimization is small.

6.1.4 Stabilizing the optimization

If we add a positive constant a_{diag} to the diagonal of the hessian matrix, i.e., $\mathbf{h} \rightarrow \mathbf{h} + a_{\text{diag}}\mathbf{I}$, where $\mathbf{I}$ is the identity matrix, then the proposed moves get smaller and rotate from the Newtonian direction to the steepest descent direction. We use this to stabilize the optimization. At each optimization iteration, we perform two kinds of MC runs. First we do a run to compute the gradient and the hessian. Second, we do a 10-times shorter MC run, that does 3 correlated sampling calculations with parameters obtained from this gradient and hessian and 3 values of a_{diag} that differ from each other by factors of 10 to get the corresponding 3 values of the

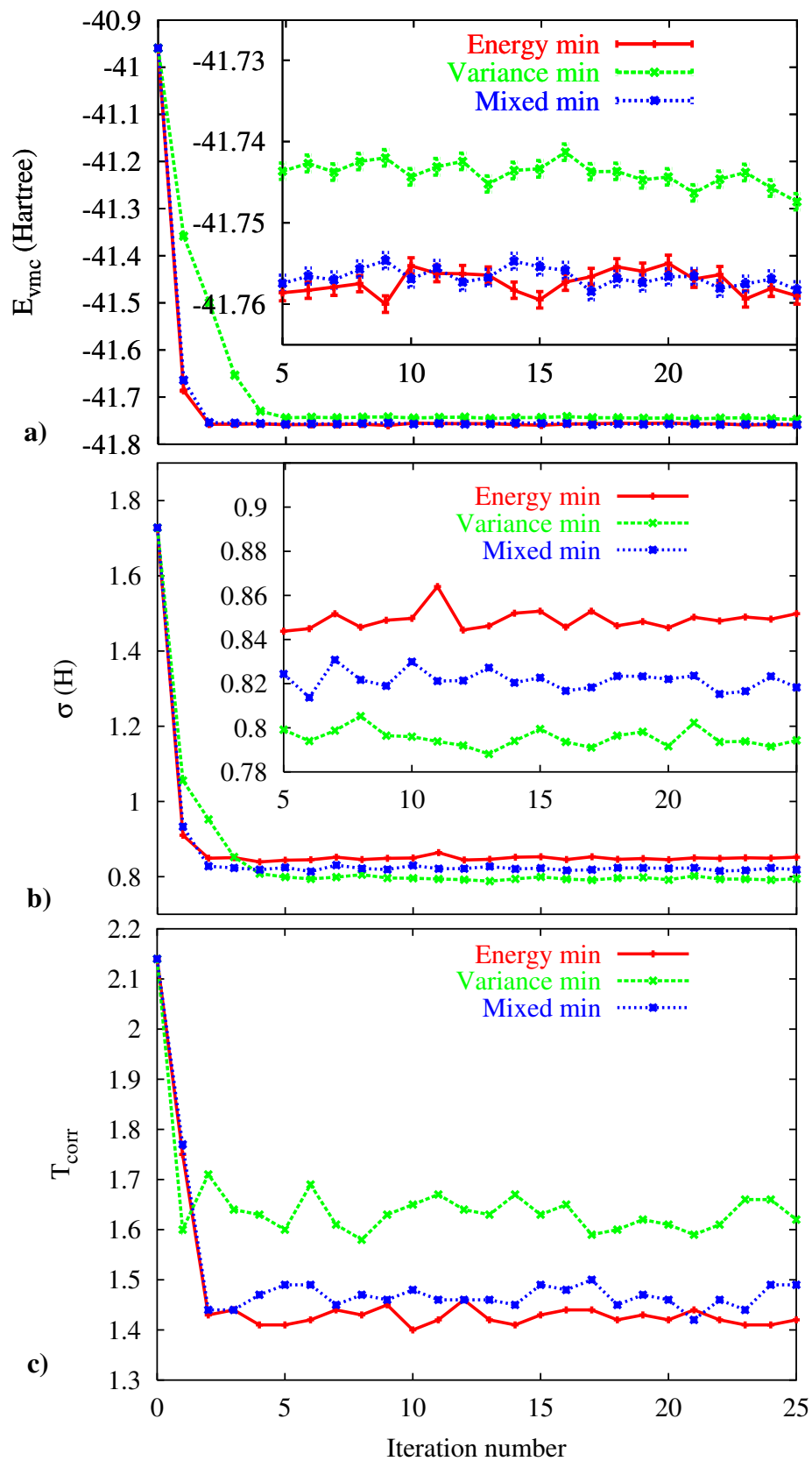


Fig. 6: Optimization of Jastrow parameters for an NO_2 molecule using energy minimization, variance minimization, and mixed energy and variance minimization. Convergence is achieved in just a few iterations. Mixed minimization lowers σ while raising the energy only slightly. T_{corr} is smallest for energy minimization.

energy (or whatever linear combination of the energy and variance that we are optimizing). The optimal value of a_{diag} is then predicted from a parabolic fit of the 3 energy values with some bounds imposed, provided the 2^{nd} derivative of the fit is positive; else the a_{diag} value that gives the lowest energy is chosen. In addition, a_{diag} is forced to increase if the parameter variations exceed some chosen thresholds, or if some parameters exit their allowed domain of variation (e.g., if a basis exponent becomes negative). Despite these precautions, the energy occasionally goes up a lot, in which case one goes back to the earlier gradient and hessian and proposes new parameter changes with a larger value of a_{diag} . This only happens for the larger systems that we study; for small systems such as NO_2 shown in Fig. 6 the optimization converges within a few iterations with no need to go back to an earlier iteration. Note that even when it is necessary to go back, the entire procedure is automatic; there is no need for human intervention. Finally we note that we do not necessarily choose the parameters from the last optimization iteration as the final best parameters. Instead we choose the parameters from the iteration for which $E_{\text{VMC}} + 3\sigma_{\text{VMC}}$ is lowest. In fact, this often is the last iteration because in order to save time, we use a small number of MC steps in the first iteration and gradually increase the number of MC steps with each new iteration (upto some maximum), so even after E_{VMC} has converged, σ_{VMC} continues to go down with iteration number.

6.2 Linear method

The linear optimization method is probably at the present time the most commonly used optimization method. It was originally developed for linear parameters [62], but was extended to nonlinear parameters [56–58]. It has the advantage that it does not require calculating the 2^{nd} derivatives of $\Psi_{\text{T}}(\mathbf{R})$ and it converges just as quickly as the Newton method. Similar to what we have described for the Newton method, there are far from obvious changes that need to be made to the straightforward version of the method in order to make it efficient. The details can be found in the original literature [62, 56–58].

6.3 Stochastic reconfiguration method

The stochastic reconfiguration method [54, 40] can be viewed as an approximation to the Newton method. Of the 3 methods mentioned in this lecture, it requires the least computation per optimization iteration, but it typically takes several times as many iterations to converge. Although it converges more slowly, when applied to heavy systems it can sometimes oscillate less than the other methods because it requires fewer derivatives and suffers less from infinite-variance estimators.

Optimization of many parameters: When the number of parameters to be optimized is large (more than a few thousand) storage of the relevant matrices (hessian for Newton method, Hamiltonian and overlap for the linear method, overlap for the stochastic reconfiguration method) becomes a problem. However, they need not be stored if iterative methods are used to solve the relevant equations. Then it becomes practical to optimize on the order of 10^5 – 10^6 parameters. Details can be found in Refs. [53, 8].

7 Outlook

I hope this lecture has given you some flavor of the sort of thinking that goes into designing accurate and efficient QMC algorithms. QMC methods have been worked upon less than some other electronic structure methods. Hence, I think there is still considerable room for improvement. An example of that would be the development of a continuous real-space PMC method that does not have a time-step error. In fact one of the earliest PMC methods invented [18–20] does not have a time-step error. However, despite this major advantage, it is not used anymore because it is much less efficient than DMC. Further, there exist, in a discrete space, efficient PMC methods that use the exponential projector with no time-step error. It seems possible that one could invent an efficient time-step error free algorithm for continuous real-space PMC as well.

Acknowledgments

I acknowledge valuable discussions with Sandeep Sharma. This work was supported by the AFOSR through grant FA9550-18-1-0095.

References

- [1] D.M. Ceperley, *Rev. Mod. Phys.* **67**, 279 (1995)
- [2] M.P. Nightingale and C.J. Umrigar (eds.): *Quantum Monte Carlo Methods in Physics and Chemistry*, NATO ASI Ser. C 525 (Kluwer, Dordrecht, 1999)
- [3] W. Foulkes, L. Mitas, R. Needs, and G. Rajagopal, *Rev. Mod. Phys.* **73**, 33 (2001)
- [4] J. Kolorenc and L. Mitas, *Rep. Prog. Phys.* **74**, 026502 (2011)
- [5] S. Zhang, in E. Pavarini, E. Koch, and U. Schollwöck (eds.) *Emergent Phenomena in Correlated Matter*, Modeling and Simulation, Vol. 3 (Forschungszentrum Jülich, 2013)
- [6] J. Toulouse, R. Assaraf, and C.J. Umrigar, p. 285 in P. Hoggan and T. Özdoğan (eds.): *Advances in Quantum Chemistry*, Vol. 73 (2015)
- [7] F. Becca and S. Sorella: *Quantum Monte Carlo Approaches for Correlated Systems* (Cambridge University Press, 2017)
- [8] J. Feldt and C. Filippi, in L. González and R. Lindh (eds.): *Quantum Chemistry and Dynamics of Excited States: Methods and Applications* (Wiley-VCH, 2019)
- [9] C. Everett and E. Cashwell.: *Third Monte Carlo sampler. Revision and extension of samplers I and II*. Tech. Rep., Los Alamos National Laboratory (1983)
- [10] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller and E. Teller, *J. Chem. Phys.* **21**, 1087 (1953)
- [11] W.K. Hastings, *Biometrika* **57**, 97 (1970)
- [12] C.J. Umrigar, *Phys. Rev. Lett.* **71**, 408 (1993)
- [13] C.J. Umrigar, p. 129 in [2]
- [14] M.P. Nightingale and H. Blöte, *Phys. Rev. Lett.* **60**, 1562 (1988)
- [15] C.J. Umrigar, M.P. Nightingale, and K.J. Runge, *J. Chem. Phys.* **99**, 2865 (1993)
- [16] J.B. Anderson, *J. Chem. Phys.* **65**, 4121 (1976)
- [17] P.J. Reynolds, D.M. Ceperley, B.J. Alder, and W.A. Lester, *J. Chem. Phys.* **77**, 5593 (1982)
- [18] M.H. Kalos, D. Levesque, and L. Verlet, *Phys. Rev. A* **9**, 2178 (1974)
- [19] D.M. Ceperley and M.H. Kalos, pp. 145–194 in K. Binder (ed.): *Monte Carlo Methods in Statistical Physics* (Springer, Berlin, 1979)

- [20] K.E. Schmidt and M.H. Kalos, p. 125 in K. Binder (ed.): *Applications of the Monte Carlo Method in Statistical Physics*, Topics in Current Physics, Vol. 36 (Springer, 1984)
- [21] G.H. Booth, A.J.W. Thom, and A. Alavi, J. Chem. Phys. **131**, 054106 (2009)
- [22] D. Cleland, G.H. Booth, and A. Alavi, J. Chem. Phys. **132**, 041103 (2010)
- [23] F.R. Petruzielo, A.A. Holmes, H.J. Changlani, M.P. Nightingale, and C.J. Umrigar, Phys. Rev. Lett. **109**, 230201 (2012)
- [24] S. Zhang and H. Krakauer, Phys. Rev. Lett. **90**, 136401 (2003)
- [25] D.M. Ceperley and L. Mitas, *Advances in Chemical Physics*, Vol. 93, 1–38 (John Wiley & Sons, 1996)
- [26] F. Schautz and C. Filippi, J. Chem. Phys. **120**, 10931 (2004)
- [27] F. Schautz, F. Buda, and C. Filippi, J. Chem. Phys. **121**, 5836 (2004)
- [28] H. Zulfikri, C. Amovilli, and C. Filippi, J. Chem. Theory Comput. **12**, 1157 (2016)
- [29] M.H. Kolodrubetz, J.S. Spencer, B.K. Clark, and W.M.C. Foulkes, J. Chem. Phys. **138**, 024110 (2013)
- [30] F.R. Petruzielo, J. Toulouse, and C.J. Umrigar, J. Chem. Phys. **134**, 064104 (2011)
- [31] C.J. Huang, C.J. Umrigar, and M.P. Nightingale, J. Chem. Phys. **107**, 3007 (1997)
- [32] T. Kato, Comm. Pure Appl. Math. **10**, 151 (1957)
- [33] R.T. Pack and W. Byers-Brown, J. Chem. Phys. **45**, 556 (1966)
- [34] C.R. Myers, C.J. Umrigar, J.P. Sethna, and J. Morgan, Phys. Rev. A **44**, 5537 (1991)
- [35] A. Scemama, A. Benali, D. Jacquemin, M. Caffarel, and P.F. Loos, J. Chem. Phys. **149**, 034108 (2018)
- [36] M. Dash, S. Moroni, A. Scemama, and C. Filippi, J. Chem. Theory Comput. **14**, 4176 (2018)
- [37] B.K. Clark, M.A. Morales, J. McMinis, J. Kim, and G.E. Scuseria, J. Chem. Phys. **135**, 244105 (2011)
- [38] C. Filippi, R. Assaraf, and S. Moroni, J. Chem. Phys. **144**, 194105 (2016)
- [39] R. Assaraf, S. Moroni, and C. Filippi, J. Chem. Theory Comput. **13**, 5273 (2017)
- [40] M. Casula, C. Attaccalite, and S. Sorella, J. Chem. Phys. **121**, 7110 (2004)
- [41] M. Bajdich, L. Mitas, L.K. Wagner, and K.E. Schmidt, Phys. Rev. B **77**, 115112 (2008)

- [42] M. Bajdich, L. Mitas, G. Drobný, L. Wagner, and K. Schmidt, Phys. Rev. Lett. **96**, 130201 (2006)
- [43] A. Mahajan and S. Sharma, J. Phys. Chem. A **123**, 3911 (2019)
- [44] G.E. Scuseria, C.A. Jiménez-Hoyos, and T.M. Henderson, J. Chem. Phys. **135**, 124108 (2011)
- [45] K.E. Schmidt, M.A. Lee, M.H. Kalos, and G.V. Chester, Phys. Rev. Lett. **47**, 807 (1981)
- [46] M.D. Brown and J.R. Trail and P. Lopez Rios and R.J. Needs, J. Chem. Phys. **126**, 224110 (2007)
- [47] P. Seth, P.L. Rios, and R.J. Needs, J. Chem. Phys. **134**, 084105 (2011)
- [48] M. Holzmann and S. Moroni, Phys. Rev. B **99**, 085121 (2019)
- [49] E. Neuscamman, Mol. Phys. **114**, 577 (2016)
- [50] H. Wei and E. Neuscamman, J. Chem. Phys. **149**, 184106 (2018)
- [51] I. Sabzeyari and S. Sharma, J. Chem. Theory Comput. **14**, 6276 (2018)
- [52] A.A. Holmes, N.M. Tubman, and C.J. Umrigar, J. Chem. Theory Comput. **12**, 3674 (2016)
- [53] E. Neuscamman, C.J. Umrigar, and G.K.L. Chan, Phys. Rev. B **85**, 045103 (2012)
- [54] S. Sorella, Phys. Rev. B **64**, 024512 (2001)
- [55] C.J. Umrigar and C. Filippi, Phys. Rev. Lett. **94**, 150201 (2005)
- [56] C.J. Umrigar, J. Toulouse, C. Filippi, S. Sorella, and R.G. Hennig, Phys. Rev. Lett. **98**, 110201 (2007)
- [57] J. Toulouse and C.J. Umrigar, J. Chem. Phys. **126**, 084102 (2007)
- [58] J. Toulouse and C.J. Umrigar, J. Chem. Phys. **128**, 174101 (2008)
- [59] R.L. Coldwell, Int. J. Quant. Chem. **11**, 215 (1977)
- [60] C.J. Umrigar, K.G. Wilson, and J.W. Wilkins, Phys. Rev. Lett. **60**, 1719 (1988)
- [61] C. Attaccalite and S. Sorella, Phys. Rev. Lett. **100**, 114501 (2008)
- [62] M.P. Nightingale and Vilen Melik-Alaverdian, Phys. Rev. Lett. **87**, 043041 (2001)

4 Optimized Quantum Monte Carlo Wave Functions

Arne Lüchow

RWTH Aachen University

Institute of Physical Chemistry, 52056 Aachen

Contents

1	Introduction	2
2	Real-space quantum Monte Carlo methods	2
2.1	Variational Monte Carlo	2
2.2	Diffusion Monte Carlo	3
3	Stochastic optimization of electronic wave functions	5
3.1	Variance and MAD minimization	6
3.2	Energy minimization	6
3.3	Terms required for parameter optimization	12
4	Examples of wave function optimizations	15
4.1	Simultaneous and alternating parameter optimizations	16
4.2	Atomic species	17
4.3	Main group compounds	19
4.4	Transition-metal compounds	19

1 Introduction

Electron structure quantum Monte Carlo (QMC) refers to a number of Monte Carlo-based techniques for electron structure calculations. In this lecture, we focus on real-space QMC where the electronic Schr  dinger equation in first quantized form is the starting point. Recently, a number of Monte Carlo methods have evolved where the Schr  dinger equation is treated in the Fock space, i.e., in second quantized form. While this equation has a number of advantages, its exact solution is defined by the finite basis set, and the physical solution is obtained only after basis set extrapolation. The electronic Schr  dinger equation

$$H\Psi = E\Psi, \quad \text{with} \quad H = T + V_{en} + V_{ee} \quad (1)$$

contains kinetic energy of the electrons T , Coulombic electron-nucleus (or electron-core) attraction V_{en} , and the electron-electron repulsion V_{ee} .

There are two seemingly unrelated Monte Carlo methods for obtaining approximate solutions for Eq. (1). The first option is the evaluation of the energy expectation value $\langle H \rangle = \langle \Psi | H | \Psi \rangle$ for a trial wave function Ψ with Monte Carlo integration. The variational principle allows the optimization of the trial wave function Ψ by minimizing $\langle H \rangle$ with respect to parameters of the wave function. This method is usually called *variational Monte Carlo* (VMC). The wave function depends on all electrons, and depends thus on the real coordinates $\mathbf{r}_1, \dots, \mathbf{r}_n$ of all n electrons and, additionally, on the spin state of each electron. Note that the Monte Carlo-based integration allows arbitrary forms for the trial wave function. Therefore, Ψ can deviate from the standard Slater determinant form and, for instance, compact correlated wave functions can be employed easily. The optimization of trial wave functions is the subject of this lecture.

The other option for solving the Schr  dinger equation with Monte Carlo methods is a stochastic projection method where the exact solution of Eq. (1) is projected out of a starting wave function. These methods are known under the name projector Monte Carlo or *diffusion Monte Carlo* (DMC) and are, due to the projection, in general more accurate than the VMC methods. Surprisingly, the DMC and the VMC methods are closely related, and the optimized VMC trial wave function can improve the DMC energies substantially as we demonstrate in section 4.

This lecture is organized as follows. The VMC and DMC methods are briefly described in the next section before the wave function optimization techniques are discussed in detail in the subsequent section. Selected applications demonstrating the optimization of trial wave functions are given in the final section.

2 Real-space quantum Monte Carlo methods

2.1 Variational Monte Carlo

In this lecture, the following notation will be used: $\mathbf{r}_i = (x_i, y_i, z_i)$ denotes the cartesian coordinates of electron i while the spin is indicated by the quantum number $m_{s,i} = \pm \frac{1}{2}$.

$$\mathbf{x}_i = (\mathbf{r}_i, m_{s,i}) = (x_i, y_i, z_i, m_{s,i}) \quad (2)$$

collects all electron variables and $\mathbf{R} = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ all electrons of the (finite) system. With this notation, the energy expectation of an unnormalized wave function $\Psi(\mathbf{R})$ is given in real space by

$$E_{\text{VMC}} = \frac{\langle \Psi(\mathbf{R}) | H | \Psi(\mathbf{R}) \rangle}{\langle \Psi(\mathbf{R}) | \Psi(\mathbf{R}) \rangle}. \quad (3)$$

This ratio can be rewritten as an integral over a probability density $p(\mathbf{R})$ and the *local energy* $E_L(\mathbf{R})$

$$E_{\text{VMC}} = \int E_L(\mathbf{R}) p(\mathbf{R}) d\mathbf{R}, \quad E_L(\mathbf{R}) = \frac{H\Psi(\mathbf{R})}{\Psi(\mathbf{R})}, \quad p(\mathbf{R}) = \frac{|\Psi(\mathbf{R})|^2}{\int |\Psi(\mathbf{R})|^2 d\mathbf{R}} \quad (4)$$

where the integration extends over all space. Note that $p(\mathbf{R})$ describes the probability density of positions and spins of all electrons simultaneously.

A sample with the probability density $p(\mathbf{R})$ can be obtained efficiently with the *Metropolis-Hastings algorithm* [1, 2] without the necessity of calculating the normalization integral. The integral in Eq. (4) can then be obtained simply as the sample mean of the local energy over a large sample of $p(\mathbf{R})$

$$E_{\text{VMC}} = \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{k=1}^K E_L(\mathbf{R}_k), \quad \text{with } \mathbf{R}_k \sim p(\mathbf{R}) \quad (5)$$

where $\sim p$ means 'distributed according to the probability density p .' It is important to note that this energy evaluation requires only the ability to calculate the Laplacian of the trial wave function with respect to all electrons but no integration other than the Monte Carlo integration itself. This obviously allows substantial freedom in the choice of Ψ , and in particular the use of correlation factors coupling the electron coordinates for optimal description of the electron-electron interaction and correlation. Variational Monte Carlo refers to the energy calculation using Monte Carlo integration and Metropolis-Hastings sampling of the probability density. Further details can be found in several review articles, for instance in [3, 4] and in references therein.

2.2 Diffusion Monte Carlo

The diffusion Monte Carlo method yields more accurate energies than the VMC method because in DMC the exact ground state wave function is projected according to

$$\lim_{t \rightarrow \infty} e^{-Ht} \Psi \propto \Psi_0 \quad (6)$$

The projection of the ground state wave function is obtained after constructing the corresponding Green function with a short-time approximation which leads to a stochastic process, called *drift-diffusion process*,

$$\mathbf{R}_{k+1} = \mathbf{R}_k + \mathbf{b}(\mathbf{R}_k)\tau + \Delta W_\tau, \quad \mathbf{R}_0 \sim p(\mathbf{R}) \quad (7)$$

with a discretized time step $\tau = t/N$ and a *drift vector* $\mathbf{b}$ defined with the trial function Ψ

$$\mathbf{b}(\mathbf{R}) = \frac{\nabla \Psi(\mathbf{R})}{\Psi(\mathbf{R})}. \quad (8)$$

ΔW_τ denotes a normal variate with mean $\mu = 0$ and variance $\sigma^2 = \tau$. The drift-diffusion process implements the antisymmetry of the projected solution because the drift is directed away from the nodes of the trial function Ψ and is, at the nodal surface, in fact a normal vector of the surface.

The drift-diffusion process is coupled with a weighting process

$$w_{k+1} = w_k e^{-\left[\frac{1}{2} \left(E_L(\mathbf{R}_k) + E_L(\mathbf{R}_{k+1}) \right) - E_{\text{ref}} \right] \tau}, \quad w_0 = 1 \quad (9)$$

where a reference energy E_{ref} is employed to stabilize the process. The weighted drift-diffusion process is implemented by a sample of *walkers* which are propagated according to Eq. (7). Each walker has a weight attached which changes in each step according to Eq. (9). Note that a walker corresponds to the positions of all electrons simultaneously. The stochastic process is therefore capable of describing the electron correlation in real space. It is usually stabilized by a branching/killing process based on the walker weights. As in molecular dynamics, the exact distribution is obtained in the limit of vanishing time step, but also in the long total time limit. With the final distribution of weighted walkers the DMC energy is obtained as a weighted mean

$$E_{\text{DMC}} = \frac{\sum_{k=1}^N w_k E_L(\mathbf{R}_k)}{\sum_{k=1}^N w_k} \quad (10)$$

with the weight w_k of the k th walker with the coordinates $\mathbf{R}_k$.

VMC and DMC are closely related as the drift-diffusion process Eq. (7) is usually employed as proposed step in the Metropolis-Hastings algorithm. Both algorithms are Markov chains, and after reaching the equilibrium distribution, the energy or other estimators are calculated as 'time' and sample average where the time is not the physical time. Each step in the Markov chain is referred to as *Monte Carlo step*. Furthermore, the VMC distribution $p(\mathbf{R})$ is usually the starting distribution for the DMC process as indicated in Eq (7). But the main connection results from the trial function Ψ whose nodes, or more precisely nodal hypersurface, $\Psi(\mathbf{R}) = 0$ is imposed on the DMC solution via the drift Eq. (8). The DMC algorithm described above is known as fixed-node DMC and results in the long t and short τ limit in the exact ground state energy of the Schr  dinger equation with the nodes of the trial function Ψ as additional boundary condition. Without this boundary condition, the stochastic process would converge toward the mathematical ground state of H which is the nodeless bosonic solution. The error resulting from the fixed node approximation is called *fixed-node error*.

The optimization of the VMC energy requires the minimization of the energy expectation value Eq. (5) with respect to the parameters of Ψ . The optimization of the DMC energy would require the variation of the nodes of Ψ . Unfortunately, no efficient method is known for the direct optimization of the nodes and hence usually DMC energies are optimized indirectly via the minimization of the VMC energy.

3 Stochastic optimization of electronic wave functions

The basis of all stochastic optimizations is a large sample of electron configurations $\{\mathbf{R}_k\}$, $k = 1, \dots, K$ with a distribution of $p(\mathbf{R}) = |\Psi(\mathbf{R})|^2 / \int |\Psi(\mathbf{R})|^2 d\tau$, see Eq. (4), obtained with the Metropolis-Hastings algorithm. The calculation of the sample is costly due to the serial correlation between successive Monte Carlo steps, and it can be attempted to do a *fixed sample optimization* which is a deterministic minimization [5]. As can be expected, the resulting parameters are biased toward this sample, and in practice a small number of iterations with new, independent samples of electron configurations is required.

The trial wave function $\Psi(\mathbf{R}, \mathbf{p}) \equiv \Psi(\mathbf{p})$ and the local energy $E_L(\mathbf{R}, \mathbf{p}) \equiv E_L(\mathbf{p})$ both depend on all electron positions and a set of parameters collected in a parameter vector $\mathbf{p}$. The VMC energy as sample mean of the local energy depends on the parameters and is simply

$$E_m(\mathbf{p}) = \frac{1}{K} \sum_{k=1}^K E_L(\mathbf{R}_k, \mathbf{p}). \quad (11)$$

As the parameter optimization proceeds from the initial parameters $\mathbf{p}_0$ to $\mathbf{p}$ the fixed sample ($\sim |\Psi(\mathbf{p}_0)|^2$) does no longer correspond to the distribution $|\Psi(\mathbf{p})|^2$ of the modified wave function. The change in the distribution can be corrected with a weighted mean

$$E_{wm}(\mathbf{p}) = \frac{\sum_{k=1}^K w_k E_L(\mathbf{R}_k, \mathbf{p})}{\sum_{k=1}^K w_k}, \quad w_k = \frac{|\Psi(\mathbf{R}_k, \mathbf{p})|^2}{|\Psi(\mathbf{R}_k, \mathbf{p}_0)|^2}. \quad (12)$$

The variance of the local energy

$$V_m(\mathbf{p}) = \frac{1}{K} \sum_{k=1}^K (E_L(\mathbf{R}_k, \mathbf{p}) - E_m)^2 \quad (13)$$

determines the standard error of the sample mean energy $\sqrt{V_m}/\sqrt{K}$. The more accurate the wave function the smaller the local energy fluctuations and thus the variance. For any exact eigenfunction of H the local energy is constant and thus the sample variance exactly zero. This *zero-variance property* of the sample mean as energy estimator allows the determination of energies with small statistical error bars with reasonable sample sizes provided that accurate (i.e. low variance) wave functions are used. It is important for an efficient optimization to construct similar estimators for instance for gradients with respect to parameters.

For the optimization, the gradients of the wave function and the local energy with respect to the parameters are required. Here, we denote the parameter derivatives as follows

$$\Psi_i = \frac{\partial \Psi(\mathbf{p})}{\partial p_i}, \quad \Psi_{ij} = \frac{\partial^2 \Psi(\mathbf{p})}{\partial p_i \partial p_j}, \quad E_{L,i} = \frac{\partial E_L(\mathbf{p})}{\partial p_i}. \quad (14)$$

Minimization of the VMC energy with respect to the parameters is desired similarly to the Hartree-Fock or configuration interaction (CI) methods. Minimization of the variance is, in principle, equivalent to energy minimization because only exact eigenfunctions of H have a variance of zero. Variance minimization is substantially simpler and more efficient because the

sample variance is a sum of squares, while the local energy is possibly not bounded from below. In earlier QMC work the variance minimization was preferred while in more recent papers the energy minimization dominates because of a higher accuracy of the results.

3.1 Variance and MAD minimization

The sample variance is often simplified with an estimate E_{ref} of the mean energy (which is easily available from a previous DFT calculation)

$$V_r(\mathbf{p}) = \frac{1}{K} \sum_{k=1}^K (E_L(\mathbf{R}_k, \mathbf{p}) - E_{\text{ref}})^2 \quad (15)$$

The variance minimization is a nonlinear least-squares minimization problem that can be solved with standard methods such as the Levenberg-Marquardt method. The minimization requires the Jacobian, the derivatives $E_{L,i}(\mathbf{R}_k)$ of the residuals $E_L(\mathbf{R}_k, \mathbf{p}) - E_{\text{ref}}$, and has been employed for many years [6]. We have obtained best results with the standard code 'nl2sol' for nonlinear least squares problems [7]. The variance minimization is very stable and suitable even for initial parameters that differ substantially from the optimal ones. It is therefore often used as a first optimization step followed by an energy minimization. Nonlinear least-squares minimizations require a number of iterations. If the fixed sample is used, the variance defined in Eq. (15) ignores the change of the distribution $|\Psi(\mathbf{p})|^2$ with the parameters which can be accounted for by using a weighted mean as in Eq. (12)

$$V_{wm}(\mathbf{p}) = \frac{\sum_{k=1}^K w_k (E_L(\mathbf{R}_k, \mathbf{p}) - E_{\text{ref}})^2}{\sum_{k=1}^K w_k} \quad (16)$$

Unfortunately, the weighted variance tends to be unstable when the weights start to deviate substantially from one. In this case, the effective sample size is substantially reduced. Investigations have shown that the unweighted variance Eq. (13) or Eq. (15) leads to more stable and efficient minimizations [8].

Alternatively, it is possible to optimize the mean absolute deviation (MAD) of the local energy

$$\text{MAD}_r(\mathbf{p}) = \frac{1}{K} \sum_{k=1}^K |E_L(\mathbf{R}_k, \mathbf{p}) - E_{\text{ref}}| \quad (17)$$

Since a sum of positive elements is involved the MAD minimization is also stable and found to be advantageous in some cases [9].

3.2 Energy minimization

Energy minimization is preferable to variance minimization when the goal is to calculate energy differences. As mentioned above, energy minimization requires more computational effort meaning substantially larger samples than variance minimization. Due to the stochastic nature of QMC most codes use specifically optimized routines based on standard methods. The main

variants for energy optimization are Newton-like methods, the linear method, and the stochastic reconfiguration method. The methods require different derivatives and matrix elements and differ in efficiency. For the stochastic optimization of wave functions the number of iterations to reach a convergence threshold is only one aspect of the efficiency, another being the computational effort for each step.

The computational effort for each optimization step depends on the variance of the estimators for the required gradient or matrix elements. In many cases, different estimators can be constructed that have the same limit but show different variances. Substantial work has been done in the past in finding low variance estimators, in particular estimators whose variance vanishes as the trial wave function approaches the exact eigenfunction of H .

In this section, we follow the notation used by Toulouse and Umrigar who gave an excellent account of energy minimization methods including a discussion of earlier work [10]. The expectation value over $p(\mathbf{R})$ is denoted

$$\langle A \rangle = \int A(\mathbf{R}) p(\mathbf{R}) d\mathbf{R}, \quad p(\mathbf{R}) = \frac{|\Psi(\mathbf{R})|^2}{\int |\Psi(\mathbf{R})|^2 d\mathbf{R}}, \quad (18)$$

and the corresponding estimator is the sample mean

$$\langle A \rangle = \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{k=1}^K A(\mathbf{R}_k), \quad \mathbf{R}_k \sim p(\mathbf{R}). \quad (19)$$

3.2.1 Newton methods

When expanding $E_{\text{VMC}}(\mathbf{p})$ around $\mathbf{p}_0$ with $\mathbf{p} = \mathbf{p}_0 + \Delta\mathbf{p}$ in a Taylor series up to second order, the Newton method is obtained leading to the parameter vector change

$$\Delta\mathbf{p} = -\mathbf{h}^{-1}\mathbf{g} \quad (20)$$

with the gradient and Hessian matrix of the VMC energy

$$g_i = \frac{\partial E_{\text{VMC}}}{\partial p_i}, \quad h_{ij} = \frac{\partial^2 E_{\text{VMC}}}{\partial p_i \partial p_j} \quad (21)$$

It is numerically more stable to solve the linear system $\mathbf{h}\Delta\mathbf{p} = -\mathbf{g}$ rather than inverting the estimator of the Hessian matrix. The Newton method has been first used for energy optimization in VMC by Lin et al. [11] and later improved by Umrigar and Filippi [12] and Sorella [13]. Newton methods are known for their limited convergence radius. One common stabilization technique is the adaptive addition of a positive constant to the diagonal of the Hessian matrix $\mathbf{h}$ which has the effect of switching smoothly with growing constant to steepest descent. For details see reference [10] and references therein.

The greatest challenge for the use of the Newton method for energy minimization is the construction of low variance estimators for the gradient and the Hessian matrix. For the gradient of the VMC energy we first observe that

$$\frac{\partial E_{\text{VMC}}}{\partial p_i} = \frac{\partial \langle E_L \rangle}{\partial p_i} \neq \langle E_{L,i} \rangle \quad (22)$$

because of the dependence of the distribution $p(\mathbf{R}) = p(\mathbf{R}, \mathbf{p})$ on the parameters. An estimator of the gradient can be derived as follows

$$\begin{aligned}
 \frac{\partial}{\partial p_i} \langle E_L \rangle &= \frac{\partial}{\partial p_i} \frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle} \\
 &= \frac{\frac{\partial}{\partial p_i} \langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle} - \frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle} \frac{\frac{\partial}{\partial p_i} \langle \Psi | \Psi \rangle}{\langle \Psi | \Psi \rangle} \\
 &= \frac{\langle \Psi_i | H | \Psi \rangle}{\langle \Psi | \Psi \rangle} + \frac{\langle \Psi | H | \Psi_i \rangle}{\langle \Psi | \Psi \rangle} - 2 \langle E_L \rangle \frac{\langle \Psi_i | \Psi \rangle}{\langle \Psi | \Psi \rangle} \\
 &= \left\langle \frac{\Psi_i}{\Psi} E_L \right\rangle + \left\langle \frac{H \Psi_i}{\Psi} \right\rangle - 2 \langle E_L \rangle \left\langle \frac{\Psi_i}{\Psi} \right\rangle
 \end{aligned} \tag{23}$$

Comparing

$$E_{L,i} = \frac{H \Psi_i}{\Psi} - E_L \frac{\Psi_i}{\Psi} \Rightarrow \langle E_{L,i} \rangle = \left\langle \frac{H \Psi_i}{\Psi} \right\rangle - \left\langle \frac{\Psi_i}{\Psi} E_L \right\rangle \tag{24}$$

one obtains

$$\frac{\partial}{\partial p_i} \langle E_L \rangle = 2 \left[\left\langle \frac{\Psi_i}{\Psi} E_L \right\rangle - \left\langle \frac{\Psi_i}{\Psi} \right\rangle \langle E_L \rangle \right] + \langle E_{L,i} \rangle \tag{25}$$

The last term vanishes due to Hermiticity of H

$$\langle E_{L,i} \rangle = \left\langle \frac{H \Psi_i}{\Psi} \right\rangle - \left\langle \frac{\Psi_i}{\Psi} \frac{H \Psi}{\Psi} \right\rangle = \frac{\langle \Psi | H | \Psi_i \rangle}{\langle \Psi | \Psi \rangle} - \frac{\langle \Psi_i | H | \Psi \rangle}{\langle \Psi | \Psi \rangle} = 0 \tag{26}$$

and we obtain finally [11]

$$\frac{\partial E_{\text{VMC}}}{\partial p_i} = 2 \left[\left\langle \frac{\Psi_i(\mathbf{R})}{\Psi(\mathbf{R})} E_L(\mathbf{R}) \right\rangle - \left\langle \frac{\Psi_i(\mathbf{R})}{\Psi(\mathbf{R})} \right\rangle \langle E_L(\mathbf{R}) \rangle \right]. \tag{27}$$

Note that this estimator of the parameter derivative of the VMC energy does not depend on the parameter derivative of the local energy! The wave function derivative Ψ_i arises only in a ratio with the wave function Ψ itself. Normalization constants, which are expensive to calculate in QMC, therefore cancel. Furthermore, this estimator has the zero-variance property because the two terms in the difference become identical as the local energy becomes a constant for the exact eigenfunction of H . The difference is calculated with correlated sampling, using the same sample for both terms such that fluctuations of $\Psi_k(\mathbf{R})/\Psi(\mathbf{R})$ cancel substantially. Additionally, the estimator has the form of a covariance, and it has been observed that the fluctuations of the covariance $\langle xy \rangle - \langle x \rangle \langle y \rangle$ are much smaller than fluctuations of either x or y [12].

In the light of this result, a number of different estimators for the Hessian matrix $\mathbf{h}$ have been proposed and employed that are based on the ratios Ψ_i/Ψ , Ψ_{ij}/Ψ and the covariance form. For the Hessian matrix, no estimator without the parameter derivative of the local energy $E_{L,i}$ is known. The estimators are a bit lengthy and not reproduced here. For a discussion of the various estimators of the Hessian matrix see reference [10].

The implementation of Newton-type methods requires therefore implementations of the following three terms for every type of trial wave function

$$E_{L,i}(\mathbf{R}), \quad \frac{\Psi_i(\mathbf{R})}{\Psi(\mathbf{R})}, \quad \frac{\Psi_{ij}(\mathbf{R})}{\Psi(\mathbf{R})} \tag{28}$$

Below, we discuss the computational effort for these terms for some common wave function and Hamiltonian types.

3.2.2 Linear method

If the wave function depends linearly on the parameters as in CI or Jastrow-CI wave functions, the parameters are obtained non-iteratively by diagonalization of the Hamiltonian matrix. This approach can be extended to nonlinear parameters by expanding the wave function to first order in the parameters

$$\Psi^{(1)}(\mathbf{p}) = \Psi^{(1)}(\mathbf{p}_0 + \Delta\mathbf{p}) = \Psi(\mathbf{p}_0) + \sum_{i=1}^{n_p} \Delta p_i \Psi_i(\mathbf{p}_0) \quad (29)$$

The idea is now to consider $\Psi_i = \Psi_i(\mathbf{p}_0)$, $i = 0, \dots, n_p$, with $\Psi_0 \equiv \Psi(\mathbf{p}_0)$ and the number of parameters n_p as a basis and obtain the energy minimum by solving the matrix eigenvalue problem

$$\mathbf{H}\Delta\mathbf{p} = E\mathbf{S}\Delta\mathbf{p} \quad (30)$$

with the Hamilton $H_{ij} = \langle \Psi_i | H | \Psi_j \rangle$ and overlap matrices $S_{ij} = \langle \Psi_i | \Psi_j \rangle$. Note that $\Delta p_0 = 1$ in Eq. (29) which defines the normalization of the resulting eigenvector corresponding to the intermediate normalization in standard quantum chemistry. In the case of linear parameters the CI method is recovered. In the case of nonlinear parameters the linear expansion is repeated iteratively until convergence.

Nightingale and Melik-Alaverdian demonstrated that it is advantageous and, in fact, leads to a zero-variance property if the finite sample estimators of the Hamiltonian matrix elements do not make use of the Hermiticity of the Hamiltonian [14, 10]. This results for finite samples in *unsymmetric* Hamiltonian matrices in the eigenvalue problem and in parameter changes with substantially reduced variances.

In practice, the dependence of the normalization constant of the wave function on the parameter change is accounted for. The normalized wavefunction

$$\tilde{\Psi}_0(\mathbf{p}) = \frac{1}{\sqrt{\langle \Psi(\mathbf{p}) | \Psi(\mathbf{p}) \rangle}} \Psi(\mathbf{p}) \quad (31)$$

results in the parameter derivatives

$$\tilde{\Psi}_i(\mathbf{p}) = \frac{1}{\sqrt{\langle \Psi(\mathbf{p}) | \Psi(\mathbf{p}) \rangle}} \left(\Psi_i(\mathbf{p}) - \left\langle \frac{\Psi_i(\mathbf{p})}{\Psi_0(\mathbf{p})} \right\rangle \Psi(\mathbf{p}) \right) \quad (32)$$

with

$$\frac{\langle \Psi(\mathbf{p}) | \Psi_i(\mathbf{p}) \rangle}{\langle \Psi(\mathbf{p}) | \Psi(\mathbf{p}) \rangle} = \left\langle \frac{\Psi_i(\mathbf{p})}{\Psi_0(\mathbf{p})} \right\rangle. \quad (33)$$

The first order expansion now reads

$$\tilde{\Psi}^{(1)}(\mathbf{p}_0 + \Delta\mathbf{p}) = \tilde{\Psi}_0(\mathbf{p}_0) + \sum_{i=1}^{n_p} \Delta p_i \tilde{\Psi}_i(\mathbf{p}_0). \quad (34)$$

The finite sample estimators for the matrices are constructed as follows. H_{00} is simply the current sample mean of the local energy $H_{00} = \langle E_L \rangle$. The H_{i0} elements have the estimator

$$H_{i0} = \langle \tilde{\Psi}_i | H | \tilde{\Psi}_0 \rangle = \int \frac{\tilde{\Psi}_i(\mathbf{R})}{\tilde{\Psi}_0(\mathbf{R})} E_L(\mathbf{R}) p(\mathbf{R}) d\mathbf{R} = \left\langle \frac{\Psi_i}{\Psi_0} E_L \right\rangle - \left\langle \frac{\Psi_i}{\Psi_0} \right\rangle \langle E_L \rangle \quad (35)$$

with

$$\frac{\tilde{\Psi}_i}{\tilde{\Psi}_0} = \frac{\Psi_i}{\Psi_0} - \left\langle \frac{\Psi_i}{\Psi_0} \right\rangle. \quad (36)$$

For the H_{0i} elements we obtain with Eq. (32) and without making use of Hermiticity

$$H_{0i} = \langle \tilde{\Psi}_0 | H | \tilde{\Psi}_i \rangle = \left\langle \left(\frac{1}{\Psi_0} H \Psi_i - \left\langle \frac{\Psi_i}{\Psi_0} \right\rangle E_L \right) \right\rangle = \langle E_{L,i} \rangle + \left\langle \frac{\Psi_i}{\Psi_0} E_L \right\rangle - \left\langle \frac{\Psi_i}{\Psi_0} \right\rangle \langle E_L \rangle \quad (37)$$

using

$$E_{L,i} = \frac{1}{\Psi_0} \left(\frac{\partial H \Psi_0}{\partial p_i} - E_L \frac{\partial \Psi_0}{\partial p_i} \right). \quad (38)$$

Analogously, one can derive for the remaining matrix elements H_{ij} with $i, j > 0$ [10]

$$\begin{aligned} H_{ij} = & \left\langle \frac{\Psi_i}{\Psi_0} \frac{\Psi_j}{\Psi_0} E_L \right\rangle - \left\langle \frac{\Psi_i}{\Psi_0} \right\rangle \left\langle \frac{\Psi_j}{\Psi_0} E_L \right\rangle - \left\langle \frac{\Psi_j}{\Psi_0} \right\rangle \left\langle \frac{\Psi_i}{\Psi_0} E_L \right\rangle \\ & + \left\langle \frac{\Psi_i}{\Psi_0} \right\rangle \left\langle \frac{\Psi_j}{\Psi_0} \right\rangle \langle E_L \rangle + \left\langle \frac{\Psi_i}{\Psi_0} E_{L,j} \right\rangle - \left\langle \frac{\Psi_i}{\Psi_0} \right\rangle \langle E_{L,j} \rangle \end{aligned} \quad (39)$$

and for the overlap matrix

$$S_{ij} = \left\langle \frac{\Psi_i}{\Psi_0} \frac{\Psi_j}{\Psi_0} \right\rangle - \left\langle \frac{\Psi_i}{\Psi_0} \right\rangle \left\langle \frac{\Psi_j}{\Psi_0} \right\rangle, \quad S_{00} = 1. \quad (40)$$

Note that the estimator for the Hamilton matrix of the linear method requires the local energy derivatives $E_{L,i}$ like the estimator for the Hessian matrix of the Newton method.

In practice, after the generalized ‘‘almost symmetric’’ eigenvalue problem in Eq. (30) has been solved, the correct right eigenvector has to be identified. For finite samples the correct eigenvector may not correspond to the lowest eigenvalue as spurious solutions may occur. Since the eigenvectors are orthogonal it is not difficult to identify the physical solution. A reliable way is a very short VMC calculation with new parameters arising from all eigenvectors of a few lowest eigenvalues. Alternatively, the overlap of the new wavefunction with the previous can be estimated, see [10, 15] for further discussions. For a large number of parameters the linear method becomes restricted by the necessity to store the Hamilton and overlap matrices. Zhao and Neuscammam have recently presented a blocked linear method that has been shown to be a very good approximation to the original linear method while requiring substantially less memory which is particularly important for the parallel implementation [16]. The linear method, in one of its variants, turns out to be the most efficient energy minimization method in most applications.

3.2.3 Stochastic reconfiguration, energy fluctuation potential, and perturbative optimization methods

Both the linear method and the Newton method require the calculation of the local energy parameter derivative $E_{L,i}$ which is computationally expensive if pseudo potentials are used. A possibly cheaper energy optimization method is the *stochastic reconfiguration method* [17] and its variants, the *energy fluctuation potential* method [18] (EFP) and the *perturbative optimization* method [10]. The stochastic reconfiguration method is based on the expansion of the DMC propagator to first order $\exp(-tH) \approx 1 - tH$ applied to the trial wave function Ψ and projected into the space of derivative functions Ψ_i including $\Psi_0 \equiv \Psi$.

$$\langle \Psi_i | 1 - tH | \Psi \rangle = \sum_{j=0}^{n_p} \alpha_j \langle \Psi_i | \Psi_j \rangle \quad (41)$$

The terms of this linear equation can be sampled analogously to the linear method, and the resulting parameter changes are $\Delta p_i = \alpha_i / \alpha_0$. For the related EFP method we refer to Ref. [18] and references therein. In this paper, Scemama and Filippi also derived an efficient perturbative EFP method that we present here in the notation by Toulouse and Umrigar used so far who discuss the various optimization methods in some detail [10]. The formula for the perturbative method is

$$\Delta p_i = -\frac{1}{\Delta \varepsilon_i} \sum_{j=1}^{n_p} (\mathbf{S}^{-1})_{ij} H_{j0} \quad (42)$$

with

$$\Delta \varepsilon_i = \frac{H_{ii}}{S_{ii}} - H_{00}. \quad (43)$$

By comparison of Eq. (35) and (27), H_{j0} is identified as half the gradient $\mathbf{g}$ of the VMC energy. Hence, Eq. (42) has the form of a Newton step with the inverse overlap matrix as approximative Hessian. In practice, $\Delta \mathbf{p}$ is obtained from the corresponding linear equation. The main difference to the expressions obtained from Eq. (41) lies in the factor $1/\Delta \varepsilon_i$ replacing the scalar t in the stochastic reconfiguration method. This adds flexibility to the method and leads to improved convergence [18]. The local energy derivative $E_{L,i}$ is required for the denominator $\Delta \varepsilon_i$ but not for H_{j0} . If $E_{L,i}$ is expensive to calculate one may replace $\Delta \varepsilon_i$ by a constant $\Delta \varepsilon$ as is done in the stochastic reconfiguration method [17, 10]. A less drastic approximation that we have successfully applied is a more approximative calculation of $\Delta \varepsilon_i$ only in the first iteration step and with a small sample. The convergence of the three variants are demonstrated in Figure 1 for the orbital optimization in C_2 with a full valence CAS and Jastrow wave function and Burkatzki pseudo potentials to remove the core electrons [19, 20]. It can be seen that the calculation of all $\Delta \varepsilon_i$ only in the first step with a small sample does not lead to a diminished accuracy or slower convergence while the convergence is substantially slower if all $\Delta \varepsilon_i$ are replaced by a constant [21].

In Figure 2, a comparison of all three optimization techniques is shown for the MO optimization of C_2 in a full valence CAS wave function with a polarized triple zeta STO basis and a Jastrow factor. In this example, all three methods converge extremely fast and lead to a significant

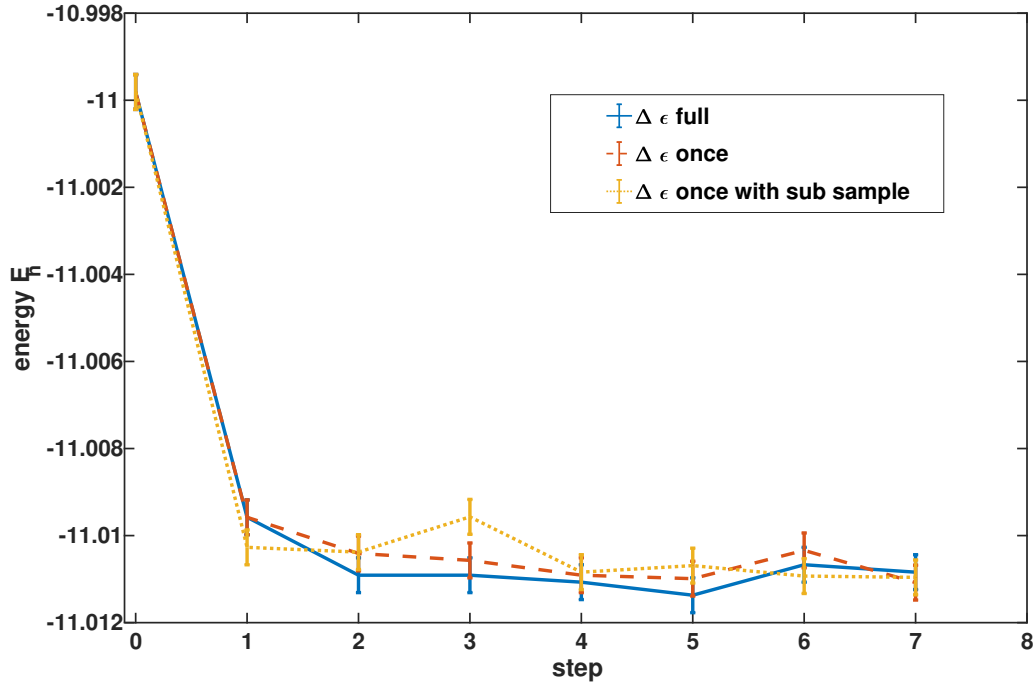


Fig. 1: *Orbital optimization in C_2 with full valence CAS and Jastrow using ECPs with the perturbative method with full $\Delta\epsilon$ calculation, calculation of $\Delta\epsilon$ only in the first iteration, and only in the first iteration with a small sample [21].*

drop in energy. While the linear method converges usually very quickly for all parameters the perturbative method does not work well for Jastrow parameters [10].

The stochastic reconfiguration method was computationally accelerated by Neuscamman et al. who observed that the overlap matrix $\langle \Psi_i | \Psi_j \rangle$ does not have to be built explicitly if stochastic estimators for matrix-vector products are used [22]. The replacement of explicit matrices by matrix-vector products is widely employed in numerical methods, often allowing to increase the dimension of systems substantially. The authors show that the calculation of matrix-vector products is only possible when the conjugate-gradient method is used to solve the linear equation Eq. (41). The accelerated stochastic reconfiguration method allowed the optimization of more than hundred thousand parameters in an massively parallel implementation.

3.3 Terms required for parameter optimization

The implementation of the energy and variance minimization requires parameter derivatives of the wave function and the local energy. We discuss here a few important aspects of these parameter derivatives with the restriction to the commonly used Slater-Jastrow wave functions

$$\Psi(\mathbf{c}, \boldsymbol{\kappa}, \mathbf{q}) = \Phi(\mathbf{c}, \boldsymbol{\kappa}) e^{U(\mathbf{q})} \quad (44)$$

with a Jastrow correlation function U with parameter vector $\mathbf{q}$ and, in general, a linear combination of Slater determinants Φ_d composed of orthogonal molecular orbitals (MO) depending

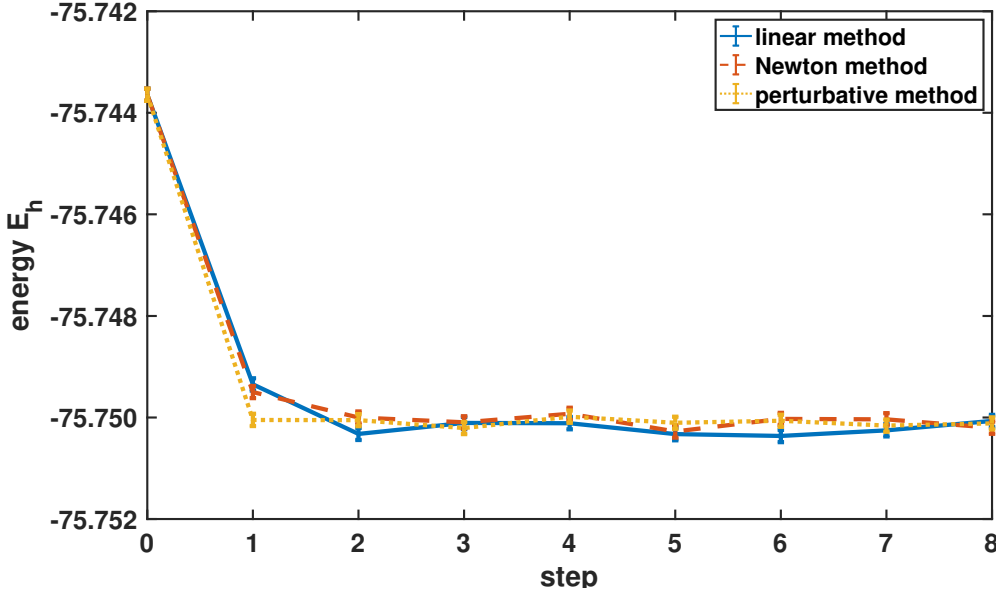


Fig. 2: Comparison of the linear method with the Newton and the perturbative method for the MO optimization in the C_2 molecule [21].

on parameters κ

$$\Phi(\mathbf{c}, \kappa) = \sum_{d=1}^{n_{\text{det}}} c_d \Phi^{(d)}(\kappa). \quad (45)$$

In the following, we stick to the above notation and indicate derivatives with respect to parameter p_i with subscript i . In the given formulae, the parameter derivatives arise only in the ratio Ψ_i/Ψ which is then given here. For Jastrow parameters we obtain

$$\frac{\Psi_i}{\Psi} = \frac{1}{\Psi} \frac{\partial \Psi(\mathbf{c}, \kappa, \mathbf{q})}{\partial q_i} = \frac{\partial U}{\partial q_i} =: U_i \quad (46)$$

and for the CI parameters c_i

$$\frac{\Psi_i}{\Psi} = \frac{1}{\Psi} \frac{\partial \Psi(\mathbf{c}, \kappa, \mathbf{q})}{\partial c_i} = \frac{\Phi^{(i)}}{\Phi}. \quad (47)$$

The derivatives with respect to the MO parameters are somewhat more involved. We obtain

$$\frac{\Psi_i}{\Psi} = \frac{1}{\Psi} \frac{\partial \Psi(\mathbf{c}, \kappa, \mathbf{q})}{\partial \kappa_i} = \frac{1}{\Phi} \sum_{d=1}^{n_{\text{det}}} c_d \Phi_i^{(d)} \quad (48)$$

where the derivative of a Slater determinant $\Phi_i^{(d)} = \partial \Phi^{(d)} / \partial \kappa_i$ is required.

All Slater determinants are formed by a set of orthogonal MOs which form, as column vectors corresponding to a basis set, a matrix Φ . The discussion here is restricted to real orbitals. Each orbital rotation $\hat{x}_{ij}$ acts on this matrix by mixing MO i and j with the angle x_{ij}

$$\hat{x}_{ij} \Phi = \cos x_{ij} \Phi + \sin x_{ij} \left(\hat{a}_i^\dagger \hat{a}_j - \hat{a}_j^\dagger \hat{a}_i \right) \Phi. \quad (49)$$

It is sufficient to calculate derivatives for the current set of orbitals corresponding to $x_{ij} = 0$. Taking the derivative with respect to the rotation angle at $x_{ij} = 0$ results in

$$\left. \frac{\partial}{\partial x_{ij}} \hat{x}_{ij} \Phi \right|_{x_{ij}=0} = \left(\hat{a}_i^\dagger \hat{a}_j - \hat{a}_j^\dagger \hat{a}_i \right) \Phi.$$

With the singlet excitation operator as defined in standard second quantized quantum chemistry [23]

$$\hat{E}_{ij}^- = \hat{a}_i^\dagger \hat{a}_j - \hat{a}_j^\dagger \hat{a}_i$$

we obtain for all Slater determinants $\Phi^{(d)}$ built from this set of MOs

$$\left. \frac{\partial}{\partial x_{ij}} \hat{x}_{ij} \Phi^{(d)} \right|_{x_{ij}=0} = \hat{E}_{ij}^- \Phi^{(d)}. \quad (50)$$

Note that the operator $\hat{E}_{ij}^-$ has to be applied for each spin orbital. Hence, the derivative of a Slater determinant with respect to an orbital rotation angle (at zero angle) is obtained simply as the value of the corresponding singly excited determinant.

The orbitals are usually partitioned in the inactive, active, and virtual orbitals where inactive orbitals are occupied in all determinants, active orbitals are occupied in some of the determinants, and virtual orbitals are not occupied in any determinant. Orbital rotations between two inactive and between two virtual orbitals do not change the Slater determinant except for possibly a sign change. In complete active space (CAS) wave functions, active-active rotations are also invariant.

The MO parameters κ are therefore built from the set of non-invariant orbital rotations x_{ij} forming a real antisymmetric matrix $\mathbf{X}$. The general treatment of orbital rotations makes use of the fact that exponentials of antisymmetric (or skew symmetric) matrices are unitary. Multiplication of the current set of MOs Φ by $\mathbf{U} = \exp(\mathbf{X})$ results in the new set of MOs. The unitary matrix $\mathbf{U}$ can be obtained with [23]

$$\mathbf{U} = \mathbf{W} \cos \tau \mathbf{W}^T + \mathbf{W} \tau^{-1} \sin \tau \mathbf{W}^T \mathbf{X} \quad (51)$$

where $\mathbf{W}$ and $-\tau^2$ are the eigenvectors and eigenvalues of the symmetric matrix $\mathbf{X}^2$, respectively. For more details on the MO optimization see Ref. [10] and references given therein.

3.3.1 Gradient of the local energy E_L

For variance minimization as well as for some of the energy minimization methods, the parameter derivative of the local energy E_L is required. We obtain straightforwardly for $H = T + V$

$$E_{L,i} := \frac{\partial E_L}{\partial p_i} = \frac{H \Psi_i}{\Psi} - E_L \frac{\Psi_i}{\Psi} = -\frac{1}{2} \frac{\nabla^2 \Psi_i}{\Psi} + \frac{1}{2} \frac{\nabla^2 \Psi}{\Psi} \frac{\Psi_i}{\Psi} + \frac{\partial V}{\partial p_i} \quad (52)$$

where the last term, the parameter derivative of the potential V , vanishes unless a localized pseudo potential is used. Note that the Laplacians $\nabla^2 \Psi$ and $\nabla^2 \Psi_i$ refer to second derivatives with respect to all electron positions, but not to parameter derivatives. The Laplacian of the

wave function derivative $\nabla^2\psi_i$ is the new term in Eq. (52). Based on the Laplacian of the wave function $\nabla^2\psi$ which is required already for the local energy evaluation, the parameter derivatives can be established for Jastrow, CI, and MO parameters similarly to the formulae given above.

More interesting is the common case of the presence of a nonlocal pseudo potential $\hat{V}_{nl}$ in addition to a local potential V_{loc} . In this case we obtain for the local energy

$$E_L = -\frac{1}{2} \frac{\nabla^2\psi}{\psi} + V_{loc} + V_{nl}, \quad V_{nl} = \frac{\hat{V}_{nl}\psi}{\psi} \quad (53)$$

with the *localized nonlocal pseudo potential* V_{nl} .

The parameter derivative becomes

$$E_{L,i} = -\frac{1}{2} \frac{\nabla^2\psi_i}{\psi} + \frac{1}{2} \frac{\nabla^2\psi}{\psi} \frac{\psi_i}{\psi} + \frac{\partial V_{nl}}{\partial p_i} \quad (54)$$

and $E_{L,i}$ differs from the case without pseudo potential only by the last term

$$\frac{\partial V_{nl}}{\partial p_i} = \frac{\hat{V}_{nl}\psi_i}{\psi} - \frac{\hat{V}_{nl}\psi}{\psi} \frac{\psi_i}{\psi} = \frac{\hat{V}_{nl}\psi_i}{\psi} - V_{nl} \frac{\psi_i}{\psi} \quad (55)$$

The localized pseudo potential is usually calculated numerically on a spherical Lebedev grid [24]. The parameter derivatives $\hat{V}_{nl}\psi_i/\psi$ can be calculated analogously. In spite of efficient update formulae the numerical integration has to be done for all electrons within a cut off radius. This is a computationally expensive step, and more so the calculation of the parameter derivative vector. The Lebedev integration is particularly costly for the MO optimization which is based on singlet excitations $\hat{E}_{ij}^-\Phi^{(d)}$ according to Eq. (50) because a large number of determinants is constructed, all of which require an evaluation at the Lebedev grid points. This step is the reason why the energy minimization for MO parameters is computationally especially expensive with methods requiring $E_{L,i}$ and why the perturbative method is the most efficient choice in this case, provided that $\Delta\varepsilon_i$ in Eq. (43) is calculated only approximately with a small sample and possibly only once.

4 Examples of wave function optimizations

In this section, simultaneous and alternating parameter optimizations are discussed before some examples of trial wave function optimizations are presented. In particular, it will be shown that the VMC energy minimization can lead to substantial and systematic improvements of the nodal surface and thus to more accurate DMC energies. Single- as well as multi-reference approaches are employed. The Jastrow, CI, and MO parameters are both partially and fully optimized in order to assess the effect on the variational energy. Results are given for atomic species, main group, and transition metal dimers. Dissociation energies and spectroscopic constants are presented. The BFD pseudo potentials by Burkatzki, Filippi, and Dolg [19, 20] which are specifically designed for QMC are used with the corresponding triple zeta basis set in all calculations. The Jastrow factors in the examples below contain two-electron and two-electron-nucleus terms as described in [29].

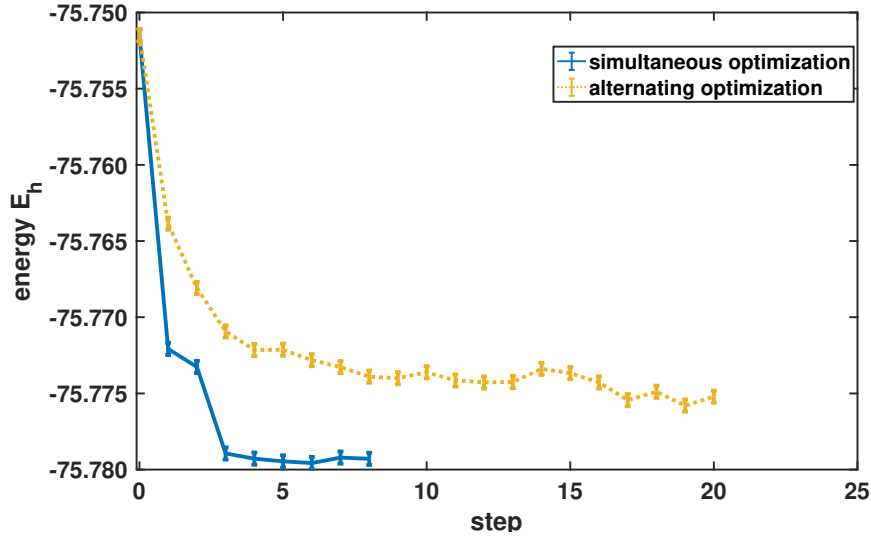


Fig. 3: Comparison of simultaneous and alternating optimization of Jastrow, MO, and CI parameters in C_2 with an all electron full valence CAS wave function with Jastrow factor using the linear method [21].

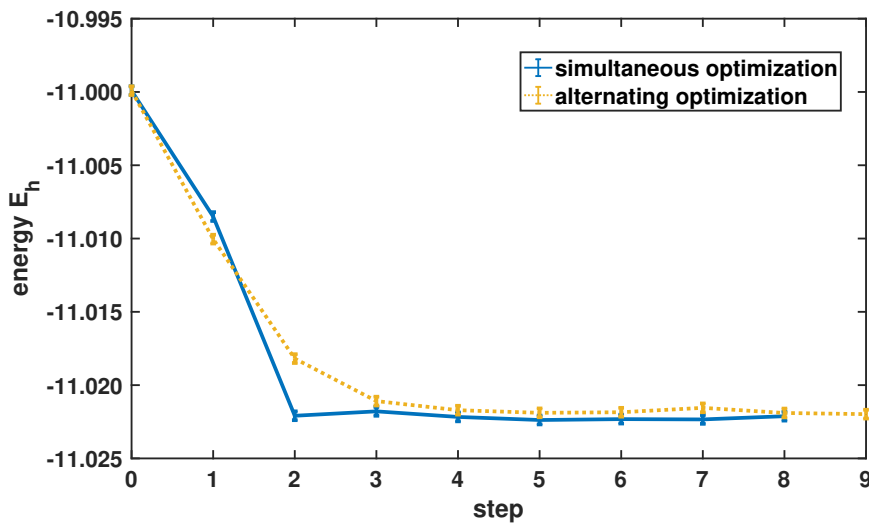


Fig. 4: Comparison of simultaneous and alternating optimization of Jastrow, MO, and CI parameters in C_2 with pseudo potentials using the linear method [21].

4.1 Simultaneous and alternating parameter optimizations

Not all optimization methods work equally well for all kinds of parameters. For instance, the linear method is expected to obtain the optimal CI parameters in one step as discussed above. On the other hand, the linear method is computationally more expensive than the perturbative method for the optimization of MO parameters when pseudo potentials are employed. It would therefore be desirable to optimize different kinds of parameters alternately with possibly different methods. Alternating optimization instead of simultaneous optimization can lead to substantially slower convergence if the parameters are strongly coupled. As an example, we show in Figure 3 the comparison of simultaneous and groupwise alternating optimization of

Jastrow, MO, and CI parameters in C_2 with an all electron full valence CAS wave function with Jastrow factor using the linear method. The slow convergence of the alternating optimizations in comparison to the very fast convergence in the simultaneous case is very obvious. On the other hand, Figure 4 demonstrates that alternating optimization can be very efficient. In this example, the same C_2 molecule is treated, the only difference being the use of pseudo potentials. In general, it has been observed that slow convergence is found in alternating optimizations for all electron calculations while the alternating optimization is efficient when pseudo potentials are used [21].

4.2 Atomic species

4.2.1 Nickel spectrum

The energy gaps between different atomic states is a quantity that can be experimentally determined in a very accurate way, therefore the opportunity to verify the accuracy of the employed method presents itself. The ground state of Nickel, as well as two excited states, are evaluated. Ni is assigned a ground state of ^3F with an occupation of $3d^84s^2$. The first excited state is a ^3D state with an occupation of $3d^94s^1$ which is energetically close to the ground state. The third considered state is of ^1S symmetry with a $3d^{10}$ occupation.

The VMC and the zero time-step extrapolated DMC energies can be found in Table 1. The calculations were performed with HF and CAS orbitals. All parameters are optimized with respect to the energy using the linear method except for the MO parameters that are optimized with the perturbative method using a small sample only for $\Delta\varepsilon_i$.

Table 1: Ni VMC and DMC energies in E_h for three states at various optimization levels, using different starting orbitals and BFD-VTZ/sm666.

State	Wave Function Ansatz	Optimization level	VMC energy	DMC energy
^3F	HF	Jas	-170.0563(4)	-170.1146(5)
		Jas+MO	-170.0645(4)	-170.1185(5)
	[10,6]-CAS	Jas+CI	-170.0554(4)	-170.1151(5)
		Jas+MO+CI	-170.0636(4)	-170.1212(5)
^3D	HF	Jas	-170.0565(4)	-170.1152(5)
		Jas+MO	-170.0641(4)	-170.1189(5)
	[10,6]-CAS	Jas	-170.0558(4)	-170.1152(5)
		Jas+MO ^a	-170.0656(4)	-170.1199(5)
^1S	HF	Jas	-169.9964(4)	-170.0500(5)
		Jas+MO	-169.9972(3)	-170.0506(5)
	[10,6]-CAS	Jas+CI	-169.9991(4)	-170.0525(6)
		Jas+MO+CI	-170.0006(4)	-170.0532(5)

^a Only one CSF is obtained.

Table 2: DMC energy gaps in E_h between the ground state and two excited states of Ni at various optimization levels, using different starting orbitals and BFD-VTZ/sm666. The experimental values are taken from the NIST Atomic Spectra Database [25].

Transition	Method	Optimization level	ΔE
$^3F - ^3D$	HF	Jas	-0.016(19)
		Jas+MO	-0.011(19)
	CAS	Jas	-0.003(19)
		Jas+MO	0.035(19)
	NIST		0.02539
$^3F - ^1S$	HF	Jas	1.759(21)
		Jas+MO	1.848(21)
	CAS	Jas+CI	1.703(21)
		Jas+MO+CI	1.850(21)
	NIST		1.82614

The MO optimization leads to an improvement of the energy, regardless of choice of initial orbitals. The lowering is however only significant for the ^3F and the ^3D state. Table 1 shows that the correct ground state of Ni cannot be reproduced at the VMC level. Even after MO optimization, the ^3D state exhibits a lower energy than ^3F . This may be rectified by choosing a more accurate Jastrow correlation function. For the Ni ^3F state, the HF trial wave function yields slightly lower VMC energies than the CAS one. This can be traced back to a loss of symmetry of the HF trial wave function compared to the CAS one. The higher flexibility of the former leads to a lowering of the energy. As for the DMC results, the MO optimization again leads to lower energies which can be explained by the improved nodal surface of the guide functions. The correct ground state can only be portrayed by the CAS guide function at the highest optimization level.

The excitation energies for the different transitions are shown in Table 2. The energy gap is severely underestimated by the HF guide function as well as by the CAS guide function without orbital optimization. The negative energy gaps mirror the fact that these approaches cannot portray the correct ground state for Nickel. One can thus deduce that the single determinant ansatz is not suitable to describe the states of the Ni atom and that the optimization of the orbital parameters is essential. The energy gap, derived from the NIST database, can however be reproduced by the CAS guide function where the orbitals are optimized in the presence of a Jastrow correlation function.

Table 3: C_2 VMC and DMC energies in E_h at various optimization levels (Jas = Jastrow only), using different starting orbitals and BFD-VTZ/SM-9t. The data are taken from Ref. [21].

Ansatz	Orbitals	Optimization level	VMC energy	DMC energy
[8,8]-CAS	CAS	Jas	-11.0680(2)	-11.0886(2)
	CAS	Jas+CI	-11.0779(3)	-11.0925(3)
	opt	Jas+MO+CI	-11.0792(2)	-11.0934(2)

Table 4: C_2 MR-DMC dissociation energies in eV at various optimization levels, using different starting orbitals and BFD-VTZ/SM-9t. The data are taken from Ref. [21].

Ansatz	Orbitals	Optimization level	D_0
CAS	CAS	Jas	6.351(9)
	CAS	Jas+CI	6.368(9)
	opt	Jas+MO+CI	6.378(7)
	exp.		6.30(2) [27]

4.3 Main group compounds

4.3.1 Carbon dimer C_2

The carbon dimer C_2 is the benchmark compound for static correlation at equilibrium bond distance, the small number of electrons making it easily feasible for multi-reference calculations. The VMC and time-step extrapolated DMC energies are given in Table 3. In contrast to the other calculations, the nine-term SM-9t Jastrow as suggested by Schmidt and Moskowitz [26] is employed for the C_2 calculations. The Jas+CI optimization improves both the VMC and DMC energies considerably. By further optimizing the molecular orbital parameters, a lowering of the energies can be observed, it is however less significant. The close DMC energies at the Jas+CI optimization level and for the fully optimized guide function indicate that the nodes must be similar.

Table 4 shows the computed MR-DMC dissociation energies at different optimization levels. The given dissociation energies D_0 are corrected for zero point energy, the core-valence correlation contribution, and spin-orbit contributions. All multi-reference dissociation energies presented here agree well with experiment, which is rather surprising since the absolute DMC energies differ significantly from one another.

4.4 Transition-metal compounds

4.4.1 Absolute energies

In this section, the VMC and DMC energies of ZnO and the effect that the parameter optimization has on them are evaluated. The VMC and the time-step extrapolated DMC energies are presented in Table 5. The non-optimized parameters are taken from the respective ab initio calculations.

Table 5: *ZnO VMC and DMC energies in E_h at various optimization levels (Jas = Jastrow only), using different starting orbitals and BFD-VTZ/sm666. The data are taken from ref. [28]*

Ansatz	Orbitals	Optimization level	VMC energy	DMC energy
Single det	HF	Jas	-242.8836(3)	-242.9931(5)
	B3LYP	Jas	-242.8944(3)	-243.0022(5)
	opt	Jas+MO	-242.9013(3)	-243.0065(6)
[16,9]-CAS	CAS	Jas	-242.8971(3)	-242.9950(5)
	CAS	Jas+CI	-242.9047(3)	-243.0023(6)
	opt	Jas+MO+CI	-242.9176(3)	-243.0111(5)

Table 5 shows a systematic lowering of the VMC energies when moving from HF over B3LYP KS to CAS orbitals. By only optimizing the Jastrow parameters, similar energies are obtained for the single- and multi-determinant trial wave functions. This shows how crucial the optimization of the antisymmetric part of the wave function is. The vast lowering in VMC energy due to the optimization of the molecular orbitals in the presence of a Jastrow correlation factor shows the substantial impact that the coupling between dynamic and static correlation has on the energies calculated with HF and CAS orbitals. For the B3LYP KS orbitals on the other hand, the MO optimization only has a small effect on the energy of the trial wave function because DFT is able to partly capture this coupling.

The MO optimization significantly improves the nodal surface of the CAS guide function. The improvement is less substantial for the KS guide function which indicates that the nodal surface was already almost optimal for this approach before the optimization. Only the fully optimized guide function is able to yield a DMC energy that is lower than the one calculated for the single-determinant approach with optimized orbitals. This emphasizes the impact that the dynamic correlation has on the quality of the nodal surface. If the molecular orbitals are not optimized, the Jastrow optimization does not change the nodal surface, due to its totally symmetric nature.

4.4.2 Dissociation energies

As example for accurate calculations of dissociation energies of small transition metal compounds we discuss the diatomics FeH, FeO, FeS, and ZnO. For all these molecules accurate experimental dissociation energies are known, while calculated dissociation energies even for accurate methods deviate substantially from the experimental data. We can demonstrate that the nodal surface obtained by VMC energy minimization of Jastrow, CI, and MO parameters in multireference Slater-Jastrow wave functions does lead to accurate dissociation energies with the DMC method. Only the VMC energy minimization for the molecule FeH is discussed here in detail as a typical example for an accurate transition metal wave function. The wave function is constructed from a [9,7]-CAS, built from the $3d$ and $4s$ orbitals on Fe and $1s$ on H, and the sm666 Jastrow factor from Ref. [29]. This results in 69 Jastrow parameters, 159 MO parameters, and 29 CI parameters. The energy minimization steps are shown in Figure 5 after

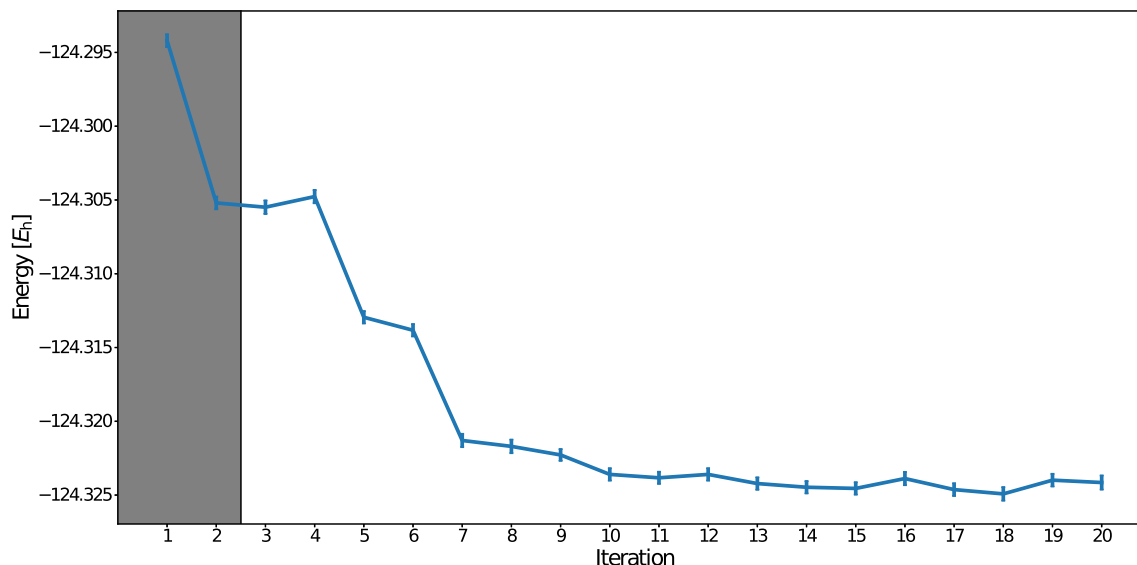


Fig. 5: VMC energy minimization of the diatomic FeH using alternating optimization of Jastrow, CI, and MO parameters using the linear method for Jastrow and CI, and the perturbative method for the MO parameters. The first two steps shown in grey refer to preoptimization of the Jastrow and MO parameters.

an initial variance optimization of the Jastrow parameters. The Jastrow and MO parameters are preoptimized with the linear and the perturbative method, respectively (shown in grey). Then the Jastrow, MO, and CI parameters are groupwise alternately optimized yielding a smooth and fast convergence.

The dissociation energies of several transition metal compounds at various optimization levels are illustrated in Table 6. The data for ZnO, FeH, and FeO are taken from Ref. [28], while the ones for FeS are presented in Ref. [21]. First of all, a systematic lowering of the deviation from the experimental dissociation energies can be observed for all compounds for the different ansätze and optimization levels. The dissociation energies of ZnO and FeO follow similar trends. The single-determinant guide function with optimized orbitals yields more accurate dissociation energies than the CAS nodes at the Jas+CI optimization level. The optimization of the CAS orbitals in the presence of a Jastrow correlation function substantially improves the dissociation energy. For both compounds, an excellent agreement with the experiment is obtained.

For FeH, the single-determinant DMC nodes fail to reproduce the dissociation energy of Schultz and Armentrout [30]. The optimization of the KS orbitals does not change the dissociation energy which implies that the nodal surface was already almost optimal for the single-determinant approach before the optimization. Both the CI and MO optimization significantly improve the dissociation energy of the CAS guide function. The MR-DMC dissociation energy computed for the fully optimized guide function agrees well with the experimental one.

Moving towards a multi-reference guide function is also necessary for the FeS system. The single-determinant approach underestimates the experimental dissociation energy, while the fully optimized MR-DMC ansatz is able to reproduce it.

Table 6: DMC dissociation energies in eV for several transition metal dimers at various optimization levels, using different starting orbitals and BFD-VTZ/sm666. The energies are corrected for core-valence correlation and spin-orbit effects.

Compound	Ansatz	Orbitals	Optimization level	D_0
ZnO	Single det	HF	Jas	1.201(19)
		B3LYP	Jas	1.449(19)
		opt	Jas+MO	1.566(21)
	CAS	CAS	Jas	1.253(19)
		CAS	Jas+CI	1.451(21)
		opt	Jas+MO+CI	1.691(19)
	exp.			1.61(4) [31]
FeO	Single det	HF	Jas	2.885(20)
		B3LYP	Jas	3.688(20)
		opt	Jas+MO	3.826(20)
	CAS	CAS	Jas	3.266(20)
		CAS	Jas+CI	3.761(20)
		opt	Jas+MO+CI	4.112(20)
	exp.			4.18(1) [32, 33]
FeH	Single det	HF	Jas	0.814(17)
		B3LYP	Jas	1.020(17)
		opt	Jas+MO	1.020(17)
	CAS	CAS	Jas	1.099(17)
		CAS	Jas+CI	1.369(17)
		opt	Jas+MO+CI	1.791(17)
	exp.			1.63(8) [30]
FeS	Single det	opt	Jas+MO	2.914(15)
	CAS	opt	Jas+MO+CI	3.159(15)
	exp.			3.31(15) [34]

4.4.3 Spectroscopic constants

The evaluation of quantities, such as the equilibrium bond distance, the harmonic frequency, and the anharmonicity allows an assessment of the accuracy of the employed method. Those spectroscopic constants are presented in Table 7 for different compounds. The data for the oxides and for FeH are taken from Ref. [28], the ones of FeS from Ref. [35].

The potential energy curves were computed for the fully optimized MR-DMC guide functions at a fixed time-step and then fitted to a Morse function, from which the spectroscopic constants could be deduced.

Table 7: *Spectroscopic constants for the different transition metal compounds. The equilibrium bond distance is given in Å, the harmonic frequency and the anharmonicity in cm⁻¹.*

System	Investigators	Method	r_e	ω_e	$\omega_e x_e$
ZnO	This work	MR-DMC	1.709	746(8)	4.4(1)
	Zack et al. [37]	Direct-absorption methods	1.7047(2)	738	4.88
	Fancher et al. [38]	Photoelectron spectrum		805(40)	
FeO	This work	MR-DMC	1.623	866(79)	4.7(7)
	Allen et al. ^a		1.619		
	Drechsler et al. [39]	anion-ZEKE		882	4
FeH	This work	MR-DMC	1.567	1842(27)	38.9(9)
	Philips et al. [40]	Near IR spectrum		1826.86	31.96
	Dulick et al. [41]			1831.8(19)	34.9(9)
FeS	This work	MR-DMC	2.031(7)	499(11)	2.53(11)
	Takano et al. [36]	Microwave spectrum	2.017		
	Wang et al. [42]	Fluorescence spectroscopy		518(5)	1.7(2)

^a derived from Allen et al. [43]

The MR-DMC bond distances of ZnO and FeO agree well with the experimental ones. The computed equilibrium bond distance of FeS is slightly larger than the one from Takano et al. [36]. The harmonic frequencies as well as the anharmonicities are in good agreement with the experiment for the four dimers. All in all, we showed that MR-DMC can yield promising results for the compounds evaluated here, provided that all sets of parameters of the trial wave functions are optimized. In particular, the difficult energy minimization of the MO parameters does lead to substantially improved accuracy of the VMC and DMC results.

Acknowledgements

Jil Ludovicy and Kaveh Haghighi Mood contributed to this lecture. Computational resources from JARA-HPC grant rwth0278 are gratefully acknowledged.

References

- [1] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953)
- [2] W.K. Hastings, Biometrika **57**, 97 (1970)
- [3] A. Lüchow, WIREs Comput. Mol. Sci. **1**, 388 (2011)
- [4] B.M. Austin, D.Y. Zubarev, and W.A. Lester, Chem. Rev. **112**, 263 (2011)
- [5] Z. Sun, S. Huang, R.N. Barnett, and W.A. Lester, J. Chem. Phys. **93**, 3326 (1990)
- [6] C.J. Umrigar, K.G. Wilson, and J.W. Wilkins, Phys. Rev. Lett. **60**, 1719 (1988)
- [7] J.E. Dennis, Jr., D.M. Gay, and R.E. Welsch, ACM Trans. Math. Softw. **7**, 369 (1981)
- [8] P.R.C. Kent, R.J. Needs, and G. Rajagopal, Phys. Rev. B **59**, 12344 (1999)
- [9] D. Bressanini, G. Morosi, and M. Mella, J. Chem. Phys. **116**, 5345 (2002)
- [10] J. Toulouse and C.J. Umrigar, J. Chem. Phys. **126**, 084102 (2007)
- [11] X. Lin, H. Zhang, and A.M. Rappe, J. Chem. Phys. **112**, 2650 (2000)
- [12] C.J. Umrigar and C. Filippi, Phys. Rev. Lett. **94**, 150201 (2005)
- [13] S. Sorella, Phys. Rev. B **71**, 241103 (2005)
- [14] M.P. Nightingale and V. Melik-Alaverdian, Phys. Rev. Lett. **87**, 043401 (2001)
- [15] J. Toulouse and C.J. Umrigar, J. Chem. Phys. **128**, 174101 (2008)
- [16] L. Zhao and E. Neuscamman, J. Chem. Theory Comput. **13**, 2604 (2017)
- [17] S. Sorella, Phys. Rev. B **64**, 024512 (2001)
- [18] A. Scemama and C. Filippi, Phys. Rev. B **73**, 241101 (2006)
- [19] M. Burkatzki, C. Filippi, and M. Dolg, J. Chem. Phys. **126**, 234105 (2007)
- [20] M. Burkatzki, C. Filippi, and M. Dolg, J. Chem. Phys. **129**, 164115 (2008)
- [21] K. Haghighi Mood: *Full Wave Function Optimization with Quantum Monte Carlo* (Ph.D. thesis, RWTH Aachen University, 2018)
- [22] E. Neuscamman, C.J. Umrigar, and G.K.-L. Chan, Phys. Rev. B **85**, 045103 (2012)
- [23] T. Helgaker, P. Jorgensen, and J. Olsen: *Molecular Electronic-Structure Theory* (Wiley, Chichester, 2002)

- [24] L. Mitáš, E.L. Shirley, and D.M. Ceperley, *J. Chem. Phys.* **95**, 3467 (1991)
- [25] A. Kramida, Y. Ralchenko, J. Reader, and N.A. Team: *NIST Atomic Spectra Database (ver. 5.5)* (<http://physics.nist.gov/asd>, 2017)
- [26] K.E. Schmidt and J.W. Moskowitz, *J. Chem. Phys.* **93**, 4172 (1990)
- [27] R.S. Urdahl, Y. Bao, and W.M. Jackson, *Chem. Phys. Lett.* **178**, 425 (1991)
- [28] J. Ludovicy, K. Haghighi Mood, and A. Lüchow, arXiv/1906.10419
- [29] A. Lüchow, A. Sturm, C. Schulte, and K. Haghighi Mood, *J. Chem. Phys.* **142**, 084111 (2015)
- [30] R.H. Schultz and P.B. Armentrout, *J. Chem. Phys.* **94**, 2262 (1991)
- [31] D.E. Clemmer, N.F. Dalleska, and P.B. Armentrout, *J. Chem. Phys.* **95**, 7263 (1991)
- [32] D.A. Chestakov, D.H. Parker, and A.V. Baklanov, *J. Chem. Phys.* **122**, 084302 (2005)
- [33] M. Li, S.-R. Liu, and P.B. Armentrout, *J. Chem. Phys.* **131**, 144310 (2009)
- [34] J. Drowart, A. Pattoret, and S. Smoes, in *Proc. Brit. Ceram. Soc.* (1967), Vol. 8, pp. 67–89
- [35] K. Haghighi Mood and A. Lüchow, *J. Phys. Chem. A* **121**, 6165 (2017)
- [36] S. Takano, S. Yamamoto, and S. Saito, *J. Mol. Spectrosc.* **224**, 137 (2004)
- [37] L.N. Zack, R.L. Pulliam, and L.M. Ziurys, *J. Mol. Spectrosc.* **256**, 186 (2009)
- [38] C.A. Fancher, H.L. de Clercq, O.C. Thomas, D.W. Robinson, and K.H. Bowen, *J. Chem. Phys.* **109**, 8426 (1998)
- [39] G. Drechsler, U. Boesl, C. Bäßmann, and E.W. Schlag, *J. Chem. Phys.* **107**, 2284 (1997)
- [40] J.G. Phillips, S.P. Davis, B. Lindgren, and W.J. Balfour, *Astrophys. J. Suppl. Ser.* **65**, 721 (1987)
- [41] M. Dulick, C.W. Bauschlicher, Jr., A. Burrows, C.M. Sharp, R.S. Ram, and P. Bernath, *Astrophys. J.* **594**, 651 (2003)
- [42] L. Wang, D.-l. Huang, J.-f. Zhen, Q. Zhang, and Y. Chen, *Chin. J. Chem. Phys.* **24**, 1 (2011)
- [43] M. Allen, L. Ziurys, and J. Brown, *Chem. Phys. Lett.* **257**, 130 (1996)

5 Variational Wave Functions for Strongly Correlated Fermionic Systems

Federico Becca

Dipartimento di Fisica, Università di Trieste

Strada Costiera 11, I-34151 Trieste, Italy

Contents

1	Introduction	2
2	The variational principle	4
3	The variational Monte Carlo method	5
3.1	General principles	5
3.2	Markov chains	7
3.3	The Metropolis-Hastings algorithm	8
4	Variational wave functions	9
4.1	The Hartree-Fock wave function	9
4.2	The Gutzwiller wave function	10
4.3	The fully-projected Gutzwiller wave function	11
4.4	The Jastrow wave function	11
4.5	The backflow wave function	13
5	Practical implementations	14
5.1	The Jastrow factor	14
5.2	Slater determinants	15
5.3	Fast computation of the determinants	17
5.4	Backflow correlations	21
6	Optimization techniques	22
6.1	Calculation of derivatives	22
6.2	The stochastic reconfiguration	24
7	Selected results	26
8	Conclusions	30

1 Introduction

Ordinary metals are characterized by wide electronic bands and large screening effects, thus implying that electron-electron interactions may be considered perturbatively or even neglected. The situation changes completely in transition-metal oxides, where the bands close to the Fermi energy are relatively narrow and the electronic interactions play a predominant role in determining low-energy properties. Indeed, while the kinetic energy favors electron delocalization, the Coulomb repulsion drives the system towards localization, whose ultimate effect is the stabilization of an insulator. This state, which is established by the strong electron-electron correlation, goes under the name of *Mott insulator* [1]. In addition, materials characterized by strong correlations possess unusual properties in the metallic phase, as well as the possibility to show unconventional superconductivity [2, 3]. The lack of a consistent microscopic description of these phenomena clearly implies that a better understanding of correlation effects is needed. Since the early pioneering works on transition-metal oxides, the theoretical approach to Mott insulators has focused on the Hubbard model, which has been independently conceived by Hubbard [4], Gutzwiller [5], and Kanamori [6]. Here, electrons on a lattice interact among each others through a simplified “Coulomb” potential that includes only the on-site term

$$\mathcal{H} = \sum_{i,j,\sigma} t_{i,j} c_{i,\sigma}^\dagger c_{j,\sigma} + \text{h.c.} + U \sum_i n_{i,\uparrow} n_{i,\downarrow}, \quad (1)$$

where $t_{i,j}$ is the hopping amplitude in a d -dimensional lattice with L sites (in the simplest case, $t_{i,j}$ is non-zero only for nearest-neighbor sites and is denoted by t) and U is the local electron-electron repulsion; $c_{j,\sigma}^\dagger$ ($c_{j,\sigma}$) creates (destroys) one electron with spin σ on a Wannier orbital residing on the site j

$$\Xi_j(\mathbf{r}) = \frac{1}{\sqrt{L}} \sum_{\mathbf{k}} e^{-i\mathbf{k} \cdot \mathbf{R}_j} \Psi_{\mathbf{k}}(\mathbf{r}), \quad (2)$$

where $\Psi_{\mathbf{k}}(\mathbf{r})$ are Bloch states constructed with the orbitals $\phi(\mathbf{r} - \mathbf{R}_i)$ centered around each site i . The operators at different sites create orthogonal states, thus satisfying the anti-commutation relations

$$\{c_{i,\sigma}, c_{j,\tau}^\dagger\} = \delta_{i,j} \delta_{\sigma,\tau}, \quad (3)$$

$$\{c_{i,\sigma}^\dagger, c_{j,\tau}^\dagger\} = 0. \quad (4)$$

The Hubbard model is defined in the Hilbert space where each site can be empty, singly occupied (with either spin up or down), or doubly occupied. Moreover, the Hamiltonian (1) commutes with the total number of particles with up or down spin (i.e., $N_\uparrow$ and $N_\downarrow$, $N_e = N_\uparrow + N_\downarrow$ being the total number of electrons), thus allowing us to consider sectors with different number of particles separately. Although very simple in its formulation, the Hubbard model is generally not solvable with the available analytical techniques, apart from the one-dimensional case [7]. Therefore, several approximate schemes have been introduced, with the support of numerical calculations.

Within the standard approaches, based upon the independent-electron approximation, it is not possible to obtain a metal-insulator transition when the band is half-filled (with one electron per site on average, i.e., $N_e = L$), unless some kind of magnetic order is imposed. As a consequence, these techniques turn the Mott insulator into a conventional band insulator, thus missing the essence of the Mott phenomenon, where a charge gap appears independently of spin order. Dynamical mean-field theory [8] offered an alternative route to this problem, giving a description of the Mott transition without the need for a symmetry breaking. However, this scheme fully neglects spatial correlations and becomes exact only in the limit of infinite dimensionality. Since charge fluctuations are very strong in low-dimensional systems and are expected to determine their physical properties, an alternative method, which allows one to take into account the role of charge fluctuations, would be very useful.

Here, we consider variational wave functions, which can be treated within Monte Carlo techniques, as a possible route to capture the low-energy properties of strongly-correlated systems. In particular, our approach is based on an approximate form for the ground-state wave function that contains the physically relevant terms for the correct description of the Mott insulating state, and, at the same time, is simple enough to allow a straightforward calculation of the physical quantities. In this way, we obtain a transparent and physically intuitive way to understand the correlation-induced localization of electrons.

In the limit of $U/t \rightarrow \infty$ (i.e., for Heisenberg or t - J models), the general form for correlated wave functions corresponds to *fully-projected* Slater determinants [9, 10], where the configurations having a finite number of double occupancies are completely removed. At half-filling, these wave functions are obviously insulating, since no charge fluctuations can occur. Within the Hubbard model, early calculations showed that the variational description of a Mott insulator is a non-trivial problem, whenever charge fluctuations are present. Indeed, the Gutzwiller on-site correlation factor [5], which is the natural extension of the full projector in the case of finite (on-site) interaction, gives an insulating state only in the limit of infinite repulsion (apart from infinite dimension), while for finite Coulomb interaction it always corresponds to a correlated metallic state. The reason for its failure has been widely discussed in the past, and an intuitive argument has been found in the lack of correlation among the charge carriers, which correspond to the empty sites (holons) and doubly occupied sites (doublons), created by charge fluctuations at finite interactions [11]. In fact, holons possess an effective positive charge, since one electron is missing, and doublons are negatively charged objects, having one more electron with respect to the average occupation number; if the system is perturbed with the insertion of an electric field, holons and doublons can move freely in opposite directions, thus leading to a metallic behavior. Variational attempts done by adding a short-range correlation term up to a distance ξ among holons and doublons, turned out to be likewise unsuccessful [11, 12]. Naively, this happens because the configurations where holons and doublons are at distances larger than ξ are not subject to any correlation term, hence they can move freely on the lattice and conduct. Following this insight, it turns out that, in order to describe a correlated insulator without breaking any symmetry, it is necessary to correlate particles over all length scales. Let us consider a more general argument in view of the above considerations. For realistic Hamiltonians,

the dynamical properties of a system reflect the long-distance behavior of the static correlation functions of its ground state. Within a variational approach, this implies that a good *Ansatz* for an insulating state requires the correct description of the density-density correlations at large distances or, equivalently, the correct behavior of the charge structure factor at small momenta. For fermionic systems, the standard form for a correlated wave function is constituted by a correlation term acting on a Slater determinant, the latter corresponding to an uncorrelated metallic state. As a consequence, a variational wave function built with a short-range correlation factor cannot change the metallic character of the determinant, unless one fully suppresses charge fluctuations, since the large distance physics remains untouched.

The above arguments suggest that a long-range correlation factor is needed in order to correctly describe the insulating state. In particular, since we are interested in the density-density correlations, a natural choice of the correlation factor contains two-body terms, which corresponds to the so-called Jastrow factor [13]. It has been widely used in the context of liquid Helium on the continuum, where it gives the correct behavior of the density structure factor [14, 15]. Here, the analytic form of the Jastrow parameters can be deduced from weak-coupling calculations. Unfortunately, for the purpose of describing an insulating state, a proper analytic form of the Jastrow parameters cannot be obtained by weak-coupling techniques. The lack of a functional form for the Jastrow term, together with the large number of variational parameters required for a long-range correlation factor, represented the main obstacle to the use of this wave function in presence of strong correlation. Nowadays, this difficulty has been overcome with the help of advanced stochastic optimization techniques, which allow one to optimize many variational parameters independently, without assuming any functional form [16, 17].

2 The variational principle

In this section, we discuss the basic aspects of the variational principle, which represents one important pillar when searching for reliable approximations of strongly-correlated systems. Given any approximate state $|\Psi\rangle$ for the exact ground state $|\mathcal{Y}_0\rangle$ of a given Hamiltonian, we can define its variational energy as

$$E = \frac{\langle\Psi|\mathcal{H}|\Psi\rangle}{\langle\Psi|\Psi\rangle}. \quad (5)$$

Since any state in the Hilbert space can be expanded in terms of the eigenfunctions $|\mathcal{Y}_i\rangle$ of the Hamiltonian (with energies E_i), the variational state can be written as

$$|\Psi\rangle = \sum_i a_i |\mathcal{Y}_i\rangle, \quad (6)$$

with $a_i = \langle\mathcal{Y}_i|\Psi\rangle$. The normalization condition reads as

$$\langle\Psi|\Psi\rangle = \sum_i |a_i|^2 = 1. \quad (7)$$

By using the expansion of Eq. (6), we easily obtain that

$$\epsilon \equiv E - E_0 = \sum_{i \neq 0} |a_i|^2 (E_i - E_0) \geq 0, \quad (8)$$

which implies that any trial state $|\Psi\rangle$ provides an upper bound of the exact energy, thus giving a controlled way to approximate the original problem. Then, all computational efforts are devoted to minimizing the variational energy E .

Let us now analyze in what sense an approximate wave function, with given “distance” in energy ϵ from the exact ground state, can be considered as a good approximation of the many-body ground state $|\mathcal{I}_0\rangle$. A crucial role is played by the gap to the first excited state, which is always finite in a system with N_e particles (apart from accidental degeneracies), namely $\Delta = E_1 - E_0 > 0$. From Eq. (8) and the fact that, for $i \neq 0$, $E_i - E_0 \geq \Delta$, it follows that

$$\epsilon \geq \Delta \sum_{i \neq 0} |a_i|^2; \quad (9)$$

then, by using the normalization condition (7), we obtain

$$\eta = 1 - |a_0|^2 \leq \frac{\epsilon}{\Delta}. \quad (10)$$

This relation tells us that, in order to have an accurate approximation of the exact ground state (i.e., $\eta \ll 1$), a sufficient condition is that the error ϵ in the variational energy has to be much smaller than the gap Δ to the first excited state.

The accuracy of generic correlation functions (i.e., expectation values of Hermitian operators, which do not commute with the Hamiltonian, over $|\Psi\rangle$) is usually worse than the one on the ground-state energy. In fact, let us consider a generic operator $\mathcal{O}$ and express the variational wave function as

$$|\Psi\rangle = a_0|\mathcal{I}_0\rangle + \sqrt{\eta}|\mathcal{I}'\rangle, \quad (11)$$

where $|\mathcal{I}'\rangle$ is orthogonal to the ground state $|\mathcal{I}_0\rangle$. Then, the difference between the expectation value calculated with the variational state and the exact one is given by

$$|\langle\Psi|\mathcal{O}|\Psi\rangle - \langle\mathcal{I}_0|\mathcal{O}|\mathcal{I}_0\rangle| = |2a_0\sqrt{\eta}\langle\mathcal{I}_0|\mathcal{O}|\mathcal{I}'\rangle + \eta\langle\mathcal{I}'|\mathcal{O}|\mathcal{I}'\rangle - \eta\langle\mathcal{I}_0|\mathcal{O}|\mathcal{I}_0\rangle|, \quad (12)$$

where, for simplicity, we have assumed real wave functions. Then, whenever the variational state is close to the exact ground state, $\eta \ll \sqrt{\eta}$ and we can neglect all the terms that are proportional to η

$$|\langle\Psi|\mathcal{O}|\Psi\rangle - \langle\mathcal{I}_0|\mathcal{O}|\mathcal{I}_0\rangle| \approx 2\sqrt{\eta}|\langle\mathcal{I}_0|\mathcal{O}|\mathcal{I}'\rangle|, \quad (13)$$

which shows that the accuracy in correlation functions is more problematic than the one on the ground-state energy, with a term proportional to $\sqrt{\eta}$.

3 The variational Monte Carlo method

3.1 General principles

Let us start by describing the general framework in which variational Monte Carlo methods are defined. First of all, we fix a complete basis set $\{|x\rangle\}$ in the Hilbert space, in which (for simplicity) the states are taken to be orthogonal and normalized such that

$$\sum_x |x\rangle\langle x| = \mathbb{I}. \quad (14)$$

Then, any quantum state $|\Psi\rangle$ can be written as

$$|\Psi\rangle = \sum_x |x\rangle \langle x|\Psi\rangle = \sum_x \Psi(x) |x\rangle. \quad (15)$$

In turn, the expectation value of an operator $\mathcal{O}$ over a given variational wave function $|\Psi\rangle$ takes the following form

$$\langle \mathcal{O} \rangle = \frac{\langle \Psi | \mathcal{O} | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \frac{\sum_x \langle \Psi | x \rangle \langle x | \mathcal{O} | \Psi \rangle}{\sum_x \langle \Psi | x \rangle \langle x | \Psi \rangle}. \quad (16)$$

The main problem in evaluating the expectation value is that the number of configurations in the sum is exponentially large with the number of particles. Therefore, for large systems, it is impossible to perform an exact enumeration of the configurations to compute $\langle \mathcal{O} \rangle$ exactly. Nevertheless, Eq. (16) can be recast into a form that can be easily treated by standard Monte Carlo methods. Indeed, we have that

$$\langle \mathcal{O} \rangle = \frac{\sum_x |\Psi(x)|^2 \mathcal{O}_L(x)}{\sum_x |\Psi(x)|^2}, \quad (17)$$

where we have defined the *local estimator* of the operator $\mathcal{O}$

$$\mathcal{O}_L(x) = \frac{\langle x | \mathcal{O} | \Psi \rangle}{\langle x | \Psi \rangle}. \quad (18)$$

The important point is that

$$\mathcal{P}(x) = \frac{|\Psi(x)|^2}{\sum_x |\Psi(x)|^2} \quad (19)$$

can be interpreted as a probability, since it is a non-negative quantity for all configurations $|x\rangle$ and is normalized, i.e., $\sum_x \mathcal{P}(x) = 1$. Therefore, the problem of computing a quantum average of the operator $\mathcal{O}$ can be rephrased into the calculation of the average of the random variable $\mathcal{O}_L(x)$ over the distribution probability $\mathcal{P}(x)$

$$\langle \mathcal{O} \rangle = \sum_x \mathcal{P}(x) \mathcal{O}_L(x). \quad (20)$$

In particular, if we consider the expectation value of the Hamiltonian, the local estimator corresponds to the so-called *local energy*, which is defined by

$$e_L(x) = \frac{\langle x | \mathcal{H} | \Psi \rangle}{\langle x | \Psi \rangle}. \quad (21)$$

An important feature of the variational Monte Carlo approach is the *zero-variance property*. Let us suppose that the variational state $|\Psi\rangle$ coincides with an exact eigenstate of $\mathcal{H}$ (not necessarily the ground state), namely $\mathcal{H}|\Psi\rangle = E|\Psi\rangle$. Then, it follows that the local energy $e_L(x)$ is constant

$$e_L(x) = \frac{\langle x | \mathcal{H} | \Psi \rangle}{\langle x | \Psi \rangle} = E \frac{\langle x | \Psi \rangle}{\langle x | \Psi \rangle} = E. \quad (22)$$

Therefore, the random variable $e_L(x)$ does not depend on $|x\rangle$, which immediately implies that its variance is zero, while its mean value E coincides with the exact eigenvalue (in other words, $e_L(x)$ is not a random variable). Clearly, this is an extreme case that is very rare for generic correlated models. However, in general, the variance of $e_L(x)$ will decrease its value whenever the variational state $|\Psi\rangle$ will approach an exact eigenstate. This fact is very important to reduce the statistical fluctuations and improve the numerical efficiency. The zero-variance property is a feature that exists only for quantum expectation values, while it is absent in classical calculations, where observables have thermal fluctuations.

3.2 Markov chains

Instead of an exact enumeration of all configurations $\{|x\rangle\}$ in the Hilbert space, the quantum average of the operator $\mathcal{O}$ is evaluated by sampling a set of configurations $\{|x_n\rangle\}$ that are distributed according to the probability $\mathcal{P}(x)$, such that

$$\langle \mathcal{O} \rangle \approx \frac{1}{N} \sum_{n=1}^N \mathcal{O}_L(x_n). \quad (23)$$

From now on, we denote configurations by using only x , dropping the ket notation. The idea to sample a generic probability distribution is to construct a non-deterministic, i.e., random, process for which a configuration x_n evolves as a function of a discrete iteration time n according to a stochastic dynamics. A particularly simple case is given by the so-called Markov chains, where the configuration at time $n+1$ just depends upon the one at time n

$$x_{n+1} = F(x_n, \xi_n), \quad (24)$$

where the function F is taken to be time independent. The stochastic nature of the dynamics (24) is due to the fact that F depends upon a random variable ξ_n that is distributed according to a probability density $\chi(\xi_n)$. Here, the main point is to define a suitable function F such that the configurations x_n will be distributed (for large enough time n) according to the probability that we want to sample. Notice that, although ξ_n and ξ_{n+1} are independent random variables, $x_n \equiv x$ and $x_{n+1} \equiv x'$ are not independent. The joint probability distribution of these variables can be decomposed into the product of the marginal and the conditional probability

$$\mathcal{P}_{\text{joint},n}(x', x) = \omega(x'|x) \mathcal{P}_n(x), \quad (25)$$

where the conditional probability is such that $\omega(x'|x) \geq 0$ for all x and x' and satisfies the following normalization

$$\sum_{x'} \omega(x'|x) = 1. \quad (26)$$

It represents the probability that, having the configuration x at the iteration n , x' appears at $n+1$; its actual form depends upon the function $F(x, \xi)$ and the probability distribution $\chi(\xi)$.

We are now in the position of deriving the so-called Master equation, associated to the Markov chain. Indeed, the marginal probability of the variable x' is given by

$$\mathcal{P}_{n+1}(x') = \sum_x \mathcal{P}_{\text{joint},n}(x', x), \quad (27)$$

so that, by using Eq. (25), we get

$$\mathcal{P}_{n+1}(x') = \sum_x \omega(x'|x) \mathcal{P}_n(x). \quad (28)$$

This equation allows us to calculate the evolution of the marginal probability $\mathcal{P}_n(x)$ as a function of n , since the conditional probability $\omega(x'|x)$ is determined by the stochastic dynamics in Eq. (24) and does not depend upon n .

The important question about the Markov process is to understand under which conditions the sequence of distributions $\mathcal{P}_n(x)$ converges to some limiting (i.e., equilibrium) distribution $\mathcal{P}(x)$ or not. In particular

1. Does a stationary distribution $\mathcal{P}(x)$ exist?
2. Is the convergence to $\mathcal{P}(x)$ guaranteed when starting from a given *arbitrary* $\mathcal{P}_0(x)$?

The first question requires that

$$\mathcal{P}(x') = \sum_x \omega(x'|x) \mathcal{P}(x). \quad (29)$$

In order to satisfy this condition, it is sufficient (but not necessary) to satisfy the so-called *detailed balance* condition

$$\omega(x'|x) \mathcal{P}(x) = \omega(x|x') \mathcal{P}(x'). \quad (30)$$

The second question requires the *ergodicity* condition, i.e., the possibility to reach any state x from any other one x' by performing a finite number of steps.

3.3 The Metropolis-Hastings algorithm

Finally, we present a practical way of constructing a conditional probability $\omega(x'|x)$ that satisfies the detailed balance condition (30), such that, for large values of n , the configurations x_n are distributed according to a given probability distribution $\mathcal{P}(x)$. Metropolis and collaborators [18] introduced a very simple scheme, which is also very general and can be applied to many different cases. Later, the so-called Metropolis algorithm has been extended to more general cases by Hastings [19] (very often, the name of “Metropolis-Hastings algorithm” is also used). As a first step, we split the transition probability $\omega(x'|x)$ into two pieces

$$\omega(x'|x) = T(x'|x)A(x'|x), \quad (31)$$

where $T(x'|x)$ defines a *trial probability* that proposes the new configuration x' from the present one x and $A(x'|x)$ is the *acceptance probability*. In the original work by Metropolis and co-workers, the trial probability has been taken symmetric, i.e., $T(x'|x) = T(x|x')$. However,

in the generalized version of the algorithm $T(x'|x)$ can be chosen with large freedom, as long as ergodicity is ensured. Then, in order to define a Markov process that satisfies the detailed balance condition, the proposed configuration x' is accepted with a probability

$$A(x'|x) = \min \left\{ 1, \frac{\mathcal{P}(x')T(x|x')}{\mathcal{P}(x)T(x'|x)} \right\}. \quad (32)$$

Without loss of generality, we can always choose $T(x|x) = 0$, namely we never propose to remain with the same configuration. Nevertheless, $\omega(x|x)$ can be finite, since the proposed move can be rejected. The actual value of $\omega(x|x)$ is fixed by the normalization condition $\sum_{x'} \omega(x'|x) = 1$.

The proof that detailed balance is satisfied by considering the acceptance probability of Eq. (32) is very simple. Indeed, let us consider the case in which x and $x' \neq x$ are such that

$$\frac{\mathcal{P}(x')T(x|x')}{\mathcal{P}(x)T(x'|x)} > 1, \quad (33)$$

in this case, we have that

$$A(x'|x) = 1, \quad (34)$$

$$A(x|x') = \frac{\mathcal{P}(x)T(x'|x)}{\mathcal{P}(x')T(x|x')}; \quad (35)$$

then, we can directly verify that the detailed balance is satisfied. A similar proof can be obtained in the opposite case where $\mathcal{P}(x')T(x|x')/[\mathcal{P}(x)T(x'|x)] < 1$.

4 Variational wave functions

4.1 The Hartree-Fock wave function

For fermionic models, the simplest example for a variational wave function is given by the Hartree-Fock approximation, where the many-body state is taken to be a product state of suitably optimized single-particle orbitals

$$|\Psi_{\text{HF}}\rangle = \prod_{\alpha=1}^{N_e} \Phi_{\alpha}^{\dagger} |0\rangle; \quad (36)$$

here, $\Phi_{\alpha}^{\dagger}$ can be expressed in terms of the original fermionic operators as

$$\Phi_{\alpha}^{\dagger} = \sum_i W_{\uparrow,\alpha,i}^* c_{i,\uparrow}^{\dagger} + \sum_i W_{\downarrow,\alpha,i}^* c_{i,\downarrow}^{\dagger}, \quad (37)$$

where $\{W_{\sigma,\alpha,i}\}$ are complex coefficients that can be optimized to get the best variational state. The condition that orbitals are normalized and orthogonal to each other implies that

$$\sum_i (W_{\uparrow,\alpha,i} W_{\uparrow,\beta,i}^* + W_{\downarrow,\alpha,i} W_{\downarrow,\beta,i}^*) = \delta_{\alpha,\beta}. \quad (38)$$

The expectation value of any Hamiltonian can be easily evaluated analytically. On the lattice, it is relatively simple to obtain a solution for the Hartree-Fock equations by using iterative methods. However, while the Hartree-Fock approximation may give reasonable results in the weak-coupling regime, its accuracy becomes questionable for moderate and strong interactions. For example, a Mott insulator, with no symmetry breaking, cannot be described within this approximation; moreover, it is also not possible to stabilize superconducting phases in purely repulsive Hamiltonians. Therefore, a step forward is needed, in order to reach a better description of strongly-correlated systems.

4.2 The Gutzwiller wave function

The simplest example of a correlated state, which goes beyond the Hartree-Fock approximation, has been proposed by Gutzwiller to describe the effect of the Hubbard- U interaction in reducing the weight of configurations with doubly-occupied sites [5]. The Gutzwiller wave function is constructed by starting from the ground state $|\Phi_0\rangle$ of the $U = 0$ model and then applying an operator $\mathcal{P}_G$ that suppresses the weight of configurations with doubly-occupied sites

$$|\Psi_G\rangle = \mathcal{P}_G|\Phi_0\rangle; \quad (39)$$

here, $\mathcal{P}_G$ is the so-called Gutzwiller factor that depends upon a single variational parameter g (e.g., $g > 0$ for the repulsive Hubbard model)

$$\mathcal{P}_G = \exp \left[-\frac{g}{2} \sum_i (n_i - n)^2 \right], \quad (40)$$

where n is the average density.

The effect of the Gutzwiller factor becomes clear once the variational state is expanded in a basis set whose elements $\{|x\rangle\}$ represent configurations with particles sitting on the lattice sites. Indeed, since the Gutzwiller factor is diagonal in this basis (it contains the density operator on each site n_i), we have that

$$\langle x|\Psi_G\rangle = \mathcal{P}_G(x)\langle x|\Phi_0\rangle, \quad (41)$$

where $\mathcal{P}_G(x) \leq 1$ is a number that depends on how many doubly-occupied sites are present in the configuration $|x\rangle$. Therefore, the amplitude of the non-interacting state $\langle x|\Phi_0\rangle$ is renormalized by $\mathcal{P}_G(x)$.

For the Hubbard model with a generic hopping amplitude $t_{i,j}$, when the particle density is $n = 1$, a metal-insulator transition is expected at finite values of U/t . However, a simple argument suggests that the Gutzwiller wave function can describe such a transition only when the variational parameter g tends to infinity. Indeed, for $n = 1$, on average, there is one particle per site and density excitations are represented by doublons and holons. In the non-interacting state $|\Phi_0\rangle$, these objects are free to move and then responsible for the conductivity (a doublon is negatively charged with respect to the average background, while the holon is positively charged). The effect of the Gutzwiller factor is to penalize the formation of such objects; however, once

created, doublons and holons are no longer correlated, thus being free to move independently. Only when the energetic penalty is infinite, an insulator is obtained; here, all the density degrees of freedom are frozen and no transport is possible, implying an oversimplified description of a true insulator, where instead density fluctuations are always present. Extensive calculations have shown that g remains finite for all values of U/t and diverges only for $U/t = \infty$ [20,21].

4.3 The fully-projected Gutzwiller wave function

Here, we briefly discuss the limit of $g = \infty$, which corresponds to the *fully-projected* Gutzwiller state. For $n = 1$, the Gutzwiller factor becomes a projector in the sub-space with only singly-occupied sites

$$\mathcal{P}_\infty = \prod_i (n_{i,\uparrow} - n_{i,\downarrow})^2, \quad (42)$$

which implies that all configurations with empty or doubly occupied sites are annihilated. In fermionic models, there is still an exponentially large number of states with one electron per site, which differ by their spin configurations. Therefore, non-trivial spin fluctuations are still allowed. These kinds of fully-projected wave functions $|\Psi_\infty\rangle = \mathcal{P}_\infty|\Phi_0\rangle$ have been considered within the so-called resonating-valence bond (RVB) approach, which has been proposed by Anderson to capture the physics of exotic magnetic systems (spin liquids) [22] in frustrated Heisenberg models [23].

The case to include empty sites, which is relevant when $n < 1$, has been widely used to describe high-temperature superconductors in the t - J model [9,10]. In this case the Gutzwiller projection is written as

$$\mathcal{P}_\infty = \prod_i (1 - n_{i,\uparrow}n_{i,\downarrow}), \quad (43)$$

which annihilates all configurations containing doubly-occupied sites, but leaving untouched configurations with only empty and singly-occupied sites.

4.4 The Jastrow wave function

As we have discussed above, the variational description of an insulator with density fluctuations is not captured by the simple Gutzwiller wave function (39) and requires a modification of the correlation term that is applied to the non-interacting wave function. A straightforward generalization of the Gutzwiller wave function is given by the inclusion of long-range terms in the correlator

$$|\Psi_J\rangle = \mathcal{J}|\Phi_0\rangle, \quad (44)$$

where $\mathcal{J}$ is the Jastrow factor [13] that has been introduced in continuum models much before the Gutzwiller wave function. On the lattice, $\mathcal{J}$ takes a simple form

$$\mathcal{J} = \exp \left[-\frac{1}{2} \sum_{i,j} v_{i,j} (n_i - n)(n_j - n) \right], \quad (45)$$

where $v_{i,j}$ is a pseudo-potential for density-density correlations in the variational state. For translationally invariant models, $v_{i,j}$ only depends upon the relative distance of the two sites i and j , i.e., $|\mathbf{R}_i - \mathbf{R}_j|$; moreover, the on-site term $v_{i,i}$ corresponds to the Gutzwiller parameter g . The Jastrow pseudo-potential can be either parametrized in some way, in order to reduce the number of variational parameters, or optimized for all possible distances, which are $O(L)$ in translationally invariant systems. The role of the long-range tail of the Jastrow factor is to create a bound state between holons and doublons, possibly impeding conduction, but still allowing local density fluctuations. Indeed, we have shown that such Jastrow terms may turn a non-interacting metallic state $|\Phi_0\rangle$ into an insulator [24, 25]. In particular, by denoting the Fourier transform of the (translationally invariant) pseudo-potential $v_{i,j}$ by v_q , the gapless (i.e., metallic) phase is described by having $v_q \approx 1/|\mathbf{q}|$ for $|\mathbf{q}| \rightarrow 0$, in any spatial dimension d ; by contrast, a fully gapped (i.e., insulating) phase is obtained in one-dimension with $v_q \approx 1/|\mathbf{q}|^2$ for $|\mathbf{q}| \rightarrow 0$ [24]. This singular behavior of the pseudo-potential induces an exponential decay of the density-density correlations. In two and three spatial dimensions, a holon-doublon bound-state is generated by $v_q \approx \beta/|\mathbf{q}|^d$ for a sufficiently large value of β [25]. However, these behaviors of the pseudo-potential, which are obtained by an energy minimization, are not sufficient to have a fully gapped phase, since a residual power-law behavior in the density-density correlations is still present.

The Jastrow wave function of Eq. (44) has been introduced to study models on the continuum [13] and has been employed to perform the first quantum Monte Carlo calculation in a many-body system [14]. More precisely, a system of interacting bosons has been considered to model the ground-state properties of ^4He in three spatial dimensions with Lennard-Jones interactions. Here, in a first-quantization notation, the wave function with N_b bosons reads

$$\Psi_J(\mathbf{r}_1, \dots, \mathbf{r}_{N_b}) = \prod_{i < j} f(r_{i,j}) = \exp \left[- \sum_{i < j} u(r_{i,j}) \right], \quad (46)$$

where $\{\mathbf{r}_i\}$ are the coordinates of the bosons and $f(r_{i,j}) = \exp[-u(r_{i,j})]$ is a function that depends upon the relative distance between two bosons $r_{i,j} = |\mathbf{r}_i - \mathbf{r}_j|$. Notice that the wave function of Eq. (46) is totally symmetric when exchanging two particles, thus having the correct symmetry for a bosonic wave function.

A suitable correlated wave function for N_e fermions can be obtained by applying the symmetric Jastrow factor $\prod_{i < j} f(r_{i,j})$ to a (anti-symmetric) Slater determinant, which, by using first-quantization notations, reads as

$$\Psi_{\text{HF}}(\mathbf{r}_1, \dots, \mathbf{r}_{N_e}) = \det\{\phi_\alpha(\mathbf{r}_j)\}, \quad (47)$$

where $\{\phi_\alpha(\mathbf{r}_j)\}$ is a set of one-particle orbitals. Then the Jastrow-Slater wave function is given by

$$\Psi_{\text{JS}}(\mathbf{r}_1, \dots, \mathbf{r}_{N_e}) = \prod_{i < j} f(r_{i,j}) \times \Psi_{\text{HF}}(\mathbf{r}_1, \dots, \mathbf{r}_{N_e}). \quad (48)$$

In total, this wave function is anti-symmetric when exchanging two particles and, therefore, has the correct symmetry for a fermionic state.

4.5 The backflow wave function

An alternative way to include some correlation inside the original variational state is to introduce a parametrization that allows the orbitals to depend upon the positions of the other particles, leading to the concept of *backflow* correlations. In quantum systems, a particle that moves is surrounded by a counter-flow generated by all the other particles; the existence of this flow pattern pushes away the particles, thus preventing a significant overlap among them. This idea has been originally introduced by Wigner [26] and then developed by Feynman [27, 28] in the context of excitations in ^4He and the effective mass of a ^3He impurity in liquid ^4He . In the fermionic case, the Slater determinant is not constructed with the actual positions of the electrons $(\mathbf{r}_1, \dots, \mathbf{r}_{N_e})$, see Eq. (47), but with new “coordinates” given by

$$\mathbf{r}_i^b = \mathbf{r}_i + \sum_{j \neq i} \eta(|\mathbf{r}_i - \mathbf{r}_j|)(\mathbf{r}_j - \mathbf{r}_i), \quad (49)$$

where $\eta(|\mathbf{r}_i - \mathbf{r}_j|)$ is a suitable function that describes the effective displacement of the i -th particle due to the j -th one. The simplest wave function is built by taking plane-waves with positions given by $\{\mathbf{r}_i^b\}$. The effect of backflow correlations introduces many-body effects inside the Slater determinant, since, when the i -th electron is moved, all the new “coordinates” are modified, such that all particles respond to the movement of the single electron, adapting their positions accordingly.

Wave functions including both Jastrow factors and backflow correlations have been used to study Helium systems within the so-called hyper-netted chain approximation [29, 30]. They also have been used in Monte Carlo calculations to compute the properties of the homogeneous electron gas in two and three spatial dimensions [31, 32]. The advantage of the backflow wave function is that a single Slater determinant is used, allowing us to perform calculations with a large number of particles.

More recently, the same idea of modifying the single-electron orbitals to improve variational wave functions has been extended for lattice models [33, 34]. Here, the transformation (49) cannot be applied, since electrons live on the lattice sites. Nevertheless, we can imagine that the amplitudes of the Hartree-Fock orbitals (37) are changed according to the many-body configuration

$$W_{\sigma,i,\alpha}^b = \eta_0 W_{\sigma,i,\alpha} + \sum_{j \neq i} \eta_j \mathcal{O}_{i,j} W_{\sigma,j,\alpha}, \quad (50)$$

where $\{\eta_j\}$ is a set of variational parameters and $\mathcal{O}_{i,j}$ is a generic many-body operator that acts on the sites i and j . For example, within the repulsive Hubbard model, the formation of holon-doublon pairs is energetically expensive for large values of U/t ; therefore, these objects tend to recombine into singly-occupied sites. In this case, we can consider a many-body operator $\mathcal{O}_{i,j} = D_i H_j$, where D_i (H_i) is the operator that gives 1 if the site i is doubly occupied (empty) and 0 otherwise. Then, the many-body state, which is constructed by taking the Slater determinant of these new “orbitals”, will contain terms with single occupation, thus releasing the energy. More complicated expressions of the new “orbital” can be considered, as described in Refs. [33, 34].

5 Practical implementations

In the Monte Carlo evaluation of quantum averages, see Eq. (20), we must compute

- The ratio of probabilities with different configurations, which implies the ratio of overlaps between the given variational state and two configurations of the basis set

$$\frac{\mathcal{P}(x')}{\mathcal{P}(x)} = \frac{|\langle x'|\Psi\rangle|^2}{|\langle x|\Psi\rangle|^2}, \quad (51)$$

as required in the Metropolis algorithm.

- The local estimator $\mathcal{O}_L(x)$, which, in turn, implies ratios of overlaps and matrix elements of the observable between states of the basis set. For example, when considering the energy, we have

$$e_L(x) = \frac{\langle x|\mathcal{H}|\Psi\rangle}{\langle x|\Psi\rangle} = \sum_{x'} \langle x|\mathcal{H}|x'\rangle \frac{\langle x'|\Psi\rangle}{\langle x|\Psi\rangle}. \quad (52)$$

Naively, the computation of the local estimator looks a tremendously hard task, since it requires a summation over all the states of the many-body Hilbert space; however, thanks to the locality of the Hamiltonian (or, similarly, any other local operator or correlation function), only few terms actually contribute to the sum. Indeed, given the configuration $|x\rangle$, the matrix element $\langle x|\mathcal{H}|x'\rangle$ is non-zero only for $O(L)$ configurations $|x'\rangle$. As an example, let us consider the fermionic Hubbard model: by using the local basis, $|x\rangle$ is connected only to few other configurations that differ for the hopping of one electron from a given site to one of its neighbors; then, the maximum number of such processes is equal to the number of sites L times the number of bonds times 2 (due to the spin). Therefore, the computation of the local estimator only requires a small number of operations, usually proportional to the number of sites/particles.

Therefore, the building block of the variational Monte Carlo approach is the computation of $\langle x|\Psi\rangle$, which is the amplitude of the variational state over a generic element of the basis set. More precisely, along all the Markov process, only ratios of these overlaps must be computed. This calculation must be done for each configuration that is visited along the Markov process and, therefore, it must be done as fast as possible. This fact imposes some constraint on the form of the variational wave function. Usually, fermionic states require the calculation of determinants. Fortunately, there are fast (i.e., polynomial) algorithms to evaluate determinants, thus allowing us to consider these states as variational *Ansätze* for electron systems.

5.1 The Jastrow factor

Let us focus on the Jastrow factor of Eq. (45). For a translational invariant model, $v_{i,j}$ only depends upon the distance between i and j , thus the number of parameters can be reduced to $O(L)$. We would like to remark that, within the Monte Carlo approach, it is possible to treat exactly (but still having statistical errors) the limit of singular Jastrow factors with $v_{i,i} = \infty$, e.g., the Gutzwiller projector that eliminates all doubly-occupied sites (this case being relevant

for an infinite Hubbard- U interaction). Indeed, this case can be easily incorporated by building a Markov chain where only configurations $|x\rangle$ that satisfy this constraint are visited.

The advantage of considering the Jastrow-Slater wave function in the variational Monte Carlo technique is that the calculations can be extremely efficient and fast. Indeed, whenever the Jastrow factor is diagonal in the chosen basis, we have that

$$\langle x|\Psi_J\rangle = \mathcal{J}(x)\langle x|\Phi_0\rangle, \quad (53)$$

where $\mathcal{J}(x)$ is the value of the Jastrow operator computed for the configuration $|x\rangle$, i.e., $\mathcal{J}|x\rangle = \mathcal{J}(x)|x\rangle$. Therefore, given the electronic configuration, $\mathcal{J}(x)$ is a number that can be evaluated in $O(L^2)$ operations, thus leading to the same complexity when computing the ratio $\mathcal{J}(x')/\mathcal{J}(x)$ that appears in the Metropolis ratio (51). However, whenever the two configurations differ only by few electron hoppings, it is possible to apply a fast computation of the ratio, which involves $O(L)$ operations.

Finally, we emphasize that, in order to have a polynomial algorithm, the Jastrow factor must only contain operators that are diagonal in the basis $|x\rangle$, otherwise $\mathcal{J}|x\rangle$ would generate an exponentially large number of states, ruling out any calculation on large systems.

5.2 Slater determinants

A generic Slater determinant can be obtained as the ground state of a suitable quadratic Hamiltonian $\mathcal{H}_0$. First of all, we contract the spin index σ and the lattice site i into a single index I running from 1 to $2L$

$$c_{i,\uparrow} \equiv d_i, \quad (54)$$

$$c_{i,\downarrow} \equiv d_{i+L}. \quad (55)$$

Then, we start from the simple case in which the non-interacting Hamiltonian is written as

$$\mathcal{H}_0 = \sum_{I,J} t_{I,J} d_I^\dagger d_J, \quad (56)$$

which contains hopping terms only, also including processes in which the spin along z is not conserved, i.e., the terms with $I \leq L$ and $J > L$ and vice-versa. In a compact form, the non-interacting Hamiltonian of Eq. (56) can be written as

$$\mathcal{H}_0 = \mathbf{d}^\dagger \mathbf{T} \mathbf{d}, \quad (57)$$

where

$$\mathbf{d}^\dagger = \left(d_1^\dagger \quad \dots \quad d_{2L}^\dagger \right) \quad (58)$$

and

$$\mathbf{T} = \begin{pmatrix} t_{1,1} & \dots & t_{1,2L} \\ \vdots & \ddots & \vdots \\ t_{2L,1} & \dots & t_{2L,2L} \end{pmatrix}. \quad (59)$$

Since $\mathcal{H}_0$ commutes with the total number of electrons $N_e = \sum_I d_I^\dagger d_I$, the eigenstates are single-particle orbitals. In practice, the $2L \times 2L$ matrix $\mathbf{T}$ can be easily diagonalized by using standard libraries (e.g., LAPACK routines)

$$\mathcal{H}_0 = \mathbf{d}^\dagger \mathbf{U} \mathbf{U}^\dagger \mathbf{T} \mathbf{U} \mathbf{U}^\dagger \mathbf{d} = \mathbf{\Phi}^\dagger \mathbf{E} \mathbf{\Phi} = \sum_{\alpha} \varepsilon_{\alpha} \phi_{\alpha}^{\dagger} \phi_{\alpha}, \quad (60)$$

where $\mathbf{U}$ is a unitary matrix (that preserves anti-commutation relations of fermionic operators), $\mathbf{E} = \text{diag}(\varepsilon_1, \dots, \varepsilon_{2L})$ is the diagonal matrix containing the $2L$ eigenvalues ε_{α} of $\mathbf{T}$, and $\mathbf{\Phi}^\dagger = (\phi_1^\dagger, \dots, \phi_{2L}^\dagger)$ is defined in terms of the eigenvectors of $\mathbf{T}$

$$\phi_{\alpha}^{\dagger} = \sum_I U_{I,\alpha} d_I^{\dagger}. \quad (61)$$

Now, the many-body state $|\Phi_0\rangle$ can be constructed by occupying the N_e lowest-energy orbitals

$$|\Phi_0\rangle = \prod_{\alpha=1}^{N_e} \phi_{\alpha}^{\dagger} |0\rangle = \left(\sum_I U_{I,1} d_I^{\dagger} \right) \dots \left(\sum_I U_{I,N_e} d_I^{\dagger} \right) |0\rangle. \quad (62)$$

The generic configuration, with N_e electrons, which is visited along the Markov process, reads

$$|x\rangle = d_{R_1}^{\dagger} \dots d_{R_{N_e}}^{\dagger} |0\rangle, \quad (63)$$

where $j = 1, \dots, N_e$ includes both up and down spins and R_j assumes values from 1 to $2L$: the positions of spin-up electrons coincide with the site number, while the positions of spin-down electrons must be shifted by L . The overlap $\langle x | \Phi_0 \rangle$ is then given by

$$\begin{aligned} \langle x | \Phi_0 \rangle &= \langle 0 | d_{R_{N_e}} \dots d_{R_1} \left(\sum_I U_{I,1} d_I^{\dagger} \right) \dots \left(\sum_I U_{I,N_e} d_I^{\dagger} \right) |0\rangle \\ &= \langle 0 | d_{R_{N_e}} \dots d_{R_1} \left[\sum_p (-1)^p \prod_{\alpha=1}^{N_e} U_{p\{R_j\},\alpha} \right] d_{R_1}^{\dagger} \dots d_{R_{N_e}}^{\dagger} |0\rangle, \end{aligned} \quad (64)$$

where the sum inside the square bracket is over all the possible permutations of the $\{R_j\}$ in $|x\rangle$; the sign appears because of the anti-commutation relations of fermionic operators. Then, we get

$$\langle x | \Phi_0 \rangle = \det\{U_{R_j,\alpha}\}. \quad (65)$$

When constructing the many-body state (62), we must pay attention to construct a unique many-body state, i.e., occupy the correct lowest-energy orbitals. When the highest-occupied orbital and the lowest-unoccupied one have different energies (*closed shell* configuration), the choice is unique. Instead, it can also happen that there is a degeneracy that does not allow a unique choice (*open shell* configuration). Whenever the non-interacting Hamiltonian $\mathcal{H}_0$ is diagonalized numerically, the eigenstates provided by standard libraries do not carry definite quantum numbers (like momentum), but are given by generic linear combinations of degenerate orbitals (which erratically depend on the numerical precision used for the computation). This is not a problem

whenever *all* the degenerate eigenstates are included in the many-body state (62), since taking any linear combination of columns in the matrix $\mathbf{U}$ will not change the value of the determinant in Eq. (65). By contrast, not including all degenerate eigenstates will cause a problem, since the determinant will depend on the particular combination of states that is considered. Therefore, whenever a numerical diagonalization is done, we must verify that a closed-shell configuration occurs. Otherwise, the problem of having a vanishing gap can be overcome by constructing suitable orbitals with definite quantum numbers, to obtain a reproducible simulation of the many-body state.

In practice, the diagonalization of the non-interacting Hamiltonian $\mathcal{H}_0$ must be performed at the beginning of the Monte Carlo calculation; then, we need to store the reduced part of the $\mathbf{U}$, obtained by keeping only the N_e columns that correspond to occupied orbitals. This is a $2L \times N_e$ matrix

$$\mathbf{U} = \begin{pmatrix} U_{1,1} & \cdots & U_{1,N_e} \\ \vdots & \ddots & \vdots \\ U_{2L,1} & \cdots & U_{2L,N_e} \end{pmatrix}. \quad (66)$$

Then, the overlap with a generic electronic configuration is given by the determinant of the matrix obtained by considering only the rows of (66) corresponding to the electron positions $\{R_j\}$, thus giving a $N_e \times N_e$ matrix.

Remarkably, the same kind of formalism can be used also if the non-interacting Hamiltonian contains an electron pairing that couples up and down spins. Indeed, let us consider a Bardeen-Cooper-Schrieffer (BCS) Hamiltonian described by

$$\mathcal{H}_0 = \sum_{i,j,\sigma} t_{i,j} c_{i,\sigma}^\dagger c_{j,\sigma} - \mu_0 \sum_{i,\sigma} c_{i,\sigma}^\dagger c_{i,\sigma} + \sum_{i,j} \Delta_{i,j} c_{i,\uparrow}^\dagger c_{j,\downarrow}^\dagger + \text{h.c.}, \quad (67)$$

where we have included a chemical potential μ_0 , which fixes, on average, the number of electrons. In this case, the total number of particles is not conserved and the concept of single-particle orbitals is not defined. Indeed, the ground state of the BCS Hamiltonian is naturally written in terms of a pairing function. Nevertheless, we can perform a particle-hole transformation on the spin-down electrons

$$c_{i,\uparrow} \rightarrow f_{i,\uparrow} \equiv d_i \quad (68)$$

$$c_{i,\downarrow} \rightarrow f_{i,\downarrow}^\dagger \equiv d_{i+L}^\dagger. \quad (69)$$

Then, apart from constant terms, the transformed BCS Hamiltonian has the form of Eq. (57). Since, after the particle-hole transformation, the number of particles (but not the z component of the spin) is conserved, the eigenstates of the BCS Hamiltonian can be expressed into “orbitals”, similarly to the ones of Eq. (61), but without having a definite spin component along z .

5.3 Fast computation of the determinants

Let us now show how to compute efficiently the ratio of determinants when the two configurations $|x\rangle$ and $|x'\rangle$ differ by one or few electron hoppings. According to Eq. (65) the overlap

$\langle x|\Phi_0\rangle$ is given by the determinant of $\tilde{U}_{j,\alpha} \equiv U_{R_j,\alpha}$

$$\tilde{\mathbf{U}} = \begin{pmatrix} U_{R_1,1} & \cdots & U_{R_1,N_e} \\ \vdots & \ddots & \vdots \\ U_{R_l,1} & \cdots & U_{R_l,N_e} \\ \vdots & \ddots & \vdots \\ U_{R_{N_e},1} & \cdots & U_{R_{N_e},N_e} \end{pmatrix}, \quad (70)$$

which is obtained taking only the rows corresponding to the occupied sites of the matrix $\mathbf{U}$ of Eq. (66).

Let us start and consider the electronic configurations in which $|x'\rangle$ is obtained from $|x\rangle$ just by hopping the one electron from R_l to R'_l , i.e., $|x'\rangle = d_{R'_l}^\dagger d_{R_l}|x\rangle$. The new matrix $\tilde{\mathbf{U}}'$ will be equal to $\tilde{\mathbf{U}}$, except that the elements of the l -th row will be changed from $U_{R_l,\alpha}$ to $U_{R'_l,\alpha}$

$$\tilde{\mathbf{U}}' = \begin{pmatrix} U_{R_1,1} & \cdots & U_{R_1,N_e} \\ \vdots & \ddots & \vdots \\ U_{R'_l,1} & \cdots & U_{R'_l,N_e} \\ \vdots & \ddots & \vdots \\ U_{R_{N_e},1} & \cdots & U_{R_{N_e},N_e} \end{pmatrix}. \quad (71)$$

Then, the ratio between two configurations that differ only by a single fermion hopping is

$$\frac{\langle x'|\Phi_0\rangle}{\langle x|\Phi_0\rangle} = \frac{\langle x|d_{R_l}^\dagger d_{R'_l}|\Phi_0\rangle}{\langle x|\Phi_0\rangle} = \frac{\det \tilde{\mathbf{U}}'}{\det \tilde{\mathbf{U}}}. \quad (72)$$

By denoting with K the new site of the l -th electron (i.e., $K \equiv R'_l$), the updated matrix elements are given by a compact form

$$\tilde{U}'_{j,\alpha} = \tilde{U}_{j,\alpha} + \delta_{j,l}(U_{K,\alpha} - \tilde{U}_{l,\alpha}) = \tilde{U}_{j,\alpha} + \delta_{j,l}v_\alpha^{K,l}, \quad (73)$$

where we have defined $v_\alpha^{K,l} \equiv U_{K,\alpha} - \tilde{U}_{l,\alpha}$; here the indices K and l are fixed, since they specify the site where the electron is hopping and the electron index, respectively. This equation can be rewritten in the following way

$$\tilde{U}'_{j,\alpha} = \sum_\beta \tilde{U}_{j,\beta} \left(\delta_{\beta,\alpha} + \tilde{U}_{\beta,l}^{-1} v_\alpha^{K,l} \right) = \sum_\beta \tilde{U}_{j,\beta} Q_{\beta,\alpha}, \quad (74)$$

where

$$Q_{\beta,\alpha} = \delta_{\beta,\alpha} + \tilde{U}_{\beta,l}^{-1} v_\alpha^{K,l}. \quad (75)$$

Therefore, $\tilde{\mathbf{U}}' = \tilde{\mathbf{U}}\mathbf{Q}$, which implies that the calculation of the ratio of the determinants of $\tilde{\mathbf{U}}'$ and $\tilde{\mathbf{U}}$ is equivalent to the calculation of the determinant of $\mathbf{Q}$

$$\frac{\det \tilde{\mathbf{U}}'}{\det \tilde{\mathbf{U}}} = \det \mathbf{Q}. \quad (76)$$

The great simplification comes from the fact that the determinant of $\mathbf{Q}$ can be easily computed. Indeed, $\mathbf{Q}$ has a particularly simple form that can be written as

$$Q_{\beta,\alpha} = \delta_{\beta,\alpha} + \mathcal{B}_\beta \mathcal{A}_\alpha, \quad (77)$$

where $\mathcal{B}_\beta = \tilde{U}_{\beta,l}^{-1}$ and $\mathcal{A}_\alpha = v_\alpha^{K,l}$. Although the matrix is not Hermitian, the eigenvalues of the matrix $\mathbf{Q}$ can be obtained from the secular equation

$$\sum_{\alpha} Q_{\beta,\alpha} v_\alpha = \lambda v_\beta; \quad (78)$$

by using the explicit form of Eq. (77), we obtain

$$v_\beta + \mathcal{B}_\beta \sum_{\alpha} \mathcal{A}_\alpha v_\alpha = \lambda v_\beta, \quad (79)$$

which implies that all vectors v_α that are orthogonal to $\mathcal{A}_\alpha$ are eigenvectors with eigenvalue $\lambda = 1$ (there are $N_e - 1$ of such vectors); in addition, $v_\alpha = \mathcal{B}_\alpha$ is also eigenvector with $\lambda = 1 + \sum_{\alpha} \mathcal{A}_\alpha \mathcal{B}_\alpha = \sum_{\alpha} U_{K,\alpha} \tilde{U}_{\alpha,l}^{-1}$. Therefore, we have that

$$\frac{\det \tilde{\mathbf{U}}'}{\det \tilde{\mathbf{U}}} = \det \mathbf{Q} = \sum_{\alpha} U_{K,\alpha} \tilde{U}_{\alpha,l}^{-1}. \quad (80)$$

Having stored (at the beginning of the simulation) the matrix $\tilde{\mathbf{U}}^{-1}$ for the configuration $|x\rangle$, this calculation requires $O(N_e)$ operations, instead of the $O(N_e^3)$ needed to evaluate a determinant. Then, once the new configuration $|x'\rangle$ is accepted along the Markov process, the matrix $\tilde{\mathbf{U}}^{-1}$ must be updated. This can be done in $O(N_e^2)$ operations. In fact, we have that $(\tilde{\mathbf{U}}')^{-1} = \mathbf{Q}^{-1} \tilde{\mathbf{U}}^{-1}$, the inverse of the matrix $\mathbf{Q}$ being given by (as easily verified)

$$Q_{\alpha,\beta}^{-1} = \delta_{\alpha,\beta} - \frac{1}{\det \mathbf{Q}} \mathcal{B}_\alpha \mathcal{A}_\beta. \quad (81)$$

Then, the updated matrix elements of $(\tilde{\mathbf{U}}')^{-1}$ are given by

$$\tilde{U}_{\alpha,j}^{-1'} = \tilde{U}_{\alpha,j}^{-1} - \frac{\tilde{U}_{\alpha,l}^{-1}}{\det \mathbf{Q}} \left(\sum_{\beta} U_{K,\beta} \tilde{U}_{\beta,j}^{-1} - \delta_{l,j} \right). \quad (82)$$

This is a closed equation for updating the matrix $\tilde{\mathbf{U}}^{-1}$.

We would like to emphasize that the previous results for the calculation of the ratio of determinants and the updating can be further simplified. Indeed, at the beginning of the calculation, we can compute and store a $2L \times N_e$ matrix $\mathbf{W}$, whose elements are given by

$$W_{I,j} = \sum_{\alpha} U_{I,\alpha} \tilde{U}_{\alpha,j}^{-1}; \quad (83)$$

then the ratio of determinants (80) costs $O(1)$ operations, since it consists in taking the element corresponding to the new site (row) and the electron performing the hopping process (column)

$$\frac{\det \tilde{\mathbf{U}}'}{\det \tilde{\mathbf{U}}} = W_{K,l}. \quad (84)$$

The evaluation of $\mathbf{W}$ requires the knowledge of the full matrix $\mathbf{U}$, which has been computed and stored once for all at the beginning of the simulation (it does not depend upon the electronic configuration), and $\tilde{\mathbf{U}}^{-1}$, which instead depends upon the configuration $|x\rangle$. Then, a simple updating scheme for $\mathbf{W}$ is possible. In fact, by multiplying both sides of Eq. (82) by $U_{I,\alpha}$ and summing over α , we obtain

$$W'_{I,j} = W_{I,j} - \frac{W_{I,l}}{W_{K,l}} (W_{K,j} - \delta_{l,j}), \quad (85)$$

where, we have used that $\det \mathbf{Q} = W_{K,l}$, according to Eq. (84) and the definition of Eq. (83). Since each matrix element must be updated with $O(1)$ operations, the total cost is $O(2LN_e)$.

Let us finish this part by generalizing the previous formalism to the case where more than one electron hop, i.e., $|x'\rangle = d_{R'_{l_1}}^\dagger d_{R_{l_1}} \dots d_{R'_{l_m}}^\dagger d_{R_{l_m}} |x\rangle$, thus leading to a modification of m rows of the $\tilde{\mathbf{U}}$ matrix; for example, this could be the case for pair-hopping or spin-flip processes. Then, Eq. (73) generalizes into

$$\tilde{U}'_{j,\alpha} = \tilde{U}_{j,\alpha} + \sum_{r=1}^m \delta_{j,l_r} (U_{K_r,\alpha} - \tilde{U}_{l_r,\alpha}) = \tilde{U}_{j,\alpha} + \sum_{r=1}^m \delta_{j,l_r} v_\alpha^{K_r,l_r}; \quad (86)$$

as before, the indices K_r and l_r (for $r = 1, \dots, m$) are fixed, because they specify the sites where the electrons are hopping and the electron indices, respectively. By performing the same algebra as before, we get

$$\tilde{U}'_{j,\alpha} = \sum_{\beta} \tilde{U}_{j,\beta} \left(\delta_{\beta,\alpha} + \sum_{r=1}^m \tilde{U}_{\beta,l_r}^{-1} v_\alpha^{K_r,l_r} \right) = \sum_{\beta} \tilde{U}_{j,\beta} Q_{\beta,\alpha}, \quad (87)$$

where now the matrix $\mathbf{Q}$ has the following form

$$Q_{\beta,\alpha} = \delta_{\beta,\alpha} + \sum_{r=1}^m \mathcal{B}_\beta^r \mathcal{A}_\alpha^r, \quad (88)$$

where $\mathcal{B}_\beta^r = \tilde{U}_{\beta,l_r}^{-1}$ and $\mathcal{A}_\alpha^r = v_\alpha^{K_r,l_r}$. As before, the determinant of $\mathbf{Q}$ can be easily computed by solving the corresponding eigenvalue problem

$$\sum_{\alpha} Q_{\beta,\alpha} v_\alpha = v_\beta + \sum_{r=1}^m \mathcal{B}_\beta^r \sum_{\alpha} \mathcal{A}_\alpha^r v_\alpha = \lambda v_\beta, \quad (89)$$

which implies that all vectors that are orthogonal to the subspace defined by the $\mathcal{A}_\alpha^r$'s are eigenvectors with $\lambda = 1$; moreover, $v_\alpha = \sum_{r=1}^m x_r \mathcal{B}_\alpha^r$ is an eigenvector provided that the coefficients x_r satisfy

$$\sum_{s=1}^m \left(\delta_{r,s} + \sum_{\alpha} \mathcal{A}_\alpha^r \mathcal{B}_\alpha^s \right) x_s = \lambda x_r. \quad (90)$$

Therefore, the m non-trivial eigenvalues of $\mathbf{Q}$ are those of the $m \times m$ matrix

$$C_{r,s} = \delta_{r,s} + \sum_{\alpha} \mathcal{A}_\alpha^r \mathcal{B}_\alpha^s = W_{K_r,l_s}. \quad (91)$$

The final expression of the ratio of the two determinants is given by

$$\frac{\det \tilde{\mathbf{U}}'}{\det \tilde{\mathbf{U}}} = \det(W_{K_r, l_s}). \quad (92)$$

Also in this case, once the move is accepted, we have to update the matrix $(\tilde{\mathbf{U}}')^{-1} = \mathbf{Q}^{-1} \tilde{\mathbf{U}}^{-1}$. As before, the inverse of the $\mathbf{Q}$ matrix can be obtained

$$Q_{\alpha, \beta}^{-1} = \delta_{\alpha, \beta} - \sum_{r, s=1}^m \mathcal{B}_{\alpha}^r C_{r, s}^{-1} \mathcal{A}_{\beta}^s. \quad (93)$$

Therefore, we get

$$\tilde{U}_{\alpha, j}^{-1'} = \tilde{U}_{\alpha, j}^{-1} - \sum_{r, s=1}^m \tilde{U}_{\alpha, l_r}^{-1} C_{r, s}^{-1} \left(\sum_{\beta} U_{K_s, \beta} \tilde{U}_{\beta, j}^{-1} - \delta_{l_s, j} \right). \quad (94)$$

Then, the updated $\mathbf{W}'$ is obtained by multiplying both sides of the previous equation by $U_{i, \alpha}$ and summing over α

$$W'_{I, j} = W_{I, j} + \sum_{r=1}^m W_{I, l_r} b_j^{(r)}, \quad (95)$$

where

$$b_j^{(r)} = - \sum_{s=1}^m C_{r, s}^{-1} (W_{K_s, j} - \delta_{l_s, j}). \quad (96)$$

5.4 Backflow correlations

Here, we would like to discuss how to implement the updating of the determinant part in the presence of backflow correlations that have been introduced on lattice problem. In practice, we consider a quadratic Hamiltonian to construct the non-interacting orbitals $\{U_{I, \alpha}\}$, see Eq. (61). In the simplest approach, in which the backflow correlations act on holon-doublon (nearest-neighbor) pairs [33], the overlap between the generic configuration $|x\rangle$ and the backflow wave function $|\Phi_0^b\rangle$ is constructed from the “correlated” orbital with backflow correction

$$U_{I, \alpha}^b = \eta_0 U_{I, \alpha} + \eta_1 \sum_{\langle j \rangle_i} D_i H_j U_{J, \alpha}, \quad (97)$$

where $I = i$ ($I = i + L$) for electrons with spin up (down), and equivalently for J and j ; $\langle j \rangle_i$ indicates the sites j that are nearest neighbors of i ; D_i (H_i) is the operator that gives 1 if the site i is doubly occupied (empty) and 0 otherwise; finally η_0 and η_1 are variational parameters. Then, the wave function for the given configuration $|x\rangle$ is obtained by taking the determinant of the matrix

$$\tilde{\mathbf{U}}^b = \begin{pmatrix} U_{R_1, 1}^b & \cdots & U_{R_1, N_e}^b \\ \vdots & \ddots & \vdots \\ U_{R_{N_e}, 1}^b & \cdots & U_{R_{N_e}, N_e}^b \end{pmatrix}, \quad (98)$$

in which the rows correspond to the sites occupied by the electrons.

In practice, we can consider more general forms for the “correlated” orbitals, by considering further terms [34], but still remaining with the spirit of considering a linear combination of non-interacting orbitals depending on the many-body configuration $|x\rangle$ (in the previous case, the linear combination is taken for configurations having holons and doublons at nearest-neighbor sites).

6 Optimization techniques

6.1 Calculation of derivatives

In this section, we consider wave functions that depend upon a set of p variational parameters, which are arranged into a vector $\alpha = (\alpha_1, \dots, \alpha_p)$ and explicitly specified in the definition of the quantum state

$$\Psi_\alpha(x) = \langle x | \Psi_\alpha \rangle, \quad (99)$$

In the following, we discuss the basics ingredients that are necessary to compute the derivatives of the variational energy with respect to a given variational parameter α_k

$$f_k = -\frac{\partial E_\alpha}{\partial \alpha_k} = -\frac{\partial}{\partial \alpha_k} \frac{\langle \Psi_\alpha | \mathcal{H} | \Psi_\alpha \rangle}{\langle \Psi_\alpha | \Psi_\alpha \rangle}. \quad (100)$$

The dependence of E_α on the variational parameters is just a consequence of the fact that the wave function $|\Psi_\alpha\rangle$ depends upon α . Thus, in order to differentiate E_α , it is convenient to expand $|\Psi_\alpha\rangle$ for small changes $\alpha_k \rightarrow \alpha_k + \delta\alpha_k$. For a given configuration $|x\rangle$, where $\Psi_\alpha(x)$ is a complex number (in case of complex parameters, we can assume that all the α_k are real, once we consider their real and imaginary parts separately), we have that

$$\Psi_{\alpha+\delta\alpha_k}(x) = \Psi_\alpha(x) + \delta\alpha_k \frac{\partial \Psi_\alpha(x)}{\partial \alpha_k} + O(\delta\alpha_k^2), \quad (101)$$

where the notation $\Psi_{\alpha+\delta\alpha_k}(x)$ means that only the component α_k of the vector α has been incremented by $\delta\alpha_k$. In the following, for simplicity, we assume that $\Psi_\alpha(x) \neq 0$ for all the configurations. For fermionic systems in the continuous space, the nodal region $\Psi_\alpha(x) = 0$ represents a negligible (i.e., with zero measure) integration domain. On the lattice, accidental configurations with $\Psi_\alpha(x) = 0$ can be removed by considering a tiny perturbation of the variational *Ansatz* (e.g., by adding a small noisy part) and considering the limit of vanishing perturbation. Then, Eq. (101) can be formally written in terms of a local operator $\mathcal{O}_k$, corresponding to the parameter α_k and defined by diagonal matrix elements $\mathcal{O}_k(x)$

$$\langle x | \mathcal{O}_k | x' \rangle = \delta_{x,x'} \mathcal{O}_k(x), \quad (102)$$

$$\mathcal{O}_k(x) = \frac{\partial \ln \Psi_\alpha(x)}{\partial \alpha_k} = \frac{1}{\Psi_\alpha(x)} \frac{\partial \Psi_\alpha(x)}{\partial \alpha_k}; \quad (103)$$

here, in principle, $\mathcal{O}_k(x)$ may depend upon the variational parameters α , however, to keep the notation simple, we prefer not to put the label α in the local operators. The important point is

that $\mathcal{O}_k(x)$ can be usually computed for the given *Ansatz* of the variational state. In this way, we can write a formal expansion of the many-body state as

$$|\Psi_{\alpha+\delta\alpha_k}\rangle = (1 + \delta\alpha_k \mathcal{O}_k) |\Psi_\alpha\rangle, \quad (104)$$

which can be readily verified by taking the overlap of both sides of the above equation with $|x\rangle$ and using Eqs. (102) and (103). Notice that the diagonal operator $\mathcal{O}_k$ is not necessarily Hermitian, as its diagonal elements are not necessarily real, for a generic complex case.

Let us now show how to obtain the explicit form of the energy derivative with respect to a given variational parameter. It is clear that the variational energy E_α (as well as any other correlation function) does not depend on the overall normalization (and global phase) of the wave function. In other words, by scaling the wave function by an arbitrary complex constant c , i.e., $|\Psi_\alpha\rangle \rightarrow c|\Psi_\alpha\rangle$, E_α remains unchanged. In order to exploit this property, it is better to consider explicitly normalized wave functions. First of all we define

$$|v_{0,\alpha}\rangle \equiv \frac{|\Psi_\alpha\rangle}{\|\Psi_\alpha\|}, \quad (105)$$

where $\|\Psi_\alpha\|$ indicates the norm of the state $|\Psi_\alpha\rangle$. Then, we define a set of states (one for each value of $k = 1, \dots, p$)

$$|v_{k,\alpha}\rangle \equiv (\mathcal{O}_k - \overline{\mathcal{O}}_k) |v_{0,\alpha}\rangle, \quad (106)$$

where

$$\overline{\mathcal{O}}_k = \langle v_{0,\alpha} | \mathcal{O}_k | v_{0,\alpha} \rangle = \frac{\langle \Psi_\alpha | \mathcal{O}_k | \Psi_\alpha \rangle}{\langle \Psi_\alpha | \Psi_\alpha \rangle}. \quad (107)$$

The states $|v_{k,\alpha}\rangle$ are orthogonal to $|v_{0,\alpha}\rangle$, as easily verified when using Eq. (107); however, they are neither normalized nor orthogonal to each other, i.e., in general $\langle v_{k,\alpha} | v_{k',\alpha} \rangle \neq \delta_{k,k'}$ for $k, k' \neq 0$. Therefore, the set of states $|v_{0,\alpha}\rangle$ and $\{|v_{k,\alpha}\rangle\}$ defines a *semi-orthogonal* basis.

In order to compute the normalized wave function when the parameter α_k is changed, we first compute the norm of $|\Psi_{\alpha+\delta\alpha_k}\rangle$

$$\begin{aligned} \|\Psi_{\alpha+\delta\alpha_k}\|^2 &= \langle \Psi_\alpha | (1 + \delta\alpha_k \mathcal{O}_k)^* (1 + \delta\alpha_k \mathcal{O}_k) | \Psi_\alpha \rangle \\ &= \|\Psi_\alpha\|^2 [1 + 2\Re(\delta\alpha_k \overline{\mathcal{O}}_k) + O(\delta\alpha_k^2)]. \end{aligned} \quad (108)$$

Then, we have that

$$\begin{aligned} |v_{0,\alpha+\delta\alpha_k}\rangle &= \frac{|\Psi_{\alpha+\delta\alpha_k}\rangle}{\|\Psi_{\alpha+\delta\alpha_k}\|} = |v_{0,\alpha}\rangle + [\delta\alpha_k \mathcal{O}_k - \Re(\delta\alpha_k \overline{\mathcal{O}}_k)] |v_{0,\alpha}\rangle + O(\delta\alpha_k^2) \\ &= [1 + i\Im(\delta\alpha_k \overline{\mathcal{O}}_k)] |v_{0,\alpha}\rangle + \delta\alpha_k |v_{k,\alpha}\rangle + O(\delta\alpha_k^2), \end{aligned} \quad (109)$$

which can be finally recast as

$$|v_{0,\alpha+\delta\alpha_k}\rangle = \exp(i\delta\phi) [|v_{0,\alpha}\rangle + \delta\alpha_k |v_{k,\alpha}\rangle] + O(\delta\alpha_k^2), \quad (110)$$

where $\delta\phi = \Im(\delta\alpha_k \overline{\mathcal{O}}_k)$.

By using the above expression, it is immediate to work out the derivative of the variational energy E_α with respect to a given variational parameter α_k

$$\begin{aligned} \frac{\partial E_\alpha}{\partial \alpha_k} &= \lim_{\delta \alpha_k \rightarrow 0} \frac{\langle v_{0,\alpha+\delta \alpha_k} | \mathcal{H} | v_{0,\alpha+\delta \alpha_k} \rangle - \langle v_{0,\alpha} | \mathcal{H} | v_{0,\alpha} \rangle}{\delta \alpha_k} = \langle v_{k,\alpha} | \mathcal{H} | v_{0,\alpha} \rangle + \langle v_{0,\alpha} | \mathcal{H} | v_{k,\alpha} \rangle \\ &= 2\Re \left[\frac{\langle \Psi_\alpha | \mathcal{H} (\mathcal{O}_k - \bar{\mathcal{O}}_k) | \Psi_\alpha \rangle}{\langle \Psi_\alpha | \Psi_\alpha \rangle} \right]. \end{aligned} \quad (111)$$

Notice also that, as expected, the phase factor $\delta\phi$ does not enter in the above expression.

In order to evaluate the force f_k by a standard Monte Carlo sampling, we introduce a completeness relation to have

$$\begin{aligned} f_k &= -2\Re \left[\frac{\sum_x \langle \Psi_\alpha | \mathcal{H} | x \rangle \langle x | (\mathcal{O}_k - \bar{\mathcal{O}}_k) | \Psi_\alpha \rangle}{\sum_x \langle \Psi_\alpha | x \rangle \langle x | \Psi_\alpha \rangle} \right] = \\ &= -2\Re \left[\sum_x \frac{e_L^*(x) (\mathcal{O}_k(x) - \bar{\mathcal{O}}_k) |\Psi_\alpha(x)|^2}{\sum_x |\Psi_\alpha(x)|^2} \right], \end{aligned} \quad (112)$$

where $e_L^*(x)$ is the complex conjugate of the local energy (here, we omitted the index α , in harmony with the notation adopted for the local operators $\mathcal{O}_k$), as it is generally a complex-valued function. Then, f_k can be evaluated by considering

$$f_k \approx -2\Re \left[\frac{1}{N} \sum_{i=1}^N e_L^*(x_i) (\mathcal{O}_k(x_i) - \bar{\mathcal{O}}_k) \right], \quad (113)$$

$$\bar{\mathcal{O}}_k \approx \frac{1}{N} \sum_{i=1}^N \mathcal{O}_k(x_i). \quad (114)$$

6.2 The stochastic reconfiguration

The knowledge of energy derivatives of Eq. (100) allows us to employ the steepest-descent method [35] to change the variational parameters $\alpha = (\alpha_1, \dots, \alpha_p)$, even when p is very large

$$\alpha'_k = \alpha_k + \delta \alpha_k, \quad (115)$$

$$\delta \alpha_k = \Delta f_k, \quad (116)$$

where Δ is an arbitrary (small) constant. In principle, its value can be optimized to reach the lowest possible energy at each iteration; however, in most applications, it is common practice to keep Δ constant along the minimization procedure. Then, the variational parameters are iteratively improved along a Markov chain procedure. In absence of noise, the steepest-descent method always converges to a minimum, where the Euler conditions $f_k = 0$ are satisfied. Indeed, let us suppose that $f_k \neq 0$, then the energy for α' is given by a Taylor expansion to linear order in Δ

$$E_{\alpha'} = E_\alpha + \sum_k \frac{\partial E_\alpha}{\partial \alpha_k} \delta \alpha_k + O(\Delta^2) = E_\alpha - \Delta \sum_k f_k^2 + O(\Delta^2), \quad (117)$$

where we used that $\partial E_\alpha / \partial \alpha_k = -f_k$ and $\delta \alpha_k = \Delta f_k$. Therefore, for small Δ , when the linear truncation is accurate enough in the Taylor expansion, we obtain that

$$\Delta E \equiv E_{\alpha'} - E_\alpha = -\Delta \sum_k f_k^2 \leq 0; \quad (118)$$

here, the equality sign holds only when $f_k = 0$. Thus, the method converges to a minimum for a large number of iterations just because the energy monotonically decreases with the number of iterations. Within the steepest-descent method only the first derivative of the energy is computed and it is certain that a small change of the parameters $\delta \alpha = (\delta \alpha_1, \dots, \delta \alpha_p)$ parallel to the force $\mathbf{f} = (f_1, \dots, f_p)$ will decrease the energy; the only issue concerns the size of Δ , which must be taken sufficiently small to make the quadratic term in Eq. (117) negligible.

Although the steepest-descent approach eventually converges to a local minimum, its effectiveness may suffer from cases where the dependence on the variational parameters is highly non-linear (e.g., in the Jastrow factors). In this case, a small change of a given variational parameter can produce very different wave functions and physical quantities, whereas another parameter may weakly affect the wave function. In order to overcome these difficulties, it is important to introduce an appropriate metric δs^2 that is used to estimate the “proximity” of two normalized (complex) wave functions $|v_{0,\alpha}\rangle$ and $|v_{0,\alpha+\delta\alpha}\rangle$

$$\delta s^2 = \min_{\delta\theta} \|\exp(-i\delta\theta)v_{0,\alpha+\delta\alpha} - v_{0,\alpha}\|^2. \quad (119)$$

Here, the minimization on the phase factor $\delta\theta$ is necessary because we do not want to distinguish between two wave functions that differ only by an overall phase factor, as they produce the same correlation functions. In other words, we want to define a distance δs^2 that vanishes when we have physically equivalent wave functions. Then, we replace in Eq. (119) the expression for $|v_{0,\alpha+\delta\alpha}\rangle$ that is obtained by generalizing Eq. (110) to the case where several parameters are changed

$$|v_{0,\alpha+\delta\alpha}\rangle = \exp(i\delta\phi) \left[|v_{0,\alpha}\rangle + \sum_k \delta\alpha_k |v_{k,\alpha}\rangle \right] + O(|\delta\alpha|^2), \quad (120)$$

where $\delta\phi = \sum_k \Im(\delta\alpha_k \bar{\mathcal{O}}_k)$. Now, the minimization over $\delta\theta$ gives $\delta\theta = \delta\phi$, thus leading to

$$\delta s^2 = \sum_{k,k'} \langle v_{k,\alpha} | v_{k',\alpha} \rangle \delta\alpha_k \delta\alpha_{k'} + O(|\delta\alpha|^2). \quad (121)$$

Since all increments $\delta\alpha_k$ are assumed real (as discussed previously, here we assume that all parameters are real), we can symmetrize the previous expression with respect to the indices k and k' and neglect the terms that are $O(|\delta\alpha|^2)$, obtaining

$$\delta s^2 = \frac{1}{2} \sum_{k,k'} (\langle v_{k,\alpha} | v_{k',\alpha} \rangle + \langle v_{k',\alpha} | v_{k,\alpha} \rangle) \delta\alpha_k \delta\alpha_{k'}. \quad (122)$$

In this way, we can finally identify a matrix $\mathbf{S}$ that fully determines the metric in the space of normalized wave functions

$$S_{k,k'} = \Re(\langle v_{k,\alpha} | v_{k',\alpha} \rangle), \quad (123)$$

which implies that the distance between two wave functions reads

$$\delta s^2 = \sum_{k,k'} S_{k,k'} \delta \alpha_k \delta \alpha_{k'}. \quad (124)$$

At this point, it is natural to improve the steepest-descent method by using the metric given by $\mathbf{S}$. The minimization of $\Delta E + \mu \delta s^2$ with the metric δs^2 given in Eq. (124) improves the convergence to the minimum of the variational energy with respect to the simple steepest-descent approach, as non-equivalent parameters can be appropriately changed with a different scale. This approach is called *stochastic reconfiguration* [16]. The minimization of $\Delta E + \mu \delta s^2$ gives

$$\sum_{k'} S_{k,k'} \delta \alpha_{k'} = \frac{f_k}{2\mu}, \quad (125)$$

which is a set of linear equations for the unknown vector $\delta \alpha$. After having solved this linear system, we can update the variational parameters until convergence is reached; as in the steepest-descent method, we can set $\Delta = 1/(2\mu)$ small enough, which may be kept fixed during the optimization. We would like to stress the fact that, since the matrix $\mathbf{S}$ is strictly positive definite, the energy is monotonically decreasing along the optimization as

$$\Delta E = -\Delta \sum_{k,k'} S_{k,k'}^{-1} f_k f_{k'} < 0. \quad (126)$$

Within a Monte Carlo procedure, the matrix $\mathbf{S}$ is evaluated by a finite sampling of N configurations $\{x_i\}$ as

$$S_{k,k'} \approx \Re \left[\frac{1}{N} \sum_{i=1}^N (\mathcal{O}_k(x_i) - \overline{\mathcal{O}}_k)(\mathcal{O}_{k'}(x_i) - \overline{\mathcal{O}}_{k'}) \right]; \quad (127)$$

note that the forces are computed in a similar way, see Eq. (113).

7 Selected results

Here, we would like to show a few selected results that have been obtained by using Jastrow-Slater wave functions, also including backflow correlations. First of all, we would like to emphasize the importance of the long-range tail in the Jastrow factor, in order to correctly reproduce the strong-coupling regime of the Hubbard model. In Fig. 1, we report the variational energies for the one-dimensional Hubbard model at half filling for two different clusters with $L = 18$ and 82 sites. The Slater determinant is obtained by filling the lowest-energy states of the quadratic Hamiltonian (56) with only nearest-neighbor hopping (dubbed as “Fermi sea”). The comparison is between the cases with (on-site) Gutzwiller and (long-range) Jastrow factors. It is well known that, the fully-projected Gutzwiller state gives very accurate energies for the Heisenberg model [36]. Instead, considering the Hubbard model with finite repulsion U/t , it turns out that, in the strong-coupling limit, the Gutzwiller state gives a rather poor variational description, missing completely the super-exchange energy generated by the virtual-hopping processes. This happens because, by increasing U/t , the Gutzwiller parameter g increases, and

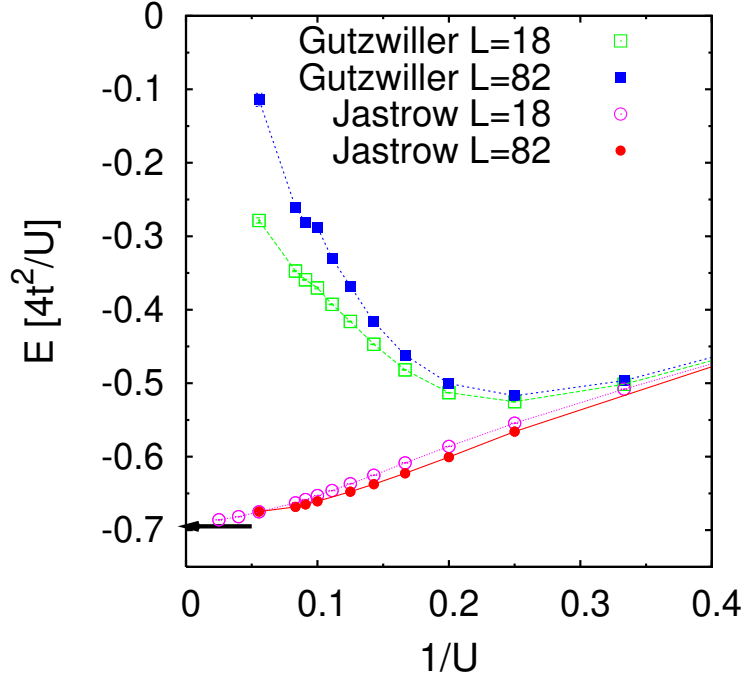


Fig. 1: Energy per site (in units of $4t^2/U$) for the one-dimensional Hubbard model at half filling. The results for wave functions in which the Gutzwiller and Jastrow factors are applied to the Fermi sea are reported for two cluster sizes ($L = 18$ and 82). The arrow indicates the energy per site for the fully-projected Fermi sea in the Heisenberg model.

the hopping processes, which create double occupancies, become less probable, with a consequent kinetic-energy loss. By contrast, the long-range Jastrow factor enables us to connect the fully-projected insulator valid in the strong-coupling limit to an insulating state at finite U/t , as demonstrated by the fact that the variational energy of $|\Psi_J\rangle$ approaches the one calculated with the fully-projected Gutzwiller wave function.

In two dimensions (or for frustrated one-dimensional cases, with next-nearest-neighbor hopping t'), the situation is more delicate: the Jastrow factor is no longer sufficient to correctly reproduce the ground-state properties in the strongly-correlated regime and, in many cases, backflow correlations are necessary to reach high accuracies. For example, we consider the Hubbard model with $t' = 0$ and $t'/t = 0.7$, for two different variational wave functions. The first is constructed by applying a Jastrow factor on top of a BCS state, which is suitable to describe non-magnetic (i.e., spin-liquid) states. The second is obtained from a quadratic Hamiltonian that explicitly includes antiferromagnetic order, to describe a magnetically ordered phase

$$\mathcal{H}_{\text{AF}} = -t \sum_{\langle i,j \rangle, \sigma} c_{i,\sigma}^\dagger c_{j,\sigma} + h.c. + \Delta_{\text{AF}} \sum_j \left(e^{i \mathbf{Q} \cdot \mathbf{R}_j} c_{j,\uparrow}^\dagger c_{j,\downarrow} + e^{-i \mathbf{Q} \cdot \mathbf{R}_j} c_{j,\downarrow}^\dagger c_{j,\uparrow} \right), \quad (128)$$

where $\langle \dots \rangle$ indicates neighboring sites and $\mathbf{Q} = (\pi, \pi)$ is the pitch vector for the Néel order. In order to have the correct spin-spin correlations at large distance, we can apply a long-range spin Jastrow factor

$$\mathcal{J}_s = \exp \left[-\frac{1}{2} \sum_{i,j} u_{i,j} S_i^z S_j^z \right], \quad (129)$$

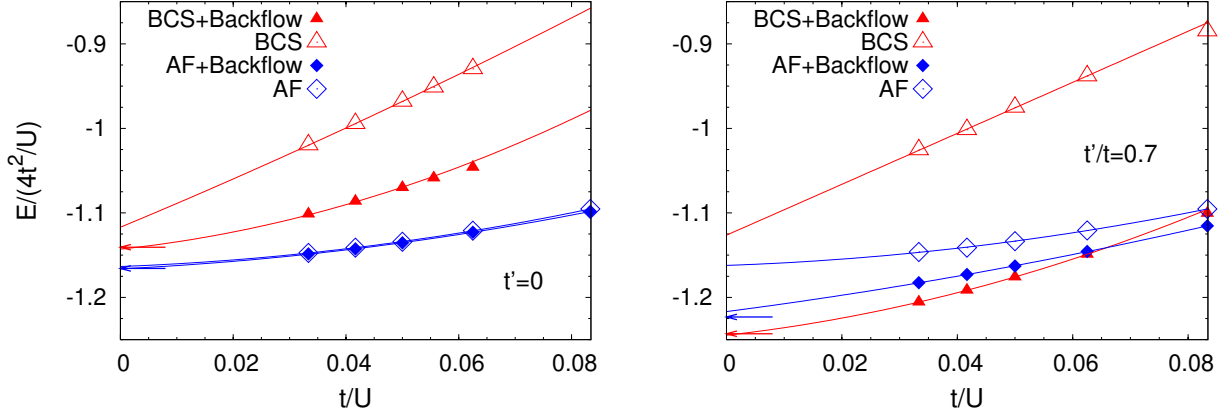


Fig. 2: Energies per site (in units of $J = 4t^2/U$) for the two-dimensional Hubbard model at half filling, for both the unfrustrated ($t' = 0$) and frustrated ($t'/t = 0.7$) case. The cases with and without backflow correlations are reported (for the BCS state). The results for the wave function with antiferromagnetic order and no BCS pairing are also shown. Arrows indicate the energies per site for the corresponding fully-projected states in the Heisenberg model.

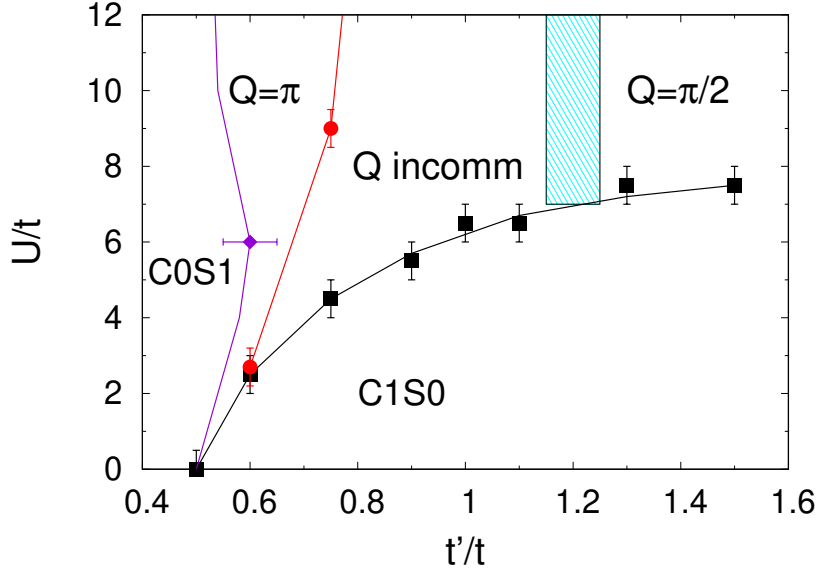


Fig. 3: Phase diagram of the t - t' Hubbard model at half-filling with the metallic phase with gapped spin excitations (C1S0) and the insulating phase with gapless spin excitations (C0S1). The insulating phase with gapped spin excitations (C0S0) has regions with commensurate ($Q=\pi$ and $Q = \pi/2$) and incommensurate (Q incomm) spin-spin correlations.

where $S_i^z = 1/2(n_{i,\uparrow} - n_{i,\downarrow})$ and $u_{i,j}$ is a pseudo-potential that can be optimized for each independent distance. In analogy with the density Jastrow factor of Eq. (45), it governs spin fluctuations orthogonal to the magnetic field Δ_{AF} [37]. It is important to stress that the uncorrelated state, obtained from Eq. (128), can easily satisfy the single-occupancy constraint by taking $\Delta_{AF} \rightarrow \infty$. In this limit, it also contains the virtual hopping processes, which are generated by the kinetic term, implying that it is possible to reproduce super-exchange processes.

The results for the energies of these two variational states are reported in Fig. 2. In general, the results for the Hubbard model (at finite interaction U/t) are not smoothly connected to the ones

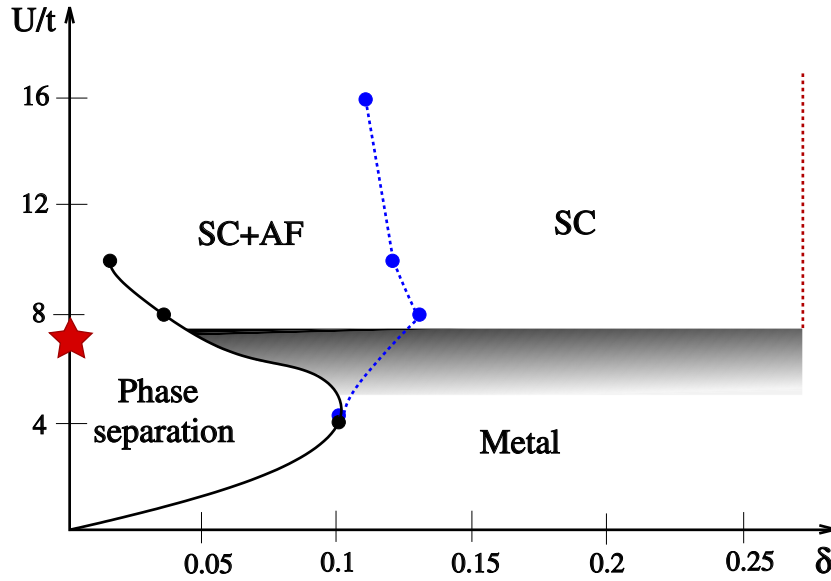


Fig. 4: Schematic phase diagram for the two-dimensional Hubbard model by varying U/t and the filling factor $\delta = 1 - n$ (where n is the density of electrons). The red star labels the location of the hidden Mott transition U_{Mott}/t at half filling. The black line with black dots denotes the boundary of the phase-separation region, that shrinks for $U/t \gtrsim U_{\text{Mott}}/t$. The dashed blue line with blue dots marks the disappearance of Δ_{AF} in the variational state. The dashed red line indicates the boundary of the region where sizable pairing correlations are detected. Finally, in the shaded gray region finite-size effects are strong and precise results cannot be obtained in the thermodynamic limit.

obtained with the fully-projected states and the Heisenberg model, except for the unfrustrated case with $t' = 0$ and the antiferromagnetic wave function. However, thanks to backflow correlations, it is possible to obtain a correct extrapolation to the infinite- U limit. The importance of backflow correlations is extremely important in the frustrated case, where they are essential also for improving the accuracy of the antiferromagnetic wave function.

Therefore, by using Jastrow-Slater states, which include backflow correlations, it is possible to get quite accurate results in a variety of models. As an example, we report the phase diagram of the t - t' Hubbard model in one dimension at half filling, see Fig. 3 [38]. There are three main phases: one is stabilized for small values of the frustrating ratio t'/t , which is gapless in the spin sector and gapped in the charge one (denoted by C0S1); another, for small values of U/t and sufficiently large t'/t , which is gapless in the charge sector and gapped in the spin one (C1S0); and the last one, for large U/t and t'/t , which is fully gapped and dimerized (C0S0).

As a final example, we show the phase diagram that is obtained for the two-dimensional Hubbard model, as a function of U/t and the filling factor $\delta = 1 - n$, see Fig. 4 [39]. Here, the antiferromagnetic phase, which is stable at half filling (i.e., $\delta = 0$), gives rise to phase separation at small values of the interaction strength. Remarkably, these results suggest that a reminiscence of the Mott transition, hidden by the antiferromagnetic phase at half-filling, emerges after a careful analysis of the BCS pairing. This hidden Mott transition is intimately related with the change from Slater to Mott antiferromagnetism, the former being related to a Fermi surface in-

stability towards antiferromagnetic order, while the latter being connected to a super-exchange mechanism. For Coulomb interactions that are smaller than this “critical” value, the system is unstable towards phase separation and there is no strong evidence that superconductivity may emerge; by contrast, for larger values of U/t , hole doping drives the Mott antiferromagnet into a homogeneous superconducting phase, with the condensation energy gain shifting from potential to kinetic by increasing U/t .

8 Conclusions

Variational wave functions, such as Jastrow-Slater states, represents a very powerful and useful tool to investigate strongly-correlated systems on the lattice. In the recent past, there has been increasing evidence that it is possible to construct many-body states that may compete with other numerical methods, such as density-matrix renormalization group or its recent developments based upon tensor networks [40]. Future developments, in which neural networks may be used to generalize the Jastrow factor and better describe correlation effects [41], may be beneficial to solve the many-body problem.

References

- [1] N. Mott, Proc. Phys. Soc. (London) **62**, 416 (1949)
- [2] M. Imada, A. Fujimori, and Y. Tokura, Rev. Mod. Phys. **70**, 1039 (1998)
- [3] P. Lee, N. Nagaosa, and X.G. Wen, Rev. Mod. Phys. **78**, 17 (2006)
- [4] J. Hubbard, Proc. Royal Soc. of London **276**, 238 (1963)
- [5] M. Gutzwiller, Phys. Rev. Lett. **10**, 159 (1963)
- [6] J. Kanamori, Prog. Theor. Phys. **30**, 275 (1963)
- [7] H. Bethe, Z. Phys. **71**, 205 (1931)
- [8] A. Georges, G. Kotliar, W. Krauth, and M. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996)
- [9] P. Anderson, Science **235**, 1196 (1987)
- [10] F. Zhang and T. Rice, Phys. Rev. B **37**, 3759 (1988)
- [11] T. Kaplan, P. Horsch, and P. Fulde, Phys. Rev. Lett. **49**, 889 (1982)
- [12] H. Yokoyama and H. Shiba, J. Phys. Soc. Jpn. **59**, 3669 (1990)
- [13] R. Jastrow, Phys. Rev. **98**, 1479 (1955)
- [14] W. McMillan, Phys. Rev. **138**, A442 (1965)
- [15] L. Reatto and G. Chester, Phys. Rev. **155**, 88 (1967)
- [16] S. Sorella, Phys. Rev. B **64**, 024512 (2001)
- [17] S. Sorella, Phys. Rev. B **71**, 241103 (2005)
- [18] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1957)
- [19] W. Hastings, Biometrika **57**, 97 (1970)
- [20] H. Yokoyama and H. Shiba, J. Phys. Soc. Jpn. **56**, 1490 (1987)
- [21] H. Yokoyama and H. Shiba, J. Phys. Soc. Jpn. **56**, 3582 (1987)
- [22] P. Anderson, Mater. Res. Bull. **8**, 153 (1973)
- [23] L. Capriotti, F. Becca, A. Parola, and S. Sorella, Phys. Rev. Lett. **87**, 097201 (2001)
- [24] M. Capello, F. Becca, M. Fabrizio, S. Sorella, and E. Tosatti, Phys. Rev. Lett. **94**, 026406 (2005)

- [25] M. Capello, F. Becca, S. Yunoki, and S. Sorella, Phys. Rev. B **73**, 245116 (2006)
- [26] E. Wigner and F. Seitz, Phys. Rev. **46**, 509 (1934)
- [27] R. Feynman, Phys. Rev. **94**, 262 (1954)
- [28] R. Feynman and M. Cohen, Phys. Rev. **102**, 1189 (1956)
- [29] V. Pandharipande and N. Itoh, Phys. Rev. A **8**, 2564 (1973)
- [30] K. Schmidt and V. Pandharipande, Phys. Rev. B **19**, 2504 (1979)
- [31] Y. Kwon, D. Ceperley, and R. Martin, Phys. Rev. B **48**, 12037 (1993)
- [32] Y. Kwon, D. Ceperley, and R. Martin, Phys. Rev. B **58**, 6800 (1998)
- [33] L. Tocchio, F. Becca, A. Parola, and S. Sorella, Phys. Rev. B **78**, 041101 (2008)
- [34] L. Tocchio, F. Becca, and C. Gros, Phys. Rev. B **83**, 195138 (2011)
- [35] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery:
Numerical recipes 3rd edition: the art of scientific computing
(Cambridge University Press, 2007)
- [36] C. Gros, Ann. Phys. **189**, 53 (1989)
- [37] F. Becca, M. Capone, and S. Sorella, Phys. Rev. B **62**, 12700 (2000)
- [38] L. Tocchio, F. Becca, and C. Gros, Phys. Rev. B **81**, 205109 (2010)
- [39] L. Tocchio, F. Becca, and S. Sorella, Phys. Rev. B **94**, 195126 (2016)
- [40] J. LeBlanc, A. Antipov, F. Becca, I. Bulik, G.L. Chan, C.M. Chung, Y. Deng, M. Ferrero, T. Henderson, C. Jiménez-Hoyos, E. Kozik, X.W. Liu, A. Millis, N. Prokof'ev, M. Qin, G.E. Scuseria, H. Shi, B. Svistunov, L. Tocchio, I. Tupitsyn, S. White, S. Zhang, B.X. Zheng, Z. Zhu, and E. Gull, Phys. Rev. X **5**, 041041 (2015)
- [41] G. Carleo and M. Troyer, Science **355**, 602 (2017)

6 Auxiliary-Field Quantum Monte Carlo at Zero- and Finite-Temperature

Shiwei Zhang

Center for Computational Quantum Physics

Flatiron Institute, New York, NY10010, USA

Department of Physics, College of William & Mary

Williamsburg, VA23185, USA

Contents

1	Introduction	2
2	Formalism	3
2.1	Hubbard-Stratonovich transformation	4
2.2	Ground-state projection	6
2.3	Finite-temperature, grand-canonical ensemble calculations	9
3	Constraining the paths in AFQMC	11
3.1	An exact boundary condition	11
3.2	Implementing the boundary condition and importance sampling	12
3.3	The phaseless formalism for complex auxiliary-fields	15
4	Overview of applications, further reading, and outlook	19
A	A few basics of Monte Carlo techniques	22
B	Properties of non-orthogonal Slater determinants	24

1 Introduction

Auxiliary-field quantum Monte Carlo (AFQMC) is increasingly becoming a general and powerful tool for studying interacting many-fermion systems, in different sub-areas including the study of correlated electron models, cold Fermi gas, electronic structure of solids, and quantum chemistry. My lecture will give an introduction to the advances, both in the zero-temperature or ground-state form and in the non-zero-temperature ($T > 0$ K) form, which have allowed AFQMC to become a systematic and scalable tool. The ground-state portion of the lecture will draw (or even copy!) from Refs. [1] and [2]; the $T > 0$ K portion is based on Refs. [3] and [4]. As we have seen repeatedly echoed through this school, the accurate treatment of interacting quantum systems is one of the grand challenges in modern science. Explicit solution of the many-body Schrödinger equation leads to rapidly growing computational cost as a function of system size (see, e.g., [5]). To circumvent the problem, most computational quantum mechanical studies of large, realistic systems rely on simpler independent-particle approaches based on density-functional theory (DFT) (see, e.g., [6, 7]), using an approximate energy functional to include many-body effects. These replace the electron-electron interaction by an effective potential, thereby reducing the problem to a set of one-electron equations. Methods based on DFT and through its Car-Parrinello molecular dynamics implementation [8] have been extremely effective in complex molecules and solids [6]. Such approaches are the standard in electronic structure, widely applied in condensed matter, quantum chemistry, and materials science.

Despite the tremendous successes of DFT, the treatment of electronic correlation is approximate. For strongly correlated systems (e.g., high-temperature superconductors, heavy-fermion metals, magnetic materials, optical lattices), where correlation effects from particle interaction crucially modify materials properties, the approximation can lead to qualitatively incorrect results. Even in moderately correlated systems when the method is qualitatively correct, the results are sometimes not sufficiently accurate. For example, in ferroelectric materials the usually acceptable 1% errors in DFT predictions of the equilibrium lattice constant can lead to almost full suppression of the ferroelectric order. Additional challenges are present to go beyond ground state and account for thermal as well as quantum fluctuations.

The development of alternatives to independent-particle theories is therefore of paramount fundamental and practical significance. To accurately capture the quantum many-body effects, the size of the Hilbert space involved often grows exponentially. Simulation methods utilizing Monte Carlo (MC) sampling [9–14] are, in principle, both non-perturbative and well-equipped to handle details and complexities in the external environment. They are a unique combination of accuracy, general applicability, favorable scaling (low-power) of computational cost with physical system size, and scalability on parallel computing platforms [15].

For fermion systems, however, a so-called sign problem [16–18] arises in varying forms in these MC simulation methods. The Pauli exclusion principle requires that the states be anti-symmetric under interchange of two particles. As a consequence, negative signs appear, which cause cancellations among contributions of the MC samples of the wave function or density matrix. In some formalisms, as we discuss below, a phase appears which leads to a continuous

degeneracy and more severe cancellations. As the temperature is lowered or the system size is increased, such cancellation becomes more and more complete. The net signal thus decays *exponentially* versus noise. The algebraic scaling is then lost, and the method breaks down.

In AFQMC, we cast the MC random walks in a space of over-complete Slater determinants, which significantly reduces the severity of the sign problem. In this space we study the properties of the paths and derive exact boundary conditions for the sign of their contributions in the ground-state wave function or $T > 0$ K density matrix. This then allows us to formulate approximate constraints on the random walk paths which are less sensitive to the details of the constraint. We then develop internal checks and constraint release methods to systematically improve the approach. These methods have come under the name of constrained path Monte Carlo (CPMC) [19] for systems where there is a sign problem (for example, Hubbard-like models where the auxiliary-fields are real due to the short-ranged interactions). For electronic systems where there is a phase problem (as the Coulomb interaction leads to complex auxiliary fields), the methods have been referred to as phaseless or phase-free AFQMC [14, 20, 21]. We will refer to the methods as AFQMC following more recent literature in *ab initio* electronic structure. When it is necessary to emphasize the constrained-path (CP) approximation, to distinguish from unconstrained zero-temperature (free-projection) or $T > 0$ K (sometimes referred to as determinantal MC) calculations, we will refer to them as CP-AFQMC.

2 Formalism

The Hamiltonian for any many-fermion system with two-body interactions (e.g., the electronic Hamiltonian under the Born-Oppenheimer approximation) can be written as

$$\hat{H} = \hat{H}_1 + \hat{H}_2 = -\frac{\hbar^2}{2m} \sum_{m=1}^M \nabla_m^2 + \sum_{m=1}^M V_{\text{ext}}(\mathbf{r}_m) + \sum_{m<n}^M V_{\text{int}}(\mathbf{r}_m - \mathbf{r}_n), \quad (1)$$

where $\mathbf{r}_m$ is the real-space coordinate of the m -th fermion. The one-body part of the Hamiltonian, $\hat{H}_1$, consists of the kinetic energy of the electrons and the effect of the ionic (and any other external) potentials. (We have represented the external potential as local, although this does not have to be the case. For example, in plane-wave calculations we will use a norm-conserving pseudopotential, which will lead to a non-local function V_{ext} .) The two-body part of the Hamiltonian, $\hat{H}_2$, contains the electron-electron interaction terms. The total number of fermions, M , will be fixed in the ground-state calculations; for $T > 0$ K, a chemical potential term will be included and the number of fermions will fluctuate. For simplicity, we have suppressed the spin-index; when the spins need to be explicitly distinguished, M_σ will denote the number of electrons with spin σ ($\sigma = \uparrow$ or $\downarrow$). We assume that the interaction is spin-independent, so the total S_z , defined by $(M_\uparrow - M_\downarrow)$, is fixed in the ground-state calculation, although it will be straightforward to generalize our discussions to treat other cases, for example, when there is spin-orbit coupling (SOC) [22].

With any chosen one-particle basis, the Hamiltonian can be written in second quantization in the general form

$$\hat{H} = \hat{H}_1 + \hat{H}_2 = \sum_{i,j}^N T_{ij} c_i^\dagger c_j + \frac{1}{2} \sum_{i,j,k,l}^N V_{ijkl} c_i^\dagger c_j^\dagger c_k c_l, \quad (2)$$

where the one-particle basis, $\{|\chi_i\rangle\}$ with $i = 1, 2, \dots, N$, can be lattice sites (Hubbard model), plane-waves (as in solid state calculations) [23], or Gaussians (as in quantum chemistry) [20, 24], etc. The operators $c_i^\dagger$ and c_i are creation and annihilation operators on $|\chi_i\rangle$, satisfying standard fermion commutation relations. The one-body matrix elements, T_{ij} , contain the effect of both the kinetic energy and external potential. For the $T > 0$ K calculations we will discuss, a term $\mu \hat{n}$ containing the chemical potential μ and density operator is included [12, 25, 26]. The two-body matrix elements, V_{ijkl} , are from the interaction. The matrix elements are expressed in terms of the basis functions, for example,

$$V_{ijkl} = \int d\mathbf{r}_1 d\mathbf{r}_2 \chi_i^*(\mathbf{r}_1) \chi_j^*(\mathbf{r}_2) V_{\text{int}}(\mathbf{r}_1 - \mathbf{r}_2) \chi_k(\mathbf{r}_2) \chi_l(\mathbf{r}_1). \quad (3)$$

In quantum chemistry calculations, these are readily evaluated with standard Gaussian basis sets. In solid state calculations with plane-waves, the kinetic and electron-electron interaction terms have simple analytic expressions, while the electron-ion potential leads to terms which are provided by the pseudopotential generation. We will assume that all matrix elements in Eq. (2) have been evaluated and are known as we begin our many-body calculations.

2.1 Hubbard-Stratonovich transformation

The two-body part in the Hamiltonian in Eq. (2), $\hat{H}_2$, can be written in the following form

$$\hat{H}_2 = \frac{1}{2} \sum_{\gamma=1}^{N_\gamma} \lambda_\gamma \hat{v}_\gamma^2, \quad (4)$$

where λ_γ is a constant, $\hat{v}_\gamma$ is a one-body operator similar to $\hat{H}_1$, and N_γ is an integer. There are various ways to achieve the decomposition in Eq. (4) for a general two-body term [27]. Below we outline the two most commonly applied cases in electronic structure: **(a)** with *plane-wave basis* and **(b)** for a more dense matrix V_{ijkl} resulting from a *general basis set* such as Gaussians. In a *plane-wave basis*, the two-body part is the Fourier transform of $1/|\mathbf{r}_m - \mathbf{r}_n|$ [23]

$$\hat{H}_2 \rightarrow \frac{1}{2\Omega} \sum_{i,j,k,l} \frac{4\pi}{|\mathbf{G}_i - \mathbf{G}_k|^2} c_i^\dagger c_j^\dagger c_l c_k \delta_{\mathbf{G}_i - \mathbf{G}_k, \mathbf{G}_l - \mathbf{G}_j} \delta_{\sigma_i, \sigma_k} \delta_{\sigma_j, \sigma_l}, \quad (5)$$

where $\{\mathbf{G}_i\}$ are planewave wave-vectors, Ω is the volume of the supercell, and σ denotes spin. Let us use $\mathbf{Q} \equiv \mathbf{G}_i - \mathbf{G}_k$, and define a density operator in momentum space

$$\hat{\rho}(\mathbf{Q}) \equiv \sum_{\mathbf{G}, \sigma} c_{\mathbf{G}+\mathbf{Q}, \sigma}^\dagger c_{\mathbf{G}, \sigma}, \quad (6)$$

where the sum is over all $\mathbf{G}$ vectors which allow both $\mathbf{G}$ and $\mathbf{G} + \mathbf{Q}$ to fall within the pre-defined kinetic energy cutoff, E_{cut} , in the planewave basis. The two-body term in Eq. (5) can then be manipulated into the form

$$\hat{H}_2 \rightarrow \sum_{\mathbf{Q} \neq 0} \frac{\pi}{\Omega Q^2} [\hat{\rho}^\dagger(\mathbf{Q}) \hat{\rho}(\mathbf{Q}) + \hat{\rho}(\mathbf{Q}) \hat{\rho}^\dagger(\mathbf{Q})], \quad (7)$$

where the sum is over all $\mathbf{Q}$'s except $\mathbf{Q} = 0$, since in Eq. (5) the $\mathbf{G}_i = \mathbf{G}_k$ term is excluded due to charge neutrality, and we have invoked $\rho^\dagger(\mathbf{Q}) = \rho(-\mathbf{Q})$. By making linear combinations of $[(\rho^\dagger(\mathbf{Q}) + \rho(\mathbf{Q}))]$ and $[(\rho^\dagger(\mathbf{Q}) - \rho(\mathbf{Q}))]$ terms, we can then readily write the right-hand side in Eq. (7) in the desired square form of Eq. (4) [23].

With a *general basis* (e.g., Gaussians basis as standard in chemistry), the most straightforward way to decompose $\hat{H}_2$ is through a direct diagonalization [20, 28, 2]. However, this is computationally costly. A modified Cholesky decomposition leads to $\mathcal{O}(N)$ fields [24, 21]. This approach, which has been commonly used in AFQMC for molecular systems with Gaussian basis sets and for downfolded Hamiltonians [29], realizes the following

$$V_{ijkl} = V_{(i,l),(j,k)} = V_{\mu\nu} \doteq \sum_{\gamma=1}^{N_\gamma} L_\mu^\gamma L_\nu^\gamma, \quad (8)$$

where $\mu = (i, l)$ and $\nu = (j, k)$ are composite indices introduced for convenience. The process is carried out recursively using a modified Cholesky algorithm [30, 31]. Recently an alternative with density-fitting has also been used [32]. The 4-index matrix elements can be represented to a pre-defined accuracy tolerance δ (which for molecular calculations is typical of δ range between 10^{-4} and 10^{-6} in atomic units [21]), such that

$$\hat{H}_2 \rightarrow \frac{1}{2} \sum_{\gamma=1}^{N_{\text{CD}}} \left(\sum_{il} L_{\mu(i,l)}^\gamma c_i^\dagger c_l \right) \left(\sum_{jk} L_{\nu(j,k)}^\gamma c_j^\dagger c_k \right) + \mathcal{O}(\delta). \quad (9)$$

Hence the form in Eq. (4) is realized, with $\hat{v}_\gamma = \sum_{il} L_{\mu(i,l)}^\gamma c_i^\dagger c_l$. The process decomposing the two-body interaction is illustrated in Fig. 1.

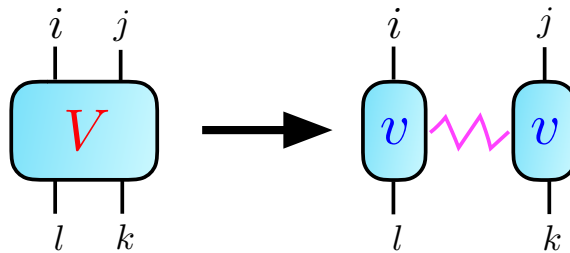


Fig. 1: Schematic illustration of the decoupling of the two-body interaction, either via Cholesky decomposition, the planewave factorization, or density fitting. The number of auxiliary index γ (magenta wiggly line) controls the number of auxiliary fields.

We can then apply the Hubbard-Stratonovich (HS) transformation to each term in Eq. (4)

$$e^{-\frac{\Delta\tau}{2} \lambda \hat{v}^2} = \int_{-\infty}^{\infty} dx \frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}} e^{x\sqrt{-\Delta\tau\lambda} \hat{v}}, \quad (10)$$

where x is an auxiliary-field variable. The key idea is that the quadratic form (in $\hat{v}$) on the left is replaced by a linear one on the right. If we denote the collection of auxiliary fields by $\mathbf{x}$ and combine one-body terms from $\hat{H}_1$ and $\hat{H}_2$, we obtain the following compact representation of the outcome of the HS transformation

$$e^{-\Delta\tau\hat{H}} = \int d\mathbf{x} p(\mathbf{x}) \hat{B}(\mathbf{x}), \quad (11)$$

where $p(\mathbf{x})$ is a probability density function (PDF), for example, a multi-dimensional Gaussian. The propagator $\hat{B}(\mathbf{x})$ in Eq. (11) has a special form, namely, a product of operators of the type

$$\hat{B} = \exp\left(\sum_{ij} c_i^\dagger U_{ij} c_j\right), \quad (12)$$

with U_{ij} depending on the auxiliary field. The matrix representation of $\hat{B}(\mathbf{x})$ will be denoted by $B(\mathbf{x})$. See Appendix B for more details.

Note that the constant in front of $\hat{v}$ in the exponent on the right-hand side of Eq. (10) can be real or imaginary. So the matrix elements of $B(\mathbf{x})$ can become complex, for example when λ in Eq. (10) is positive, which occurs in both of the forms discussed above. Sometimes we will refer to this situation as having complex auxiliary fields, but it should be understood that this is interchangeable with $\hat{B}(\mathbf{x})$ being complex, and the relevant point is whether the Slater determinant develops complex matrix elements which evolve stochastically.

In essence, the HS transformation replaces the two-body interaction by one-body interactions with a set of random external auxiliary fields. In other words, it converts an interacting system into many *non-interacting* systems living in fluctuating external auxiliary-fields. The sum over all configurations of auxiliary fields recovers the interaction.

Different forms of the HS transformation can affect the performance of the AFQMC method. For example, it is useful to subtract a mean-field “background” from the two-body term prior to the decomposition [33, 34, 20]. The idea is that using the HS to decompose any constant shifts in the two-body interaction will necessarily result in more statistical noise. In fact, it has been shown [35, 21] that the mean-field background subtraction can not only impact the statistical accuracy, but also lead to different quality of approximations under the CP methods that we discuss in the next section.

2.2 Ground-state projection

Most ground-state quantum MC (QMC) methods are based on iterative projection

$$|\Psi_0\rangle \propto \lim_{\tau \rightarrow \infty} e^{-\tau\hat{H}} |\Psi_T\rangle; \quad (13)$$

that is, the ground state $|\Psi_0\rangle$ of a many-body Hamiltonian $\hat{H}$ can be projected from any known trial state $|\Psi_T\rangle$ that satisfies $\langle\Psi_T|\Psi_0\rangle \neq 0$. This can be achieved by numerical iteration

$$|\Psi^{(n+1)}\rangle = e^{-\Delta\tau\hat{H}}|\Psi^{(n)}\rangle, \quad (14)$$

where $|\Psi^{(0)}\rangle = |\Psi_T\rangle$. The ground-state expectation $\langle\hat{O}\rangle$ of a physical observable $\hat{O}$ is given by

$$\langle\hat{O}\rangle = \lim_{n \rightarrow \infty} \frac{\langle\Psi^{(n)}|\hat{O}|\Psi^{(n)}\rangle}{\langle\Psi^{(n)}|\Psi^{(n)}\rangle}. \quad (15)$$

For example, the ground-state energy can be obtained by letting $\hat{O} = \hat{H}$. A so-called mixed estimator exists, however, which is exact for the energy (or any other $\hat{O}$ that commutes with $\hat{H}$) and can lead to considerable simplifications in practice

$$E_0 = \lim_{n \rightarrow \infty} \frac{\langle\Psi_T|\hat{H}|\Psi^{(n)}\rangle}{\langle\Psi_T|\Psi^{(n)}\rangle}. \quad (16)$$

QMC methods carry out the iteration in Eq. (14) by Monte Carlo (MC) sampling. At the simplest conceptual level, one can understand the difference between different classes of QMC methods as what space is used to represent the wave function or density matrix and to carry out the integration. The AFQMC methods work in second quantized representation and in a non-orthogonal space of Slater determinants. Ground-state AFQMC represents the many-body wave function stochastically in the form

$$|\Psi^{(n)}\rangle = \sum_{\phi} \alpha_{\phi} |\phi\rangle, \quad (17)$$

where $|\phi\rangle$ is a Slater determinant

$$|\phi\rangle \equiv \hat{\phi}_1^{\dagger} \hat{\phi}_2^{\dagger} \cdots \hat{\phi}_M^{\dagger} |0\rangle. \quad (18)$$

The Slater determinants evolve with n via rotations of the orbitals, as do their coefficients, which are represented by the weights in the MC sampling. The $T > 0$ AFQMC formalism, as we will see later, is closely related and works in the same space.

From Thouless' theorem, the operation of $e^{-\tau\hat{H}_1}$ on a Slater determinant simply yields another determinant. Thus for an independent-particle Hamiltonian, where $\hat{H}_2$ is replaced by a one-body operator, the ground-state projection would therefore turn into the propagation of a single Slater determinant. In DFT, for example, under the local density approximation (LDA), $\hat{H}_2$ is replaced by $\hat{H}_{\text{LDA}} = \hat{H}_1 + \hat{V}_{xc}$, where $\hat{V}_{xc}$ contains the density operator in real-space, with matrix elements given by the exchange-correlation functional which is computed with the local density from the current Slater determinant in the self-consistent process. An iterative procedure can be carried out, following Eq. (14), to project the solution using the approximate Hamiltonians, as an imaginary-time evolution of a single Slater determinant [36]. This is illustrated by the blue line in Fig. 2. Note that this procedure is formally very similar to time-dependent DFT (TDDFT), except for the distinction of imaginary versus real time.

With Eq. (11), we can now turn the many-body projection into a linear combination of iterative projections, mimicking the evolution of an ensemble of the corresponding non-interacting systems subjected to fluctuating external (auxiliary) fields. For simplicity, we will take $|\Psi^{(0)}\rangle$ (i.e., $|\Psi_T\rangle$) as a single Slater determinant. Using the mixed estimator in Eq. (16), we can visualize the calculation of the energy as carrying out the projection of the ket $|\Psi^{(n)}\rangle$ by an open-ended random walk. Aside from technical steps such as importance sampling, population control, numerical stabilization, etc. to make this process practical and more efficient [19, 2], we could think of the calculation at the conceptual level as: (i) start a population of walkers $\{|\phi_k^{(0)}\rangle\}$ with $k = 1, \dots, N_w$ labeling the walkers, (ii) sample an $\mathbf{x}$ from $p(\mathbf{x})$ for each walker, (iii) propagate it by $B(\mathbf{x})$, (iv) sweep through the population to advance to $n=1$, and repeat steps (ii) and (iii) to iterate n . The ideas are illustrated in Fig. 2. [Note that, if we use a linear combination of Slater determinants of the form of Eq. (17) for $|\Psi^{(0)}\rangle$, we can sample the determinants to initialize the population in (i).] In Appendix A we include a brief review of MC and random walks.

We now expand on the formalism a bit more to make a more explicit connection with the $T > 0$ formalism that we will discuss in Sec. 2.3. There we will look more deeply into the origin of the sign problem. By making the formal connection, the discussion on the sign problem at finite- T can be straightforwardly connected to the ground-state situation here. For computing

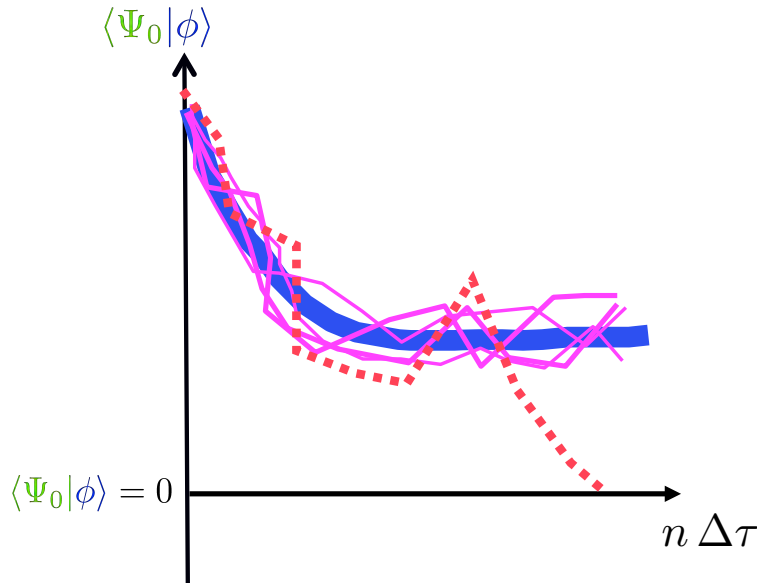


Fig. 2: Illustration of the iterative imaginary-time projection to the ground state. The overlap of the Slater determinants with a test wave function (e.g., the exact ground state $|\Psi_0\rangle$) is plotted vs. imaginary time $n\Delta\tau$. The thick blue line indicates a projection using $e^{-\Delta\tau\hat{H}_{\text{LDA}}(\phi^{(n)})}$ which converges to the LDA ground state (or a local minimum). The wiggly magenta lines indicate an AFQMC projection which captures the many-body effect beyond LDA as a stochastic linear superposition. The propagator is obtained by expanding the two-body part of the $\hat{H}$, namely $\hat{H}_2 - \hat{V}_{xc}$, by a Hubbard-Stratonovich transformation as discussed in Sec. 2.1. The dotted redline indicates a path which can lead to a sign problem (Sec. 3.1).

the ground-state expectation $\langle \hat{O} \rangle$ in Eq. (15), the denominator is

$$\begin{aligned} \langle \psi^{(0)} | e^{-n\Delta\tau\hat{H}} e^{-n\Delta\tau\hat{H}} | \psi^{(0)} \rangle &= \int \langle \psi^{(0)} | \prod_{l=1}^{2n} d\mathbf{x}^{(l)} p(\mathbf{x}^{(l)}) \hat{B}(\mathbf{x}^{(l)}) | \psi^{(0)} \rangle \\ &= \int \prod_l d\mathbf{x}^{(l)} p(\mathbf{x}^{(l)}) \det \left([\Psi^{(0)}]^\dagger \prod_l B(\mathbf{x}^{(l)}) \Psi^{(0)} \right). \end{aligned} \quad (19)$$

In the early forms of ground-state AFQMC calculations [13, 37] (which in the presence of a sign problem are transient estimates that have exponential scaling), a value of n is first chosen and fixed throughout the calculation. If we use X to denote the collection of the auxiliary-fields $X = \{\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(2n)}\}$ and $D(X)$ to represent the integrand in Eq. (19), we can write an estimator of the expectation value of Eq. (15) as

$$\langle \hat{O} \rangle = \frac{\int \overline{\langle \hat{O} \rangle} D(X) dX}{\int D(X) dX} = \frac{\int \overline{\langle \hat{O} \rangle} |D(X)| s(X) dX}{\int |D(X)| s(X) dX}, \quad (20)$$

where

$$s(X) \equiv D(X)/|D(X)| \quad (21)$$

measures the “sign” of $D(X)$, and $\langle s \rangle = \langle s(X) \rangle_{|D|}$ gives an indication of the severity of the sign problem, as further discussed in Sec. 3.1 under the $T > 0$ formalism. The non-interacting expectation for a given X is $\langle \hat{O} \rangle \equiv \langle \hat{O} \rangle_{\phi_L \phi_R}$ as defined in Eq. (55) in Appendix B, where

$$\begin{aligned} \langle \phi_L | &= \langle \psi^{(0)} | \hat{B}(\mathbf{x}^{(2n)}) \hat{B}(\mathbf{x}^{(2n-1)}) \dots \hat{B}(\mathbf{x}^{(n+1)}) \\ | \phi_R \rangle &= \hat{B}(\mathbf{x}^{(n)}) \hat{B}(\mathbf{x}^{(n-1)}) \dots \hat{B}(\mathbf{x}^{(1)}) | \psi^{(0)} \rangle, \end{aligned}$$

which are both Slater determinants. In Appendix B, basic properties of Slater determinants and the computation of expectation values are reviewed.

$D(X)$ as well as $\langle \phi_L |$ and $| \phi_R \rangle$ are completely determined by the path X in auxiliary-field space. The expectation in Eq. (20) is therefore in the form of Eq. (49), with $f(X) = |D(X)|$ and $g(X) = \langle \hat{O} \rangle$. The important point is that, for each X , $|D(X)|$ is a number and $g(X)$ can be evaluated using Eqs. (56) and (57). Often the Metropolis Monte Carlo algorithm [38] is used to sample auxiliary-fields X from $|D(X)|$. Any $\langle \hat{O} \rangle$ can then be computed following the procedure described by Eq. (48) in Appendix A.

2.3 Finite-temperature, grand-canonical ensemble calculations

At a temperature $T > 0$ K, the expectation value of any physical observable

$$\langle \hat{O} \rangle \equiv \frac{\text{Tr}(\hat{O} e^{-\beta\hat{H}})}{\text{Tr}(e^{-\beta\hat{H}})}, \quad (22)$$

which is a weighted average in terms of the partition function Z in the denominator

$$Z \equiv \text{Tr}(e^{-\beta\hat{H}}) = \text{Tr}[\underbrace{e^{-\Delta\tau\hat{H}} \dots e^{-\Delta\tau\hat{H}} e^{-\Delta\tau\hat{H}}}_L], \quad (23)$$

where $\beta = 1/kT$ is the inverse temperature, $\Delta\tau = \beta/L$ is the time-step, and L is the number of “time slices.” Substituting Eq. (11) into Eq. (23) gives

$$Z = \sum_X \int dX P(X) \text{Tr}[B(\mathbf{x}_L) \cdots B(\mathbf{x}_2)B(\mathbf{x}_1)], \quad (24)$$

where $X \equiv \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_L\}$ denotes a complete *path* in auxiliary-field space, and $P(X) = \prod_{l=1}^L p(\mathbf{x}_l)$. Note that, similar to the $T = 0$ K case, a Trotter error can be present in Eq. (11), which is controllable (by, e.g., extrapolating results with different values of $\Delta\tau$). The trace over fermion degrees of freedom can be performed analytically [12, 39], which yields

$$Z = \int dX P(X) \det[I + B(\mathbf{x}_L) \cdots B(\mathbf{x}_2)B(\mathbf{x}_1)], \quad (25)$$

where I is the $N \times N$ unit matrix. Note that the trace in Eq. (24), which is over numbers of particles and initial states, is now replaced by the fermion determinant

$$D(X) \equiv P(X) \det[I + B(\mathbf{x}_L) \cdots B(\mathbf{x}_2)B(\mathbf{x}_1)], \quad (26)$$

which can be readily computed for each path X . The sum over all paths is evaluated by standard Monte Carlo techniques. Relating L to $2n$ in the ground-state discussions around Eq. (20), we can now connect these two classes of methods. Below we will rely on the finite- T form in understanding the origin of the sign problem. This will serve as a common framework for thinking about the sign problem, and then subsequently the phase problem, by analyzing the nature and behavior of the paths in X space.

The symptom of the sign problem is that $D(X)$ is not always positive. In practice, the MC samples of X are drawn from the probability distribution defined by $|D(X)|$. As β increases, $D(X)$ approaches an antisymmetric function and its average sign,

$$\langle s \rangle = \sum_X D(X) / \sum_X |D(X)|, \quad (27)$$

vanishes exponentially, as illustrated in Fig. 3. In the limit of $T \rightarrow 0$ or $\beta \rightarrow \infty$, the distribution becomes fully antisymmetric. (In ground-state calculations discussed in Sec. 2.2, the situation depicted in Fig. 3 corresponds to the equilibration phase at shorter projection times τ . When the ground state is reached after equilibration, there will be no “green” part left.)

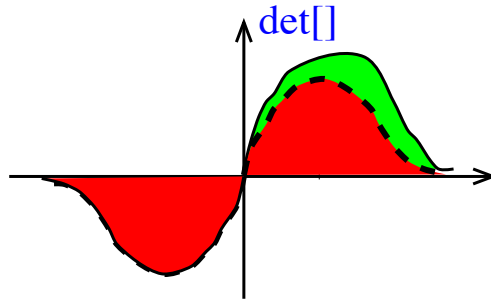


Fig. 3: Schematic illustration of the sign problem. The horizontal axis denotes an abstraction of the many-dimensional auxiliary-field paths X . The sign problem occurs because the contributing part (shaded area) is exponentially small compared to what is sampled, namely $|D(X)|$. The origin of the symptoms shown here is explained in Sec. 3.1.

3 Constraining the paths in AFQMC

The sign/phase problem is a generic feature, and is only absent in special cases where the single-particle propagator $\hat{B}(\mathbf{x})$ satisfies particular symmetries (see, for example, Ref. [40]). In these cases, $D(x)$ is real and non-negative, and $s(X)$ remains strictly 1. The symmetries are affected by the choice of the basis and the form of the decomposition, so it is possible to “engineer away” the sign problem in some cases. However, these are still limited to very special classes of Hamiltonians. In general, a sign problem arises if $\hat{B}(\mathbf{x})$ is real, and a phase problem arises if $\hat{B}(\mathbf{x})$ is complex.

For real $\hat{B}(\mathbf{x})$ (e.g. Hubbard-type of short-range repulsive interactions with the spin decomposition), the sign problem occurs because of the fundamental symmetry between the fermion ground-state $|\Psi_0\rangle$ and its negative $-|\Psi_0\rangle$ [18, 41]. In $T = 0$ K calculations, for any ensemble of Slater determinants $\{|\phi\rangle\}$ which gives a MC representation of the ground-state wave function, as in Eq. (17), this symmetry implies that there exists another ensemble $\{-|\phi\rangle\}$ which is also a correct representation. In other words, the Slater determinant space can be divided into two degenerate halves (+ and −) whose bounding surface $\mathcal{N}$ is defined by $\langle\Psi_0|\phi\rangle = 0$. This dividing surface is unknown. (In the cases with special symmetry mentioned above, the two sides separated by the surface are both positive.) In the illustration in Fig. 2, the surface is the horizontal axis; the “−” ensemble is given by the mirror images of the paths shown, i.e., by reflecting them with respect to the horizontal axis.

3.1 An exact boundary condition

To gain further insight on the origin of the sign problem, we conduct a thought experiment in which we generate all the complete paths X by L successive steps, from $\mathbf{x}_1$ to $\mathbf{x}_L$ [25]. We use the $T > 0$ K formalism, whose understanding will provide a straightforward connection to the ground-state method. We consider the contribution in Z of an individual *partial path* $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_l\}$ at step l

$$\mathcal{P}_l(\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_l\}) \equiv \text{Tr}[\underbrace{\mathcal{B}\mathcal{B}\cdots\mathcal{B}}_{L-l} B(\mathbf{x}_l) \cdots B(\mathbf{x}_2)B(\mathbf{x}_1)], \quad (28)$$

where $\mathcal{B} \equiv e^{-\Delta\tau H}$, which in general is not a single-particle propagator. In particular, we consider the case when $\mathcal{P}_l = 0$. This means that, after the remaining $L - l$ steps are finished, the collective contribution from *all* complete paths that result from the partial path will be precisely zero. In other words, complete paths whose first l elements are $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_l\}$ do not contribute in Z ; the sum over all possible configurations of $\{\mathbf{x}_{l+1}, \mathbf{x}_{l+2}, \dots, \mathbf{x}_L\}$ simply reproduces the $\mathcal{B}$ ’s in (28), leading to zero by definition.

Thus, in our thought experiment any partial path that reaches the axis in Fig. 4 immediately turns into noise, regardless of what it does at future l ’s. A complete path which is in contact with the axis at any point belongs to the “antisymmetric” part of $D(X)$ in Fig. 3, whose contributions cancel. The “noise” paths, which become an increasingly larger portion of all paths as β increases, are the origin of the sign problem.

Since $\mathcal{P}_0$ is positive and $\mathcal{P}_l$ changes continuously with l at the limit $\Delta\tau \rightarrow 0$, a complete path contributes **iff** it stays entirely *above* the axis in Fig. 4. Thus, in our thought experiment, imposition of the boundary condition (BC)

$$\mathcal{P}_1(\{\mathbf{x}_1\}) > 0, \quad \mathcal{P}_2(\{\mathbf{x}_1, \mathbf{x}_2\}) > 0, \quad \dots, \quad \mathcal{P}_L(\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_L\}) > 0 \quad (29)$$

will ensure all contributing complete paths to be selected while eliminating all noise paths. *The axis acts as an infinitely absorbing boundary.* A partial path is terminated and discarded as soon as it reaches the boundary. By discarding a path, we eliminate all of its future paths, including the ones that would eventually make positive contributions. The BC makes the distribution of complete paths vanish at the axis, which accomplishes complete cancellation of the negative and the corresponding positive contributions in the antisymmetric part of $D(X)$. Calculation of Z from our thought experiment remains exact.

3.2 Implementing the boundary condition and importance sampling

3.2.1 Implementation of the boundary condition (BC) at finite $\Delta\tau$

In actual simulations $\Delta\tau$ is finite and paths are defined only at a discrete set of imaginary times. The BC on the underlying continuous paths is the same, namely that the probability distribution must vanish at the axis in Fig. 4.

In Fig. 5, we illustrate how the BC is imposed under the discrete representation. The “contact” point is likely to be between time slices and not well defined, i.e., $\mathcal{P}_l$ may be zero at a non-integer value of l . To the lowest order, we can terminate a partial path when its $\mathcal{P}_l$ first turns negative. That is, we still impose Eq. (29) in our thought experiment to generate paths. In Fig. 5, this means terminating the path at $l = n$ (point B) and thereby discarding all its future paths (represented by the dashed lines ‘BS...’ and ‘BT...’).

We actually use a higher order approximation, by terminating at either $l = n - 1$ or $l = n$, i.e., either point B or point A. The probability for terminating at A is chosen such that it approaches 1 smoothly as $\mathcal{P}_{n-1} \rightarrow 0$, for example, $p_A = 1/[1 + \mathcal{P}_{n-1}/|\mathcal{P}_n|]$. If A is chosen, all future

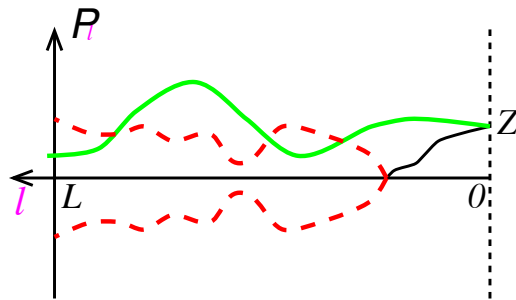


Fig. 4: Schematic illustration the boundary condition to control the sign problem. $\mathcal{P}_l$ (Eq. (28)) is shown as a function of the length of the partial path, l , for several paths. When $\mathcal{P}_l$ becomes 0, ensuing paths (dashed lines) collectively contribute zero. Only complete paths with $\mathcal{P}_l > 0$ for all l (solid line) contribute in Z .

paths from A are discarded (represented by ‘AR...’ and ‘AB...’); otherwise we terminate at B as above. This is in the spirit of the so-called mirror correction (see, e.g., Ref. [42, 19]).

It is important to note that, in both approaches, the finite- $\Delta\tau$ error in imposing the BC vanishes as $\Delta\tau \rightarrow 0$. The first method, terminating at B, makes a fluctuating error of $\mathcal{O}(\Delta\tau)$ in the location of the absorbing boundary. The method we actually use ensures that the location of the boundary is correct to the next order, and is a second order algorithm.

3.2.2 Approximating the boundary condition

In practice $\mathcal{B}$ is not known. We replace it by a known trial propagator B_T . The BC now yields approximate results, which become exact if B_T is exact. If B_T is in the form of a single-particle propagator, we can evaluate Eq. (28) and combine it with Eq. (29) to obtain the approximate BC: For each l ,

$$\mathcal{P}_l^T(\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_l\}) = \det[I + (\prod_{m=1}^{L-l} B_T) B(\mathbf{x}_l) \cdots B(\mathbf{x}_1)] > 0. \quad (30)$$

Different forms of $\mathcal{B}$ are possible, and of course it is valuable and important to think about improving $\mathcal{B}$. Recently self-consistent constraints formulated in ground-state AFQMC [43] have been generalized to the $T > 0$ method [4].

3.2.3 Importance sampling algorithm—automatic imposition of the constraint and “nudge” in the random walks

(1) *Automatic imposition of the constraint.* To implement the constraint in a path-integral framework of Eq. (20) at $T = 0$ K and Eq. (25) at $T > 0$ K would have severe difficulties with ergodicity. We wish to generate MC samples of X which satisfy the conditions in (30). The most natural way to accomplish this is to incorporate the boundary conditions on the path as an additional acceptance condition [44]. However, the BC is *non-local*; it also breaks translational invariance in imaginary time, since the condition selects an $l = 1$ to start. Updating the auxiliary-fields at a particular time l can cause violation of the constraint in the future (at

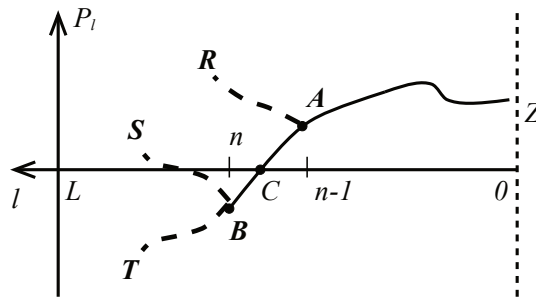


Fig. 5: Imposition of the boundary condition at finite $\Delta\tau$. Paths are discrete. The point of contact, C (see Fig. 4), must be approximated, either by B (low order algorithm) or by interpolation between B and A (higher order).

larger l) or in the past (when sweeping the path backwards in l). Without a scheme to propose paths that incorporates information on future contributions, it is difficult to find complete paths which satisfy all the constraints, especially as β increases. Global updating is difficult, because the number of green paths is exponentially small compared to the total number of possible paths, as illustrated in Fig. 4. The formulation of the open-ended, branching random walk approach [19, 25] in imaginary-time solves this problem. An additional advantage is that it is straightforward to project to longer imaginary-time in order to approach the ground state for $T = 0$ K. Moreover, when we carry out constraint release [35], the formalism will rely on the open-ended random walk.

(2) *Nudging the random walk.* The goal of the branching random walk algorithm is to generate MC samples of X which *both* satisfy the conditions in (30) *and* are distributed according to $D(X)$. The basic idea is to carry out the thought experiment stochastically. We construct an algorithm which builds directly into the sampling process both the constraints and some knowledge of the projected contribution. In terms of the *trial* projected partial contributions $\mathcal{P}_l^T \equiv \mathcal{P}_l^T(\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_l\})$ in the $T > 0$ K method, the fermion determinant $D(X)$ can be written as

$$D(X) = \frac{\mathcal{P}_L^T}{\mathcal{P}_{L-1}^T} \frac{\mathcal{P}_{L-1}^T}{\mathcal{P}_{L-2}^T} \dots \frac{\mathcal{P}_2^T}{\mathcal{P}_1^T} \frac{\mathcal{P}_1^T}{\mathcal{P}_0^T} \mathcal{P}_0^T. \quad (31)$$

As illustrated in Fig. 6, we construct the path X by a random walk of L steps, corresponding to stochastic representations of the L ratios in Eq. (31). At the l -th step, we sample $\mathbf{x}_l$ from the conditional probability density function defined by $(\mathcal{P}_l^T / \mathcal{P}_{l-1}^T)$. This allows us to select $\mathbf{x}_l$ according to the best estimate of its projected contribution in Z . In other words, we sample from a probability distribution function of the contributing paths only (solid lines in Fig. 4), instead of *all* possible paths. Note that the probability distribution for $\mathbf{x}_l$ vanishes smoothly as $\mathcal{P}_l^T$ approaches zero, naturally imposing the BC in Eq. (30) as anticipated in part (1) of this section. Ref. [4] contains a more detailed outline of the procedure.

(3) *Connecting the importance sampling algorithms between $T = 0$ and $T > 0$ K methods.* In the above we have used the $T > 0$ K framework to describe the basic idea of importance sampling. At each step, the importance-sampling transformation uses $\mathcal{P}_l^T / \mathcal{P}_{l-1}^T$ to modify the probability distribution from which $\mathbf{x}_l$ is sampled. The ground-state version, conceptually, uses $\langle \Psi_T | \phi^{(n)} \rangle / \langle \Psi_T | \phi^{(n-1)} \rangle$. In Sec. 3.3, we show how this is actually realized when the auxiliary-fields are continuous (complex), by the use of a force bias. To connect back to the $T > 0$ K form, all we need is to invoke the formal equivalence

$$\frac{\langle \Psi_T | \phi'(\mathbf{x}) \rangle}{\langle \Psi_T | \phi \rangle} \Longleftrightarrow \frac{\det[I + (\prod_{m=1}^{L-l} B_T) B(\mathbf{x}) B(\mathbf{x}_{l-1}) \dots B(\mathbf{x}_1)]}{\det[I + (\prod_{m=1}^{L-l+1} B_T) B(\mathbf{x}_{l-1}) \dots B(\mathbf{x}_1)]}. \quad (32)$$

On the right of Eq. (32), the finite- T path segment $\{\mathbf{x}_1, \dots, \mathbf{x}_{l-1}\}$ has been generated already, and the step in question is l , where we wish to generate $\mathbf{x}_l$, denoted by $\mathbf{x}$.

3.3 The phaseless formalism for complex auxiliary-fields

When the many-body Hamiltonian leads to a decomposition with $\lambda > 0$ in Eq. (4), the resulting HS transformation will have complex auxiliary-fields. This is the case for the electron-electron repulsion. (As mentioned earlier, when we refer loosely to having complex auxiliary-fields, what we really mean is that the propagator resulting from the two-body Hamiltonian is complex. Incidentally, it is always possible to have real auxiliary-fields, for example by making a negative shift to the positive potential, but that simply leads to many fluctuating fields to recover a constant background, and a much more severe sign problem [14,35].) In this situation a phase problem arises, as illustrated in Fig. 7. The random walks now extend into a complex plane, and the boundary condition discussed in Sec. 3.1 must be generalized. This generalization is not straightforward, since the orbitals (or one-particle propagators) contain random phases which are entangled with the phases in the coefficients in front of the determinants. We next describe the formalism to deal with the problem [14], using the $T = 0$ K framework. See Eq. (32) in the previous section for how to “translate” this to $T > 0$ K.

With a continuous auxiliary-field, importance sampling is achieved with a force bias [14,45]. To sketch a derivation we write the full many-body propagator as

$$\int \left(\frac{1}{\sqrt{2\pi}} \right)^{N_\gamma} e^{-\mathbf{x}^2/2} B(\mathbf{x}) d\mathbf{x}, \quad (33)$$

where for formal convenience we will assume that we can combine products of one-body propagators, as in Eq. (12), into a joint form by summing the one-body operators in the exponent, or vice versa, possibly incurring additional Trotter errors which will be dealt with separately. For

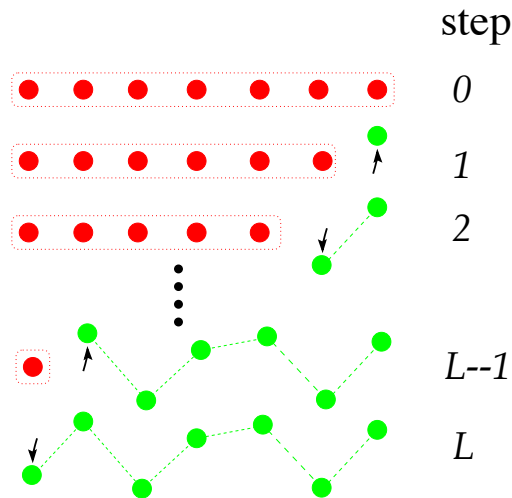


Fig. 6: Illustration of the sampling procedure in the algorithm. Circles represent auxiliary-fields $\mathbf{x}_l$. A row shows the field configuration at the corresponding step number shown on the right. Within each row, the imaginary-time index l increases as we move to the left, i.e., the first circle is $\mathbf{x}_1$ and the last $\mathbf{x}_L$. Red circles indicate fields which are not “activated” yet, i.e., B_T is still in place of B . Green circles indicate fields that have been sampled, with the arrow indicating the one being sampled in the current step.

compactness we will also omit the normalization factor, $(1/\sqrt{2\pi})^{N_\gamma}$, of the Gaussian probability density function from here on.

We introduce a shift in the integral in Eq. (33), which leads to an alternative propagator

$$\int e^{-\mathbf{x}^2/2} e^{\mathbf{x} \cdot \bar{\mathbf{x}} - \bar{\mathbf{x}}^2/2} B(\mathbf{x} - \bar{\mathbf{x}}) d\mathbf{x}. \quad (34)$$

The new propagator is exact for any choice of the shift $\bar{\mathbf{x}}$, which can be complex in general.

We recall that the random walk is supposed to sample the coefficient α_ϕ in Eq. (17)

$$|\Psi_0\rangle \doteq \sum_{\{\phi\}} w_\phi |\phi\rangle. \quad (35)$$

The sum in Eq. (35) is over the population of walkers after equilibration and is over the Monte Carlo samples, typically much smaller than the sum in Eq. (17). The weight of each walker $|\phi\rangle$, w_ϕ , can be thought of as 1 (all walkers with equal weight); it is allowed to fluctuate only for efficiency considerations.

Using the idea of importance sampling, we seek to replace Eq. (35) by the following to sample Eq. (17):

$$|\Psi_0\rangle = \sum_{\phi} w_\phi \frac{|\phi\rangle}{\langle\Psi_T|\phi\rangle}, \quad (36)$$

where any overall phase of the walker $|\phi\rangle$ is cancelled in the numerator and denominator on the right-hand side [14]. This implies a modification to the propagator in Eq. (34):

$$\int \langle\Psi_T|\phi'(\mathbf{x})\rangle e^{-\mathbf{x}^2/2} e^{\mathbf{x} \cdot \bar{\mathbf{x}} - \bar{\mathbf{x}}^2/2} B(\mathbf{x} - \bar{\mathbf{x}}) \frac{1}{\langle\Psi_T|\phi\rangle} d\mathbf{x}, \quad (37)$$

where $|\phi'(\mathbf{x})\rangle = B(\mathbf{x} - \bar{\mathbf{x}})|\phi\rangle$ and the trial wave function $|\Psi_T\rangle$ represents the best guess to $|\Psi_0\rangle$. Now the weight of the waker under importance sampling becomes

$$w_{\phi'(\mathbf{x})} = w_\phi \frac{\langle\Psi_T|\phi'(\mathbf{x})\rangle}{\langle\Psi_T|\phi\rangle} e^{\mathbf{x} \cdot \bar{\mathbf{x}} - \bar{\mathbf{x}}^2/2}. \quad (38)$$

We can minimize the fluctuation of the factor on the right-hand side with respect to $\mathbf{x}$, by evaluating the ratio $\langle\Psi_T|\phi'(\mathbf{x})\rangle/\langle\Psi_T|\phi\rangle$ in Eqs. (37) and (38), or correspondingly, the ratio on the right-hand side of Eq. (32) for $T > 0$ K. Expanding the propagators in $\Delta\tau$, and rearranging terms [1, 21], we obtain the optimal choice for the force bias

$$\bar{\mathbf{x}} = \bar{\mathbf{v}} \equiv -\frac{\langle\Psi_T|\hat{\mathbf{v}}|\phi\rangle}{\langle\Psi_T|\phi\rangle} \sim \mathcal{O}(\sqrt{\Delta\tau}) \quad (39)$$

for ground state and a similar expectation value $\bar{\mathbf{v}}$ evaluated at time-step l for $T > 0$. The weight factor in Eq. (38) can then be manipulated into a more compact form

$$w_{\phi'(\mathbf{x})} = w_\phi \exp(-\Delta\tau E_L(\phi)), \quad (40)$$

where E_L is the local energy of the Slater determinant

$$E_L(\phi) \equiv \frac{\langle\Psi_T|\hat{H}|\phi\rangle}{\langle\Psi_T|\phi\rangle}. \quad (41)$$

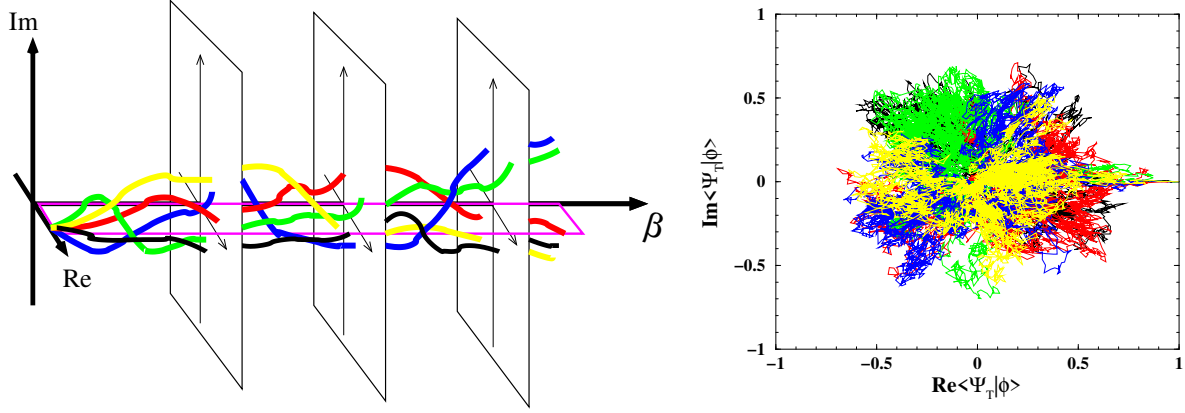


Fig. 7: Schematic illustration of the phase problem and the constraint to control it, using the ground-state formalism. The left panel shows, as a function of projection time $\beta \equiv n\Delta\tau$, trajectories of 5 walkers characterized by the real (Re) and imaginary (Im) parts of their overlap with the ground-state wave function. The right panel shows the walker distribution integrated over imaginary time, i.e., the different frames in the left panel stacked together along β . The phase problem occurs because the initial phase “coherence” of the random walkers rapidly deteriorates with β , as they become uniformly distributed in the Re-Im-plane. The idea of the phase constraint [14] is to apply a gauge transformation such that confining the random walk in the single magenta plane (left) is a good approximation.

We now have the full propagator of Eq. (37)

$$\int e^{-\mathbf{x}^2/2} B(\mathbf{x} - \bar{\mathbf{v}}) \exp(-\Delta\tau E_L(\phi)) d\mathbf{x}. \quad (42)$$

Projection with Eq. (42) will in principle lead to the ground-state wave function in the form of Eq. (36). The weight of the walker is determined by E_L , which is independent of any phase factor of the determinant. We will refer to this as the local energy form of AFQMC. As an alternative, referred to as the hybrid form, we could evaluate the weight of each walker directly according to Eq. (38) after the propagation. (In this form $\bar{\mathbf{x}}$ can in principle be anything, but of course poor choices will lead to larger fluctuations in the weights.)

In the limit of an exact $|\Psi_T\rangle$, E_L is a *real* constant, and the weight of each walker remains real. The mixed estimate for the energy from Eq. (16) is phaseless

$$E_0^c = \frac{\sum_{\phi} w_{\phi} E_L(\phi)}{\sum_{\phi} w_{\phi}}. \quad (43)$$

With a general $|\Psi_T\rangle$ which is not exact, a natural approximation is to replace E_L in Eq. (42) by its real part, $\text{Re } E_L$. The same replacement is then necessary in Eq. (43).

In early tests in the electron gas and in simple solids, we found that the phase of E_L could be carried for very long imaginary-times (and then resetting) without causing any noticeable effect on the phase problem. Keeping the phase also did not appear to affect the total energy computed from the mixed estimator. For computing observables in the ground state, back-propagation (BP) [19, 45] is needed. It was found that restoring the phases of E_L in the BP paths helped

improve the computed observables and correlations [46]. More tests for $T > 0$ are necessary but it is likely that keeping the phase will be preferable (especially because the paths are much shorter than in ground-state calculations). At $T > 0$, it was also found that restoring time-translation symmetry after the complete path has been sampled improves the results [4].

This formalism is all that is needed to handle the sign problem in the case of a *real* $\hat{v}$. For any $\hat{v}$ the shift $\bar{x}$ diverges as a walker approaches the origin in the complex plane shown in the right panel of Fig. 7, i.e., as $\langle \Psi_T | \phi' \rangle \rightarrow 0$. The effect of the divergence is to move the walker away from the origin. With a *real* $\hat{v}$, $\Delta\theta = 0$ and the random walkers move only on the real axis. If they are initialized to have positive overlaps with $|\Psi_T\rangle$, $\bar{x}$ will ensure that the overlaps remain positive throughout the random walk. Thus in this case our formalism above reduces to the CP methods ground state [42, 19] and finite- T [25].

For a general case with a complex $\hat{v}$, however, the phaseless formalism alone is not sufficient to remove the phase problem. To illustrate this we consider the phase of $\langle \Psi_T | \phi'(\mathbf{x} - \bar{\mathbf{x}}) \rangle / \langle \Psi_T | \phi \rangle$, or the equivalent form for $T > 0$ K given by Eq. (32). This phase, which we shall denote by $\Delta\theta$, is in general non-zero: $\Delta\theta \sim \mathcal{O}(-\mathbf{x} \text{Im}(\bar{\mathbf{x}}))$. The walkers will thus undergo a random walk in the complex plane defined by $\langle \Psi_T | \phi' \rangle$. At large β they will therefore populate the complex plane symmetrically, independent of their initial positions. This is illustrated in the right panel of Fig. 7, which shows $\langle \Psi_T | \phi \rangle$ for three-dimensional jellium with two electrons at $r_s = 10$ for a total projection time of $\beta = 250$. The distribution of the walkers is seen to be symmetric about the phase angle, and any signal that the walkers are all real initially (and $\langle \Psi_T | \phi^{(0)} \rangle = 1$) is lost in the statistical noise.

In other words, for a complex $\hat{v}$, the random walk is “rotationally invariant” in the complex plane, and the divergence of $\bar{x}$ is not enough to prevent the build-up of a finite density at the origin. Near the origin the local energy E_L diverges, which causes diverging fluctuations in the weights of walkers. To address this we make an additional approximation. We project the random walk to “one-dimension” and multiply the weight of each walker in each step by $\cos(\Delta\theta)$

$$w_{\phi'} = w_{\phi'} \max\{0, \cos(\Delta\theta)\} \quad (44)$$

in addition to Eq. (40) (or Eq. (38) if the hybrid form is used). This is only a good approximation in the presence of the similarity transformation that we have already preformed in what we have been calling importance sampling. There is a subtle but fundamental distinction between the formalism we have introduced and “traditional” importance sampling. In the latter, only the sampling efficiency is affected, not the expectation, because the transformation involves only real and non-negative functions. Here, in contrast, the functions, as given in Eq. (32), are *complex* and determine a gauge choice for our random walks. The proper choice of the force bias ensures that the leading order in the overall phase of $|\phi\rangle$ in the propagator in Eq. (37) is eliminated.

Several alternatives to Eq. (44) were tested [14, 47, 34]. One that seemed to work as well was $\exp(-(\text{Im}(\bar{\mathbf{x}}))^2/2)$, which is the same as $\cos(\Delta\theta)$ in the limit of small values of $\Delta\theta$. Another was to impose $\text{Re} \langle \Psi_T | \phi' \rangle > 0$, which gave similar results, but with somewhat larger variance.

4 Overview of applications, further reading, and outlook

Due to space limitations, we will not include any examples of applications here. The AFQMC method has been applied to lattice models, ultracold atoms, solids, and molecular systems.

In lattice models, most of the applications involve “only” a sign problem, because of the short-range nature of the interaction, although in multi-band models there can be tradeoffs between decompositions which lead to a sign or a phase problem [48]. A large body of results exist, including recent benchmark results [49]. Systems of $\mathcal{O}(1000)$ electrons have been treated quite routinely. The AFQMC method has demonstrated excellent capabilities and accuracy, illustrating its potential as a general many-body computational paradigm. A key recent development [43] is to use the density or density matrix computed from AFQMC as a feedback into a mean-field calculation. The trial wave function or density matrix obtained from the mean-field is then fed back into the AFQMC as a constraint, and a self-consistent constraining condition is achieved. This has led to further improvement in the accuracy and robustness of the calculation [43, 50]. This development is also seeing more applications in molecules and solids.

A related area involves ultracold atoms, where many valuable calculations have been performed with AFQMC. In addition to the physics advances these calculations have led to, this has proved a fertile test ground for methodological developments, including better sampling methods [51, 52], computation of excitations and dynamical correlations [53], the use of BCS [or antisymmetrized geminal power (AGP)] trial wave functions [54, 55] and projection/path-integral in Hartree-Fock-Bogoliubov (HFB) space [56], treatment of spin-orbit coupling (SOC) [22], Bose systems and Fermi-Bose mixtures [45, 57], and achieving linear scaling in lattice or basis size at $T > 0$ K [58]. Many of these developments have direct applications in correlated systems of molecules and solids.

For molecular systems, a recent review article [21] describes in more detail the application of AFQMC in quantum chemistry. The formulation of AFQMC with Gaussian basis sets [20, 24] has been extremely valuable. Direct comparisons can be made with high-level QC results, which have provided valuable benchmark information and have been crucial in developing the method. Many calculations have been performed using AFQMC as a “post-DFT” or “post-HF” approach for molecules and solids by simply feeding in the solution from standard DFT or HF. Several other systematic benchmarks have appeared, for example on the G1 set [59], on a set of 44 3d transition-metal containing diatomics [60], on singlet-triplet gaps in organic biradicals [61], etc. Major recent methodological advances include the computation of observables [46], geometry optimization [62], and speedups using GPUs [32], low-rank tensor decomposition [63], embedding [64] and localization/downfolding [65], and correlated sampling [66], etc.

For solids, calculations have been done using planewaves and pseudopotentials including recent implementation of multiple-projector ones [67], with downfolding [29], with Gaussian-type orbitals [68], etc. A benchmark study [69] was recently carried out involving a large set of modern many-body methods. AFQMC was found to be comparable in accuracy to CCSD(T), the gold standard in chemistry [70, 71], near equilibrium geometry. For bond breaking, AFQMC was able to maintain systematic accuracy. The AFQMC method can also be used to study

excited states. Excited states distinguished by different symmetry from the ground state can be computed in a manner similar to the ground state. For other excited states, prevention of collapse into the ground state and control of the fermion sign/phase problem are accomplished by a constraint using an excited state trial wave function [72]. An additional orthogonalization constraint is formulated to use virtual orbitals in solids for band structure calculations [73].

The AFQMC method is a computational framework for many-body calculations which combines a field-theoretic description with stochastic sampling. AFQMC has shown strong promise with its scalability (with system size and with parallel computing platforms), capability (total energy computation and beyond), and accuracy. The method is just coming into form, and rapid advances in algorithmic development and in applications are on-going. The literature is growing and I only listed a portion of it above. In addition, there is a pedagogical code for lattice models written in Matlab [74] that will be very useful for getting into the method.

We have discussed both ground-state and $T > 0$ auxiliary-field-based methods for correlated electron systems. We have outlined the formalism in a rather general way which allows for a systematic understanding of the origin of the sign/phase problem as well as the underlying theory for a scalable method capable of handling large molecules and bulk systems.

Often in the study of correlated models in condensed matter or cold atoms in optical lattices, the comment “But there is a sign problem” is made, to imply “so QMC calculations cannot be used here.” I hope that these lectures and the large number of results in the literature with AFQMC will change this mindset. Calculations indeed can be done in systems where a sign problem is present—often with some of the most accurate results that are presently possible.

The AFQMC method has low-polynomial (cubic) scaling with system size, similar to DFT calculations. The structure of the open-ended random walk makes it ideally suited for modern high-performance computing platforms, with exceptional capacity for parallel scaling [15]. The rapid growth of high-performance computing resources will thus provide a strong boost to the application of AFQMC in the study of molecules and solids. The connection with independent-electron calculations, as we have highlighted, makes it straightforward to build AFQMC on top of DFT or HF codes, and take advantage of the many existing technical machineries developed over the past few decades in materials modeling. It also gives AFQMC much versatility as a general many-body method, offering, for example, straightforward treatment of heavier elements and spin-orbit coupling, computation of dynamical correlations, and the capability to embed AFQMC naturally and seamlessly in a calculation at the independent-particle level.

The development of AFQMC is entering an exciting new phase. We expect the method and the concept discussed here to see many applications, and to significantly enhance the capability of quantum simulations in interacting fermion systems. A large number of possible directions can be pursued, including many opportunities for algorithmic improvements and speedups. These will be spurred forward and stimulated by growth in applications, which we hope will in turn allow more rapid realization of a general many-body computational framework for correlated quantum materials.

Acknowledgments

I am indebted to numerous colleagues and outstanding students and postdocs. Length restriction does not allow the full list here, but I would be remiss without thanking Y. He, H. Krakauer, F. Ma, M. Motta, W. Purwanto, M. Qin, J. Shee, H. Shi, and E. Vitali for their contributions. My group has been a part of the Simons Foundation's Many-Electron collaboration, from which we have benefitted greatly. Support from the Department of Energy (DOE) and the National Science Foundation (NSF) is also gratefully acknowledged. Computing was done on the Oak Ridge Leadership Computing Facilities via an INCITE award, and on the HPC facilities at William & Mary. The Flatiron Institute is a division of the Simons Foundation.

Appendices

A A few basics of Monte Carlo techniques

We list a few key elements from standard Monte Carlo (MC) techniques which are important to our discussions on QMC. For an introduction on MC methods, see, e.g., Ref. [38].

MC methods are often used to compute many-dimensional integrals of the form

$$G = \frac{\int_{\Omega_0} f(\mathbf{x})g(\mathbf{x})d\mathbf{x}}{\int_{\Omega_0} f(\mathbf{x})d\mathbf{x}}, \quad (45)$$

where $\mathbf{x}$ is a vector in a many-dimensional space and Ω_0 is a domain in this space. We will assume that $f(\mathbf{x}) \geq 0$ on Ω_0 and that it is normalizable, i.e., the denominator is finite. A familiar example of the integral in Eq. (45) comes from classical statistical physics, where $f(\mathbf{x})$ is the Boltzmann distribution.

To compute G by MC, we *sample* $\mathbf{x}$ from a probability density function (PDF) proportional to $f(\mathbf{x})$: $\bar{f}(\mathbf{x}) \equiv f(\mathbf{x}) / \int_{\Omega_0} f(\mathbf{x})d\mathbf{x}$. This means to generate a sequence $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_i, \dots\}$ so that the probability that any $\mathbf{x}_i$ is in the sub-domain $(\mathbf{x}, \mathbf{x} + d\mathbf{x})$ is

$$\text{Prob}\{\mathbf{x}_i \in (\mathbf{x}, \mathbf{x} + d\mathbf{x})\} = \bar{f}(\mathbf{x})d\mathbf{x}. \quad (46)$$

There are different techniques to sample a many-dimensional function $f(\mathbf{x})$. The most general and perhaps most commonly used technique to sample $f(\mathbf{x})$ (i.e., the PDF $\bar{f}(\mathbf{x})$) is the Metropolis algorithm, which creates a Markov chain random walk [38] in $\mathbf{x}$ -space whose equilibrium distribution is the desired function. We will also use a branching random walk, in which case there can be a weight w_i associated with each sampled $\mathbf{x}_i$. (In Metropolis, $w_i = 1$.) The MC samples provide a formal representation of f

$$f(\mathbf{x}) \propto \sum_i w_i \delta(\mathbf{x} - \mathbf{x}_i). \quad (47)$$

Given $\mathcal{M}$ independent samples from $f(\mathbf{x})$, the integral in Eq. (45) is estimated by

$$G_{\mathcal{M}} = \sum_{i=1}^{\mathcal{M}} w_i g(\mathbf{x}_i) / \sum_{i=1}^{\mathcal{M}} w_i \quad (48)$$

The error in the estimate decays algebraically with the number of samples: $|G_{\mathcal{M}} - G| \propto 1/\sqrt{\mathcal{M}}$. Using the results above, we can compute

$$G' = \frac{\int_{\Omega_0} f(\mathbf{x})g(\mathbf{x})h(\mathbf{x})d\mathbf{x}}{\int_{\Omega_0} f(\mathbf{x})h(\mathbf{x})d\mathbf{x}}, \quad (49)$$

if the function $h(\mathbf{x})$ is such that both the numerator and denominator exist. Formally

$$G'_{\mathcal{M}} = \frac{\sum_{i=1}^{\mathcal{M}} w_i g(\mathbf{x}_i) h(\mathbf{x}_i)}{\sum_{i=1}^{\mathcal{M}} w_i h(\mathbf{x}_i)}, \quad (50)$$

although, as we will see, difficulties arise when $h(\mathbf{x})$ can change sign and is rapidly oscillating. Integral equations are another main area of applications of MC methods. For example [38], the integral equation

$$\Psi'(\mathbf{x}) = \int_{\Omega_0} K(\mathbf{x}, \mathbf{y}) w(\mathbf{y}) \Psi(\mathbf{y}) d\mathbf{y}, \quad (51)$$

can be viewed in terms of a *random walk* if it has the following properties: $\Psi(\mathbf{y})$ and $\Psi'(\mathbf{x})$ can be viewed as PDF's (in the sense of f in Eq. (45)), $w(\mathbf{y}) \geq 0$, and $K(\mathbf{x}, \mathbf{y})$ is a PDF for $\mathbf{x}$ conditional on $\mathbf{y}$. Then, given an ensemble $\{\mathbf{y}_i\}$ sampling $\Psi(\mathbf{y})$, the following two steps will allow us to generate an ensemble that samples $\Psi'(\mathbf{x})$. First an absorption/branching process is applied to each $\mathbf{y}_i$ according to $w(\mathbf{y}_i)$. For example, we can make $\text{int}(w(\mathbf{y}_i) + \xi)$ copies of $\mathbf{y}_i$, where ξ is a uniform random number on $(0, 1)$. Second we randomly walk each new $\mathbf{y}_j$ to an $\mathbf{x}_j$ by sampling the PDF $K(\mathbf{x}_j, \mathbf{y}_j)$. The resulting $\{\mathbf{x}_j\}$ are MC samples of $\Psi'(\mathbf{x})$. An example of this is the one-dimensional integral equation

$$\Psi(x) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi}} e^{-(x-y)^2} \sqrt{2} e^{-y^2/2} \Psi(y) dy, \quad (52)$$

which has a solution $\Psi(x) = e^{-x^2/2}$. The random walks, starting from an arbitrary distribution, will iteratively converge to a distribution sampling $\Psi(x)$.

B Properties of non-orthogonal Slater determinants

In Eq. (18), the operator $\hat{\varphi}_m^\dagger \equiv \sum_i c_i^\dagger \varphi_{i,m}$, with m taking an integer value among $1, 2, \dots, M$, creates an electron in a single-particle orbital φ_m : $\hat{\varphi}_m^\dagger |0\rangle = \sum_i \varphi_{i,m} |\chi_i\rangle$. The content of the orbital can thus be conveniently expressed as an N -dimensional vector $\{\varphi_{1,m}, \varphi_{2,m}, \dots, \varphi_{N,m}\}$. The Slater determinant $|\phi\rangle$ in Eq. (18) can then be expressed as an $N \times M$ matrix

$$\Phi \equiv \begin{pmatrix} \varphi_{1,1} & \varphi_{1,2} & \cdots & \varphi_{1,M} \\ \varphi_{2,1} & \varphi_{2,2} & \cdots & \varphi_{2,M} \\ \vdots & \vdots & & \vdots \\ \varphi_{N,1} & \varphi_{N,2} & \cdots & \varphi_{N,M} \end{pmatrix}.$$

Each column of this matrix represents a single-particle orbital that is completely specified by its N -dimensional vector. For convenience, we will think of the different columns as all properly orthonormalized, which is straightforward to achieve by, for example, modified Gram-Schmidt (see e.g., [2,21]). For example the occupied manifold in a DFT solution forms a “wave function” which is a Slater determinant.

A key property of Slater determinants is the *Thouless Theorem*: any one-particle operator $\hat{B}$ of the form in Eq. (12), when acting on a Slater determinant, simply leads to another Slater determinant [75], i.e.,

$$\hat{B}|\phi\rangle = \hat{\phi}'_1^\dagger \hat{\phi}'_2^\dagger \cdots \hat{\phi}'_M^\dagger |0\rangle \equiv |\phi'\rangle \quad (53)$$

with $\hat{\phi}'_m^\dagger = \sum_j c_j^\dagger \Phi'_{jm}$ and $\Phi' \equiv e^U \Phi$, where U is a square matrix whose elements are given by U_{ij} and $B \equiv \exp(U)$ is therefore an $N \times N$ square matrix as well. In other words, the operation of $\hat{B}$ on $|\phi\rangle$ simply involves multiplying an $N \times N$ matrix to the $N \times M$ matrix representing the Slater determinant.

In standard quantum chemistry (QC) methods, the many-body ground-state wave function is also represented by a sum of Slater determinants. However, there is a key difference between it and the AFQMC representation. In QC methods, the different Slater determinants are orthogonal, because they are obtained by excitations based on a fixed set of orbitals (for instance, the occupied and virtual orbitals from a reference calculation such as Hartree-Fock). In AFQMC, the Slater determinants are non-orthogonal because they are generated by “rotating” (only) the occupied orbitals through the operations in Eq. (53).

Several properties of non-orthogonal Slater determinants are worth mentioning. The overlap between two of them is given by

$$\langle \phi | \phi' \rangle = \det(\Phi^\dagger \Phi'). \quad (54)$$

We can define the expectation of an operator $\hat{O}$ with respect to a pair of non-orthogonal Slater determinants

$$\langle \hat{O} \rangle_{\phi\phi'} \equiv \frac{\langle \phi | \hat{O} | \phi' \rangle}{\langle \phi | \phi' \rangle}, \quad (55)$$

for instance single-particle Green function $G_{ij} \equiv \langle c_i c_j^\dagger \rangle_{\phi\phi'}$

$$G_{ij} \equiv \frac{\langle \phi | c_i c_j^\dagger | \phi' \rangle}{\langle \phi | \phi' \rangle} = \delta_{ij} - [\Phi' (\Phi^\dagger \Phi')^{-1} \Phi^\dagger]_{ij}. \quad (56)$$

Given the Green function matrix G , the general expectation defined in Eq. (55) can be computed for most operators of interest. For example, we can calculate the expectation of a general two-body operator, $\hat{O} = \sum_{ijkl} O_{ijkl} c_i^\dagger c_j^\dagger c_k c_l$, under the definition of Eq. (55)

$$\langle \hat{O} \rangle_{\phi\phi'} = \sum_{ijkl} O_{ijkl} (G'_{jk} G'_{il} - G'_{ik} G'_{jl}), \quad (57)$$

where the matrix G' is defined as $G' \equiv I - G$.

There are several generalizations of the formalism we have discussed which extend the capability and/or accuracy of the AFQMC framework. These can be thought of as generalizing one or both of the Slater determinants in Eqs. (54), (55), and (56). From the viewpoint of AFQMC, the “bra” in these equations represents the trial wave function, and the “ket” represents the random walker:

- The first generalization is to replace $\langle \phi |$ by a projected Bardeen-Cooper-Schrieffer (BCS) wave function, that is, to use a projected BCS as a trial wave function, which can be advantageous for systems with pairing order. The corresponding overlap, Green functions, and two-body mixed expectations have been worked out [54].
- The second is to have both $\langle \phi |$ and $|\phi' \rangle$ in generalized HF (GHF) form, which is necessary to treat systems with SOC. The required modification to the formalism outlined above is given by [22].
- The third generalization is to have both sides in Hartree-Fock-Bogoliubov (HFB) form, for example, to treat Hamiltonians with pairing fields. This will also be useful when using AFQMC as an impurity solver in which the embedding induces pairing order. The corresponding AFQMC formalism has been described [56].

References

- [1] S. Zhang, in W. Andreoni and S. Yip (eds.): *Handbook of Materials Modeling* (Springer, Cham, 2018)
- [2] S. Zhang, in E. Pavarini, E. Koch, and U. Schollwöck (eds.): *Emergent Phenomena in Correlated Matter*, Modeling and Simulation, Vol. 3 (Forschungszentrum Jülich, 2013)
- [3] S. Zhang, Comput. Phys. Commun. **127**, 150 (2000)
- [4] Y.Y. He, M. Qin, H. Shi, Z.Y. Lu, and S. Zhang, Phys. Rev. B **99**, 045108 (2019)
- [5] A. Szabo and N. Ostlund: *Modern quantum chemistry* (McGraw-Hill, New York, 1989)
- [6] W. Kohn, Rev. Mod. Phys. **71**, 1253 (1999) and references therein
- [7] R.M. Martin: *Electronic Structure: Basic theory and practical methods* (Cambridge University Press, 2004)
- [8] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985)
- [9] M.H. Kalos, D. Levesque, and L. Verlet, Phys. Rev. A **9**, 2178 (1974)
- [10] W.M.C. Foulkes, L. Mitas, R.J. Needs, and G. Rajagopal, Rev. Mod. Phys. **73**, 33 (2001) and references therein
- [11] D.M. Ceperley, Rev. Mod. Phys. **67**, 279 (1995) and references therein
- [12] R. Blankenbecler, D.J. Scalapino, and R.L. Sugar, Phys. Rev. D **24**, 2278 (1981)
- [13] G. Sugiyama and S.E. Koonin, Ann. Phys. (NY) **168**, 1 (1986)
- [14] S. Zhang and H. Krakauer, Phys. Rev. Lett. **90**, 136401 (2003)
- [15] K.P. Esler, J. Kim, D.M. Ceperley, W. Purwanto, E.J. Walter, H. Krakauer, S. Zhang, P.R.C. Kent, R.G. Hennig, C. Umrigar, M. Bajdich, J. Kolorenc, L. Mitas, and A. Srinivasan, J. Phys.: Conf. Series **125**, 012057 (2008)
- [16] K.E. Schmidt and M.H. Kalos, in K. Binder (ed.): *Applications of the Monte Carlo Method in Statistical Physics* (Springer, Heidelberg, 1984)
- [17] E.Y. Loh Jr., J.E. Gubernatis, R.T. Scalettar, S.R. White, D.J. Scalapino, and R. Sugar, Phys. Rev. B **41**, 9301 (1990)
- [18] S. Zhang, in M.P. Nightingale and C.J. Umrigar (eds.): *Quantum Monte Carlo Methods in Physics and Chemistry* (Kluwer Academic Publishers, 1999)
- [19] S. Zhang, J. Carlson, and J.E. Gubernatis, Phys. Rev. B **55**, 7464 (1997)

- [20] W.A. Al-Saidi, S. Zhang, and H. Krakauer, J. Chem. Phys. **124**, 224101 (2006)
- [21] M. Motta and S. Zhang, Wiley Interdiscip. Rev. Comput. Mol. Sci. **8**, e1364 (2018)
- [22] P. Rosenberg, H. Shi, and S. Zhang, J. Phys. Chem. Solids **128**, 161 (2019)
- [23] M. Suewattana, W. Purwanto, S. Zhang, H. Krakauer, and E.J. Walter, Phys. Rev. B **75**, 245123 (2007)
- [24] W. Purwanto, H. Krakauer, Y. Virgus, and S. Zhang, J. Chem. Phys. **135**, 164105 (2011)
- [25] S. Zhang, Phys. Rev. Lett. **83**, 2777 (1999)
- [26] Y. Liu, M. Cho, and B. Rubenstein, J. Chem. Theory and Comput. **14**, 4722 (2018)
- [27] J.W. Negele and H. Orland: *Quantum Many-Particle Systems* (Perseus Books, Reading, MA, 1998)
- [28] W.A. Al-Saidi, H. Krakauer, and S. Zhang, J. Chem. Phys. **126**, 194105 (2007)
- [29] F. Ma, W. Purwanto, S. Zhang, and H. Krakauer, Phys. Rev. Lett. **114**, 226401 (2015)
- [30] H. Koch, A.S. de Merás, and T.B. Pedersen, J. Chem. Phys. **118**, 9481 (2003)
- [31] F. Aquilante, L. De Vico, N. Ferre, G. Ghigo, P.Å Malmqvist, P. Neogady, T. Pedersen, M. Pitonak, M. Reiher, B. Roos, L. Serrano-Andres, M. Urban, V. Veryazov, and R. Lindh, J. Comp. Chem. **31**, 224 (2010)
- [32] J. Shee, E.J. Arthur, S. Zhang, D.R. Reichman, and R.A. Friesner, J. Chem. Theory Comput. **14**, 4109 (2018)
- [33] R. Baer, M. Head-Gordon, and D. Neuhauser, J. Chem. Phys. **109**, 6219 (1998)
- [34] W. Purwanto and S. Zhang, Phys. Rev. A **72**, 053610 (2005)
- [35] H. Shi and S. Zhang, Phys. Rev. B **88**, 125132 (2013)
- [36] S. Zhang and D.M. Ceperley, Phys. Rev. Lett. **100**, 236404 (2008)
- [37] S. Sorella, S. Baroni, R. Car, and M. Parrinello, Europhys. Lett. **8**, 663 (1989)
- [38] M.H. Kalos and P.A. Whitlock: *Monte Carlo methods*, Vol I (Wiley, 1986)
- [39] J.E. Hirsch, Phys. Rev. B **31**, 4403 (1985)
- [40] Z.C. Wei, C. Wu, Y. Li, S. Zhang, and T. Xiang, Phys. Rev. Lett. **116**, 250601 (2016)
- [41] S. Zhang and M.H. Kalos, Phys. Rev. Lett. **67**, 3074 (1991)
- [42] S. Zhang, J. Carlson, and J.E. Gubernatis, Phys. Rev. Lett. **74**, 3652 (1995)

- [43] M. Qin, H. Shi, and S. Zhang, *Phys. Rev. B* **94**, 235119 (2016)
- [44] S.B. Fahy and D.R. Hamann, *Phys. Rev. Lett.* **65**, 3437 (1990)
- [45] W. Purwanto and S. Zhang, *Phys. Rev. E* **70**, 056702 (2004)
- [46] M. Motta and S. Zhang, *J. Chem. Theory Comput.* **13**, 5367 (2017)
- [47] S. Zhang, H. Krakauer, W.A. Al-Saidi, and M. Suewattana, *Comput. Phys. Commun.* **169**, 394 (2005)
- [48] H. Hao, B.M. Rubenstein, and H. Shi, *Phys. Rev. B* **99**, 235142 (2019)
- [49] J.P.F. LeBlanc, A.E. Antipov, F. Becca, I.W. Bulik, G.K.L. Chan, C.M. Chung, Y. Deng, M. Ferrero, T.M. Henderson, C.A. Jiménez-Hoyos, E. Kozik, X.W. Liu, A.J. Millis, N.V. Prokof'ev, M. Qin, G.E. Scuseria, H. Shi, B.V. Svistunov, L.F. Tocchio, I.S. Tupitsyn, S.R. White, S. Zhang, B.X. Zheng, Z. Zhu, and E. Gull, *Phys. Rev. X* **5**, 041041 (2015)
- [50] B.X. Zheng, C.M. Chung, P. Corboz, G. Ehlers, M.P. Qin, R.M. Noack, H. Shi, S.R. White, S. Zhang, and G.K.L. Chan, *Science* **358**, 1155 (2017)
- [51] H. Shi, S. Chiesa, and S. Zhang, *Phys. Rev. A* **92**, 033603 (2015)
- [52] H. Shi and S. Zhang, *Phys. Rev. E* **93**, 033303 (2016)
- [53] E. Vitali, H. Shi, M. Qin, and S. Zhang, *Phys. Rev. B* **94**, 085140 (2016)
- [54] J. Carlson, S. Gandolfi, K.E. Schmidt, and S. Zhang, *Phys. Rev. A* **84**, 061602 (2011)
- [55] E. Vitali, P. Rosenberg, and S. Zhang, arXiv:1905.05012
- [56] H. Shi and S. Zhang, *Phys. Rev. B* **95**, 045144 (2017)
- [57] B.M. Rubenstein, S. Zhang, and D.R. Reichman, *Phys. Rev. A* **86**, 053606 (2012)
- [58] Y.Y. He, H. Shi, and S. Zhang, arXiv e-prints arXiv:1906.02247 (2019)
- [59] E.J. Landinez Borda, J. Gomez, and M.A. Morales, *J. Chem. Phys.* **150**, 074105 (2019)
- [60] J. Shee, B. Rudsteyn, E.J. Arthur, S. Zhang, D.R. Reichman, and R.A. Friesner, *J. Chem. Theory Comput.* **15**, 2346 (2019)
- [61] J. Shee, E.J. Arthur, S. Zhang, D.R. Reichman, and R.A. Friesner, arXiv:1905.13316
- [62] M. Motta and S. Zhang, *J. Chem. Phys.* **148**, 181101 (2018)
- [63] M. Motta, J. Shee, S. Zhang, and G. Kin-Lic Chan, arXiv:1810.01549
- [64] Y. Virgus, W. Purwanto, H. Krakauer, and S. Zhang, *Phys. Rev. Lett.* **113**, 175502 (2014)

- [65] B. Eskridge, H. Krakauer, and S. Zhang, *J. Chem. Theory Comput.* **15**, 3949 (2019)
- [66] J. Shee, S. Zhang, D.R. Reichman, and R.A. Friesner, *J. Chem. Theory Comput.* **13**, 2667 (2017)
- [67] F. Ma, S. Zhang, and H. Krakauer, *Phys. Rev. B* **95**, 165103 (2017)
- [68] S. Zhang, F.D. Malone, and M.A. Morales, *J. Chem. Phys.* **149**, 164102 (2018)
- [69] M. Motta, D.M. Ceperley, G.K.L. Chan, J.A. Gomez, E. Gull, S. Guo, C.A. Jiménez-Hoyos, T.N. Lan, J. Li, F. Ma, A.J. Millis, N.V. Prokof'ev, U. Ray, G.E. Scuseria, S. Sorella, E.M. Stoudenmire, Q. Sun, I.S. Tupitsyn, S.R. White, D. Zgid, and S. Zhang, *Phys. Rev. X* **7**, 031059 (2017)
- [70] R.J. Bartlett and M. Musiał, *Rev. Mod. Phys.* **79**, 291 (2007)
- [71] T.D. Crawford and H.F. Schaefer III, *Rev. Comput. Chem.* **14**, 33 (2000)
- [72] W. Purwanto, S. Zhang, and H. Krakauer, *J. Chem. Phys.* **130**, 094107 (2009)
- [73] F. Ma, S. Zhang, and H. Krakauer, *New J. Phys.* **15**, 093017 (2013)
- [74] H. Nguyen, H. Shi, J. Xu, and S. Zhang, *Comput. Phys. Commun.* **185**, 3344 (2014)
- [75] D.R. Hamann and S.B. Fahy, *Phys. Rev. B* **41**, 11352 (1990)

7 Exact Diagonalization and Lanczos Method

Erik Koch

Jülich Supercomputer Centre

Forschungszentrum Jülich

Contents

1	Many-electron states	2
2	Second quantization	5
2.1	Creation and annihilation operators	5
2.2	Representation of Slater determinants	7
2.3	Representation of n -body operators	8
2.4	Transforming the orbital basis	12
3	Exact diagonalization	13
3.1	Variational principles	13
3.2	Matrix eigenvalue problem	15
3.3	Dimension of the Hilbert space and sparseness	16
4	Lanczos Method	18
4.1	Krylov space	20
4.2	Spectral functions	24
A	Appendices	26
A.1	Non-orthonormal basis	26
A.2	Useful commutation relations	27
A.3	Downfolding	28

1 Many-electron states

One of the great surprises of quantum mechanics is the existence of indistinguishable objects. Classically this is not possible: objects can always be distinguished at least by their position in space, meaning that indistinguishable objects must be identical. This is Leibniz' Principle of the Identity of Indiscernibles [1]. For quantum objects, however, the uncertainty principle makes the distinction of particles by their position impossible. This allows for the existence of elementary particles. They form the basic units of all matter. So, quite remarkably, all the different objects we know are made of indistinguishable building blocks.

In the formalism of quantum mechanics, indistinguishability means that no observable lets us distinguish one of these particles from the other. This means that every observable for, e.g., electrons, must treat each electron in the same way. Thus, in principle, observables must act on all electrons in the universe. In practice we can, of course, distinguish electrons localized on the moon from those in our lab to an excellent approximation. Thus, for all practical purposes, we can restrict our description to the electrons in the system under consideration, assuming that the differential overlap with all other electrons vanishes. Any observable $M(x_1, \dots, x_N)$ for the N electrons in our system must then be symmetric under permutations of the variables x_i .

The consequences are straightforward: An observable $M(x)$ acting on a single-particle degree of freedom x must act on all indistinguishable particles in the same way, i.e., $\sum_i M(x_i)$. Likewise, a two-body observable $M(x, x')$ must act on all pairs in the same way, $\sum_{i,j} M(x_i, x_j)$ with $M(x, x') = M(x', x)$. We can thus write any observable in the form

$$M(\mathbf{x}) = M^{(0)} + \sum_i M^{(1)}(x_i) + \frac{1}{2!} \sum_{i \neq j} M^{(2)}(x_i, x_j) + \frac{1}{3!} \sum_{i \neq j \neq k} M^{(3)}(x_i, x_j, x_k) + \dots \quad (1)$$

$$= M^{(0)} + \sum_i M^{(1)}(x_i) + \sum_{i < j} M^{(2)}(x_i, x_j) + \sum_{i < j < k} M^{(3)}(x_i, x_j, x_k) + \dots, \quad (2)$$

where the summations can be restricted since the operators must be symmetric in their arguments, while for two or more identical coordinates the operator is really one of lower order: $M^{(2)}(x_i, x_i)$, e.g., only acts on a single coordinate and should be included in $M^{(1)}$.

For the many-body wave functions $\Psi(x_1, x_2, \dots)$ the situation is slightly more complex. Since the probability density $|\Psi(x_1, x_2, \dots)|^2$ is an observable, the wave function should transform as one-dimensional (irreducible) representations of the permutation group. Which irreducible representation applies to a given type of elementary particle is determined by the spin-statistics theorem [2, 3]: The wave functions of particles with integer spin are symmetric, those of particles with half-integer spin change sign when two arguments are exchanged. From an arbitrary N -particle wave function we thus obtain a many-electron wavefunction by antisymmetrizing

$$\mathcal{A}\Psi(x_1, \dots, x_N) := \frac{1}{\sqrt{N!}} \sum_P (-1)^P \Psi(x_{p(1)}, \dots, x_{p(N)}) , \quad (3)$$

where $(-1)^P$ is the parity of the permutation P that maps $n \rightarrow p(n)$. Since there are $N!$ different permutations, this can easily become an extremely costly operation. Remarkably, a product of

N single-electron states φ_α can be antisymmetrized much more efficiently (in $\mathcal{O}(N^3)$ steps) by writing it in the form of a determinant

$$\Phi_{\alpha_1, \dots, \alpha_N}(x_1, \dots, x_N) := \mathcal{A} \varphi_{\alpha_1}(x_1) \cdots \varphi_{\alpha_N}(x_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \varphi_{\alpha_1}(x_1) & \varphi_{\alpha_2}(x_1) & \cdots & \varphi_{\alpha_N}(x_1) \\ \varphi_{\alpha_1}(x_2) & \varphi_{\alpha_2}(x_2) & \cdots & \varphi_{\alpha_N}(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{\alpha_1}(x_N) & \varphi_{\alpha_2}(x_N) & \cdots & \varphi_{\alpha_N}(x_N) \end{vmatrix}. \quad (4)$$

For $N=1$ the Slater determinant is simply the one-electron orbital $\Phi_\alpha(x) = \varphi_\alpha(x)$ while for $N=2$ we get the familiar expression $\Phi_{\alpha, \alpha'}(x, x') = (\varphi_\alpha(x)\varphi_{\alpha'}(x') - \varphi_{\alpha'}(x)\varphi_\alpha(x'))/\sqrt{2}$ for the two-electron Slater determinant.

Slater determinants are important because they can be used to build a basis of the many-electron Hilbert space. To see how, we consider a complete set of orthonormal single-electron states

$$\sum_n \overline{\varphi_n(x)} \varphi_n(x') = \delta(x - x') \text{ (complete)} \quad \int dx \overline{\varphi_n(x)} \varphi_m(x) = \delta_{n,m} \text{ (orthonormal)}. \quad (5)$$

To expand an arbitrary N -particle function $a(x_1, \dots, x_N)$, we start by considering it as a function of x_1 with $x_2, \dots, x_N$ kept fixed. We can then expand it in the complete set $\{\varphi_n\}$ as

$$a(x_1, \dots, x_N) = \sum_{n_1} a_{n_1}(x_2, \dots, x_N) \varphi_{n_1}(x_1)$$

with expansion coefficients that depend on the remaining coordinates

$$a_{n_1}(x_2, \dots, x_N) = \int dx_1 \overline{\varphi_{n_1}(x_1)} a(x_1, x_2, \dots, x_N).$$

These, in turn, can be expanded as a functions of x_2

$$a_{n_1}(x_2, \dots, x_N) = \sum_{n_2} a_{n_1, n_2}(x_3, \dots, x_N) \varphi_{n_2}(x_2).$$

Repeating this, we obtain the expansion of a in product states

$$a(x_1, \dots, x_N) = \sum_{n_1, \dots, n_N} a_{n_1, \dots, n_N} \varphi_{n_1}(x_1) \cdots \varphi_{n_N}(x_N).$$

When the N -particle function is antisymmetric, applying the antisymmetrizer (3) will leave it unchanged, i.e., the expansion coefficients will be antisymmetric under permutation of the indices: $a_{n_{p(1)}, \dots, n_{p(N)}} = (-1)^P a_{n_1, \dots, n_N}$. Fixing some particular order of the indices, e.g., $n_1 < n_2 < \dots < n_N$, we thus get an expansion in Slater determinants

$$\Psi(x_1, \dots, x_N) = \sum_{n_1 < \dots < n_N} a_{n_1, \dots, n_N} \sqrt{N!} \Phi_{n_1, \dots, n_N}(x_1, \dots, x_N).$$

Since we can write any antisymmetric function as such a configuration-interaction expansion, the set of Slater determinants

$$\left\{ \Phi_{n_1, \dots, n_N}(x_1, \dots, x_N) \mid n_1 < n_2 < \dots < n_N \right\} \quad (6)$$

forms a basis of the N -electron Hilbert space. Since the overlap of two Slater determinants

$$\begin{aligned} \int d\mathbf{x} \overline{\Phi_{\alpha_1, \dots, \alpha_N}(\mathbf{x})} \Phi_{\beta_1, \dots, \beta_N}(\mathbf{x}) &= \frac{1}{N!} \sum_{P, P'} (-1)^{P+P'} \prod_n \int dx_n \overline{\varphi_{\alpha_{P(n)}}(x_n)} \varphi_{\alpha_{P'(n)}}(x_n) \\ &= \begin{vmatrix} \langle \varphi_{\alpha_1} | \varphi_{\beta_1} \rangle & \cdots & \langle \varphi_{\alpha_1} | \varphi_{\beta_N} \rangle \\ \vdots & \ddots & \vdots \\ \langle \varphi_{\alpha_N} | \varphi_{\beta_1} \rangle & \cdots & \langle \varphi_{\alpha_N} | \varphi_{\beta_N} \rangle \end{vmatrix} \end{aligned} \quad (7)$$

is the determinant of the overlap of the constituent orbitals, the Slater determinants (6) form a complete orthonormal basis of the N -electron Hilbert space when the orbitals $\varphi_n(x)$ are a complete orthonormal basis of the one-electron Hilbert space.

While we use a set of N one-electron orbitals $\varphi_n(x)$ to define an N -electron Slater determinant $\Phi_{\alpha_1, \dots, \alpha_N}(\mathbf{x})$, this representation is not unique: Any unitary transformation among the N occupied orbitals will not change the determinant. Thus, strictly, a Slater determinant is not determined by the set of indices we usually give, but, up to a phase, by the N -dimensional subspace spanned by the orbitals $\varphi_1, \dots, \varphi_N$ in the single-electron Hilbert space. The projector to this space is the one-body density matrix

$$\Gamma^{(1)}(x, x') = N \int dx_2 \cdots dx_N \overline{\Phi(x, x_2, \dots, x_N)} \Phi(x', x_2, \dots, x_N). \quad (8)$$

To see this, we expand the Slater determinant along its first row

$$\Phi_{\alpha_1 \dots \alpha_N}(x_1, \dots, x_N) = \frac{1}{\sqrt{N}} \sum_{n=1}^N (-1)^{1+n} \varphi_{\alpha_n}(x_1) \Phi_{\alpha_{i \neq n}}(x_2, \dots, x_N), \quad (9)$$

where $\Phi_{\alpha_{i \neq n}}(x_2, \dots, x_N)$ is the determinant with the first row and the n -th column removed, which can be written as $N-1$ -electron Slater determinants with orbital α_n removed. Inserting this into (8) we find

$$\Gamma_{\Phi}^{(1)}(x, x') = \sum_{n=1}^N \overline{\varphi_{\alpha_n}(x)} \varphi_{\alpha_n}(x'), \quad (10)$$

which is the expansion of the one-body density matrix in eigenfunctions (natural orbitals), with eigenvalues (natural occupation numbers) either one or zero. Any many-electron wave function $\Psi(\mathbf{x})$ with the same one-body density matrix $\Gamma_{\Phi}^{(1)}$ equals the Slater determinant $\Phi(\mathbf{x})$ up to a phase, i.e., $|\langle \Psi | \Phi \rangle| = 1$.

We can generalize this procedure and calculate higher order density matrices by introducing the generalized Laplace expansion

$$\Phi_{\alpha_1 \dots \alpha_N}(\mathbf{x}) = \frac{1}{\sqrt{\binom{N}{p}}} \sum_{n_1 < \dots < n_p} (-1)^{1+\sum_i n_i} \Phi_{\alpha_{n_1} \dots \alpha_{n_p}}(x_1, \dots, x_p) \Phi_{\alpha_{i \notin \{n_1, \dots, n_p\}}}(x_{p+1}, \dots, x_N),$$

which is obtained by writing the permutation of all N indices as a permutation of $N-p$ indices and the remaining p indices separately summing over all distinct sets of p indices. This allows us to evaluate arbitrary matrix elements and higher order density matrices [4]. But as can be seen from the above expansion, the expressions very quickly get quite cumbersome. Fortunately there is a representation that is much better suited to handling antisymmetric wave functions. It is called second quantization.

2 Second quantization

While originally introduced for quantizing the electromagnetic field, we can use the formalism of second quantization just as a convenient way of handling antisymmetric wave functions [5,6]. The idea behind this approach is remarkably simple: When writing Slater determinants in the form (4) we are working in a real-space basis. It is, however, often simpler to consider abstract states: Instead of a wave function $\varphi_\alpha(x)$, we write a Dirac state $|\alpha\rangle$. Second quantization allows us to do the same for Slater determinants.

Let us consider a Slater determinant for two electrons, one in state $\varphi_\alpha(x)$, the other in state $\varphi_\beta(x)$. It is simply the antisymmetrized product of the two states

$$\Phi_{\alpha\beta}(x_1, x_2) = \frac{1}{\sqrt{2}} (\varphi_\alpha(x_1)\varphi_\beta(x_2) - \varphi_\beta(x_1)\varphi_\alpha(x_2)). \quad (11)$$

This expression is quite cumbersome because we explicitly specify the coordinates. We can try to get rid of the coordinates by defining a two-particle Dirac state

$$|\alpha, \beta\rangle := \frac{1}{\sqrt{2}} (|\alpha\rangle|\beta\rangle - |\beta\rangle|\alpha\rangle).$$

While the expression is somewhat simpler, we still have to keep track of the order of the particles by specifying the position of the kets. The idea of second quantization is to specify the states using operators

$$|\alpha, \beta\rangle = c_\beta^\dagger c_\alpha^\dagger |0\rangle. \quad (12)$$

Now the order of the particles is specified by the order of the operators. To ensure the antisymmetry of the wave function the operators have to change sign when they are reordered

$$|\alpha, \beta\rangle = c_\beta^\dagger c_\alpha^\dagger |0\rangle = -c_\alpha^\dagger c_\beta^\dagger |0\rangle = -|\beta, \alpha\rangle. \quad (13)$$

2.1 Creation and annihilation operators

To arrive at the formalism of second quantization we postulate a set of operators that have certain reasonable properties. We then verify that we can use these operators to represent Slater determinants. But first we consider a few simple states to motivate what properties the new operators ought to have.

To be able to construct many-electron states, we start from the simplest such state: $|0\rangle$ the *vacuum state* with no electron, which we assume to be normalized $\langle 0|0\rangle = 1$. Next we introduce for each single-electron state $|\alpha\rangle$ an operator $c_\alpha^\dagger$ such that $c_\alpha^\dagger |0\rangle = |\alpha\rangle$. We call them *creation operators* since they add an electron (in state α) to the state that they act on: in $c_\alpha^\dagger |0\rangle$ the creation operator adds an electron to the vacuum state ($N=0$), resulting in a single-electron state ($N=1$). Applying another creation operator produces a two-electron state $c_\beta^\dagger c_\alpha^\dagger |0\rangle$, ($N=2$). To ensure the antisymmetry of the two electron state, the product of creation operators has to change sign when they are reordered: $c_\alpha^\dagger c_\beta^\dagger = -c_\beta^\dagger c_\alpha^\dagger$. This is more conveniently written as $\{c_\alpha^\dagger, c_\beta^\dagger\} = 0$ by introducing the *anti-commutator*

$$\{A, B\} := AB + BA. \quad (14)$$

As we have seen, the simplest state we can produce with the creation operators is the single-electron state $|\alpha\rangle = c_\alpha^\dagger|0\rangle$. When we want to calculate its norm, we have to consider the adjoint of $c_\alpha^\dagger|0\rangle$, formally obtaining $\langle\alpha|\alpha\rangle = \langle 0|c_\alpha c_\alpha^\dagger|0\rangle$, or, more generally, $\langle\alpha|\beta\rangle = \langle 0|c_\alpha c_\beta^\dagger|0\rangle$. This implies that c_α , the adjoint of a creation operator, must remove an electron from the state, otherwise the overlap of $c_\alpha c_\beta^\dagger|0\rangle$ with the vacuum state $\langle 0|$ would vanish. We therefore call the adjoint of the creation operator an *annihilation operator*. We certainly cannot take an electron out of the vacuum state, so $c_\alpha|0\rangle = 0$. To obtain the overlap of one-electron states we postulate the anticommutation relation $\{c_\alpha, c_\beta^\dagger\} = \langle\alpha|\beta\rangle$, giving $\langle 0|c_\alpha c_\beta^\dagger|0\rangle = \langle 0|\{c_\alpha, c_\beta^\dagger\} - c_\beta^\dagger c_\alpha|0\rangle = \langle\alpha|\beta\rangle$. For completeness, taking the adjoint of the anticommutation relation for the creation operators, we obtain the corresponding anticommutator of the annihilators: $\{c_\alpha, c_\beta\} = 0$. Thus, we are led to define the vacuum state $|0\rangle$ and the set of operators c_α related to single-electron states $|\alpha\rangle$ with the properties

$$\boxed{\begin{array}{ll} c_\alpha|0\rangle = 0 & \{c_\alpha, c_\beta\} = 0 = \{c_\alpha^\dagger, c_\beta^\dagger\} \\ \langle 0|0\rangle = 1 & \{c_\alpha, c_\beta^\dagger\} = \langle\alpha|\beta\rangle \end{array}} \quad (15)$$

As a direct consequence we obtain the Pauli principle in the form $c_\alpha c_\alpha = 0 = c_\alpha^\dagger c_\alpha^\dagger$.

We note that the creators transform in the same way as the single-electron states they represent

$$|\tilde{\alpha}_i\rangle = \sum_\mu |\alpha_\mu\rangle U_{\mu i} \quad \rightsquigarrow \quad c_{\tilde{\alpha}_i}^\dagger|0\rangle = \sum_\mu c_{\alpha_\mu}^\dagger|0\rangle U_{\mu i} = \left(\sum_\mu c_{\alpha_\mu}^\dagger U_{\mu i}\right)|0\rangle. \quad (16)$$

The creators and annihilators are clearly not operators in a Hilbert space, but transfer states from an N -electron to an $N\pm 1$ -electron Hilbert space, i.e., they are operators defined on *Fock space*. It is also remarkable that the mixed anti-commutator is the only place where the orbitals that distinguish different operators enter.

To make contact with the notation of first quantization, we introduce the *field operators* $\hat{\Psi}^\dagger(x)$, with $x = (r, \sigma)$, that create an electron of spin σ at position r , i.e., in state $|x\rangle = |r, \sigma\rangle$. Given a complete, orthonormal set of orbitals $\{\varphi_n\}$, we can expand $|x\rangle$

$$\hat{\Psi}^\dagger(x)|0\rangle = |x\rangle = \sum_n |\varphi_n\rangle \langle\varphi_n|x\rangle = \sum_n c_{\varphi_n}^\dagger|0\rangle \langle\varphi_n|x\rangle \quad (17)$$

from which we obtain

$$\hat{\Psi}^\dagger(x) = \sum_n \overline{\langle x|\varphi_n\rangle} c_{\varphi_n}^\dagger = \sum_n \overline{\varphi_n(x)} c_{\varphi_n}^\dagger. \quad (18)$$

The anticommutators then follow from (15) for an orthonormal and complete set, e.g.,

$$\{\hat{\Psi}(x), \hat{\Psi}^\dagger(x')\} = \sum_{n,m} \langle x|\varphi_n\rangle \underbrace{\{c_{\varphi_n}, c_{\varphi_m}^\dagger\}}_{=\delta_{n,m}} \langle\varphi_m|x'\rangle = \sum_n \langle x|\varphi_n\rangle \langle\varphi_n|x'\rangle = \langle x|x'\rangle = \delta(x-x'),$$

resulting in the anticommutation relations for the field operators

$$\{\hat{\Psi}(x), \hat{\Psi}(x')\} = 0 = \{\hat{\Psi}^\dagger(x), \hat{\Psi}^\dagger(x')\} \quad \text{and} \quad \{\hat{\Psi}(x), \hat{\Psi}^\dagger(x')\} = \langle x|x'\rangle. \quad (19)$$

We can, of course, expand the field operators also in a non-orthogonal set of orbitals $\{|\chi_i\rangle\}$, as long as it is complete, $\sum_{i,j} |\chi_i\rangle \langle S^{-1} \rangle_{ij} \langle \chi_j| = \mathbb{1}$, where $S_{ij} = \langle \chi_i | \chi_j \rangle$ is the overlap matrix,

$$\hat{\Psi}^\dagger(x) = \sum_{i,j} c_i^\dagger (S^{-1})_{ij} \langle \chi_j | x \rangle. \quad (20)$$

Conversely, given any single-electron wave functions in real space $\varphi(x)$, we can express the corresponding creation operator in terms of the field operators

$$c_\varphi^\dagger = \int dx \varphi(x) \hat{\Psi}^\dagger(x). \quad (21)$$

Its anticommutator with the field annihilator just gives back the single-electron wave function

$$\{\hat{\Psi}(x), c_\varphi^\dagger\} = \int dx' \varphi(x') \{\hat{\Psi}(x), \hat{\Psi}^\dagger(x')\} = \varphi(x). \quad (22)$$

2.2 Representation of Slater determinants

We have now all the tools in place to write the Slater determinant (4) in second quantization, using the creation operators to specify the occupied orbitals and the field operators to give the coordinates for the real-space representation:

$$\Phi_{\alpha_1 \alpha_2 \dots \alpha_N}(x_1, x_2, \dots, x_N) = \frac{1}{\sqrt{N!}} \langle 0 | \hat{\Psi}(x_1) \hat{\Psi}(x_2) \dots \hat{\Psi}(x_N) c_{\alpha_N}^\dagger \dots c_{\alpha_2}^\dagger c_{\alpha_1}^\dagger | 0 \rangle. \quad (23)$$

Note how writing the Slater determinant as an expectation value of annihilation and creation operators nicely separates the coordinates on the left from the orbitals on the right. This is just the desired generalization of the Dirac notation $\varphi(x) = \langle x | \varphi \rangle$.

Not surprisingly, the proof of (23) is by induction. As a warm-up we consider the case of a single-electron wave function ($N=1$). Using the anticommutation relation (22), we see that

$$\langle 0 | \hat{\Psi}(x_1) c_{\alpha_1}^\dagger | 0 \rangle = \langle 0 | \varphi_{\alpha_1}(x_1) - c_{\alpha_1}^\dagger \hat{\Psi}(x_1) | 0 \rangle = \varphi_{\alpha_1}(x_1). \quad (24)$$

For the two-electron state $N=2$, we anticommute $\hat{\Psi}(x_2)$ in two steps to the right

$$\begin{aligned} \langle 0 | \hat{\Psi}(x_1) \hat{\Psi}(x_2) c_{\alpha_2}^\dagger c_{\alpha_1}^\dagger | 0 \rangle &= \langle 0 | \hat{\Psi}(x_1) (\varphi_{\alpha_2}(x_2) - c_{\alpha_2}^\dagger \hat{\Psi}(x_2)) c_{\alpha_1}^\dagger | 0 \rangle \\ &= \langle 0 | \hat{\Psi}(x_1) c_{\alpha_1}^\dagger | 0 \rangle \varphi_{\alpha_2}(x_2) - \langle 0 | \hat{\Psi}(x_1) c_{\alpha_2}^\dagger \hat{\Psi}(x_2) c_{\alpha_1}^\dagger | 0 \rangle \\ &= \varphi_{\alpha_1}(x_1) \varphi_{\alpha_2}(x_2) - \varphi_{\alpha_2}(x_1) \varphi_{\alpha_1}(x_2). \end{aligned} \quad (25)$$

We see how anticommuting automatically produces the appropriate signs for the antisymmetric wave function. Dividing by $\sqrt{2}$, we obtain the desired two-electron Slater determinant.

The general case of an N -electron state works just the same. Anti-commuting $\hat{\Psi}(x_N)$ all the way to the right produces $N-1$ terms with alternating sign

$$\begin{aligned} \langle 0 | \hat{\Psi}(x_1) \dots \hat{\Psi}(x_{N-1}) \hat{\Psi}(x_N) c_{\alpha_N}^\dagger c_{\alpha_{N-1}}^\dagger \dots c_{\alpha_1}^\dagger | 0 \rangle &= \\ &+ \langle 0 | \hat{\Psi}(x_1) \dots \hat{\Psi}(x_{N-1}) c_{\alpha_{N-1}}^\dagger \dots c_{\alpha_1}^\dagger | 0 \rangle \varphi_{\alpha_N}(x_N) \\ &- \langle 0 | \hat{\Psi}(x_1) \dots \hat{\Psi}(x_{N-1}) \prod_{n \neq N-1} c_{\alpha_n}^\dagger | 0 \rangle \varphi_{\alpha_{N-1}}(x_N) \\ &\vdots \\ &(-1)^{N-1} \langle 0 | \hat{\Psi}(x_1) \dots \hat{\Psi}(x_{N-1}) c_{\alpha_N}^\dagger \dots c_{\alpha_2}^\dagger | 0 \rangle \varphi_{\alpha_1}(x_N). \end{aligned}$$

Using (23) for the $N-1$ -electron states, this is nothing but the Laplace expansion of

$$D = \begin{vmatrix} \varphi_{\alpha_1}(x_1) & \varphi_{\alpha_2}(x_1) & \cdots & \varphi_{\alpha_N}(x_1) \\ \varphi_{\alpha_1}(x_2) & \varphi_{\alpha_2}(x_2) & \cdots & \varphi_{\alpha_N}(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{\alpha_1}(x_N) & \varphi_{\alpha_2}(x_N) & \cdots & \varphi_{\alpha_N}(x_N) \end{vmatrix}$$

along the N th row. Dividing by $\sqrt{N!}$ we see that we have shown (23) for N -electron states, completing the proof by induction.

Given this representation of Slater determinants it is easy to eliminate the coordinates so we can work with N -electron states rather than N -electron wave functions—just as in Dirac notation. In particular we can rewrite the basis of Slater determinants (6) into a basis of product states

$$\{c_{n_N}^\dagger \cdots c_{n_1}^\dagger |0\rangle \mid n_1 < \cdots < n_N\}, \quad (26)$$

which allows us to express any N -electron state as

$$|\Psi\rangle = \sum_{n_1 < \cdots < n_N} a_{n_1, \dots, n_N} c_{n_N}^\dagger \cdots c_{n_1}^\dagger |0\rangle. \quad (27)$$

2.3 Representation of n -body operators

To work with N -electron states rather than Slater determinants, we also have to rewrite the N -electron operators $M(x)$ appropriately. This is easily done by incorporating the coordinates that we have separated from the Slater determinants into the operators such that the expectation values remain unchanged. This is, again, analogous to the Dirac formalism:

$$\int dx \overline{\varphi_n(x)} M(x) \varphi_m(x) = \langle \varphi_n | \underbrace{\int dx |x\rangle M(x) \langle x|}_{=: \hat{M}} \varphi_m \rangle = \langle \varphi_n | \hat{M} | \varphi_m \rangle. \quad (28)$$

For N -electron Slater determinants this becomes

$$\begin{aligned} & \int dx_1 \cdots dx_N \overline{\Phi_{\beta_1 \dots \beta_N}(x_1, \dots, x_N)} M(x_1, \dots, x_N) \Phi_{\alpha_1 \dots \alpha_N}(x_1, \dots, x_N) \\ &= \int dx_1 \cdots dx_N \langle 0 | c_{\beta_1} \cdots c_{\beta_N} \hat{\Psi}^\dagger(x_N) \cdots \hat{\Psi}^\dagger(x_1) |0\rangle M(x_1, \dots, x_N) \langle 0 | \hat{\Psi}(x_1) \cdots \hat{\Psi}(x_N) c_{\alpha_N}^\dagger \cdots c_{\alpha_1}^\dagger |0\rangle \\ &= \langle 0 | c_{\beta_1} \cdots c_{\beta_N} \hat{M} c_{\alpha_N}^\dagger \cdots c_{\alpha_1}^\dagger |0\rangle \end{aligned}$$

with the representation of the n -body operator in terms of field operators

$$\hat{M} := \frac{1}{N!} \int dx_1 \cdots dx_N \hat{\Psi}^\dagger(x_N) \cdots \hat{\Psi}^\dagger(x_1) M(x_1, \dots, x_N) \hat{\Psi}(x_1) \cdots \hat{\Psi}(x_N). \quad (29)$$

Note that this particular form of the operator is only valid when applied to N -electron states, since we have used that the N annihilation operators bring us to the zero-electron space, where $|0\rangle\langle 0| = \mathbb{1}$. Keeping this in mind, we can work entirely in terms of our algebra (15).

To see what (29) means, we look, in turn, at the different n -body parts of $M(\mathbf{x})$, (2):

$$M(\mathbf{x}) = M^{(0)} + \sum_i M^{(1)}(x_i) + \sum_{i < j} M^{(2)}(x_i, x_j) + \sum_{i < j < k} M^{(3)}(x_i, x_j, x_k) + \cdots \quad (30)$$

We start with the simplest case, the zero-body operator, which, up to a trivial prefactor, is $M^{(0)} = 1$. Operating on an N -electron wave function, it gives

$$\begin{aligned} \hat{M}^{(0)} &= \frac{1}{N!} \int dx_1 dx_2 \cdots x_N \hat{\Psi}^\dagger(x_N) \cdots \hat{\Psi}^\dagger(x_2) \hat{\Psi}^\dagger(x_1) \hat{\Psi}(x_1) \hat{\Psi}(x_2) \cdots \hat{\Psi}(x_N) \\ &= \frac{1}{N!} \int dx_2 \cdots x_N \hat{\Psi}^\dagger(x_N) \cdots \hat{\Psi}^\dagger(x_2) \quad \hat{N} \quad \hat{\Psi}(x_2) \cdots \hat{\Psi}(x_N) \\ &= \frac{1}{N!} \int dx_2 \cdots x_N \hat{\Psi}^\dagger(x_N) \cdots \hat{\Psi}^\dagger(x_2) \quad 1 \quad \hat{\Psi}(x_2) \cdots \hat{\Psi}(x_N) \\ &\vdots \\ &= \frac{1}{N!} 1 \cdot 2 \cdots N = 1, \end{aligned} \quad (31)$$

where we have used that the operator

$$\int dx \hat{\Psi}^\dagger(x) \hat{\Psi}(x) = \hat{N}$$

counts the number of electrons: Applied to the vacuum state it gives $\hat{N}|0\rangle = 0$, while its commutator with any creation operator produces that operator

$$[\hat{N}, c_n^\dagger] = \int dx [\hat{\Psi}^\dagger(x) \hat{\Psi}(x), c_n^\dagger] = \int dx \hat{\Psi}^\dagger(x) \{\hat{\Psi}(x), c_n^\dagger\} = \int dx \hat{\Psi}^\dagger(x) \varphi_n(x) = c_n^\dagger. \quad (32)$$

where we have used the simple relation $[AB, C] = A\{B, C\} - \{A, C\}B$. Commuting with an annihilator we pick up a minus sign $[\hat{N}, \hat{\Psi}(x')] = -\hat{\Psi}(x')$. Thus, commuting $\hat{N}$ through a general product state, we obtain for each creation operator that we encounter a copy of the state, while for each annihilator we obtain minus that state, giving in total the original state times the difference in the number of creation and annihilation operators.

Remarkably, while we started from an operator acting on N -electron states, the resulting operator in second quantized form is independent of the number of electrons. We will see that this is an important general feature of operators in second quantization which makes working in Fock spaces amazingly simple.

We note that (31) just means that the overlap of two Slater determinants (7) is equal to that of the corresponding product states

$$\int d\mathbf{x} \overline{\Phi_{\alpha_1, \dots, \alpha_N}(\mathbf{x})} \Phi_{\beta_1, \dots, \beta_N}(\mathbf{x}) = \langle 0 | c_{\alpha_1} \cdots c_{\alpha_N} c_{\beta_N}^\dagger \cdots c_{\beta_1}^\dagger | 0 \rangle. \quad (33)$$

2.3.1 One-body operators

Next we consider one-body operators $\sum_j M^{(1)}(x_j)$

$$\begin{aligned}
 \hat{M}^{(1)} &= \frac{1}{N!} \int dx_1 \cdots dx_N \hat{\Psi}^\dagger(x_N) \cdots \hat{\Psi}^\dagger(x_1) \sum_j M^{(1)}(x_j) \hat{\Psi}(x_1) \cdots \hat{\Psi}(x_N) \\
 &= \frac{1}{N!} \sum_j \int dx_j \hat{\Psi}^\dagger(x_j) M^{(1)}(x_j) (N-1)! \hat{\Psi}(x_j) \\
 &= \frac{1}{N} \sum_j \int dx_j \hat{\Psi}^\dagger(x_j) M^{(1)}(x_j) \hat{\Psi}(x_j) \\
 &= \int dx \hat{\Psi}^\dagger(x) M^{(1)}(x) \hat{\Psi}(x)
 \end{aligned} \tag{34}$$

Here we have first anticommutated $\hat{\Psi}^\dagger(x_j)$ all the way to the left and $\hat{\Psi}(x_j)$ to the right. Since these take the same numbers of anticommutations, there is no sign involved. In between these field operators we are left with a zero-body operator for $N-1$ electrons, producing, when $\hat{M}^{(1)}$ acts on an N -electron state, a factor of $(N-1)!$. Again we notice that we obtain an operator that no longer depends on the number of electrons, i.e., that is valid in the entire Fock space.

Expanding the field-operators in a complete orthonormal set $\hat{\Psi}(x) = \sum_n \varphi_n(x) c_n$ gives

$$\hat{M}^{(1)} = \sum_{n,m} \int dx \overline{\varphi_n(x)} M(x) \varphi_m(x) c_n^\dagger c_m = \sum_{n,m} \langle \varphi_n | M^{(1)} | \varphi_m \rangle c_n^\dagger c_m = \sum_{n,m} c_n^\dagger M_{nm}^{(1)} c_m. \tag{35}$$

The matrix elements $M_{nm}^{(1)} = \langle \varphi_n | M^{(1)} | \varphi_m \rangle$ transform like a single-electron matrix $\mathbf{M}^{(1)}$: From (16) and writing the annihilation operators as a column vector $\mathbf{c}$ we see that

$$\hat{M}^{(1)} = \mathbf{c}^\dagger \mathbf{M}^{(1)} \mathbf{c} = \mathbf{c}^\dagger \mathbf{U}^\dagger \mathbf{U} \mathbf{M}^{(1)} \mathbf{U}^\dagger \mathbf{U} \mathbf{c} = \tilde{\mathbf{c}}^\dagger \tilde{\mathbf{M}}^{(1)} \tilde{\mathbf{c}}. \tag{36}$$

Once we have arrived at the representation in terms of orbitals, we can restrict the orbital basis to a non-complete set. This simply gives the operator in the variational (Fock) subspace spanned by the orbitals.

We note that the expression (35) not only works for local operators but also for differential operators like the momentum or kinetic energy: we have taken care not to exchange the order of $M^{(1)}$ and one of its field operators. We can write truly non-local operators in a similar way. As an example, the one-body density operator is given by

$$\hat{I}^{(1)}(x; x') = \hat{\Psi}^\dagger(x) \hat{\Psi}(x') \tag{37}$$

so that one coordinate is not integrated over, rather setting it to x in the bra and x' in the ket. In an orthonormal basis it becomes

$$\hat{I}^{(1)}(x; x') = \sum_{n,m} \overline{\varphi_n(x)} \varphi_m(x') c_n^\dagger c_m. \tag{38}$$

2.3.2 Two-body operators

For the two-body operators $\sum_{i<j} M^{(2)}(x_i, x_j)$ we proceed in the familiar way, anti-commuting first the operators with the coordinates involved in $M^{(2)}$ all the way to the left and right, respectively. This time we are left with a zero-body operator for $N-2$ electrons:

$$\begin{aligned}
 \hat{M}^{(2)} &= \frac{1}{N!} \int dx_1 \cdots dx_N \hat{\Psi}^\dagger(x_N) \cdots \hat{\Psi}^\dagger(x_1) \sum_{i<j} M^{(2)}(x_i, x_j) \hat{\Psi}(x_1) \cdots \hat{\Psi}(x_N) \\
 &= \frac{1}{N!} \sum_{i<j} \int dx_i dx_j \hat{\Psi}^\dagger(x_j) \hat{\Psi}^\dagger(x_i) M^{(2)}(x_i, x_j) (N-2)! \hat{\Psi}(x_i) \hat{\Psi}(x_j) \\
 &= \frac{1}{N(N-1)} \sum_{i<j} \int dx_i dx_j \hat{\Psi}^\dagger(x_j) \hat{\Psi}^\dagger(x_i) M^{(2)}(x_i, x_j) \hat{\Psi}(x_i) \hat{\Psi}(x_j) \\
 &= \frac{1}{2} \int dx dx' \hat{\Psi}^\dagger(x') \hat{\Psi}^\dagger(x) M^{(2)}(x, x') \hat{\Psi}(x) \hat{\Psi}(x')
 \end{aligned}$$

Expanding in an orthonormal basis, we get

$$\begin{aligned}
 \hat{M}^{(2)} &= \frac{1}{2} \sum_{n,n',m,m'} \int dx dx' \overline{\varphi_{n'}(x') \varphi_n(x)} M^{(2)}(x, x') \varphi_m(x) \varphi_{m'}(x') c_n^\dagger c_{n'}^\dagger c_m c_{m'} \\
 &= \frac{1}{2} \sum_{n,n',m,m'} \langle \varphi_n \varphi_{n'} | M^{(2)} | \varphi_m \varphi_{m'} \rangle c_n^\dagger c_{n'}^\dagger c_m c_{m'} \quad (39)
 \end{aligned}$$

where the exchange of the indices in the second line is a consequence of the way the Dirac state for two electrons is usually written: first index for the first coordinate, second index for the second, while taking the adjoint of the operators changes their order. $M_{nn',mm'} = \langle \varphi_n \varphi_{n'} | M^{(2)} | \varphi_m \varphi_{m'} \rangle$ transforms like a fourth-order tensor: Transforming to a different basis (16) gives

$$\tilde{M}_{\nu\nu',\mu\mu'}^{(2)} = \sum_{n,n',m,m'} U_{\nu n}^\dagger U_{\nu' n'}^\dagger M_{nn',mm'} U_{m\mu} U_{m'\mu'} \quad (40)$$

Form the symmetry of the two-body operator $M^{(2)}(x, x') = M^{(2)}(x', x)$ follows $M_{nn',mm'} = M_{n'n,mm'}$. Moreover, $M_{nn,mm'}$ will not contribute to $\hat{M}^{(2)}$ since $c_n^\dagger c_n^\dagger = \{c_n^\dagger, c_n^\dagger\}/2 = 0$, and likewise for $M_{n'n',mm}$.

Note that the representation (39) is not quite as efficient as it could be: The terms with n and n' and/or m and m' exchanged connect the same basis states. Collecting these terms by introducing an ordering of the operators and using the symmetry of the matrix elements we obtain

$$\hat{M}^{(2)} = \sum_{n'>n, m'>m} c_{n'}^\dagger c_n^\dagger \underbrace{\left(M_{nn',mm'}^{(2)} - M_{n'n,mm'}^{(2)} \right)}_{=: \check{M}_{nn',mm'}^{(2)}} c_m c_{m'} \quad (41)$$

Since the states $\{c_{n'}^\dagger c_n^\dagger | 0 \rangle \mid n' > n\}$ form a basis of the two-electron Hilbert space, considering nn' as the index of a basis state, the $\check{M}_{nn',mm'}^{(2)}$ form a two-electron matrix $\check{\mathbf{M}}^{(2)}$.

The procedure of rewriting operators in second quantization obviously generalizes to observables acting on more than two electrons in the natural way. We note that, while we started from

a form of the operators (30) that was explicitly formulated in an N -electron Hilbert space, the results (31), (35), and (39) are of the same form no matter what value N takes. Thus these operators are valid not just on some N -electron Hilbert space, but on the entire Fock space. This is a particular strength of the second-quantized formalism.

2.4 Transforming the orbital basis

We noted in (16) that the creators transform in the same way as the orbitals they represent

$$|\beta_i\rangle = U|\alpha_i\rangle = \sum_j |\beta_j\rangle \langle \alpha_j | \alpha_i \rangle = \sum_\mu |\alpha_\mu\rangle \underbrace{\langle \alpha_\mu | U | \alpha_i \rangle}_{=: U_{\mu i}} \rightsquigarrow c_{\beta_i}^\dagger = \sum_\mu c_{\alpha_\mu}^\dagger U_{\mu i}, \quad (42)$$

so the “operators” really transform like states. Writing the transformation matrix as $U = e^M$, where M is anti-Hermitian, $M^\dagger = -M$ when U is unitary, but can be any matrix when U is merely invertible, we can write the basis transformation in a form appropriate for operators:

$$c_{\beta_i}^\dagger = e^{c^\dagger M c} c_{\alpha_\mu}^\dagger e^{-c^\dagger M c}. \quad (43)$$

To see this, we use the Baker-Campbell-Hausdorff formula in the form

$$e^{\lambda A} B e^{-\lambda A} = B + \lambda [A, B] + \frac{\lambda^2}{2!} [A, [A, B]] + \frac{\lambda^3}{3!} [A, [A, [A, B]]] + \dots, \quad (44)$$

where the expansion coefficients follow by taking the derivatives of the left hand side at $\lambda = 0$, together with the commutator

$$[c_{\alpha_\mu}^\dagger c_{\alpha_\nu}, c_{\alpha_\kappa}^\dagger] = c_{\alpha_\mu}^\dagger \delta_{\nu, \kappa} \quad (45)$$

from which we obtain for the repeated commutators

$$\left[\sum_{\mu, \nu} M_{\mu\nu} c_{\alpha_\mu}^\dagger c_{\alpha_\nu}, \sum_{\kappa} c_{\alpha_\kappa}^\dagger (M^n)_{\kappa i} \right] = \sum_{\mu, \nu, \kappa} c_{\alpha_\mu}^\dagger M_{\mu\nu} \delta_{\nu, \kappa} (M^n)_{\kappa i} = \sum_{\mu} c_{\alpha_\mu}^\dagger (M^{n+1})_{\mu i}. \quad (46)$$

To keep the derivation simple, we have chosen to transform an operator from the orthonormal basis that we also used to write the exponential operator. Being linear, the transform works, of course, the same for an arbitrary creation operator.

Using this form of the basis transformation and noticing that $e^{-c^\dagger M c} |0\rangle = |0\rangle$, we immediately see that acting with the exponential of a one-body operator on a product state results in another product state

$$e^{c^\dagger M c} \prod c_{\alpha_n}^\dagger |0\rangle = \prod e^{c^\dagger M c} c_{\alpha_n}^\dagger e^{-c^\dagger M c} |0\rangle = \prod c_{\beta_n}^\dagger |0\rangle. \quad (47)$$

This is, e.g., used when working in the interaction picture. Anticommutators with transformed operators, (42), are simply $\{c_{\alpha_j}, e^{-c^\dagger M c} c_{\alpha_i}^\dagger e^{-c^\dagger M c}\} = \langle \alpha_j | e^M | \alpha_i \rangle$.

Annihilation operators, being the adjoint of the creators, transform in just the expected way

$$c_{\beta_i} = e^{-c^\dagger M^\dagger c} c_{\alpha_\mu} e^{c^\dagger M^\dagger c}, \quad (48)$$

which means that for unitary transformations, where M is anti-Hermitian, creators and annihilators transform in the same way. Note that in the imaginary-time formalism the annihilators are, via analytic continuation, chosen to transform in the same way as the creators, making them different from the adjoint of the creators.

3 Exact diagonalization

We have worked, so far, with complete, i.e., infinite bases. This is, of course, not possible in actual computer simulations, where we have to confine ourselves to finite basis sets. Such calculations on subspaces are based on the variational principle.

3.1 Variational principles

The variational principle and the Schrödinger equation are equivalent. Consider the energy expectation value as a wave-function functional

$$E[\Psi] = \frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle}. \quad (49)$$

Its variation is

$$E[\Psi + \delta\Psi] = E[\Psi] + \frac{\langle \delta\Psi | H | \Psi \rangle + \langle \Psi | H | \delta\Psi \rangle}{\langle \Psi | \Psi \rangle} - \langle \Psi | H | \Psi \rangle \frac{\langle \delta\Psi | \Psi \rangle + \langle \Psi | \delta\Psi \rangle}{\langle \Psi | \Psi \rangle^2} + \mathcal{O}^2. \quad (50)$$

The first-order term vanishes for $H|\Psi\rangle = E[\Psi]|\Psi\rangle$, which is the Schrödinger equation. Since the eigenfunctions

$$H|\Psi_n\rangle = E_n|\Psi_n\rangle, \quad (51)$$

can be chosen to form an orthonormal basis, we can expand any wavefunction as

$$|\Psi\rangle = \sum_n |\Psi_n\rangle \langle \Psi_n | \Psi \rangle \quad (52)$$

and determine, as long as $\langle \Psi | \Psi \rangle \neq 0$, its energy expectation value

$$\frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \frac{\sum_{m,n} \langle \Psi | \Psi_m \rangle \langle \Psi_m | H | \Psi_n \rangle \langle \Psi_n | \Psi \rangle}{\sum_{m,n} \langle \Psi | \Psi_m \rangle \langle \Psi_m | \Psi_n \rangle \langle \Psi_n | \Psi \rangle} = \frac{\sum_n E_n |\langle \Psi_n | \Psi \rangle|^2}{\sum_n |\langle \Psi_n | \Psi \rangle|^2}. \quad (53)$$

Since by definition no eigenenergy can be lower than the ground state energy E_0 , we immediately see that the energy expectation value can never drop below the ground state energy

$$\frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \frac{\sum_n E_n |\langle \Psi_n | \Psi \rangle|^2}{\sum_n |\langle \Psi_n | \Psi \rangle|^2} \geq \frac{\sum_n E_0 |\langle \Psi_n | \Psi \rangle|^2}{\sum_n |\langle \Psi_n | \Psi \rangle|^2} = E_0. \quad (54)$$

We can use the same argument to generalize this variational principle: Assume we have arranged the eigenenergies in ascending order, $E_0 \leq E_1 \leq \dots$, then the energy expectation value for a wavefunction that is orthogonal to the n lowest eigenstates, can not drop below E_n

$$\frac{\langle \Psi_{\perp n} | H | \Psi_{\perp n} \rangle}{\langle \Psi_{\perp n} | \Psi_{\perp n} \rangle} \geq E_n \quad \text{if } \langle \Psi_i | \Psi_{\perp n} \rangle = 0 \text{ for } i = 0, \dots, n-1. \quad (55)$$

This generalized variational principle is, of course, only of practical use if we know something about the eigenstates, e.g., when we can use symmetries to ensure orthogonality.

For an ab-initio Hamiltonian of N electrons in the field of nuclei of charge Z_α at position $\mathbf{R}_\alpha$,

$$H = -\frac{1}{2} \sum_i \Delta_i - \sum_{i,\alpha} \frac{Z_\alpha}{|\mathbf{r}_i - \mathbf{R}_\alpha|} + \sum_{i<j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{\alpha<\beta} \frac{Z_\alpha Z_\beta}{|\mathbf{R}_\alpha - \mathbf{R}_\beta|}, \quad (56)$$

the Schrödinger equation is a partial differential equation. In second quantization it becomes a linear-algebra problem: We introduce an orbital basis set $\{\varphi_k | k\}$, which for simplicity we assume here to be orthonormal, from which we construct an orthonormal basis of N -electron product states, $\{\Phi_{k_1, \dots, k_N} | k_1 < \dots < k_N\}$. To simplify the notation we sort the basis states, e.g., lexicographically in the orbital indices $\mathbf{k} = (k_1, \dots, k_N)$ and define the row vector of basis states $|\Phi\rangle := (|\Phi_1\rangle, |\Phi_2\rangle, \dots)$. The expansion of a state $|\Psi\rangle$ in this basis can then be written as

$$|\Psi\rangle = \sum_{k_1 < \dots < k_N} a_{k_1, \dots, k_N} |\Phi_{k_1, \dots, k_N}\rangle = \sum_i a_i |\Phi_i\rangle = |\Phi\rangle \mathbf{a}, \quad (57)$$

where $\mathbf{a}$ is the vector of expansion coefficients. Likewise we can write the Schrödinger equation as a matrix eigenvalue problem

$$\mathbf{H}\mathbf{a} = \langle \Phi | \hat{H} | \Phi \rangle \mathbf{a} = \begin{pmatrix} \langle \Phi_1 | \hat{H} | \Phi_1 \rangle & \langle \Phi_1 | \hat{H} | \Phi_2 \rangle & \dots \\ \langle \Phi_2 | \hat{H} | \Phi_1 \rangle & \langle \Phi_2 | \hat{H} | \Phi_2 \rangle & \dots \\ \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \end{pmatrix} = E \begin{pmatrix} a_1 \\ a_2 \\ \vdots \end{pmatrix} = E\mathbf{a}. \quad (58)$$

From the eigenvectors of the matrix $\mathbf{H}$ we easily recover the eigenstates of the Hamiltonian

$$\mathbf{H}\mathbf{a}_n = E_n \mathbf{a}_n \quad \rightsquigarrow \quad \hat{H}|\Psi_n\rangle = E_n |\Psi_n\rangle \quad \text{with } |\Psi_n\rangle = |\Phi\rangle \mathbf{a}_n. \quad (59)$$

Unfortunately, for an ab-initio Hamiltonian like (56) we need an infinite orbital basis set, so that the Hamiltonian matrix $\mathbf{H}$ is infinite dimensional. A pragmatic approach to allow for computer simulations is to simply restrict the calculation to a finite basis $|\tilde{\Phi}\rangle := (|\Phi_1\rangle, \dots, |\Phi_{\tilde{L}}\rangle)$, i.e., work with a finite matrix $\tilde{\mathbf{H}} := \langle \tilde{\Phi} | \hat{H} | \tilde{\Phi} \rangle$ of dimension $\tilde{L}$. The crucial question is then how the eigenvectors

$$\tilde{\mathbf{H}}\tilde{\mathbf{a}}_n = \tilde{E}_n \tilde{\mathbf{a}}_n \quad \rightsquigarrow \quad |\tilde{\Psi}_n\rangle := |\tilde{\Phi}\rangle \tilde{\mathbf{a}}_n \quad (60)$$

are related to those of H . The answer is surprisingly simple [7]: The eigenvalues of $\tilde{\mathbf{H}}$, ordered as $\tilde{E}_0 \leq \tilde{E}_1 \leq \dots \leq \tilde{E}_{\tilde{L}-1}$, are variational with respect to those of H :

$$E_n \leq \tilde{E}_n \quad \text{for } n \in \{0, \dots, \tilde{L}-1\}. \quad (61)$$

To show this, we construct a state in $\text{span}(|\tilde{\Psi}_0\rangle, \dots, |\tilde{\Psi}_n\rangle)$, which by construction has an energy expectation value $\leq \tilde{E}_n$, that is orthogonal to the exact eigenstates $|\Psi_0\rangle, \dots, |\Psi_{n-1}\rangle$, so that by the generalized variational principle its expectation value is $\geq E_n$. Being the non-zero solution of $n+1$ linear equations with $n+1$ variables, such a state certainly exists, hence $E_n \leq \tilde{E}_n$.

To get reliable results, we simply have to systematically increase the basis until the change in the desired eigenvalues becomes smaller than the accuracy required by the physical problem. The art is, of course, to devise clever basis sets such that this is achieved already for bases of low dimensions.

The convergence of the matrix eigenvalues with increasing basis size is surprisingly regular. Let us extend our original basis of $\tilde{L}$ states by an additional $L - \tilde{L}$ states. Then, repeating the above argument with the L -dimensional problem taking the role of $\hat{H}$, we obtain (61) with E_n being the eigenvalues of the L -dimensional Hamiltonian matrix $\mathbf{H}$. Since $\mathbf{H}$ now is finite, we can use the same argument for $-\mathbf{H}$, obtaining

$$-E_{L-i} \leq -\tilde{E}_{\tilde{L}-i} \quad \text{for } i \in \{1, \dots, \tilde{L}\}. \quad (62)$$

Taking the two inequalities together we obtain

$$E_n \leq \tilde{E}_n \leq E_{n+(L-\tilde{L})} \quad \text{for } n \in \{0, \dots, \tilde{L}-1\}. \quad (63)$$

For the special case $L = \tilde{L}+1$ of adding a single basis state, this is the Hylleraas-Undheim/MacDonald nesting property for eigenvalues in successive approximations

$$E_1 \leq \tilde{E}_1 \leq E_2 \leq \tilde{E}_2 \leq \dots \leq \tilde{E}_L \leq E_{L+1}. \quad (64)$$

3.2 Matrix eigenvalue problem

For practical calculations we have to set up the Hamiltonian matrix $\tilde{\mathbf{H}} = \langle \tilde{\Phi} | \hat{H} | \tilde{\Phi} \rangle$ and the state vectors $\tilde{\mathbf{a}}$ for the chosen basis. This is particularly easy for a basis of Slater determinants constructed from a basis set of K orbitals $\{\varphi_k \mid k = 0, \dots, K-1\}$. The basis states are then the N -electron product states of $|\Phi_{k_1, \dots, k_N}\rangle = c_{k_N}^\dagger \dots c_{k_1}^\dagger |0\rangle$ with $k_1 < \dots < k_N$. We can write them more computer friendly as

$$|n_{K-1}, \dots, n_0\rangle = \prod_{k=0}^{K-1} (c_k^\dagger)^{n_k} |0\rangle \quad (65)$$

which is the occupation number representation with $n_k \in \{0, 1\}$ and $\sum n_k = N$. It is natural to interpret the vector of occupation numbers as the binary representation of the integer $\sum_k 2^{n_k}$. This implies a natural ordering of the basis functions $|\Phi_l\rangle$. For the simple case of $K=4$ orbitals and $N=2$ electrons we obtain

i	(n_3, n_2, n_1, n_0)	state	l
0	0000		
1	0001		
2	0010		
3	0011	$c_1^\dagger c_0^\dagger 0\rangle = \Phi_1\rangle$	1
4	0100		
5	0101	$c_2^\dagger c_0^\dagger 0\rangle = \Phi_2\rangle$	2
6	0110	$c_2^\dagger c_1^\dagger 0\rangle = \Phi_3\rangle$	3
7	0111		
8	1000		
9	1001	$c_3^\dagger c_0^\dagger 0\rangle = \Phi_4\rangle$	4
10	1010	$c_3^\dagger c_1^\dagger 0\rangle = \Phi_5\rangle$	5
11	1011		
12	1100	$c_3^\dagger c_2^\dagger 0\rangle = \Phi_6\rangle$	6
13	1101		
14	1110		
15	1111		

The bit representation of the basis states also simplifies setting up the Hamiltonian matrix. Given the Hamiltonian in second quantization

$$\hat{H} = \sum_{n,m} T_{nm} c_n^\dagger c_m + \sum_{n'>n, m'>m} \underbrace{(U_{nn',mm'} - U_{n'n,mm'})}_{=\check{U}_{nn',mm'}} c_{n'}^\dagger c_n^\dagger c_m c_{m'} \quad (66)$$

the matrix element $\langle \Phi_l | \hat{H} | \Phi_{l'} \rangle$, with $|\Phi_{l'}\rangle = c_{k'_N}^\dagger \cdots c_{l'_1}^\dagger |0\rangle$, is given by

$$\sum_{n,m} T_{nm} \langle 0 | c_{l_1} \cdots c_{l_N} c_n^\dagger c_m c_{l'_N}^\dagger \cdots c_{l'_1}^\dagger | 0 \rangle + \sum_{\substack{n'>n \\ m'>m}} \check{U}_{nn',mm'} \langle 0 | c_{l_1} \cdots c_{l_N} c_n^\dagger c_{n'}^\dagger c_m c_{m'} c_{l'_N}^\dagger \cdots c_{l'_1}^\dagger | 0 \rangle.$$

Anticommuting the operators coming from the Hamiltonian, the matrix elements become overlaps of $N+1$ and $N+2$ -electron product states, which, by (33) and (7), are just the determinants of the overlap matrices of the corresponding orbitals. When $\hat{H}$ is written in the same orbitals as these $|\Phi_l\rangle$, the overlap matrices simplify to permutation matrices with determinant ± 1 . In the occupation number representation, calculating this Fermi sign reduces to counting set bits. As an example we consider a simple hopping of an electron:

$$\begin{aligned} c_6^\dagger c_2 |\Phi_{l(181)}\rangle &= c_6^\dagger c_2 c_7^\dagger c_5^\dagger c_4^\dagger c_2^\dagger c_0^\dagger |0\rangle \\ &= (-1)^3 c_6^\dagger c_7^\dagger c_5^\dagger c_4^\dagger c_2^\dagger c_0^\dagger |0\rangle \\ &= (-1)^3 c_6^\dagger c_7^\dagger c_5^\dagger c_4^\dagger (1 - c_2^\dagger c_2) c_0^\dagger |0\rangle \\ &= (-1)^3 c_6^\dagger c_7^\dagger c_5^\dagger c_4^\dagger c_0^\dagger |0\rangle \\ &= + |\Phi_{l(241)}\rangle = (-1)^2 c_7^\dagger c_6^\dagger c_5^\dagger c_4^\dagger c_0^\dagger |0\rangle \end{aligned}$$

In the occupation number representation this becomes

$$\begin{array}{ccccccc} & \overbrace{}^{\downarrow} & & & & & \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 \end{array} = (-1)^c \begin{array}{ccccccc} 1 & 1 & 1 & 1 & 0 & 0 & 1 \end{array}$$

where c is the count of set bits between the orbitals of the electron hop. Note that a dedicated machine instruction, `popcnt`, for counting set bits is part of the x86 SSE4 instruction set, see also [8].

3.3 Dimension of the Hilbert space and sparseness

Setting up basis states and Hamiltonian matrix in this way, we can easily solve the many-body problem on our variational space by using any linear algebra library. This is the exact diagonalization approach. As discussed above, it gives us variational estimates of the ground and excited states. But there is a serious practical problem: the dimension of the many-body Hilbert space. For an N -electron problem with a basis set of K orbitals there are $K(K-1)(K-2)\cdots(K-(N-1))$ ways of picking N occupied orbitals out of K . Since we only use one specific ordering of these indices, we have to divide by $N!$ to obtain the number of such determinants:

$$\dim \mathcal{H}_K^{(N)} = \frac{K!}{N!(K-N)!} = \binom{K}{N}. \quad (67)$$

Using Stirling's formula we see that for an N -electron problem this increases faster than exponentially with the size K of the basis set. This is the problem we face when converging the basis set for a finite system, e.g., a molecule. For solids we usually keep the number of orbitals per lattice site fixed, but scale to the thermodynamic limit, increasing the system size M while keeping the electron density N/M fixed. Also here the Hilbert space increases faster than exponentially. To give an impression of the problem we note that for $N = 25$ electrons and $K = 100$ orbitals the dimension already exceeds 10^{23} .

For exact diagonalization the problem gets even worse. Assuming we have a machine with 1 TeraBytes = 2^{40} Bytes of RAM available. Using single precision (4 bytes) for the matrix elements, storing a matrix of dimension $(2^{40}/4)^{1/2} = 524\,288$ would already use up all memory. The dimension problem can be somewhat mitigated by exploiting symmetries: When the Hamiltonian commutes with the projection of the total spin, the number of up- and down-spin electrons is conserved separately. The N -electron Hamiltonian is then block diagonal in the sectors with fixed $N_\uparrow$ and $N_\downarrow$. The dimension of these blocks is significantly smaller than that of the full N -electron Hilbert space. Using the same orbital basis for each spin

$$\dim \mathcal{H}_{2K}^{(N_\uparrow, N_\downarrow)} = \binom{K}{N_\uparrow} \times \binom{K}{N_\downarrow}. \quad (68)$$

The S_z symmetry can be very easily implemented using the same ideas as introduced for the general case: just use bit representations for the up- and down-spin electrons separately. In fact, when the total spin projection is conserved, we can distinguish electrons of different spin. Still, the Hilbert space of the single-band, half-filled Hubbard model with just 12 sites has dimension 853 776. Using further symmetries, if they exist, we could bring down the dimension somewhat further, however at the expense of considerable and problem-specific effort.

The key to going to larger systems is the realization that the vast majority of the elements of the Hamiltonian matrix is zero. This is quite easy to see. For the ab-initio Hamiltonian (56) with electron-electron repulsion, matrix elements between configurations that differ in more than two electron occupations vanish. Thus, for each configuration there may only be the diagonal element, $N \times (K-1)$ hopping terms, and $N(N-1)/2 \times (K-N)(K-N-1)/2$ pair-hopping terms. Thus the fraction of non-zero matrix elements of $\tilde{\mathbf{H}}$ to the total number is

$$\left(1 + N \left(1 + \frac{N-1}{2} \frac{(K-N-1)}{2} \right) (K-N) \right) / \binom{K}{N} \quad (69)$$

which, with increasing problem size, rapidly approaches zero. For the example of $N = 25$ electrons in $K = 100$ orbitals only 834 376 of the (over 10^{23}) matrix elements per row can be non-zero. This is the worst case. The sparsity of many-body Hamiltonians is even more pronounced when working in a tight-binding basis with short-ranged hopping and local electron-electron repulsion. Thus, many-body Hamiltonians are exceedingly sparse and the more so the larger the problem. They are therefore ideally suited for approaches like the Lanczos method, that are based on matrix-vector products, which for the sparse matrices scale close to linearly in the matrix dimension.

4 Lanczos Method

As we have seen, we can find the ground-state $|\Psi_0\rangle$ and its energy E_0 for a Hamiltonian H from the variational principle. The wavefunction-functional

$$E[\Psi] = \frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle} \quad (70)$$

is minimized for $\Psi = \Psi_0$, with $E[\Psi_0] = E_0$. The functional gradient

$$\frac{\delta E[\Psi]}{\delta \langle \Psi |} = \frac{H|\Psi\rangle - E[\Psi]|\Psi\rangle}{\langle \Psi | \Psi \rangle} = |\Psi_a\rangle \quad (71)$$

gives the direction of steepest-ascent of the functional from the point $|\Psi\rangle$. Moving in the opposite direction will thus result in a wavefunction with lower energy expectation value: $E[\Psi - \alpha\Psi_a] < E[\Psi]$ for small, positive α .

To find the optimum value of α , we minimize the quadratic form $E[\Psi - \alpha\Psi_a]$. For this, it is convenient to introduce an orthogonal basis in the space spanned by the two vectors $|\Psi\rangle$ and $|\Psi_a\rangle$. From (71) we see that $\text{span}(|\Psi\rangle, |\Psi_a\rangle) = \text{span}(|\Psi\rangle, H|\Psi\rangle)$. As first basis vector, we normalize $|\Psi\rangle$

$$|v_0\rangle = |\Psi\rangle / \sqrt{\langle \Psi | \Psi \rangle},$$

for the second vector we orthogonalize $H|v_0\rangle$ to $|v_0\rangle$

$$|\tilde{v}_1\rangle = H|v_0\rangle - |v_0\rangle\langle v_0|H|v_0\rangle \quad (72)$$

and normalize to obtain $|v_1\rangle$. With $a_n := \langle v_n | H | v_n \rangle$ and $b_1^2 := \langle \tilde{v}_1 | \tilde{v}_1 \rangle$, eq. (72) becomes

$$H|v_0\rangle = b_1|v_1\rangle + a_0|v_0\rangle \quad (73)$$

from which we see that $\langle v_1 | H | v_0 \rangle = b_1$.

We can then write any normalized wavefunction in $\text{span}(|\Psi\rangle, H|\Psi\rangle) = \text{span}(|v_0\rangle, |v_1\rangle)$ as

$$|v\rangle = \cos(\theta)|v_0\rangle + \sin(\theta)|v_1\rangle. \quad (74)$$

Minimizing the expectation value

$$\langle v | H | v \rangle = a_0 \cos^2(\theta) + 2b_1 \sin(\theta) \cos(\theta) + a_1 \sin^2(\theta), \quad (75)$$

with respect to θ , we obtain, dividing by $\cos^2(\theta)$, the quadratic equation

$$b_1 \tan^2(\theta) + (a_0 - a_1) \tan(\theta) - b_1 = 0. \quad (76)$$

Solving for θ we find the lowest-energy state on the subspace spanned by $|v_0\rangle$ and $H|v_0\rangle$. Alternatively, we can diagonalize the Hamiltonian matrix on the two-dimensional subspace, which in the basis $|v_0\rangle, |v_1\rangle$ is given by

$$H_{\text{span}(|\Psi\rangle, H|\Psi\rangle)} = \begin{pmatrix} a_0 & b_1 \\ b_1 & a_1 \end{pmatrix}. \quad (77)$$

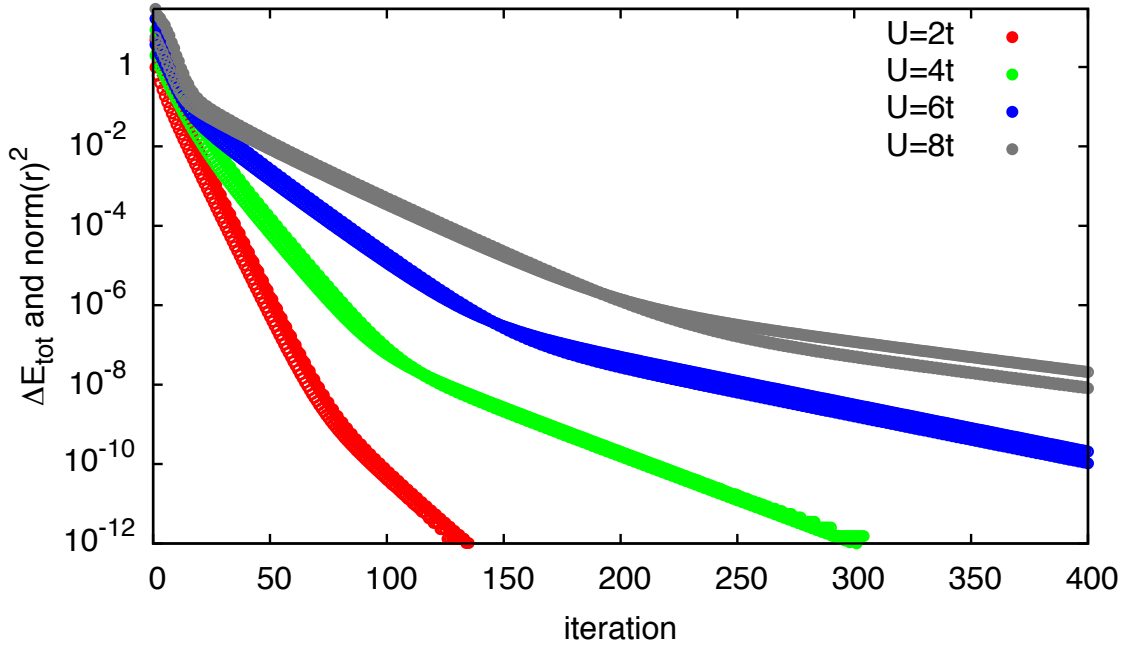


Fig. 1: Convergence of the residual (filled circles) and the corresponding lowest eigenvalue (open circles) for a steepest-descent minimization of a Hubbard-chain of 10 sites at half-filling, starting from a random initial vector.

Naturally, we can use the variational state of lowest energy

$$|\Psi^{(2)}\rangle = \cos(\theta_{\min})|v_0\rangle + \sin(\theta_{\min})|v_1\rangle \quad (78)$$

as the starting point for another steepest-descent minimization. Doing this repeatedly, we obtain a series of vectors with decreasing energy expectation value, which rapidly converge to a minimum. For a generic functional such a steepest-descent minimization would usually end up in a local, not the global minimum, which makes the optimization of high-dimensional functions such a hard problem. The energy functional (70), however, has only minima for the ground-states, all other stationary points are saddle points. We can thus expect rapid convergence to the ground state, examples given in figure 1, except when the starting vector $|v_0\rangle$ is orthogonal to the ground state. In this case also $\langle\Psi_0|H|v_0\rangle = 0$ so that we only converge to the lowest state that overlaps with $|v_0\rangle$. Should $|v_0\rangle$ happen to be an exact eigenvector, $H|v_n\rangle$ does not add a new dimension so that the algorithm terminates with $b_1 = 0$.

For checking convergence of this steepest-descent method, introduced by Kantorovich [9] and, independently, by Hestenes and Karush [10], we can monitor the change in the energy expectation value or determine when the residual

$$r[\Psi] = \|(H - E[\Psi])|\Psi\rangle\|^2 = \langle\Psi|H^2|\Psi\rangle - E[\Psi]^2, \quad (79)$$

which measures the quality of the eigenstate, becomes sufficiently small. As shown in Fig. 1, both are closely related.

4.1 Krylov space

If we apply the method of steepest-descent L times, starting from a vector $|v_0\rangle$, the resulting vector will lie in

$$\mathcal{K}^L(|v_0\rangle) = \text{span}(|v_0\rangle, H|v_0\rangle, H^2|v_0\rangle, \dots, H^L|v_0\rangle), \quad (80)$$

the $L+1$ -dimensional *Krylov space* [11] of H over $|v_0\rangle$. Instead of repeatedly minimizing the energy in two-dimensional subspaces, we could directly find the state of lowest energy in $\mathcal{K}^L(|v_0\rangle)$. Having more degrees of freedom for the minimization will lead to even faster convergence.

To implement this idea, we construct an orthonormal basis $|v_n\rangle$ of the Krylov space. We start with the normalized vector $|v_0\rangle$. The second basis vector $|v_1\rangle$ is constructed as in the steepest-descent method (72):

$$b_1|v_1\rangle = |\tilde{v}_1\rangle = H|v_0\rangle - a_0|v_0\rangle. \quad (81)$$

The next basis vector is likewise constructed as $H|v_n\rangle$ orthogonalized to all previous vectors, and normalized

$$b_2|v_2\rangle = |\tilde{v}_2\rangle = H|v_1\rangle - \sum_{i=0}^1 |v_i\rangle\langle v_i|H|v_1\rangle = H|v_1\rangle - a_1|v_1\rangle - b_1|v_0\rangle. \quad (82)$$

where $a_n = \langle v_n|H|v_n\rangle$ and $b_n^2 = \langle \tilde{v}_n|\tilde{v}_n\rangle$. The fourth basis vector is

$$b_3|v_3\rangle = |\tilde{v}_3\rangle = H|v_2\rangle - \sum_{i=0}^2 |v_i\rangle\langle v_i|H|v_2\rangle = H|v_2\rangle - a_2|v_2\rangle - b_2|v_1\rangle. \quad (83)$$

The last term in the orthogonalization vanishes, because (81) together with the orthogonality of the already constructed basis vectors for $n < 3$ implies $\langle v_2|H|v_0\rangle = b_1\langle v_2|v_1\rangle + a_0\langle v_2|v_0\rangle = 0$. The construction of the further basis vectors follows the same scheme

$$b_{n+1}|v_{n+1}\rangle = |\tilde{v}_{n+1}\rangle = H|v_n\rangle - \sum_{i=0}^n |v_i\rangle\langle v_i|H|v_n\rangle = H|v_n\rangle - a_n|v_n\rangle - b_n|v_{n-1}\rangle$$

with $a_n = \langle v_n|H|v_n\rangle$ and $b_n^2 = \langle \tilde{v}_n|\tilde{v}_n\rangle$. Rearranging shows that H is tridiagonalized

$$H|v_n\rangle = b_n|v_{n-1}\rangle + a_n|v_n\rangle + b_{n+1}|v_{n+1}\rangle$$

which in turn implies that $H|v_i\rangle$ is orthogonal to all basis states, except $|v_i\rangle$ and $|v_{i\pm 1}\rangle$. This tridiagonalization of H is the essence of the *Lanczos method* [12].

After L steps the Hamiltonian on the $L+1$ -dimensional Krylov space is given by

$$H_{\mathcal{K}^L(|v_0\rangle)} = \begin{pmatrix} a_0 & b_1 & 0 & 0 & & 0 & 0 \\ b_1 & a_1 & b_2 & 0 & \cdots & 0 & 0 \\ 0 & b_2 & a_2 & b_3 & & 0 & 0 \\ 0 & 0 & b_3 & a_3 & & 0 & 0 \\ & \vdots & & & \ddots & \vdots & \\ 0 & 0 & 0 & 0 & & a_{L-1} & b_L \\ 0 & 0 & 0 & 0 & \cdots & b_L & a_L \end{pmatrix} \quad (84)$$

v=init	
b0=norm2(v)	not part of tridiagonal matrix
scal(1/b0,v)	$v = v_0\rangle$
w=0	
w=w+H*v	$w = H v_0\rangle$
a[0]=dot(v,w)	
axpy(-a[0],v,w)	$w = \tilde{v}_1\rangle = H v_0\rangle - a_0 v_0\rangle$
b[1]=norm2(w)	
for n=1,2,...	
if abs(b[n])<eps then exit	invariant subspace
scal(1/b[n],w)	$w = v_n\rangle$
scal(-b[n],v)	$v = -b_n v_{n-1}\rangle$
swap(v,w)	
w=w+H*v	$w = H v_n\rangle - b_n v_{n-1}\rangle$
a[n]=dot(v,w)	$a[n] = \langle v_n H v_n\rangle - b_n\langle v_n v_{n-1}\rangle$
axpy(-a[n],v,w)	$w = \tilde{v}_{n+1}\rangle$
b[n+1]=norm2(w)	
diag(a[0]..a[n], b[1]..b[n])	getting a_{n+1} needs another $H v\rangle$
if converged then exit	
end	

Table 1: The implementation of the Lanczos iteration requires only two N -dimensional vectors for tridiagonalizing H and thus for calculating the ground-state energy. Constructing the Lanczos-approximation of the ground-state vector requires a second iteration and one additional N -dimensional vector. The by far most expensive operation is the matrix-vector product, which requires a problem specific implementation, while the vector operations use the BLAS.

If we do not normalize the basis vectors, we obtain an iteration of the form

$$|\Phi_{n+1}\rangle = H|\Phi_n\rangle - \frac{\langle\Phi_n|H|\Phi_n\rangle}{\langle\Phi_n|\Phi_n\rangle}|\Phi_n\rangle - \frac{\langle\Phi_n|\Phi_n\rangle}{\langle\Phi_{n-1}|\Phi_{n-1}\rangle}|\Phi_{n-1}\rangle \quad (85)$$

where $|\Phi_n\rangle = \prod_{i=1}^n b_i |v_n\rangle$ in terms of which we have

$$a_n = \frac{\langle\Phi_n|H|\Phi_n\rangle}{\langle\Phi_n|\Phi_n\rangle}, \quad b_n^2 = \frac{\langle\Phi_n|\Phi_n\rangle}{\langle\Phi_{n-1}|\Phi_{n-1}\rangle}. \quad (86)$$

In this unnormalized basis the Hamiltonian appears non-Hermitian

$$H|\Phi_n\rangle = b_n^2|\Phi_{n-1}\rangle + a_n|\Phi_n\rangle + |\Phi_{n+1}\rangle, \quad (87)$$

but, of course, it actually is

$$\langle\Phi_{n+1}|H|\Phi_n\rangle = \langle\Phi_{n+1}|\Phi_{n+1}\rangle = b_{n+1}^2\langle\Phi_n|\Phi_n\rangle = \langle\Phi_n|H|\Phi_{n+1}\rangle. \quad (88)$$

The numerical implementation only requires keeping two N -dimensional vectors in memory. It is shown in table 1.

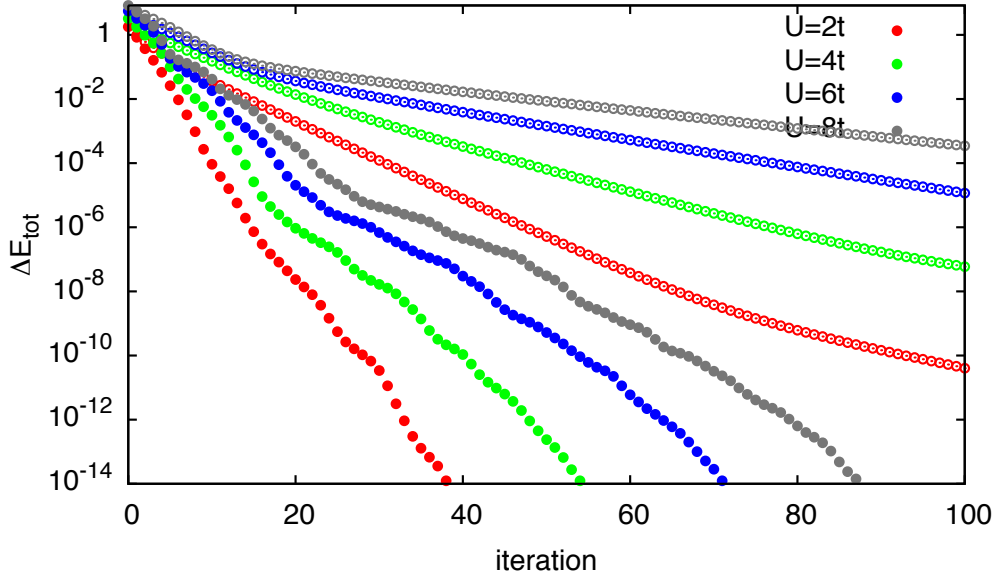


Fig. 2: Convergence of the lowest eigenvalue for a Lanczos iteration (full circles) compared to steepest-descent minimization (open circles) of a 10-site Hubbard-chain at half-filling, starting from a random initial vector. Due to the additional variational degrees of freedom, Lanczos converges significantly faster. Overall, convergence for the half-filled system gets harder for larger U , as the distance to the lowest excited states is reduced ($\sim t^2/U$) and the spectrum widens ($\sim U$). In all cases, convergence is reached after less than $L \approx 100$ Lanczos iterations, to be compared to the dimension $N=63\,504$ of the Hilbert space.

Diagonalizing (84), after a few tens to hundred iterations, the lowest eigenvalue of the tridiagonal representation of H on the Krylov space gives an excellent approximation to the ground-state energy of H in the full Hilbert space (Fig. 2). A formal estimate of the convergence was given by Kaniel and Paige [13]: For an $N+1$ -dimensional, symmetric matrix H with eigenvalues E_n , the lowest eigenvalue $\check{E}_0$ of the tridiagonal representation of H on the $(L+1)$ -dimensional Krylov space over $|v_0\rangle$ fulfills

$$\frac{\check{E}_0 - E_0}{E_N - E_0} \leq \left(\frac{\tan(\arccos(\langle \check{\Psi}_0 | \Psi_0 \rangle))}{T_L \left(1 + 2 \frac{E_1 - E_0}{E_N - E_1} \right)} \right)^2 \quad (89)$$

where $T_L(x)$ is the Chebyshev polynomial of order L and $\langle \check{\Psi}_0 | \Psi_0 \rangle$ the overlap of the Lanczos approximation to the ground-state $\check{\Psi}_0$ with the ground-state of H . Thus, if the initial state $|v_0\rangle$ is not orthogonal to the non-degenerate ground-state, convergence is exponential with a rate roughly increasing with the square root of the gap to the first excited state measured in units of the width of the spectrum.

The approximate ground-state vector is given by the linear combination

$$|\check{\Psi}_0\rangle = \sum_{n=0}^L \check{\psi}_{0,n} |v_n\rangle, \quad (90)$$

where $\check{\psi}_0$ is the ground-state vector of the $L+1$ -dimensional tridiagonal matrix (84). Instead of storing all $L+1$ basis vectors $|v_n\rangle$, we can restart the Lanczos iteration from the same $|v_0\rangle$,

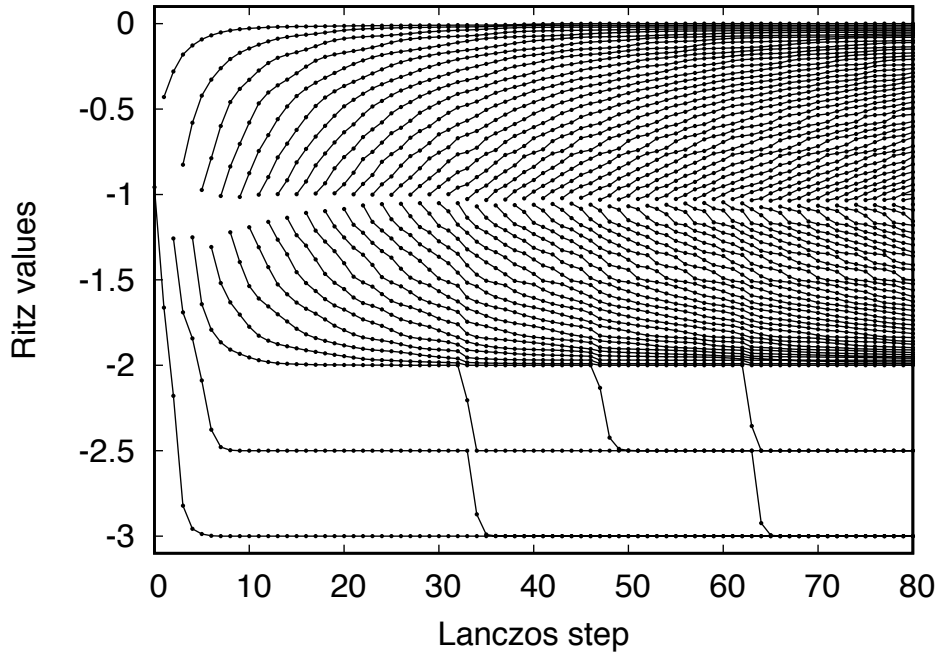


Fig. 3: *Lanczos method for a matrix with eigenvalues $-3, -3, -2.5, -2, -1.99, -1.98, \dots, -0.01, 0$. For clarity the n -th lowest/highest Lanczos eigenvalues are connected by lines. Note how the eigenvalues of successive steps are nested as described by (64). Because of the gap the lowest eigenvalue is reached much faster than the highest. For the degenerate eigenspace of value -3 only the state proportional to the projection of $|v_0\rangle$ on that space is found. Well after convergence of the low eigenvalues orthogonality is lost and ghost states appear.*

accumulating the sum (90) iteration by iteration. This only requires keeping one additional N -dimensional vector in memory.

So far we have tacitly assumed that the Krylov vectors $H^n|v_0\rangle$ are linearly independent. If not, there will be a vector $H|\tilde{v}_m\rangle$ that vanishes when orthogonalized to the previous states, i.e., $b_n = 0$. This means that the Krylov space $\text{span}(|v_0\rangle, |v_1\rangle, \dots, |v_m\rangle)$ is invariant under H , i.e., we have found an exact eigenspace of H . For a large matrix H it is quite unlikely to be that lucky. Still, as the Lanczos iteration approaches the ground-state, we encounter a similar situation: Close to an eigenstate, the functional (70) becomes almost stationary, i.e., the coefficients b_n almost vanish. Normalization of the very short vector $|\tilde{v}_n\rangle$ then amplifies numerical noise in that vector. This makes the numerical $|v_n\rangle$, which in theory should automatically be orthogonal to all $|v_m\rangle$ with $m < n-2$, actually have finite overlaps with these vectors. This loss of orthogonality manifests itself in the appearance of multiple copies of eigenvectors (ghost states) which are unrelated to the actual multiplicities of the eigenvalues. This makes the Lanczos method unpractical for tridiagonalizing dense matrices. For the ground-state the variational principle prevents severe problem from the loss of orthogonality. An example of the appearance of ghost states is shown in figure 3.

If we want to reliably obtain excited states, we need to explicitly orthogonalize to the previous basis states. This leads to the Lanczos method with complete reorthogonalization [13]. A similar orthogonalization is performed in the Arnoldi method [14], which, however, is devised for unsymmetric matrices.

4.2 Spectral functions

The Lanczos method gives excellent approximations to the largest and smallest eigenvalues, but, as seen in Fig. 3, not for the bulk of the spectrum. It therefore looks ill suited for determining functions that depend on the excited states, like the Lehmann representation

$$G_c(z) = \left\langle \Psi_c \left| \frac{1}{z - H} \right| \Psi_c \right\rangle = \sum_{n=0}^N \frac{\langle \Psi_c | \Psi_n \rangle \langle \Psi_n | \Psi_c \rangle}{z - E_n} \quad (91)$$

which, in terms of the eigenstates on the Krylov space $\mathcal{K}^L(|\Psi_c\rangle)$, would be written as

$$\check{G}_c(z) = \left\langle \Psi_c \left| \frac{1}{z - \check{H}_c} \right| \Psi_c \right\rangle = \sum_{n=0}^L \frac{\langle \Psi_c | \check{\Psi}_n \rangle \langle \check{\Psi}_n | \Psi_c \rangle}{z - \check{E}_n}. \quad (92)$$

This is straightforward to calculate: We run L Lanczos iterations, starting from the (normalized) vector $|\Psi_c\rangle$, to create the tridiagonal $\check{H}_c$. The matrix element of the resolvent is the top left matrix element of the inverse of

$$z - \check{H}_c = \begin{pmatrix} z - a_0 & -b_1 & 0 & 0 & \cdots & 0 & 0 \\ -b_1 & z - a_1 & -b_2 & 0 & \cdots & 0 & 0 \\ 0 & -b_2 & z - a_2 & -b_3 & \cdots & 0 & 0 \\ 0 & 0 & -b_3 & z - a_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & z - a_{L-1} & -b_L \\ 0 & 0 & 0 & 0 & \cdots & -b_L & z - a_L \end{pmatrix}. \quad (93)$$

This is easily determined, partitioning the matrix as indicated

$$z - \check{H}_c = \begin{pmatrix} z - a_0 & B^{(1)T} \\ B^{(1)} & z - \check{H}_c^{(1)} \end{pmatrix} \quad (94)$$

and inverting the block-matrix, giving, see appendix A.3,

$$[(z - \check{H}_c)^{-1}]_{00} = \left(z - a_0 - B^{(1)T} (z - \check{H}_c^{(1)})^{-1} B^{(1)} \right)^{-1} = (z - a_0 - b_1^2 [(z - \check{H}_c^{(1)})^{-1}]_{00})^{-1}.$$

Repeating inversion by partitioning for the submatrices $\check{H}_c^{(n)}$ we obtain the continued fraction

$$\check{G}_c(z) = [(z - \check{H}_c)^{-1}]_{00} = \frac{1}{z - a_0 - \frac{b_1^2}{z - a_1 - \frac{b_2^2}{z - a_2 - \cdots}}}, \quad (95)$$

which terminates with $-b_L^2/(z - a_L)$. We find the spectral representation (92) by diagonalizing the Lanczos matrix $\check{H}_c$ giving us the $L+1$ eigenvalues $\check{E}_n$ and eigenvectors $\check{\psi}_n$. Since

$$|\check{\Psi}_n\rangle = \sum_{l=0}^L \check{\psi}_{n,l} |v_l\rangle \quad (96)$$

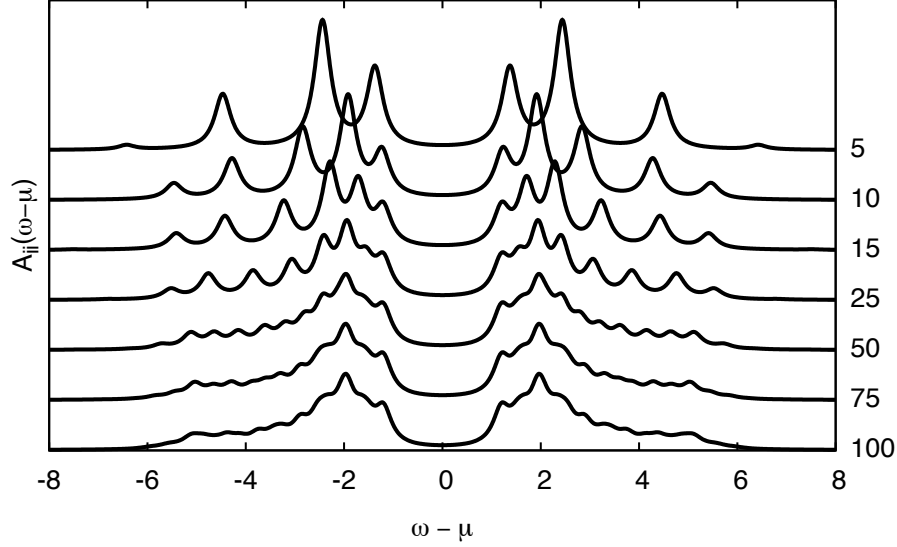


Fig. 4: Convergence of the spectral function with increasing number of Lanczos steps, $L=5, 10, 15, 25, 50, 75$, and 100 , for a 14 -site Hubbard chain with $U = 5t$ at half filling. With increasing L , more and more moments of the photoemission and inverse photoemission part of the spectrum are reproduced correctly.

the matrix elements are given by $\langle \check{\Psi}_n | \Psi_c \rangle = \check{\psi}_{n,0}$. Thus

$$\check{G}_c(z) = \sum_{n=0}^L \frac{|\check{\psi}_{n,0}|^2}{z - \check{E}_n}. \quad (97)$$

The spectral function $\check{A}(\omega \pm i\eta) = \mp \frac{1}{\pi} \text{Im } \check{G}(\omega \pm i\eta)$ obtained this way, surprisingly, converges very quickly. An example is shown in figure 4.

To understand how the $L+1$ eigenstates of $\check{H}$ can represent the full spectrum so well, we consider the moments of the spectral function

$$\int_{-\infty}^{\infty} d\omega \omega^m \check{A}(\omega) = \sum_{n=0}^L |\check{\psi}_{n,0}|^2 \check{E}_n^m = \sum_{n=0}^L \langle \Psi_c | \check{\Psi}_n \rangle \langle \check{\Psi}_n | \Psi_c \rangle \check{E}_n^m = \langle \Psi_c | \check{H}^m | \Psi_c \rangle. \quad (98)$$

Since $\check{H}$ is the projection of H onto the Krylov space $\mathcal{K}^L(|\Psi_c\rangle)$, we have $\check{H}^m |\Psi_c\rangle = H^m |\Psi_c\rangle$ for $m \leq L$. Thus the Lanczos representation $\check{A}(z)$ correctly reproduces the first $2L+1$ moments of the spectral function $A(z)$. A further Lanczos step adds one new level to the continued fraction (95), leaving all previous terms unchanged. $b_m^2 = 0$ then implies that the continued fraction terminates, and all moments are given correctly. A near vanishing $b_m^2 \approx 0$, which gives rise to the loss of orthogonality of the Lanczos vectors, for the spectral function merely means that further terms in the continued fraction hardly contribute any more.

So far we have considered diagonal elements of the resolvent. Off-diagonal matrix elements

$$G_{c_1, c_2}(z) = \left\langle \Psi_{c_2} \left| \frac{1}{z - H} \right| \Psi_{c_1} \right\rangle \quad (99)$$

are easily obtained by considering the diagonal elements for the linear combinations

$$\left\langle \Psi_{c_1} \pm \Psi_{c_2} \left| \frac{1}{z - H} \right| \Psi_{c_1} \pm \Psi_{c_2} \right\rangle = G_{c_1, c_1}(z) \pm G_{c_1, c_2}(z) \pm G_{c_2, c_1}(z) + G_{c_2, c_2}(z). \quad (100)$$

A Appendices

A.1 Non-orthonormal basis

A general one-electron basis of functions $|\chi_n\rangle$ will have an overlap matrix $S_{nm} = \langle\chi_n|\chi_m\rangle$ that is positive definite (and hence invertible) and hermitian. The completeness relation is

$$\mathbb{1} = \sum_{k,l} |\chi_k\rangle (S^{-1})_{kl} \langle\chi_l|. \quad (101)$$

With it we can easily write the Schrödinger equation $\hat{H}|v\rangle = \varepsilon|v\rangle$ in matrix form

$$\sum_k \underbrace{\langle\chi_i|H|\chi_k\rangle}_{=:H_{ik}} \underbrace{\sum_l (S^{-1})_{kl} \langle\chi_l|v\rangle \langle\chi_i|\hat{H}|v\rangle}_{=:v_k} = \varepsilon \langle\chi_i|v\rangle = \varepsilon \sum_k \underbrace{\langle\chi_i|\chi_k\rangle}_{=:S_{ik}} \underbrace{\sum_l (S^{-1})_{kl} \langle\chi_l|v\rangle}_{=:v_k}. \quad (102)$$

Collecting all components this becomes the generalized eigenvalue problem $\mathbf{H}\mathbf{v} = \varepsilon\mathbf{S}\mathbf{v}$. From the solution $\mathbf{v}$ we can easily construct $|v\rangle = \sum v_k |\chi_k\rangle$ [15]. It is, however, often more convenient to have an orthonormal basis, so that we do not have to deal with the overlap matrices in the definition of the second quantized operators or the generalized eigenvalue problem.

To orthonormalize the basis $\{|\chi_n\rangle\}$, we need to find a basis transformation $\mathbf{T}$ such that

$$|\varphi_n\rangle := \sum_m |\chi_m\rangle T_{mn} \quad \text{with} \quad \langle\varphi_n|\varphi_m\rangle = \delta_{mn}. \quad (103)$$

This implies that $\mathbf{T}^\dagger \mathbf{S} \mathbf{T} = \mathbb{1}$, or equivalently $\mathbf{S}^{-1} = \mathbf{T} \mathbf{T}^\dagger$. This condition does not uniquely determine $\mathbf{T}$. In fact there are many orthonormalization techniques, e.g., Gram-Schmidt orthonormalization or Cholesky decomposition.

Usually we will have chosen the basis functions $|\chi_n\rangle$ for a physical reason, e.g., atomic orbitals, so that we would like the orthonormal basis functions to be as close to the original basis as possible, i.e., we ask for the basis transformation $\mathbf{T}$ that minimizes

$$\begin{aligned} \sum_n \left\| |\varphi_n\rangle - |\chi_n\rangle \right\|^2 &= \sum_n \left\| \sum_m |\chi_m\rangle (T_{mn} - \delta_{mn}) \right\|^2 \\ &= \text{Tr} (\mathbf{T}^\dagger - \mathbb{1}) \mathbf{S} (\mathbf{T} - \mathbb{1}) = \text{Tr} (\underbrace{\mathbf{T}^\dagger \mathbf{S} \mathbf{T}}_{=\mathbb{1}} - \mathbf{T}^\dagger \mathbf{S} - \mathbf{S} \mathbf{T} + \mathbf{S}). \end{aligned} \quad (104)$$

Given an orthonormalization $\mathbf{T}$, we can obtain any other orthonormalization $\tilde{\mathbf{T}}$ by performing a unitary transformation, i.e., $\tilde{\mathbf{T}} = \mathbf{T} \mathbf{U}$. Writing $\mathbf{U} = \exp(i\lambda \mathbf{M})$ with $\mathbf{M}$ a Hermitian matrix, we obtain the variational condition

$$0 \stackrel{!}{=} \text{Tr} (+i\mathbf{M} \mathbf{T}^\dagger \mathbf{S} - i\mathbf{S} \mathbf{T} \mathbf{M}) = i \text{Tr} (\mathbf{T}^\dagger \mathbf{S} - \mathbf{S} \mathbf{T}) \mathbf{M}, \quad (105)$$

which is fulfilled for $\mathbf{S} \mathbf{T} = \mathbf{T}^\dagger \mathbf{S}$, i.e., $\mathbf{S} \mathbf{T}^2 = \mathbf{T}^\dagger \mathbf{S} \mathbf{T} = \mathbb{1}$. The second variation at $\mathbf{T} = \mathbf{S}^{-1/2}$

$$\frac{1}{2} \text{Tr} (\mathbf{M}^2 \mathbf{S}^{1/2} + \mathbf{S}^{1/2} \mathbf{M}^2) > 0 \quad (106)$$

is positive, since $\mathbf{S}$ and the square of the hermitian matrix $\mathbf{M}$ are both positive definite. Hence the Löwdin symmetric orthogonalization [16]

$$\mathbf{T}_{\text{Löwdin}} = \mathbf{S}^{-1/2} \quad (107)$$

minimizes the modification of the basis vectors due to orthogonalization.

A.2 Useful commutation relations

Expressions of commutators of products of operators can be derived by adding and subtracting terms that differ only in the position of one operator, e.g.,

$$\begin{aligned}
 [A_1 A_2 \cdots A_N, B] &= A_1 A_2 \cdots A_N B - B A_1 A_2 \cdots A_N \\
 &= A_1 A_2 \cdots A_N B - A_1 A_2 \cdots B A_N \\
 &\quad + A_1 A_2 \cdots B A_N - A_1 \cdots B A_{N-1} A_N \\
 &\quad + \cdots \\
 &\quad + A_1 B A_2 \cdots A_N - B A_1 A_2 \cdots A_N \\
 &= \sum_i A_1 \cdots A_{i-1} [A_i, B] A_{i+1} \cdots A_N
 \end{aligned}$$

The following special cases are particularly useful

$$\begin{aligned}
 [AB, C] &= A [B, C] + [A, C] B \\
 &= A\{B, C\} - \{A, C\}B \\
 [A, BC] &= B [A, C] + [A, B] C \\
 &= [A, B] C + B [A, C] \\
 &= \{A, B\}C - B\{A, C\} \\
 [AB, CD] &= A [B, C] D + AC [B, D] + [A, C] DB + C [A, D] B \\
 &= A\{B, C\}D - AC\{B, D\} + \{A, C\}DB - C\{A, D\}B
 \end{aligned}$$

Important examples are

$$\begin{aligned}
 [c_i^\dagger c_j, c_\gamma^\dagger] &= \langle j|\gamma\rangle c_i^\dagger \\
 [c_i^\dagger c_j, c_\gamma] &= -\langle i|\gamma\rangle c_j
 \end{aligned}$$

For the commutator of products of creation and annihilation operators appearing in one- and two-body operators we find

$$[c_i^\dagger c_j, c_\alpha^\dagger c_\beta] = [c_i^\dagger c_j, c_\alpha^\dagger] c_\beta + c_\alpha^\dagger [c_i^\dagger c_j, c_\beta] = \langle j|\alpha\rangle c_i^\dagger c_\beta - \langle \beta|i\rangle c_\alpha^\dagger c_j$$

and

$$[c_i^\dagger c_j^\dagger c_k c_l, c_\alpha^\dagger c_\beta] = \langle l|\alpha\rangle c_i^\dagger c_j^\dagger c_k c_\beta + \langle k|\alpha\rangle c_i^\dagger c_j^\dagger c_\beta c_l - \langle \beta|j\rangle c_i^\dagger c_\alpha^\dagger c_k c_l - \langle \beta|i\rangle c_\alpha^\dagger c_j^\dagger c_k c_l$$

A.3 Downfolding

To integrate-out high-energy degrees of freedom, we partition the Hilbert space of the full system into states of interest (low-energy states) and ‘other’ states, which will be integrated out. The Hamiltonian is then written in blocks

$$H = \begin{pmatrix} H_{00} & T_{01} \\ T_{10} & H_{11} \end{pmatrix}, \quad (108)$$

where H_{00} is the Hamiltonian restricted to the states of interest (reduced Hilbert space), H_{11} the Hamiltonian for the ‘other’ states, and the T matrices describe transitions between the two subspaces. The resolvent is partitioned likewise

$$G(\omega) = (\omega - H)^{-1} = \begin{pmatrix} \omega - H_{00} & -T_{01} \\ -T_{10} & \omega - H_{11} \end{pmatrix}^{-1}. \quad (109)$$

Its elements are easily determined by solving the system of two linear matrix equations

$$\begin{pmatrix} \omega - H_{00} & -T_{01} \\ -T_{10} & \omega - H_{11} \end{pmatrix} \begin{pmatrix} G_{00} & G_{01} \\ G_{10} & G_{11} \end{pmatrix} = \begin{pmatrix} \mathbb{1} & \mathbb{0} \\ \mathbb{0} & \mathbb{1} \end{pmatrix}, \quad (110)$$

keeping track of the order of the sub-matrix products. The resolvent on the reduced Hilbert space is thus given by

$$G_{00}(\omega) = \left(\omega - \underbrace{(H_{00} + T_{01}(\omega - H_{11})^{-1}T_{10})}_{=H_{\text{eff}}(\omega)} \right)^{-1}. \quad (111)$$

This expression looks just like the resolvent for a Hamiltonian H_{eff} on the reduced Hilbert space. This effective Hamiltonian describes the physics of the full system, but operates only on the small reduced Hilbert space: For an eigenvector $H|\Psi\rangle = E|\Psi\rangle$ on the full Hilbert space

$$H|\Psi\rangle = \begin{pmatrix} H_{00} & T_{01} \\ T_{10} & H_{11} \end{pmatrix} \begin{pmatrix} |\Psi_0\rangle \\ |\Psi_1\rangle \end{pmatrix} = E \begin{pmatrix} |\Psi_0\rangle \\ |\Psi_1\rangle \end{pmatrix} \quad (112)$$

its projection $|\Psi_0\rangle$ onto the reduced Hilbert space is an eigenstate of $H_{\text{eff}}(E)$. On the other hand, we can construct the full eigenstate from a solution $H_{\text{eff}}(E)|\Psi_0\rangle = E|\Psi_0\rangle$ on the reduced Hilbert space by upfolding $|\Psi\rangle \propto (\mathbb{1} + (E - H_{11})^{-1}T_{10})|\Psi_0\rangle$.

Of course, this drastic simplification comes at a price: the effective Hamiltonian is energy dependent. If the hopping matrix elements in T_{01} are small, and/or the states in the part of the Hilbert space that has been integrated out are energetically well-separated from the states that are explicitly considered, this energy dependence can, to a good approximation, be neglected. We can then replace ω by some characteristic energy ε_0 for the states in the reduced Hilbert space to obtain an energy-independent Hamiltonian

$$H_{\text{eff}}(\omega) = H_{00} + T_{01}(\omega - H_{11})^{-1}T_{10} \approx H_{00} + T_{01}(\varepsilon_0 - H_{11})^{-1}T_{10} = H_{\text{eff}}(\varepsilon_0) \quad (113)$$

that gives a good description of the electrons in the reduced Hilbert space, i.e., the states with an energy close to ε_0 . Expanding $(\omega - H_{11})^{-1}$ about ε_0 , we can systematically improve the approximation (linear and higher-order methods).

References

- [1] P. Forrest: *The Identity of Indiscernibles*, in The Stanford Encyclopedia of Philosophy <http://plato.stanford.edu/entries/identity-indiscernible/>
- [2] W. Pauli: *The connection between spin and statistics*, Phys. Rev. **58**, 715 (1940)
- [3] R.F. Streater and A.S. Wightman: *PCT, Spin and Statistics, and All That* (Benjamin, New York, 1964)
- [4] P.-O. Löwdin: *Quantum Theory of Many-Particle Systems I*, Phys. Rev. **97**, 1474 (1955)
- [5] P. Jordan and O. Klein: *Zum Mehrkörperproblem in der Quantenmechanik*, Z. Physik **45**, 751 (1927)
- [6] P. Jordan and E. Wigner: *Über das Paulische Äquivalenzverbot*, Z. Physik **47**, 631 (1928)
- [7] R. Peierls: *Surprises in Theoretical Physics* (Princeton University Press, 1979)
- [8] H.S. Warren: *Hacker's Delight*, 2nd ed. (Addison-Wesley, 2013)
- [9] L.V. Kantorovich: *Functional analysis and applied mathematics*, Uspekhi Mat. Nauk, **3**, 89 (1948)
- [10] M.R. Hestenes and W. Karush: *A method of gradients for the calculation of the characteristic roots and vectors of a real symmetric matrix*, J. Res. Nat. Bureau Standards **47**, 45 (1951)
- [11] A.N. Krylov: *On the numerical solution of the equation by which in technical questions frequencies of small oscillations of material systems are determined*, Izv. Akad. Nauk SSSR, Otd. Mat. Estest. **7**, 491 (1931)
- [12] C. Lanczos: *An Iteration Method for the Solution of the Eigenvalue Problem of Linear Differential and Integral Operators*, J. Res. Nat. Bur. Stand. **49**, 255 (1950)
- [13] G.H. Golub and C.F. van Loan: *Matrix Computations* (Johns Hopkins Univ. Press, 1996)
- [14] W.E. Arnoldi: *The Principle of Minimized Iteration in the Solution of the Matrix Eigenproblem*, Quarterly of Applied Mathematics **9**, 17 (1951)
- [15] A. Szabo and N.S. Ostlund: *Modern Quantum Chemistry* (McGraw Hill, New York, 1989)
- [16] P.-O. Löwdin: *On the Non-Orthogonality Problem Connected with the Use of Atomic Wave Functions in the Theory of Molecules and Crystals*, J. Chem. Phys. **18**, 365 (1950)

8 Quantum Chemistry DMRG in a Local Basis

E. Miles Stoudenmire

Center for Computational Quantum Physics

Flatiron Institute, New York, NY 10010 USA

Contents

1	Introduction	2
2	DMRG and matrix product states	2
2.1	Matrix product state form of the wavefunction	3
2.2	Overview of the DMRG algorithm	5
2.3	MPO forms of Hamiltonians	8
3	Quantum chemistry: brief overview and discretization methods	9
3.1	Basis set discretization	10
3.2	Grid discretization	13
3.3	Comparison of basis set and grid discretization	14
4	Local bases for quantum chemistry DMRG	14
4.1	Approach 1: hybrid grid and basis set, or “sliced basis” approach	15
4.2	Approach 2: multi-sliced gausslet basis	19
4.3	DMRG with sliced-basis or multi-sliced discretization	27
5	Conclusions and future directions	29

1 Introduction

The electronic structure problem, and the quantum many-body problem in general, is an exponentially hard problem. A full accounting of the quantum mechanical behavior of N electrons would take truly astronomical resources once N is greater than about 30 electrons. This is far fewer than in most chemical or solid state systems! Electronic structure is also a continuum problem, so to apply most computational methods one must discretize the Hamiltonian. The approximations needed to handle the exponentially large many-body wavefunction and the multi-scale details of continuum discretization involve subtle tradeoffs.

By now there are many computational approaches for simulating systems of interacting electrons, but each has major limitations. In quantum chemistry, many computational approaches start from a non-interacting picture, where the electronic wavefunction is a Slater determinant—the fermionic analogue of a product state. Interactions are treated by summing multiple Slater determinants. Such approaches break down when the system becomes *strongly correlated*, meaning that exponentially many Slater determinants are necessary to express the wavefunction accurately. Other approaches like quantum Monte Carlo are limited by the complicated sign structure of the wavefunction. Thus a demand remains for complementary approaches to electronic structure, so that a computational technique exists to treat every type of system.

The computational method which is the focus of this chapter is the *density matrix renormalization group* (DMRG) algorithm. DMRG is interesting for a number of reasons. It can handle strong correlation very naturally, often working better when the system is strongly correlated versus weakly correlated. DMRG is a controlled method with an arbitrary accuracy level determined by the user, though it may come at a high cost depending on the specific system. Finally, DMRG is just one of a number of methods for optimizing *tensor network* wavefunctions, in particular *matrix product states* (MPS), which we define and discuss in more detail below. The tensor network / MPS perspective has become instrumental in adding new capabilities to DMRG and pushing its range of applications, and has opened up an entire field of study into tensor network methods. We will see later how a tensor network representation of the electronic structure Hamiltonian enables very favorable scaling of DMRG calculations for chemistry.

In what follows, we first introduce and review the basics of DMRG, what it accomplishes and how it works. Then we review the essentials of quantum chemistry, defining the electronic structure problem and discussing various ways of discretizing it for computational approaches. We finally turn to recently developed techniques for transforming the electronic structure problem into a form especially suitable for DMRG, built around spatially local choices of basis.

2 DMRG and matrix product states

The *density matrix renormalization group* (DMRG) algorithm is a method for optimizing a particular class of wavefunctions, primarily with the goal of finding ground states of many-body quantum systems [1–4]. The class of wavefunctions DMRG optimizes are known as *matrix product states* (MPS) [5, 4]; these form a very powerful class that can represent wavefunctions

of widely different systems and notably do not suffer from issues related to strong correlation. Though DMRG works best for studying one-dimensional lattice model systems, it can be successfully applied to narrow two-dimensional lattice models [6] and to ab initio Hamiltonians such as in quantum chemistry [7, 8]. DMRG also has extensions which can treat excited states, time evolution, finite-temperature systems, and open systems. Much can be said about technical aspects of DMRG and properties of MPS, such as the number of variational parameters needed for an MPS to represent ground states of various systems accurately. The extensive review Ref. [4] provides a detailed review of MPS techniques in the DMRG context, and Ref. [9] discusses mathematical aspects of the class of MPS wavefunctions and how their complexity scales with properties such as the entanglement entropy of the wavefunction they represent.

2.1 Matrix product state form of the wavefunction

Before introducing the DMRG algorithm itself, let us first define and discuss matrix product states (MPS). Because the focus of this chapter is quantum chemistry, let us focus the discussion to wavefunctions of electrons, which are just fermions that have a spin. Thus the many-body Hilbert space will be a product of single-site (or single-orbital) spaces which are four dimensional, corresponding to the states $\{|0\rangle, |\uparrow\rangle, |\downarrow\rangle, |\uparrow\downarrow\rangle\}$. Indices $s_j = 1, 2, 3, 4$ will refer to these four states on site j of a discrete lattice system (which could be a discretized form of a continuum quantum chemistry system as we will see later in Section 3).

The most general form of a many-body wavefunction on an N site system is

$$|\Psi\rangle = \sum_{\{s\}} \Psi^{s_1 s_2 s_3 \dots s_N} |s_1 s_2 s_3 \dots s_N\rangle. \quad (1)$$

All of the parameters of this wavefunction are stored in the amplitudes $\Psi^{s_1 s_2 s_3 \dots s_N}$ which have the form of an N -index tensor. The fact that this tensor has 4^N distinct components is one manifestation of the exponential many-body problem. For readers not used to the second quantization formalism for describing fermions, note that any choice of amplitudes in Eq. (1) yields a properly antisymmetrized fermionic wavefunction, even if the amplitude tensor has no particular symmetry properties itself. This is because all operators acting on this wavefunction and the basis states $|s_1 s_2 s_3 \dots s_N\rangle$ are defined in terms of fundamental raising and lowering operators $\hat{c}_{j\sigma}^\dagger$ and $\hat{c}_{j\sigma}$ (where $\sigma = \uparrow, \downarrow$) which anti-commute with each other.

2.1.1 Matrix product states

The challenge in dealing with the wavefunction Eq. (1) is finding a manageable representation of the amplitude tensor $\Psi^{s_1 s_2 s_3 \dots s_N}$. Fortunately, for any N -index tensor there is a powerful factorization known as the matrix product state, having the following form:

$$\Psi^{s_1 s_2 s_3 s_4 s_5 s_6} = \sum_{\{\alpha\}} A_{\alpha_1}^{s_1} A_{\alpha_1 \alpha_2}^{s_2} A_{\alpha_2 \alpha_3}^{s_3} A_{\alpha_3 \alpha_4}^{s_4} A_{\alpha_4 \alpha_5}^{s_5} A_{\alpha_5}^{s_6} \quad (2)$$

where the above equation shows the example of an MPS for a $N = 6$ index tensor. Note how each of the factor tensors $A_{\alpha_{j-1}\alpha_j}^{s_j}$ carries exactly one of the original *physical* indices s_j . Note

also how each factor tensor carries two *bond* indices α_i , except for the first and last tensor which carry only one bond index.¹

The range, m , of the bond indices $\alpha_i = 1, 2, \dots, m$ is called the *bond dimension* of the MPS. Although this dimension can vary for each bond index, for simplicity we will sometimes refer to a single overall bond dimension by which we mean the maximum or typical one. The bond dimension of an MPS controls both the expressivity the MPS, meaning its ability to accurately represent complicated wavefunctions, and the cost of computations with the MPS, such as steps of the DMRG algorithm which scale as m^3 . For a large enough bond dimension, an MPS can represent any wavefunction, but in the worst case this requires the bond dimension to be exponentially large. What makes MPS so useful is that in practice they can accurately capture ground states and low-lying excited states of one-dimensional (1D) or quasi-one-dimensional systems for modest bond dimensions, usually in just the many hundreds or few thousands. And in many interesting cases, such as 1D Hamiltonians with finite-range interactions, the accuracy corresponding to a given bond dimension is essentially independent of system size.²

The name *matrix product state* comes from the fact that any single amplitude of the quantum state the MPS represents can be computed as a product of matrices. Say we want to know the amplitude for the state $|0\ 0\ \uparrow\ 0\ \downarrow\ 0\rangle$ in other words the tensor component $\Psi^{0\ 0\ \uparrow\ 0\ \downarrow\ 0}$. We can obtain this by fixing the physical indices in Eq. (2):

$$\Psi^{0\ 0\ \uparrow\ 0\ \downarrow\ 0} = \sum_{\{\alpha\}} A_{\alpha_1}^0 A_{\alpha_1\alpha_2}^0 A_{\alpha_2\alpha_3}^\uparrow A_{\alpha_3\alpha_4}^0 A_{\alpha_4\alpha_5}^\downarrow A_{\alpha_5}^0. \quad (3)$$

Note that after the physical indices are set to fixed values, the factor tensors have at most two free indices. So the above expression can be computed by just treating $A_{\alpha_1}^0$ as a vector, multiplying it with matrix $A_{\alpha_1\alpha_2}^0$, and so on until all of the bond indices are contracted, resulting in the amplitude as a scalar with a computational effort scaling as m^2 .

If a certain wavefunction, such as a ground state, can be successfully approximated by an MPS with a modest bond dimension independent of system size, then one has obtained a hugely compressed representation. Observe that an MPS of typical bond dimension m has a number of parameters which scales as $4Nm^2$ versus 4^N for an arbitrary uncompressed wavefunction. So if m does not depend on N , or depends only very weakly on N , then the number of parameters grows only linearly with system size which is very manageable. In many interesting cases the number of parameters and thus the cost of calculation can be reduced even more using symmetries such as spin symmetry or particle number conservation.

2.1.2 Tensor diagrams

Expressions in traditional tensor notation such as Eq. (2) are tedious to write and difficult to read for computations involving MPS. Fortunately there is a fully rigorous graphical notation for expressing tensor computations known as *diagram notation* [10]. In diagram notation, tensors

¹Another more common name for an MPS or tensor network bond index is *virtual index*.

²There is a logarithmic dependence of bond dimension with system size for critical 1D systems, for which MPS still work very well in practice.

are represented as shapes and indices as lines. Connecting two lines implies those indices are summed over or *contracted*.

As an example, consider two tensors A_{ij} and B_{jkl} . The tensor C_{ilk} resulting from contracting A and B over the index j can be notated diagrammatically as

$$\begin{array}{c} l \\ | \\ i - \text{orange circle} - k \\ | \\ C \end{array} = i - \text{blue circle} - \text{green circle} - k \begin{array}{c} l \\ | \\ j \\ | \\ B \end{array} = \sum_j A_{ij} B_{jkl}$$

Common matrix operations can be expressed as tensor diagrams as:

$$\begin{array}{c} \text{green circle} - \text{purple circle} \\ | \quad | \\ i \quad j \end{array} = \sum_j M_{ij} v_j$$

$$\text{green circle} - \text{orange circle} = A_{ij} B_{jk} = AB$$

$$\text{red circle} - \text{blue circle} = A_{ij} B_{ji} = \text{Tr}[AB]$$

Tensor diagram notation has many advantages, such as removing the need to name every index in tensor expressions. Diagrams also make it easy to see how many indices the result of a computation will have. For example, if all of the index lines in a complicated diagram are contracted with another line, then the result must be a scalar, as in the $\text{Tr}[AB]$ example above. The advantage of tensor diagrams becomes most apparent for complicated networks of contracted tensors. For example, the diagram for the MPS of Eq. (2) is just

$$\Psi^{s_1 s_2 s_3 s_4 s_5 s_6} = \begin{array}{c} s_1 \quad s_2 \quad s_3 \quad s_4 \quad s_5 \quad s_6 \\ | \quad | \quad | \quad | \quad | \quad | \\ \text{blue circle} - \text{blue circle} - \text{blue circle} - \text{blue circle} - \text{blue circle} - \text{blue circle} \end{array} \quad (4)$$

which is much simpler than the expression in Eq. (2). Note that not only the α_i bond indices but even the physical s_j indices can be suppressed when using diagrams, but we have shown the physical indices above to make comparison to Eq. (2) easier.

2.2 Overview of the DMRG algorithm

The goal of the DMRG algorithm is to find the ground state of a given Hamiltonian in MPS form. For DMRG to be efficient, not only must the wavefunction be represented efficiently as

an MPS, but the Hamiltonian has to be represented efficiently too. As we will discuss later, finding a compact representation of the Hamiltonian is one of the central challenges in applying DMRG to problems in quantum chemistry. But for the purpose of this section, let us just assume the Hamiltonian is a sum of strictly local terms, such as the Hubbard model in one dimension. Local Hamiltonians have the property that they can be represented as a tensor network known as a *matrix product operator* (MPO).³ This means the Hamiltonian H , viewed as a tensor with $2N$ indices, can be written as a contracted product of factor tensors as follows

$$H_{s_1 s_2 s_3 s_4 s_5 s_6}^{s'_1 s'_2 s'_3 s'_4 s'_5 s'_6} = \begin{array}{c} s'_1 \quad s'_2 \quad s'_3 \quad s'_4 \quad s'_5 \quad s'_6 \\ | \quad | \quad | \quad | \quad | \quad | \\ \text{---} \bigcirc \text{---} \bigcirc \text{---} \bigcirc \text{---} \bigcirc \text{---} \bigcirc \text{---} \bigcirc \text{---} \\ | \quad | \quad | \quad | \quad | \quad | \\ s_1 \quad s_2 \quad s_3 \quad s_4 \quad s_5 \quad s_6 \end{array} \quad (5)$$

In Section 2.3 we briefly discuss some details of how to define the MPO tensors which represent a given Hamiltonian, but for this section let us assume the Hamiltonian is given in MPO form and focus on optimizing the MPS approximation to the ground state.

Finding an MPS approximation to the ground state of H means that the MPS obeys the eigenvalue equation

$$\begin{array}{c} | \quad | \quad | \quad | \quad | \quad | \\ \text{---} \bigcirc \text{---} \bigcirc \text{---} \bigcirc \text{---} \bigcirc \text{---} \bigcirc \text{---} \bigcirc \text{---} \\ | \quad | \quad | \quad | \quad | \quad | \\ \text{---} \bigcirc \text{---} \bigcirc \text{---} \bigcirc \text{---} \bigcirc \text{---} \bigcirc \text{---} \\ | \quad | \quad | \quad | \quad | \quad | \end{array} \simeq E_0 \begin{array}{c} | \quad | \quad | \quad | \quad | \quad | \\ \text{---} \bigcirc \text{---} \bigcirc \text{---} \bigcirc \text{---} \bigcirc \text{---} \bigcirc \text{---} \\ | \quad | \quad | \quad | \quad | \quad | \end{array} \quad (6)$$

where E_0 is the smallest extremal eigenvalue of H .

A very efficient way to find extremal eigenvalues of Hermitian matrices is by using *iterative eigensolver algorithms* such as the Lanczos or Davidson algorithms. Without going into a detailed description of these algorithms, the main operation needed to perform them is the multiplication of the current approximate ground state wavefunction by H . The key idea of DMRG is to use an iterative eigensolver algorithm to improve the approximate ground state in MPS form, but *only one or two tensors at a time*—here for pedagogical reasons we will discuss the case of just improving one MPS tensor at a time.

When only improving one tensor of an MPS, the other temporarily frozen MPS tensors can be interpreted as defining a sub-basis of the full Hilbert space in which the unfrozen MPS tensor is defined. For a correct implementation of DMRG, transformations of the MPS tensors need to be carried out to ensure this sub-basis is an orthonormal basis, but here we will omit these steps and refer the interested reader to Ref. [4] for further discussion of this important point.

Let us say that we are currently improving the third tensor of an MPS defined on six sites. Then the step of multiplying H by this tensor in order to use an iterative algorithm such as Lanczos

³Hamiltonians which are not strictly local, such as in systems with long-range Coulomb interactions, can still be approximated well by MPOs using appropriate compression techniques.

to improve it is equivalent to performing the following tensor contractions

Diagrammatic equation (7) shows a 3x6 grid of nodes. The nodes are colored: blue, grey, and dark blue. The grid is connected by horizontal and vertical lines. The equation is set equal to a single dark blue node with three external lines (top, left, and right).

$$\text{Diagram} = \text{Diagram} \quad (7)$$

Note that the resulting tensor on the right-hand side above has the same index structure as the third MPS tensor, so that it is possible for the MPS tensor being optimized to obey an eigenvalue equation, when interpreting all of the other tensors—both the frozen MPS tensors and the MPO tensors—as a square “matrix” multiplying the unfrozen third MPS tensor. The reason the frozen MPS tensors appear on both the top and bottom of the left-hand side of the above expression is that they are acting on the Hamiltonian MPO as a transformation which changes the basis of H from the full Hilbert space to just the subspace of the indices of the third MPS tensor.

For the contractions in Eq. (7) to be efficient, it is important in practice to perform them in a certain order. To begin with, the projection of the Hamiltonian MPO tensors into the frozen MPS tensors is performed efficiently by iterating the following pattern of contractions, showing just the case of tensors to the right of the site being optimized:

The diagrammatic proof shows the simplification of a nested sum of diagrams. It starts with a diagram consisting of a blue circle connected to a vertical bar, which is then summed with a diagram consisting of a white circle connected to a vertical bar, which is then summed with a diagram consisting of a blue circle connected to a vertical bar. This is followed by a series of equalities leading to a final simplified diagram consisting of a vertical bar.

$$\text{Diagram 1} = \text{Diagram 2} = \text{Diagram 3} = \text{Diagram 4} = \text{Diagram 5} = \text{Diagram 6} \quad (8)$$

A similar pattern is carried involving frozen MPS tensors to the left of the site to be optimized. The resulting “Hamiltonian projection” tensors are saved in memory for reuse in the iterative optimization loop and in later steps of DMRG when the optimization returns to the current site. Then, within the iterative eigensolver algorithm, an efficient pattern of contractions for carrying out one step of multiplication by H as in Eq. (7) is given by

The diagram illustrates a transformation process. On the left, a 3D lattice structure is shown, consisting of a 3x3x3 cube of nodes. The top and bottom faces are blue, the middle face is gray, and the central node is dark blue. An arrow points to the right, where a 2D representation is shown. This representation consists of two vertical gray bars, each with three horizontal lines extending from them. A central gray node is connected to these bars, and a dark blue node is connected to the central gray node by a dotted line. An arrow points down from this 2D representation to another 2D representation below it. This second 2D representation shows two vertical gray bars, each with three horizontal lines. A central gray node is connected to these bars, and a dark blue node is connected to the central gray node by a dotted line. An arrow points right from this 2D representation to a third 2D representation, which shows two vertical gray bars, each with three horizontal lines. A central gray node is connected to these bars, and a dark blue node is connected to the central gray node by a dotted line. Finally, an arrow points right from this 2D representation to a single dark blue node.

where in the first transformation above the saved Hamiltonian projection tensors were recalled from memory.

Having performed just a few iterations of the eigensolver algorithm for the third MPS tensor (the iterative eigensolver should *not* be fully converged, since the other MPS tensors are not fully optimized), the DMRG algorithm continues by freezing the improved third tensor and next optimizing the fourth MPS tensor, then the fifth tensor, etc. until reaching the end of the system. Then the tensors are optimized one at a time in reverse order, until returning to the first site, completing what is called one sweep of DMRG. In cases where DMRG is very well suited for the problem, very accurate results can often be obtained in fewer than ten sweeps. But for challenging systems many more sweeps may be needed.

One other note about the convergence of DMRG is that the single-site algorithm outlined above may get stuck in a local minimum unless extra steps are included in the algorithm, such as using a noise term [11] or a subspace expansion step [12]. Both of these approaches usually lead to very robust convergence for a wide variety of systems. Another important and frequently used variant of DMRG involves optimizing two neighboring MPS tensors at a time. In addition to helping with convergence, optimizing two tensors together allows one to easily adapt the dimension of the bond between them, letting it grow or shrink as necessary to reach a desired accuracy goal while using a few parameters as possible.

Finally, it is important to consider the scaling of the DMRG algorithm when applying it to a given system. DMRG scales as $p_1 m^3 k + p_2 m^2 k^2$ where m is the wavefunction MPS bond dimension, k is the Hamiltonian MPO bond dimension, and p_1, p_2 are constant prefactors which depend on further implementation details. So although the leading cost is driven by the complexity of the wavefunction MPS, a very important driver of cost can also be the complexity of the Hamiltonian and how efficiently it can be represented as an MPO.

2.3 MPO forms of Hamiltonians

Though it is not necessary or always advantageous to represent the Hamiltonian in MPO form in order to carry out DMRG optimization, it can be very convenient to do so. For applications such as quantum chemistry, MPO techniques can also offer huge efficiency gains when used to compress the long-range Coulomb interaction terms, as discussed further in Section 4.3.2. While full discussion of how to construct MPO Hamiltonian representations is beyond the scope of this chapter, let us discuss one illustrative case to motivate MPO constructions.

Consider the one-dimensional Hubbard model, which shares some similarities with the Hamiltonians one encounters in quantum chemistry calculations, such as a Hilbert space of mobile electrons, as well as orbital-hopping and Coulomb interaction terms, though admittedly very local versions of such terms. Recall that the Hamiltonian of this model is

$$\hat{H} = -t \sum_{j\sigma} (\hat{c}_{j\sigma}^\dagger \hat{c}_{j+1\sigma} + \hat{c}_{j+1\sigma}^\dagger \hat{c}_{j\sigma}) + U \sum_j \hat{n}_{j\uparrow} \hat{n}_{j\downarrow}. \quad (10)$$

An exact MPO representation of the above Hamiltonian can be obtained by choosing the MPO

tensor on site j to have the form

$$\hat{W}_{r_{j-1}c_j} = \begin{bmatrix} \hat{I}_j \\ \hat{c}_{j\uparrow} \\ \hat{c}_{j\uparrow}^\dagger \\ \hat{c}_{j\downarrow} \\ \hat{c}_{j\downarrow}^\dagger \\ U\hat{n}_{j\uparrow}\hat{n}_{j\downarrow} \quad -t\hat{c}_{j\uparrow}^\dagger \quad t\hat{c}_{j\uparrow} \quad -t\hat{c}_{j\downarrow}^\dagger \quad t\hat{c}_{j\downarrow} \quad \hat{I}_j \end{bmatrix} \quad (11)$$

where matrix elements which are not shown are equal to zero. By saying an MPO tensor is equal to an operator-valued matrix as in the expression above, what is meant is that fixing the bond entries to a given pair of values (r, c) makes the resulting tensor, now carrying only two physical indices, equal to the operator listed in the operator-valued matrix. A few examples of fixing the bond indices on the MPO tensor defined by Eq. (11) above are

$$\begin{array}{c} s'_j \\ | \\ 1 \text{---} \bigcirc \text{---} 1 \\ | \\ s_j \end{array} = \hat{I}_{s'_j s_j} \quad \begin{array}{c} s'_j \\ | \\ 2 \text{---} \bigcirc \text{---} 1 \\ | \\ s_j \end{array} = \hat{c}_{\uparrow s_j}^{s'_j} \quad \begin{array}{c} s'_j \\ | \\ 6 \text{---} \bigcirc \text{---} 2 \\ | \\ s_j \end{array} = -t\hat{c}_{\uparrow s_j}^{\dagger s'_j} \quad (12)$$

To apply the particular MPO above to a finite system with open boundary conditions, the open bond index on the left of the first MPO tensor is contracted with the standard basis vector $\mathbf{e}_6$ and the open index on the right of the last MPO tensor is contracted with the standard basis vector $\mathbf{e}_1$. The motivated reader can verify by constructing and fully contracting the MPO described above on small systems that it indeed reproduces the Hubbard model Hamiltonian.

Note that the size of the MPO tensor Eq. (11) is tied very closely to the number of distinct terms making up the Hamiltonian, such as up-spin hopping versus down-spin hopping terms. Terms with distinct coefficients that act on sites different distances apart also count as distinct terms for the purposes of determining the minimum size of an MPO representation. Since we will see below that discretized quantum chemistry Hamiltonians involve many such distinct terms with different ranges and coefficients, the size of MPO needed to represent them can grow very quickly with system size unless additional techniques or approximations are used.

For many more details about MPOs, their construction, and their use in quantum chemistry see Ref. [13].

3 Quantum chemistry: brief overview and discretization methods

There is a wide range methods in chemistry for studying atoms and molecules computationally, with large variations in the degree of approximation made. In the setting of *quantum chemistry*, one attempts a fully quantum treatment of either all the electrons, or at least those electrons most important for chemical processes.

It is common to work within the Born-Oppenheimer approximation, which we will do here. This approximation treats each atomic nucleus as a classical point object with positive charge Z , which is the atomic number. For each set of nuclear positions $\{\mathbf{R}_a\}$, with a indexing each atom, one solves the Schrödinger equation $\hat{H}|\Psi_0\rangle = E_0|\Psi_0\rangle$ to find the instantaneous ground state $|\Psi_0\rangle$ of the electrons. Here $\hat{H}$, $|\Psi_0\rangle$, and E_0 are all functions of the nuclear coordinates, with $\hat{H}$ given by

$$\hat{H} = \sum_{\sigma} \int_{\mathbf{r}} \hat{\psi}_{\mathbf{r}\sigma}^{\dagger} \left(-\frac{1}{2} \nabla^2 + v(\mathbf{r}) \right) \hat{\psi}_{\mathbf{r}\sigma} + \frac{1}{2} \sum_{\sigma\sigma'} \int_{\mathbf{r}\mathbf{r}'} \hat{\psi}_{\mathbf{r}\sigma}^{\dagger} \hat{\psi}_{\mathbf{r}'\sigma'}^{\dagger} u(\mathbf{r}, \mathbf{r}') \hat{\psi}_{\mathbf{r}'\sigma'} \hat{\psi}_{\mathbf{r}\sigma} . \quad (13)$$

with the summation over $\sigma = \uparrow, \downarrow$ taken over up and down spin states, and with all integrations taken over the entire 3D space, unless stated otherwise.

The nuclear positions $\mathbf{R}_a$ and atomic numbers Z_a enter through the one-body potential $v(\mathbf{r})$

$$v(\mathbf{r}) = \sum_a \frac{-Z_a}{|\mathbf{r} - \mathbf{R}_a|} , \quad (14)$$

parameterizing the Coulomb attraction of the electrons to the protons of each atom a . The function $u(\mathbf{r}, \mathbf{r}')$ parameterizes the Coulomb repulsion between electrons

$$u(\mathbf{r}, \mathbf{r}') = \frac{1}{|\mathbf{r} - \mathbf{r}'|} . \quad (15)$$

This Hamiltonian is known as the *electronic structure* Hamiltonian, and solving for its ground state or other properties known as the *electronic structure problem*.

In general, the electronic structure problem is very difficult. One reason is that it involves many strongly interacting fermions. It is also a three dimensional problem in the continuum, raising the challenging issue of how best to discretize it for methods which operate in a discrete Hilbert space, with DMRG being one such method.

A wide variety of methods have been devised to study electronic structure. Density functional theory is one common approach, especially when the electronic structure problem must be solved at a relatively low cost within other algorithms such as molecular dynamics simulations. Another method is coupled cluster, which often serves as the standard for high-accuracy studies of small molecules. Other important methods include variants of configuration-interaction [14, 15] and quantum Monte Carlo within a fixed-node approximation [16, 17]. And of course there is DMRG which is the focus of this chapter [7, 8].

One common approach in electronic structure is to divide valence electrons from core electrons—those whose orbitals are occupied with probability essentially equal to one—then remove the core electrons and appropriately modify the one-electron potential $v(\mathbf{r})$ by adding a so-called pseudopotential. But in what follows, we will only consider an all-electron approach for simplicity, meaning we will give every electron a fully quantum mechanical treatment.

3.1 Basis set discretization

A very common way to discretize the electronic structure problem is to project the single-particle basis onto a finite set of functions. The set of functions used is called the basis set. If

the initial set of functions are not normalized and orthogonal to each other (under the 2-norm), which is often the case for standard basis sets, then one typically expresses the discretized Hamiltonian in a set of orthonormal functions obtained by an transformation of the original functions. Here by basis set, we will usually mean the final set of orthonormal functions obtained after the transformation.

Consider a basis set given by N orthonormal functions $\{\phi_i(\mathbf{r})\}$ where $i = 1, \dots, N$. Then this basis set can be used to discretize the electronic structure Hamiltonian by defining the integrals t_{ij} and V_{ijkl} , where

$$t_{ij} = \int_{\mathbf{r}\mathbf{r}'} \phi_i(\mathbf{r}) \left(-\frac{1}{2}\nabla^2 + v(\mathbf{r}) \right) \phi_j(\mathbf{r}') \quad (16)$$

$$V_{ijkl} = \int_{\mathbf{r}\mathbf{r}'} \frac{\phi_i(\mathbf{r})\phi_j(\mathbf{r}')\phi_k(\mathbf{r}')\phi_l(\mathbf{r})}{|\mathbf{r} - \mathbf{r}'|} . \quad (17)$$

One also defines orbital annihilation operators

$$\hat{c}_i = \int_{\mathbf{r}} \phi_i(\mathbf{r}) \hat{\psi}_{\mathbf{r}} \quad (18)$$

and conjugate orbital creation operators $\hat{c}_i^\dagger$.

With these definitions, the electronic structure Hamiltonian projected into the basis set is

$$\hat{H} = \frac{1}{2} \sum_{ij\sigma} t_{ij} \hat{c}_{i\sigma}^\dagger \hat{c}_{j\sigma} + \frac{1}{2} \sum_{ijkl\sigma\sigma'} V_{ijkl} \hat{c}_{i\sigma}^\dagger \hat{c}_{j\sigma'}^\dagger \hat{c}_{k\sigma'} \hat{c}_{l\sigma} . \quad (19)$$

In this form, the Hamiltonian is in principle straightforwardly treatable by lattice methods such as DMRG, though it is crucial to make sure that the very large number of terms involving the V_{ijkl} interaction integrals are treated efficiently. To apply DMRG to a complicated lattice Hamiltonian such as the one above, one must choose some one-dimensional ordering of the basis. Within this ordering, the system can be viewed as a one-dimensional chain with non-local interactions. Which ordering to choose is not obvious, but different heuristics can be used to determine very good orderings, such as an analysis of the entanglement of the Hartree-Fock approximation to the ground state.

Regarding scaling, observe there are N^4 interaction integrals V_{ijkl} for a basis set of size N . Thus quantum chemistry calculations using basis sets have a cost that grows at least as N^4 , and often much more rapidly, with the total number of basis functions N . However we will see in Section 4 that one can obtain much better scaling using local bases within DMRG for certain classes of systems.

3.1.1 Gaussian basis sets

Because integrals involving Gaussian functions admit exact closed-form expressions, even when the integrand involves Gaussians, derivatives of Gaussians, and additional polynomial factors, basis sets built from three-dimensional Gaussians are a very common choice in quantum chemistry due to the efficiency they offer when performing integrals, Eqs. (16) and (17), to construct the Hamiltonian. For computational scientists used to working with lattice models such as the

Hubbard model, it can be surprising to learn that just constructing the Hamiltonian for a quantum chemistry calculation can take significant time, but note that for $N = 100$ basis functions, there are a hundred million V_{ijkl} terms, each defined by a six-dimensional integral!

Despite their name, the basis functions in a Gaussian basis set are not necessarily themselves Gaussians, but are made by summing (or “contracting”) a small number of Gaussian functions together, then possibly multiplying them by integer powers of x , y , and z for those functions designated as capturing P or D orbitals. The purely Gaussian functions which get summed together are called primitive Gaussians.

Standard and popular Gaussian basis sets are designated by acronyms such as STO-3G or cc-pVDZ (standing for “Slater-type orbitals fit to three Gaussians” or “correlation-consistent, polarized valence double-zeta”). These standard basis sets differ not only in the specific Gaussian functions they contain, but also how these functions were chosen and for what purpose. For example, the correlation consistent (cc-) basis sets were constructed for the purpose of extrapolating energies smoothly to the continuum limit, which in chemistry parlance is called the *complete basis set* limit.

The typical procedure for using a Gaussian basis set, which generally consists of groups of non-orthogonal functions centered on each atom, is to compute a set of orthogonal functions $\phi_j(\mathbf{r})$ from each of the functions $b_i(\mathbf{r})$ in the initial non-orthogonal basis. For example, one can diagonalize the overlap matrix $\mathcal{O}_{ij} = \int_{\mathbf{r}} b_i(\mathbf{r})b_j(\mathbf{r})$ and use the resulting orthogonal matrix of eigenvectors to define the $\phi_j(\mathbf{r})$ basis. Then the same transformation can be used to transform the integrals t_{ij} and V_{ijkl} to the orthonormal basis. It is important to note that although the original $b_i(\mathbf{r})$ are somewhat local and are centered on one of the atoms, the orthonormal $\phi_j(\mathbf{r})$ functions typically end up being much less localized. Though various strategies can be used to ensure the ϕ_j are as local as possible, in general they will have significant overlap with many other basis functions ϕ_k , which has negative consequences when used as a starting point for methods based on locality such as DMRG.

3.1.2 Quantum chemistry DMRG with Gaussian basis sets

The use of DMRG for quantum chemistry originates from the proposal of White and Martin [7] to use DMRG in combination with standard Gaussian basis sets, as well as an earlier study [18]. Since then, DMRG has become a powerful technique for certain quantum chemistry problems, in large part due to its ability to handle strongly correlated systems, as it does not rely on an expansion of the wavefunction in Slater determinants unlike many other quantum chemistry methods. Some notable examples of applying DMRG within Gaussian basis are an accurate study of the strongly-correlated Mn_4CaO_5 cluster in the photosystem II protein complex [19], and calculations of the challenging Cr_2 dimer [20].

Many technical improvements to the DMRG algorithm have been developed in the setting of chemistry with Gaussians or adapted to this setting. These include using the “complementary operator” technique to handle very large Hamiltonians with non-local interactions [21, 7], exploiting $SU(2)$ symmetries [22], and using matrix product operators (MPO) to compress

Hamiltonians and simplify calculations [13]. Another important step often used in conjunction with DMRG is the selection of an active space of orbitals, which are a subset of the full basis set most important for estimating chemical properties [19]. Advanced DMRG and MPS techniques for going beyond ground-state calculations have also been developed and refined in the quantum chemistry setting, such as tangent-space linear response methods [23]. For review articles about using DMRG in quantum chemistry with Gaussian basis sets, see Refs. [24], [8], and [25].

3.2 Grid discretization

An alternative to using a basis set to discretize the continuum Hamiltonian Eq. (13) is to use a *grid discretization*. Within a grid approximation to the continuum, one introduces grid points $\mathbf{r}_n$ indexed by a set of integers $\mathbf{n} = (n_x, n_y, n_z)$ where $n_x = 1, 2, \dots, N_x$ and similar for y and z . Rather than associating individual functions to these grid points, one thinks of them as locations at which to sample from any smooth function. Operators in the continuum Hamiltonian are replaced by operators defined only on grid points, such that for any sufficiently smooth wavefunction, the expected values of the grid operators accurately approximate that of the continuum operators.

For terms not involving derivatives, the replacement of continuum operators $\hat{\psi}_{\mathbf{r}\sigma}$ with grid operators $\hat{c}_{\mathbf{n}\sigma}$ is

$$\hat{\psi}_{\mathbf{r}\sigma} \rightarrow \sum_{\mathbf{n}} \delta(\mathbf{r} - \mathbf{r}_n) \hat{c}_{\mathbf{n}\sigma} . \quad (20)$$

Thus for terms such as the one-body potential energy, the transformation from continuum to grid form is

$$\frac{1}{2} \sum_{\sigma} \int_{\mathbf{r}} v(\mathbf{r}) \hat{\psi}_{\mathbf{r}\sigma}^{\dagger} \hat{\psi}_{\mathbf{r}\sigma} \rightarrow \frac{1}{2} \sum_{\mathbf{nn}'\sigma} \int_{\mathbf{r}} v(\mathbf{r}) \delta(\mathbf{r} - \mathbf{r}_n) \delta(\mathbf{r} - \mathbf{r}'_{n'}) \hat{c}_{\mathbf{n}\sigma}^{\dagger} \hat{c}_{\mathbf{n}'\sigma} = \frac{1}{2} \sum_{\mathbf{n}\sigma} v(\mathbf{r}_n) \hat{c}_{\mathbf{n}\sigma}^{\dagger} \hat{c}_{\mathbf{n}\sigma} \quad (21)$$

from which we see that the basic replacement is just to evaluate the coefficient functions at grid points and replace continuum with grid operators.

For the Coulomb interaction term, the transformation to the grid is

$$\frac{1}{2} \sum_{\sigma\sigma'} \int_{\mathbf{r}\mathbf{r}'} u(\mathbf{r}, \mathbf{r}') \hat{\psi}_{\mathbf{r}\sigma}^{\dagger} \hat{\psi}_{\mathbf{r}'\sigma'}^{\dagger} \hat{\psi}_{\mathbf{r}'\sigma'} \hat{\psi}_{\mathbf{r}\sigma} \rightarrow \frac{1}{2} \sum_{\mathbf{n} \neq \mathbf{n}'} u(\mathbf{r}_n, \mathbf{r}_{n'}) \hat{n}_{\mathbf{n}} \hat{n}_{\mathbf{n}'} \quad (22)$$

which gives a particularly simple and compact form for this computationally expensive term. Crucially, unlike the basis set approach which results in N_b^4 discrete Coulomb terms for a basis set consisting of N_b functions, the number of Coulomb terms in the grid approach is just N^2 where $N = N_x N_y N_z$ is the number of grid points. So the scaling of the number of Coulomb integrals is much more favorable in the grid approach than in the basis-set approach, though this is not the entire story as we will see just below.

Finally, for the kinetic energy term of the Hamiltonian, in the grid approach one replaces derivatives by finite-difference approximations over neighboring grid points. There are many finite-difference approximations for the second derivative operator, which can be obtained by approaches such as fitting low-order polynomials or using wavelet techniques. For a regular grid

with grid spacing a , the resulting transformation of the kinetic energy takes the form

$$-\frac{1}{2} \sum_{\sigma} \int_{\mathbf{r}} \hat{\psi}_{\mathbf{r}\sigma}^{\dagger} \nabla^2 \hat{\psi}_{\mathbf{r}\sigma} \rightarrow -\frac{1}{2a^2} \sum_{\mathbf{n}\mathbf{n}'\sigma} \Delta_{\mathbf{n}\mathbf{n}'} \hat{c}_{\mathbf{n}\sigma}^{\dagger} \hat{c}_{\mathbf{n}'\sigma} \quad (23)$$

where the $\Delta_{\mathbf{n}\mathbf{n}'}$ are finite-difference coefficients which fall exactly to zero once $\mathbf{n}$ and $\mathbf{n}'$ are more than a certain distance apart, depending on the order of the approximation.

3.3 Comparison of basis set and grid discretization

Having just reviewed two of the major approaches to discretizing continuum Hamiltonians—Gaussian basis set discretization and grid discretization—let us contrast the key aspects of each approach.

One important consideration is whether the resulting approximation is variational, meaning whether the ground-state energy of the approximate Hamiltonian is always greater than or equal to the continuum ground-state energy. A key advantage of the basis set approach is that it is variational. Grid discretization is typically not variational.

The other key consideration is the overall computational cost, with the number of Coulomb interaction terms being a major driver of cost. The number of Coulomb terms scales much better within the grid approach: just N_g^2 Coulomb terms for N_g grid points, versus N_b^4 Coulomb terms when using a basis set of N_b functions. However, Gaussian basis sets often need many fewer functions N_b compared to the number of grid points N_g needed to reach a similar accuracy. For example, approximating a small molecule with about $N_b = 100$ Gaussian basis functions requires a grid of linear size roughly $N_{x,y,z} = 100$ giving $N_g = 10^6$ grid points in total, making the grid approach more costly overall, despite its better scaling.

Thus at least for the study of small molecules using high-accuracy quantum chemistry techniques, Gaussian basis set discretization is typically the preferred approach due to its variational nature and relatively low costs.

But as we will see in the next section, one can successfully combine the best features of the grid and basis set approaches—low cost and good scaling—for use within methods such as DMRG.

4 Local bases for quantum chemistry DMRG

We turn now to a promising approach of using basis sets made from *local* functions for DMRG and tensor network methods for quantum chemistry. These bases are distinct from standard Gaussian basis sets, though the first one we will discuss uses Gaussian bases as an ingredient. Although choosing a local basis often results in the basis being larger than the number of Gaussians needed to resolve the continuum to the same accuracy, we will see that the tradeoff is worthwhile because tensor networks very naturally exploit real-space locality.

The motivation for developing alternative bases for DMRG calculations is that calculations using Gaussian basis sets scale poorly compared to DMRG calculations of lattice systems, such as of the 1D Hubbard model, which have local Hamiltonian terms only. For example, quantum

chemistry DMRG calculations of one-dimensional chains of 3D hydrogen atoms can only handle about a hundred atoms using Gaussian basis sets, whereas 1D Hubbard DMRG calculations can scale to thousands of sites. Yet both quantum chemistry calculations and Hubbard model calculations use the same local single-site or single-orbital basis: $\{|0\rangle, |\uparrow\rangle, |\downarrow\rangle, |\uparrow\downarrow\rangle\}$. Thus it follows that scaling is not about whether the degrees of freedom are electrons, but instead must originate from either the range or complexity of the Hamiltonian.

To be more precise, there are three significant factors which determine the cost of DMRG calculations for quantum chemistry:

1. the *size of the Hamiltonian* within the particular basis and format used within DMRG
2. the *bond dimension* of the MPS needed to accurately represent the ground state wavefunction in the basis used
3. the *number of sweeps*, or passes, over the system necessary to reach convergence

Crucially, factors number (1) and (2) are both tied very strongly to the size and locality of the basis one uses. For example, if one uses a basis set which is local in the sense that any function ϕ_i only overlaps with a small, finite number of other functions ϕ_j , then the number of non-zero interaction integrals V_{ijkl} scales only as N^2 , not as N^4 . This better scaling has immediate implications for the costs (1) coming from the size of the Hamiltonian. Regarding the costs (2) associated with the MPS bond dimension, a ground state in a spatially local basis can be accurately captured by an MPS of a bond dimension growing only logarithmically with the largest direction of a system extended along one dimension. In contrast, a basis that is not spatially local can result in MPS bond dimensions which are orders of magnitude larger, with the most extreme example being that of an interacting system in a non-local plane-wave basis, for which an accurate MPS representation must have a bond dimension growing exponentially with system size!

Yet one does not immediately gain from just choosing a basis set of local functions because such local bases can require many more functions N to capture the continuum than a non-local basis. But with a smart choice of local functions (Sections 4.1 and 4.2), and technology for representing the electronic structure in a compressed form (Section 4.3.2), local basis sets can be a very beneficial choice.

4.1 Approach 1: hybrid grid and basis set, or “sliced basis” approach

One idea to develop a local basis for DMRG is to combine grid discretization with Gaussian basis set discretization, with the goal of obtaining the best aspects of each. A successful way to combine a grid and basis set is to use a grid only along one spatial direction, such as the z direction, then a Gaussian basis set along the other two directions x and y . This idea, called *sliced-basis*, was recently developed for use with DMRG in Ref. [26] and successfully applied to DMRG calculations of one-dimensional, strongly correlated chains of hydrogen atoms in Refs. [26, 27].

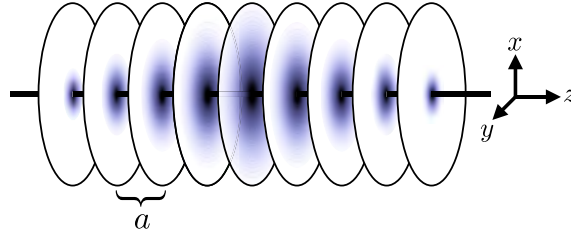


Fig. 1: Visualization of a sliced-basis representation of the continuum, where the z direction is discretized using a grid spacing a and the (x, y) directions are captured using a set of functions which can be derived from a 3D Gaussian basis.

To carry out the grid discretization along just the z direction, define a regular grid spacing a and grid points $z_n = n \cdot a$. Each value of n defines an infinite two-dimensional plane (x, y, z_n) we will call a *slice* of three-dimensional space. Next, within each continuum slice, discretize using a basis set of functions $\{\varphi_{nj}(x, y)\}$ with $j = 1, 2, \dots, N_o$. The number of functions N_o used on each slice could differ from slice to slice in principle, but here we consider just the case where N_o the same for every slice. Fig. 1 shows schematically how a sliced basis might represent a three-dimensional continuum function such as an electron orbital as linear combinations of two-dimensional functions arranged on planes separated by a small spacing a in the z direction. Although such a discretization does not strictly speaking correspond to using a basis set, it is still pedagogically useful to think of it as involving a basis set given by

$$\phi_{nj}(\mathbf{r}) = \varphi_{nj}(x, y) \delta^{1/2}(z - a \cdot n) \quad (24)$$

such that these “sliced basis functions” are ultra-local along the z -direction. The $1/2$ power on the delta function simply indicates that the ϕ are square normalized. The reason the sliced-basis approach is not technically a basis set approach is that it uses a finite-difference approximation for the z -direction kinetic energy as we will discuss below. Thus it lacks the guarantee of a variational energy. But it can still be useful to picture it in terms of the functions Eq. (24).

After introducing the grid and basis set approximations, the discrete electronic structure Hamiltonian takes the form

$$\hat{H} = \frac{1}{2} \sum_{nn'ij\sigma} t_{ij}^{nn'} \hat{c}_{ni\sigma}^\dagger \hat{c}_{n'j\sigma} + \frac{1}{2} \sum_{nn'ijkl\sigma\sigma'} V_{ijkl}^{nn'} \hat{c}_{ni\sigma}^\dagger \hat{c}_{j\sigma'}^\dagger \hat{c}_{k\sigma'} \hat{c}_{l\sigma} . \quad (25)$$

where the slice or grid indices run over $n, n' = 1, 2, \dots, N_z$ and the orbital indices over $i, j, k, l = 1, 2, \dots, N_o$. Note the resemblance to Eq. (19), except that the one-body integrals $t_{ij}^{nn'}$ and two-body interaction integrals $V_{ijkl}^{nn'}$ carry extra indices n, n' labeling pairs of slices along the z direction. The one-body integrals are defined as

$$t_{ij}^{nn'} = \delta_{nn'} \int_{\rho} \varphi_{ni}(\rho) \left(-\frac{1}{2} \nabla_{\rho}^2 + v(\rho, z_n) \right) \varphi_{n'j}(\rho) - \delta_{ij} \frac{1}{2a^2} \Delta_{nn'} , \quad (26)$$

where $\rho = (x, y)$ is a convenient shorthand, and $\Delta_{nn'}$ are second-derivative finite-difference coefficients. The two-body interaction integrals are defined as

$$V_{ijkl}^{nn'} = \int_{\rho, \rho'} \frac{\varphi_{ni}(\rho) \varphi_{n'j}(\rho') \varphi_{n'k}(\rho') \varphi_{nl}(\rho)}{\sqrt{|\rho - \rho'|^2 + (z_n - z_{n'})^2}} . \quad (27)$$

4.1.1 Scaling of the sliced-basis approach

Despite the resemblance to the basis-set form of the electronic structure Hamiltonian Eq. (19), a key distinctive of the sliced-basis approach is that the number of interaction integrals $V_{ijkl}^{nn'}$ scales as $N_z^2 N_o^4$, with N_z the number of slices and N_o the number of orbitals on each slice. When treating quasi-1D, chainlike molecules it is only the number of slices N_z that grows with the number of atoms, not N_o . Therefore the overall scaling of the method as a function of the number of atoms is quadratic, which is a huge improvement over the quartic scaling with the number of atoms incurred by the basis set approach. We discuss in Section 4.3.2 how the z -locality of the sliced-basis approach allows its quadratic scaling to be improved to just *linear* scaling when taking advantage MPO compression techniques. Thus the sliced-basis approach realizes some of the best aspects of a grid while avoiding the pitfall of the huge number of grid points arising from a fully 3D grid. By using a grid only along one direction, the overall number of interaction integrals that must be treated remains tractable.

4.1.2 Deriving the transverse basis functions

A crucial step in setting up a sliced basis for a chemistry calculation is selecting or deriving the transverse functions $\{\varphi_{nj}(x, y)\}$ which define the basis of each slice. While many approaches could be conceived to do this, the approach taken in Ref. [26] derives the transverse functions from a standard, atom-centered Gaussian basis set.

In the simplest setting, all of the atoms are of the same type (such as hydrogen atoms) and the “parent” Gaussian basis set consists purely of Gaussians (that is, the basis functions are “uncontracted”). This can be made to be true by just using the primitive Gaussians from a set without contracting them. Because we assumed the atoms are the same, the entire basis consists of the same set of Gaussian functions centered on each atom. Then for a given basis function $b_i(x, y, z)$ the associated transverse function intersecting a slice is just $\tilde{\varphi}_{ni}(x, y) = b_i(x, y, z_n)$. Because in this case $b_i(\mathbf{r})$ is a Gaussian $b_i(\mathbf{r}) \propto \exp(-\zeta_i |\mathbf{r} - \mathbf{r}_i|^2)$, its restriction to slice n will also be a Gaussian with the same length scale ζ_i . Thus although every basis function for every atom intersects with a given slice, there will only be as many unique Gaussian functions as in the basis set of just a single atom. The final transverse basis of each slice is formed by just creating an orthonormal linear combination of the $\tilde{\varphi}_{ni}$.

More generally, though, the atoms will be of different types and some of the basis functions will be contracted, so their restrictions to slice will result in large set of functions, much larger than was needed to represent any one atom in the original Gaussian basis. Using all of these functions would be costly and would not give enough extra accuracy relative to the other approximations being made to be worthwhile. So for this more general case one must truncate the functions on a slice in some sensible way. A very reasonable and effective approach is to use *principal component analysis* (PCA). Here this just means the following: say that we obtain a set of functions $\tilde{b}_{ni}(x, y) = b_i(x, y, z_n)$ by restricting each function b_i from a Gaussian basis set to the

n^{th} slice. Compute the overlap matrix

$$\mathcal{O}_{ij} = \int_{x,y} \tilde{b}_{ni}(x,y) \tilde{b}_{nj}(x,y) . \quad (28)$$

Now diagonalize this matrix and sort its eigenvalues from largest to smallest. Keep the eigenvectors v_k^j corresponding to the largest $k = 1, 2, \dots, N_o$ eigenvalues, where for the case of molecules consisting of only one atom type, N_o can be chosen to be the number of functions in the basis set of a single atom. The resulting eigenfunctions take the form

$$\tilde{\varphi}_{nk}(x,y) = \sum_j v_k^j \tilde{b}_{nj}(x,y) . \quad (29)$$

The final transverse basis $\{\varphi_{ni}(x,y)\}_{i=1}^{N_o}$ of slice n is then found by normalizing the $\tilde{\varphi}$.

4.1.3 The sliced-basis approach in practice

To test the sliced-basis approach outlined above, Refs. [26, 27] applied sliced bases to finding the ground state of one-dimensional chains of hydrogen atoms with their nuclei evenly spaced by a distance R . Although hydrogen atoms lack core orbitals, solving hydrogen chains still involves reckoning with most of the issues that make quantum chemistry challenging. These issues include scaling to large numbers of atoms, converging to the continuum or complete basis set limit, and dealing with strongly correlated wavefunctions.

One goal of studying hydrogen molecules was to compare to results obtained with standard Gaussian basis sets such as those in Ref. [27]. For this purpose, it is sufficient to study finite hydrogen chains of ten atoms or H_{10} —see the results in Fig. 2. The energy obtained using a sliced basis derived from a given Gaussian basis is similar to the energy obtained just using that Gaussian basis without slicing. This is encouraging to see, since it demonstrates that the sliced-basis maintains favorable aspects of the parent Gaussian basis like the ability to smoothly extrapolate to the continuum or complete basis set limit. On the other hand, a sliced basis is much more scalable to long hydrogen chains as we will see below. Another observation about the results in Fig. 2 is that the sliced-basis energies are generally slightly lower than the energy of the corresponding Gaussian basis, at least for larger basis sets such as cc-pVDZ (double ζ) and cc-pVTZ (triple ζ). This can be readily understood as a consequence of finer resolution of a sliced basis along the z direction, allowing electrons more freedom to avoid costly Coulomb interactions.

Another goal of studying hydrogen chains with the sliced-basis approach was to test that it can scale to very long systems extended along the z direction, with a cost that is only linear in system length. Here a crucial technical step for achieving good scaling is the compression of the Coulomb interaction using matrix product operator techniques, which we describe later in Section 4.3.2 below. Taking a fixed inter-atomic spacing $R = 3.6$ and working with a sliced basis derived from the STO-6G basis set, Fig. 3 shows results for the timing and energy of calculations up to 1000 hydrogen atoms. From the inset of the figure, one can see very close to linear scaling of the method with number of atoms, while getting consistently accurate energies across all system sizes.

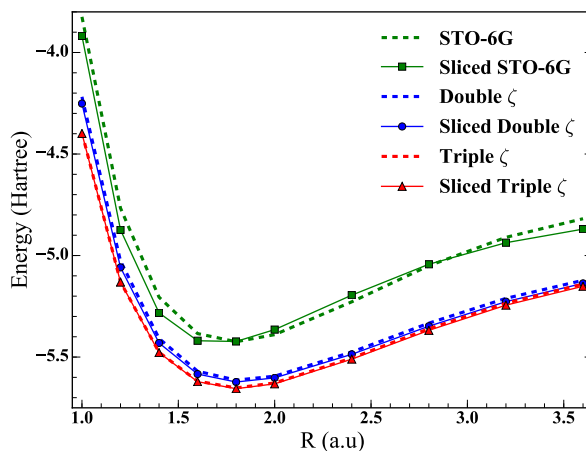


Fig. 2: Ground-state energies of H_{10} chains as a function of inter-atomic spacing R calculated using DMRG within standard Gaussian basis sets (dashed curves) and sliced basis sets (solid curves and points) using a uniform grid spacing of $a = 0.1$ atomic units [26].

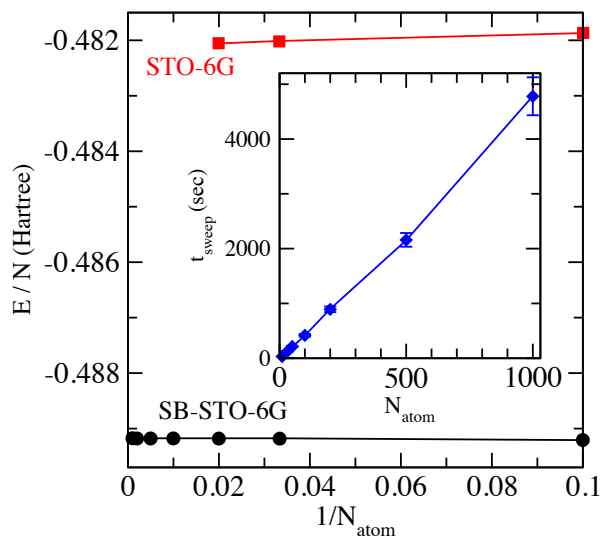


Fig. 3: Scaling with number of atoms of sliced-basis calculations up to 1000 hydrogen atoms. The inter-atomic spacing is fixed to $R = 3.6$ and a sliced basis derived from the STO-6G Gaussian basis was used. The outer plot shows the ground state energy from DMRG using the standard STO-6G basis and the sliced version (SB-STO-6G). The inset shows the average time per DMRG sweep, taking a bond dimension of $m = 100$.

4.2 Approach 2: multi-sliced gausslet basis

The sliced basis approach to discretizing the electronic structure Hamiltonian demonstrates a successful marriage of the grid and basis set approaches to quantum chemistry. Counterintuitively, it demonstrates that using *more* functions to represent the Hamiltonian can result in a more affordable calculation overall, by choosing the functions to be local (at least along one direction), so that the DMRG algorithm and MPO methods for compressing the Hamiltonian (Section 4.3.2) can perform to their full potential.

There are some drawbacks to the sliced-basis approach though. It is very oriented toward one-dimensional systems, and only scales well along the z direction: the direction discretized with a grid. The cost of including more transverse functions within each (x, y) slice is high. The sliced-basis also requires a small grid spacing a to get good accuracy. Finally, because sliced-basis transverse functions are derived from standard Gaussian bases, they sometimes inherit their weaknesses when dealing with certain molecules for which they were not designed. This can include molecules where the nuclei are extremely close together or for which electrons occupy very spatially extended orbitals.

Because these issues mostly stem from the use of Gaussian bases in making the transverse functions, a better approach would be to start from a completely different set of basis functions. Ideally such functions would behave like grid points in terms of the simplicity of the Hamiltonian formed from them, yet a relatively small number of them would be able to capture continuum details. Fortunately such functions have been recently developed for quantum chemistry applications and are called *gausslets* [28]. Gausslet basis sets have been extended to real 3D chemistry calculations based on DMRG through an approach called *multi-slicing* [29]. We will discuss both gausslets and multi-slicing below, with examples of the improvements they give over the sliced-basis approach. A key reason for these improvements is not only that multi-sliced gausslets capture the continuum with relatively few functions, but that their properties enable *diagonal approximations* for the costly Coulomb interaction terms.

4.2.1 Gausslet functions and their properties

In the one-dimensional case, a gausslet is a function $G(x)$ which is symmetric about $x = 0$ and which is orthonormal $\int_{x,x'} G(x) G(x') = \delta(x - x')$. Gausslets are also smooth, in the precise sense of being orthogonal to a certain hierarchy of oscillatory functions; they are local in the sense of falling rapidly to zero past a certain length scale; they have excellent completeness properties, meaning that linear combinations of neighboring gausslets can represent any polynomial up to a certain very high order (such as order 10); and finally gausslets have an important property of integrating like a delta function when integrated with sufficiently smooth functions. What this means is that for any polynomial $p(x)$ that is not too high-order

$$\int_{-\infty}^{\infty} dx G(x - x') p(x) = p(x') . \quad (30)$$

Thus integration against a gausslet “plugs in” the coordinate where the gausslet is centered. All of these desirable properties of gausslets resemble those of wavelets (technically wavelet scaling functions), yet underneath a gausslet is defined as a weighted sum of Gaussian functions, hence the name gausslet. Defining gausslets in terms of underlying Gaussians makes them very convenient and efficient for the integrations necessary to perform when constructing quantum chemistry Hamiltonians.

Figure 4 shows an array of one-dimensional gausslet functions with a length scale of 1.0, such that their centers are arranged on a grid with a 1.0 spacing. Gausslets can be constructed in various ways involving trade-offs in their favorable properties listed above. For more information

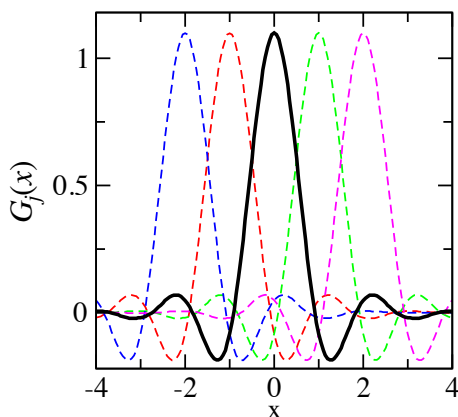


Fig. 4: One-dimensional array of gausslet functions with a length scale of 1.0. The gausslet centered at the origin is highlighted in solid black to emphasize details.

about the details of constructing them, see Ref. [28] which proposes and constructs gausslets, building upon the development in Ref. [30] of compact and symmetric families of orthogonal wavelets.

Importantly, after constructing a grid of gausslets like in Fig. 4 and using them to discretize quantum chemistry Hamiltonians, one can obtain very accurate continuum results using a gausslet spacing of about 1.0, in contrast to grid discretization which requires about an order of magnitude smaller spacing to obtain similar accuracy [28]. Better yet, we will see next that by adapting the grid on which the gausslets are centered, one can even better resolve the continuum using small numbers of gausslet functions.

4.2.2 Adapted grid of gausslets

Using an even-spaced grid of gausslet functions, as in Fig. 4 to discretize quantum chemistry Hamiltonians is more efficient than using a simple grid, yet still requires more functions than are actually needed. The reason is that while high resolution is required to capture details of the electronic wavefunction near atomic nuclei, much less resolution is needed away from nuclei. A straightforward way to reduce the number of functions needed while preserving high resolution near nuclei is to perform a coordinate mapping on the gausslet functions so that they form an adapted grid, with a finer spacing near nuclei and a coarser spacing otherwise.

Such a coordinate mapping may be defined via a function $x(u)$ which maps from a fictitious space u where the gausslets are defined to have a regular grid spacing into the actual space x used for the quantum chemistry calculation. Let $u(x)$ be defined as the inverse of the mapping $x(u)$. For the case of a single atom, a sensible coordinate mapping is

$$u(x) = \frac{1}{s} \sinh^{-1}(x/a) \quad (31)$$

defined by a *scale* parameter s and a *core cutoff* parameter a . Figure 5(a) shows how this mapping takes an evenly spaced grid along the y direction of the plot into a variable spaced grid along the x direction.

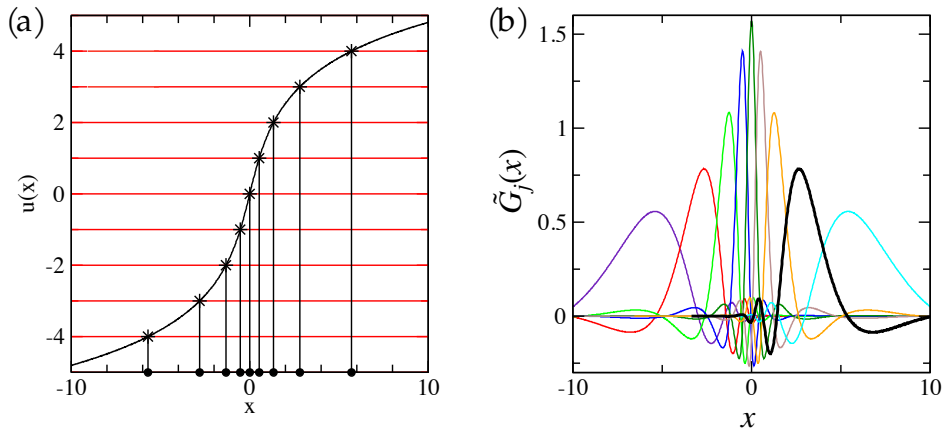


Fig. 5: Illustration of (a) a coordinate mapping Eq. (31) which maps a regularly spaced grid in the u coordinates to an adapted grid in the x coordinates. The resulting adapted gausslets shown in (b) remain orthonormal, but achieve a higher resolution near the origin. One gausslet is shown with a bold line to highlight details.

Having chosen a coordinate mapping, the transformation of the gausslets which moves their centers onto the adapted, variable-spacing grid while preserving their orthogonality and other good properties is

$$\tilde{G}_j(x) = G_j(u(x)) \sqrt{u'(x)} \quad (32)$$

where $G_j(u)$ is the gausslet centered at the integer grid point j in the u space. Figure 5(b) shows the adapted gausslets $\tilde{G}_j(x)$ resulting from using the coordinate mapping $u(x)$ defined in Eq. (31) above. One of the adapted gausslets is highlighted with a bolder line, and you can observe that it takes a distorted shape compared to the unadapted gausslets in Fig. 4. The placement of more and finer-sized gausslets near the origin gives better resolution there.

Note that when adapting gausslets for systems of multiple atoms, there are modifications of the transformation Eq. (31) which make it better suited for treating molecules. The supplemental information of Ref. [29] discusses such multi-atom coordinate transformations.

4.2.3 Multi-sliced grid

To apply the above ideas of gausslet basis sets to 3D systems, the most straightforward approach is to define basis functions as products $G_i(x) G_j(y) G_k(z)$ of 1D gausslets. But how to maintain this product form while also adapting the gausslet spacing near atomic nuclei is less obvious; for example, performing the coordinate transformations in a radially symmetric way destroys the product form of the 3D functions, resulting in integrals which are too costly when constructing the discrete Hamiltonian.

Fortunately, there is a simple way around this problem that only incurs a modest overhead in the total number of functions needed. This workaround is called *multi-slicing* and is just the idea of performing the coordinate transformation sequentially: first in the z direction (the direction along the greatest extent of the system), then in the y direction, and finally the in x direction.

In more detail, one starts by first defining a coordinate transformation $u^z(z)$ to determine an

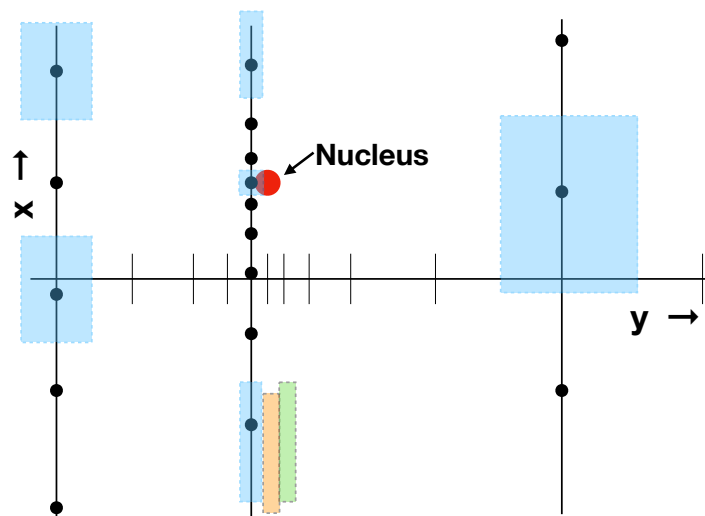


Fig. 6: Illustration of a fixed- z cut through a multi-sliced grid, showing the primary non-zero support of selected gausslet functions as colored rectangles. The vertical lines are selected x -slices and points are centers of adapted gausslets. The position of a nucleus is shown, illustrating how the multi-sliced grid bunches more gausslets of a smaller size nearby.

adapted grid of discrete z values z_k , $k = 1, 2, \dots, N_z$ which are more closely spaced whenever planes $z = z_k$ pass nearby atomic nuclei. This transformation defines planes $z = z_k$ with adapted functions $\tilde{G}_k(z)$ centered on them, forming what is called a z -slice. Now within each z -slice, a coordinate transformation is applied to the y coordinates, defining an adapted grid through a function $u^y(y)$ which yields discrete y values y_{kj} with $j = 1, 2, \dots, N_y$. The transformation $u^y(y)$ is also chosen to make the y_{kj} values more closely bunched whenever lines of fixed (y_{kj}, z_k) pass nearby an atomic nucleus. This second step defines y -slices as lines of fixed (y_{kj}, z_k) with functions $\tilde{G}_{kj}(y) \tilde{G}_j(z)$ centered on them. Finally discrete x points x_{kji} with $i = 1, 2, \dots, N_x$ are defined through a transformation $u^x(x)$ such that points (x_{kji}, y_{kj}, z_k) are more densely spaced the closer they are to nuclei. All these transformations taken together define 3D functions $\tilde{G}_{kji}(x) \tilde{G}_{kj}(y) \tilde{G}_j(z)$ centered on points (x_{kji}, y_{kj}, z_k) . To make the basis finite, only functions whose centers lie within a certain distance of at least one of the atoms are kept in the basis. By construction these final basis functions maintain a product form and orthonormality, while being adapted for higher resolution near nuclei. For more technical discussion on how to make a good choice for the u^z , u^y , and u^x coordinate transformations, see the supplemental information section of Ref. [29]. Essentially these coordinate transformation functions are chosen to have a form like that of Eq. (31), but with the s and a parameters varying according to the 3D distance of a particular z or y slice from the nearest atomic nucleus.

Figure 6 shows a 2D cut through a multi-sliced grid, with boxes illustrating the non-zero support of selected gausslet functions. From the figure, one can observe that a downside of multi-slicing is that it results in many long and thin functions far away from any nucleus which provide more resolution than is actually needed. But by using an approximate wavefunction as a guide, such as a wavefunction from a Hartree-Fock calculation, one can take the additional step of combining such redundant functions together into single functions to reduce the basis size.

4.2.4 Diagonal approximations

A final and very important step when using multi-sliced gausslet bases is to make a *diagonal approximation* to the Coulomb repulsion term when discretizing the Hamiltonian. Being able to use a diagonal form for this very expensive term is a crucial way to reduce calculation costs. A diagonal approximation to the Coulomb term means a discretization of the form

$$\frac{1}{2} \sum_{\sigma\sigma'} \int_{\mathbf{r}\mathbf{r}'} u(\mathbf{r}, \mathbf{r}') \hat{\psi}_{\mathbf{r}\sigma}^\dagger \hat{\psi}_{\mathbf{r}'\sigma'}^\dagger \hat{\psi}_{\mathbf{r}'\sigma'} \hat{\psi}_{\mathbf{r}\sigma} \rightarrow \frac{1}{2} \sum_{\mathbf{m} \neq \mathbf{m}'} V_{\mathbf{m}, \mathbf{m}'} \hat{n}_{\mathbf{m}} \hat{n}_{\mathbf{m}'} \quad (33)$$

where $\mathbf{m} = (i, j, k)$ are the grid points on which the gausslets are centered and the $\hat{n}_{\mathbf{m}} = \hat{c}_{\mathbf{m}}^\dagger \hat{c}_{\mathbf{m}}$ operators measure the occupation of the gausslet basis functions $\tilde{G}_{kji}(x)\tilde{G}_{ji}(y)\tilde{G}_k(z)$ at position $\mathbf{r}_{\mathbf{m}} = \mathbf{r}_{ijk}$. Such a form is called diagonal because it only involves N^2 terms, with N being the total number of 3D gausslet functions, rather than N^4 terms as in the most general form of the discrete Coulomb interaction Eq. (19).

The justification for using a diagonal approximation with multi-sliced gausslets is that gausslets have the ability to represent a wide variety of smooth functions while also integrating like a delta function, meaning:

$$\int_x \tilde{G}_i(x) f(x) = w_i f(x_i) \quad (34)$$

for any smooth function f where x_i is the center of the adapted gausslet $\tilde{G}_i(x)$ and $w_i = \int_x \tilde{G}_i(x)$. Note that this delta-function relation differs from Eq. (30) by the inclusion of the weight w_i . This is because, although adapted gausslets remain square-normalized, they generally no longer integrate to 1.0.

To see how the representability and delta function properties of gausslets lead to a diagonal approximation, consider just the single-particle potential term $v(x)$ for a 1D Hamiltonian in the first-quantization formalism. The action of this Hamiltonian term on a single-particle wavefunction $\psi(x)$ is $v(x)\psi(x)$. Define the resulting function to be $\phi(x) = v(x)\psi(x)$. Now if we assume that both $\psi(x)$ and $\phi(x)$ are smooth enough that they can be approximated by gausslets, that means there exist coefficients ψ_i and ϕ_i such that

$$\psi(x) = \sum_i \psi_i \tilde{G}_i(x) \quad (35)$$

$$\phi(x) = \sum_i \phi_i \tilde{G}_i(x), \quad (36)$$

where in fact

$$\psi_i = \int_x \tilde{G}_i(x) \psi(x) = w_i \psi(x_i) \quad (37)$$

$$\phi_i = \int_x \tilde{G}_i(x) \phi(x) = w_i \phi(x_i) \quad (38)$$

because of the delta-function integration property Eq. (30) of gausslets. The discrete form of the Hamiltonian we seek are the coefficients v_{ij} which are defined as mapping

$$\phi_i = \sum_j v_{ij} \psi_j. \quad (39)$$

Now observe that

$$\phi_i = \int_x \tilde{G}_i(x) \phi(x) = \int_x \tilde{G}_i(x) (v(x) \psi(x)) = w_i v(x_i) \psi(x_i) = v(x_i) \psi_i, \quad (40)$$

where we used the delta-function property of the gausslets and Eq. (37) above to obtain the last expression. By inspection of the above equation and from the definition of v_{ij} in Eq. (39), we finally see that $v_{ij} = \delta_{ij} v(x_i)$ which is the diagonal approximation we seek (in a first-quantized form). Note that the most general expression for v_{ij} in a gausslet basis would be $v_{ij} = \int_x \tilde{G}_i(x) v(x) \tilde{G}_j(x) = w_i v(x_i) \tilde{G}_j(x_i)$ which is non-zero for $i \neq j$ and thus non-diagonal. So for the diagonal approximation to be justified, one is making additional smoothness assumptions about the functions being transformed by the Hamiltonian and not just invoking properties of the Hamiltonian itself. By similar arguments one can make a diagonal approximation to the two-body Coulomb interaction of the form $V_{ijkl} = \delta_{il} \delta_{jk} u(x_i, x_j)$ when using gausslet basis functions.

The form of diagonal approximation we just outlined is called the point-wise approximation. Another type of diagonal approximation that can be derived is the *integral approximation*:

$$V_{ijkl} = \frac{\delta_{il} \delta_{jk}}{w_i w_j} \int_{x,x'} \tilde{G}_i(x) u(x, x') \tilde{G}_j(x'), \quad (41)$$

which is much more accurate than the point-wise approximation. There is also a summed diagonal approximation which is discussed in Ref. [28]. A straightforward generalization of one of these diagonal approximations to the case of 3D gausslet bases leads to the expression Eq. (33) at the beginning of this section.

4.2.5 The multi-sliced gausslet approach in practice

To test multi-sliced gausslet bases for quantum chemistry, Ref. [29] considered systems of hydrogen atoms, much like in the Ref. [26] studies of sliced-bases. But whereas Ref. [26] primarily emphasizes scalability to very long systems, Ref. [29] emphasizes the ability of multi-sliced gausslets to reach the complete basis set limit (or continuum limit). Yet multi-sliced gausslets are also scalable to very long systems when used in DMRG.

To study the effect of the scale parameter s controlling the typical spacing between neighboring gausslets, Fig. 7 shows Hartree-Fock calculations of 10-atom hydrogen chains with inter-atomic spacing $R = 1$ in standard Gaussian basis sets (horizontal lines) such as quadruple-zeta (cc-pVQZ) and 5-zeta (cc-pV5Z), as well as continuum extrapolations of Gaussian bases. In contrast, the more jagged set of points shows results of converged multi-sliced gausslet (MSG) calculations as a function of the scale or gausslet spacing (x axis of plot). For a scale below $s = 0.6$ in atomic units, the MSG results converge smoothly and systematically until reaching close agreement with the best Gaussian basis extrapolation to the complete basis set limit.

To test the multi-sliced gausslet approach within high-accuracy DMRG calculations, Fig. 8 shows DMRG calculations of 10-atom hydrogen chains as a function of the inter-atomic spacing R , including strongly correlated stretched or larger- R chains. The figure shows results

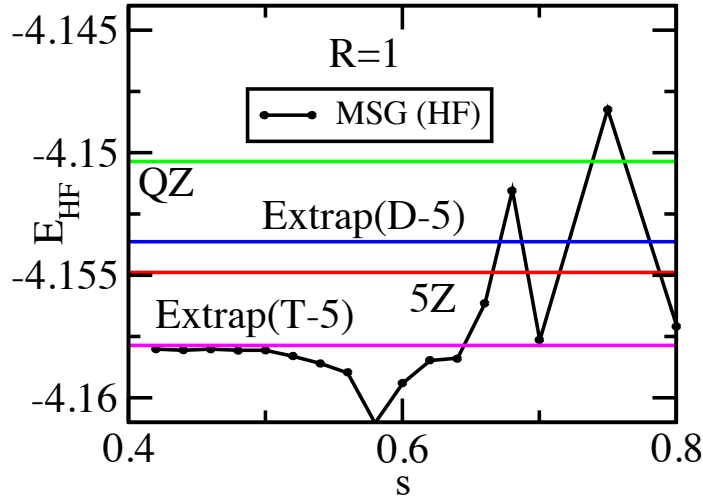


Fig. 7: Hartree-Fock energies (in units of Hartree) of 10-atom hydrogen chains with inter-atomic spacing $R = 1$. The points are results from multi-sliced gausslet (MSG) bases with varying scale parameter controlling the spacing between gausslets. The lines labeled QZ and 5Z are energies obtained with the cc-pVQZ and cc-pv5Z Gaussian basis sets, while the lines labeled Extrap are extrapolations to the complete basis set limit using either double-zeta or triple-zeta up through 5-zeta.

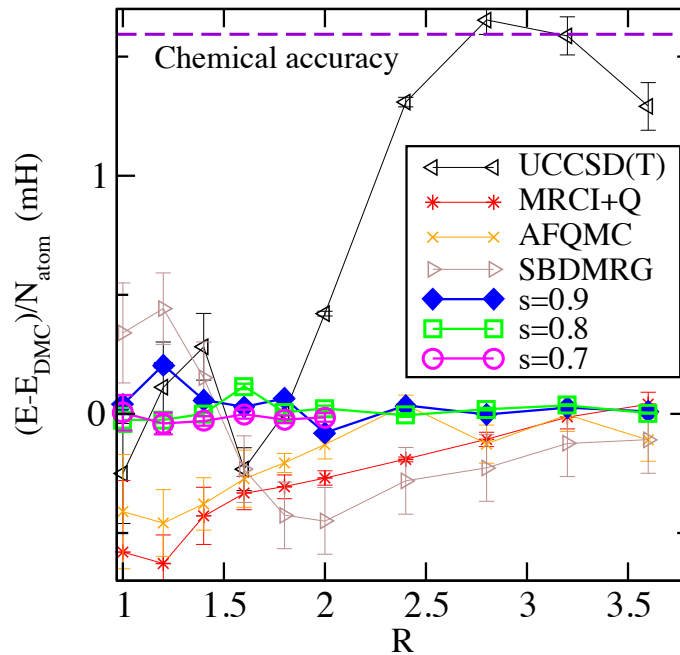


Fig. 8: Energies of 10-atom hydrogen chains computed by various methods relative to those obtained by diffusion Monte Carlo (E_{DMC}). Energy differences are shown in milli-Hartree. Results are from Ref. [27].

from competitive quantum chemistry methods including coupled cluster (UCCSD(T)), multi-reference configuration interaction (MRCI+Q), auxiliary field quantum Monte Carlo (AFQMC), DMRG using a sliced-basis (SBDMRG), and finally diffusion quantum Monte Carlo (DMC), which is used as the reference energy for the figure. These results were first obtained and dis-

cussed in much more detail in Ref. [27]. Multi-sliced gausslet DMRG (MSG-DMRG) results are labeled by the scale factor s they used. The figure shows that for scale factors $s = 0.8$ and especially $s = 0.7$, the MSG-DMRG results have the closest agreement with the DMC results on a scale of less than 0.1 milli-Hartree. Because the sources of error in DMC versus MSG-DMRG have a very different origin, their close agreement strongly suggests they are both resolving the continuum limit to the highest accuracy of the methods shown.

To obtain such accurate results, it should be noted for completeness that additional technical steps are used in Ref. [29] which are beyond the scope of the discussion here. These include a delta-function correction to the single-particle potential and combining the single-particle part of the energy with Hartree-Fock energies on a finer scale to reduce finite-scale errors in the MSG-DMRG calculations.

4.3 DMRG with sliced-basis or multi-sliced discretization

Having introduced two types of local bases which allow quantum chemistry DMRG to scale to larger numbers of atoms and better resolve the continuum limit, let us now discuss some of the technical steps involved in using these bases in actual DMRG calculations.

4.3.1 Splitting terms into multiple MPOs

One simple but very effective optimization of DMRG when treating discrete quantum chemistry Hamiltonians of the form Eq. (19) is to split different types of Hamiltonian terms into separate matrix product operators (MPOs). The reason for doing this is that even after making the most efficient possible MPO representation of the entire Hamiltonian as a single MPO, one can observe that this MPO consists of disjoint blocks of terms which do not mix with one another, such that its bond dimension is the sum of the dimensions of each of the blocks.

To illustrate why splitting the terms associated with each block into separate MPOs is more efficient, consider an example where there are four different blocks, each contributing a size k to the bond dimension. Because DMRG scales as the sum of squares of the bond dimensions of each MPO used, storing these blocked terms in separate MPOs will have a cost $4k^2$ within DMRG whereas combining them into a single MPO will have a much higher cost of $(4k)^2 = 16k^2$ in DMRG.

An example of a possible splitting of terms into separate Hamiltonians which are summed to make the total Hamiltonian could be

$$H = H_{\uparrow} + H_{\downarrow} + H_V \quad (42)$$

$$H_{\uparrow} = \frac{1}{2} \sum_{ij} t_{ij} \hat{c}_{i\uparrow}^{\dagger} \hat{c}_{j\uparrow} \quad (43)$$

$$H_{\downarrow} = \frac{1}{2} \sum_{ij} t_{ij} \hat{c}_{i\downarrow}^{\dagger} \hat{c}_{j\downarrow} \quad (44)$$

$$H_V = \frac{1}{2} \sum_{ijkl\sigma\sigma'} V_{ijkl} \hat{c}_{i\sigma}^{\dagger} \hat{c}_{j\sigma'}^{\dagger} \hat{c}_{k\sigma'} \hat{c}_{l\sigma} . \quad (45)$$

This is only one possible splitting because additional structure may be present in the V_{ijkl} terms resulting from different choices of local bases. This structure can suggest additional further splittings, such as separating terms which connect orbitals entirely in one slice versus across two different slices within the sliced-basis approach.

4.3.2 Compressing long-range interactions with MPOs

A crucial optimization that can be applied when using spatially local bases in DMRG is the compression of the two-particle Coulomb terms in the electronic structure Hamiltonian using MPO techniques. In the discussion of the sliced-basis and multi-sliced gausslet bases above, we have alluded to this compression a few times since it is so important, but delayed discussing it to this section since it is also very technical and requires a detailed familiarity with the construction of MPOs. A very detailed discussion of the compression algorithm and MPO structure is given in the appendix of Ref. [26].

Due to the limited technical scope of this chapter, let us first describe what this MPO compression accomplishes, treating the method itself as just a “black box” algorithm. In the most general case, given a set of Coulomb interaction integrals V_{ijkl} running over N orbitals, meaning $i, j, k, l = 1, 2, \dots, N$, the minimum size MPO which exactly represents the discrete Hamiltonian Eq. (19) has a bond dimension which scales as N^2 [13]. Because the DMRG algorithm scales quadratically in the bond dimension of the Hamiltonian MPO, using an uncompressed MPO for quantum chemistry results in an N^4 scaling which is typical for quantum chemistry but nevertheless very costly. However, for systems extended primarily along one dimension (taken to be the z direction) and represented using either a sliced-basis or multi-sliced gausslet basis which have the important property of consisting of local functions, exploiting both sparsity and an off-diagonal low-rank structure in the Coulomb integrals V_{ijkl} allows one to numerically construct an MPO approximation of the Coulomb interaction terms whose bond dimension grows at most logarithmically with the number of atoms in the system. In practice a compressed bond dimension of a few hundreds in size gives very good accuracy even for very large systems consisting of hundreds or thousands of atoms.

To briefly describe how the MPO compression is accomplished, consider a purely diagonal form of the Coulomb interaction $\frac{1}{2} \sum_{ij} V_{ij} \hat{n}_i \hat{n}_j$ which could result from the diagonal approximation within a gausslet basis, for example. Empirically, one can observe that *all* of the blocks $V_{ij}^{(p)}$ of the symmetric matrix V defined by restricting $i \leq p$ and $j \geq i$ are approximately low rank, implying they can be approximated well by a truncated singular value decomposition (SVD). This block-low-rank property can be understood as resulting from the smoothness of the Coulomb interaction for electrons far apart from one another. Furthermore, the unitary matrices computed in the SVD factorizations of each block $V^{(p)}$ can be related to each other by auxiliary unitary maps of a fixed size, related to the number of singular values kept. By defining MPO tensors which implement these maps, the MPO can reconstruct any of the matrix elements V_{ij} while having a bond dimension determined only by the number of singular values kept in the factorization of V , which depends very weakly on the size of the system [26].

5 Conclusions and future directions

In this chapter, we reviewed the DMRG algorithm for optimizing many-body wave functions in matrix product state form and the application of DMRG to quantum chemistry calculations. A key consideration in making DMRG efficient is choosing a basis that allows the Hamiltonian, typically represented as a matrix product operator (MPO) tensor network within DMRG, to have a manageable size. This issue becomes extremely important when applying DMRG to the electronic structure Hamiltonian used in quantum chemistry.

Although the approach of discretizing the electronic structure Hamiltonian using Gaussian basis sets has been very successful for applying DMRG to chemistry, here we reviewed two recent alternative basis constructions which are much more spatially local than standard 3D Gaussian basis sets. Though locality makes these bases much larger than Gaussian bases, these local bases can nevertheless be very advantageous for DMRG calculations whose costs are tied much more strongly to the spatial locality or smooth spatial decay of Hamiltonian terms than the number of sites or total size of the Hilbert space used in the calculation.

Looking ahead, it would be extremely welcome if a basis such as multi-sliced gausslets could be used for PEPS tensor network calculations, which are tensor networks that are scalable along two dimensions in contrast to MPS which are only scalable in one dimension. One of the key challenges is developing efficient Coulomb interaction representations suitable for PEPS, though significant progress was recently made in Ref. [31]. Another interesting direction would be to use the sliced-basis or multi-sliced gausslet bases in quantum chemistry methods very different from DMRG, such as auxiliary field quantum Monte Carlo (AFQMC). Ideas similar to the compression of the Coulomb terms (Section 4.3.2) would have to be adapted to AFQMC, but then it could benefit from the high continuum resolution of gausslets, for example.

Acknowledgements

Thanks to Steve R. White for many helpful discussions and supporting information regarding both the sliced-basis and multi-sliced Gausslet approach to quantum chemistry with DMRG.

References

- [1] S.R. White, Phys. Rev. Lett. **69**, 2863 (1992)
- [2] S. White, Phys. Rev. B **48**, 10345 (1993)
- [3] U. Schollwöck, Rev. Mod. Phys. **77**, 259 (2005)
- [4] U. Schollwöck, Annals of Physics **326**, 96 (2011)
- [5] S. Östlund and S. Rommer, Phys. Rev. Lett. **75**, 3537 (1995)
- [6] E. Stoudenmire and S.R. White, Ann. Rev. Condens. Matter Phys. **3**, 111 (2012)
- [7] S.R. White and R.L. Martin, J. Chem. Phys. **110**, 4127 (1999)
- [8] G.K.-L. Chan and S. Sharma, Ann. Rev. Phys. Chem. **62**, 465 (2011)
- [9] D. Perez-Garcia, F. Verstraete, M.M. Wolf, and J.I. Cirac, Quantum Inf. Comput. **7**, 401 (2007)
- [10] R. Penrose, Combinatorial mathematics and its applications **1**, 221 (1971)
- [11] S.R. White, Phys. Rev. B **72**, 180403 (2005)
- [12] C. Hubig, I.P. McCulloch, U. Schollwöck, and F.A. Wolf, Phys. Rev. B **91**, 155115 (2015)
- [13] G.K.-L. Chan, A. Keselman, N. Nakatani, Z. Li, and S.R. White, J. Chem. Phys. **145**, 014102 (2016)
- [14] O.L. Polyansky, A.G. Császár, S.V. Shirin, N.F. Zobov, P. Barletta, J. Tennyson, D.W. Schwenke, and P.J. Knowles, Science **299**, 539 (2003)
- [15] N.M. Tubman, J. Lee, T.Y. Takeshita, M. Head-Gordon, and K.B. Whaley, J. Chem. Phys. **145**, 044112 (2016)
- [16] D.M. Ceperley, Rev. Mod. Phys. **67**, 279 (1995)
- [17] J. Shee, B. Rudsteyn, E.J. Arthur, S. Zhang, D.R. Reichman, and R.A. Friesner, J. Chem. Theory Comput. **15**, 2346 (2019)
- [18] G. Fano, F. Ortolani, and L. Ziosi, J. Chem. Phys. **108**, 9246 (1998)
- [19] Y. Kurashige, G.K.-L. Chan, and T. Yanai, Nat. Chem. **5**, 660 (2013)
- [20] Y. Kurashige and T. Yanai, J. Chem. Phys. **135**, 094104 (2011)
- [21] T. Xiang, Phys. Rev. B **53**, R10445 (1996)
- [22] S. Sharma and G.K.-L. Chan, J. Chem. Phys. **136**, 124121 (2012)

- [23] N. Nakatani, S. Wouters, D. Van Neck, and G.K.-L. Chan, J. Chem. Phys. **140**, 024108 (2014)
- [24] G.K.-L. Chan and M. Head-Gordon, J. Chem. Phys. **116**, 4462 (2002)
- [25] R. Olivares-Amaya, W. Hu, N. Nakatani, S. Sharma, J. Yang, and G.K.-L. Chan, J. Chem. Phys. **142**, 034102 (2015)
- [26] E.M. Stoudenmire and S.R. White, Phys. Rev. Lett. **119**, 046401 (2017)
- [27] M. Motta, D.M. Ceperley, G.K.-L. Chan, J.A. Gomez, E. Gull, S. Guo, C.A. Jiménez-Hoyos, T.N. Lan, J. Li, F. Ma, A.J. Millis, N.V. Prokof'ev, U. Ray, G.E. Scuseria, S. Sorella, E.M. Stoudenmire, Q. Sun, I.S. Tupitsyn, S.R. White, D. Zgid, and S. Zhang, Phys. Rev. X **7**, 031059 (2017)
- [28] S.R. White, J. Chem. Phys. **147**, 244102 (2017)
- [29] S.R. White and E.M. Stoudenmire, Phys. Rev. B **99**, 081110 (2019)
- [30] G. Evenbly and S.R. White, Phys. Rev. A **97**, 052314 (2018)
- [31] M.J. O'Rourke, Z. Li, and G.K.-L. Chan, Phys. Rev. B **98**, 205127 (2018)

9 Density Matrix Renormalization

Karen Hallberg

Centro Atómico Bariloche and Instituto Balseiro

San Carlos de Bariloche, Argentina

Contents

1	Introduction	2
2	Calculating ground states with the DMRG	2
2.1	Definitions and method	3
2.2	Quantum information analysis	5
2.3	Standard algorithm	5
3	Calculating dynamical quantities with the DMRG	7
3.1	Lanczos dynamics	8
3.2	Correction vector dynamics	8
4	Using the DMRG as the impurity solver for dynamical mean-field theory	9
4.1	Implementation for multi-site and multi-orbital problems	10
5	Conclusions	14

1 Introduction

The density matrix renormalization group (DMRG) has evolved to become one of the most reliable and versatile numerical methods in modern computational physics. It allows for a very precise calculation of ground states and excitations of strongly correlated fermionic and bosonic systems. After its original inception by S. White in 1992 [1, 2], when it was developed to solve problems in low-dimensional quantum condensed matter, it has been extended to other fields as well, and it is now being successfully used in quantum chemistry, statistical mechanics, quantum information theory, nuclear and high-energy physics. It has been applied to a great variety of systems and problems such as spin chains and ladders, fermionic and bosonic systems, disordered models, impurities, molecules, nanoscopic systems as well as 2D electrons in high magnetic fields. And extensions to the method include two dimensional (2D) classical systems, stochastic models, phonons, quantum chemistry, field theory, finite temperature and the calculation of dynamical and time-dependent properties. Some calculations have also been performed in 2D quantum systems.

In this chapter we will introduce the basic formulation of the DMRG. We will also delve on the extension of this method to calculate dynamical behavior and show how this was implemented to enhance the performance of one of the most reliable methods to solve correlated matter, dynamical mean-field theory (DMFT).

Among the several reviews, I recommend the basic textbook [3] and reviews such as [4–6]. For newcomers to the field, it is advisable to visit the ALPS code library, containing state-of-the-art methods for solving interacting quantum systems [8] and the website containing updated information on the DMRG and publications [9].

The development DMRG has paved the way to the rich and promising world of tensor networks (see the chapter by Miles Stoudenmire in this book). A comprehensive set of lectures is given in [10]. There is a very useful library for tensor network calculations in [11].

2 Calculating ground states with the DMRG

When working with quantum systems one encounters the exponential problem, i.e., the total number of states of the systems grows exponentially with system size. In non-interacting or weakly interacting systems, one can make approximations and solve for one particle, assuming that the result will depend very weakly on the number of particles. For more strongly interacting systems, however, this is not possible and a many-body calculation is necessary. Let's consider, for example, a quantum spin chain with spins $S = 1/2$ in each of the N sites of the chain. The total Hilbert space grows exponentially as 2^N , and this is the size of the operators, in particular of the Hamiltonian, to be diagonalized to solve the system. This means that the problem becomes intractable very quickly (currently, using exact diagonalization one cannot solve for more than around $N = 30$ sites).

The first successful renormalization of a correlated systems was done by K. Wilson when he developed the Numerical Renormalization Group for the single impurity Anderson model [12].

However, later attempts to apply real space renormalization techniques to quantum correlated models led to poor results, as for example, the attempts to solve the 1D Hubbard model [13], for which the system was separated into several blocks and only the lowest-lying energy levels were kept for the new iterations. White and Noack [14] realized that one of the main problems with these real space renormalizations was the separation of the basic blocks into separate entities, so they tried solving the problem by including varying boundary conditions between them. This basic idea led White to think about using the reduced density matrix defined in part of the system as a new criterion to choose the relevant states. As we will show below, this idea solved the block boundary problem and provided a straightforward way to discard non relevant states. This led to the development of the DMRG, which proved to be one of the most accurate numerical methods to solve interacting quantum problems.

The DMRG is based on a systematic truncation of the Hilbert space by keeping the most probable states describing one or several wave functions. These wave functions will be called the target states since they will be the ones we are aiming to describe in an accurate way (for example, they could be the ground state, of some excited states). The truncations are done by calculating the reduced density matrix in part of the system and keeping only its eigenvectors with the highest eigenvalues. As we will show below, the eigenvalue is the weight of its corresponding eigenvector in the target wave function. So from here it is clear that this is a good criterion to trim the Hilbert space.

We will start with the standard DMRG as it was originally developed by S. White [1,2]. Since then there have been other ways of implementing the same basic idea of using the information provided by the density matrix to reduce the Hilbert space. These methods include the matrix product state (MPS) representation [15]

2.1 Definitions and method

Before writing the algorithm explicitly we need to define some basic concepts. The method is based on the partitioning the whole lattice into two parts, S being the system and E the environment. This way, states $|i\rangle$ are part of S and states $|j\rangle$ form E . Any state of the whole lattice $|\psi\rangle$ can be written as

$$|\psi\rangle = \sum_{i,j} \psi_{ij} |i\rangle |j\rangle.$$

Given this bipartition, any operator acting only on the system S can be calculated as

$$\langle\psi|O^S|\psi\rangle = \sum_{ij,i'j'} \psi_{ij}^* \psi_{ij} \langle i'j'|O^S|ij\rangle = \sum_{ij,i'} \psi_{ij}^* \psi_{ij} \langle i'|O^S|i\rangle = \sum_{ii'} O_{ii'}^S \underbrace{\sum_j \psi_{ij} \psi_{ij}^*}_{\rho_{ii'}^S} = \text{Tr} \rho^S O^S.$$

Here we have straightforwardly defined the reduced density matrix $\rho^S = \text{Tr}_E |\psi\rangle\langle\psi|$ which is defined in S and has the following properties:

- it is Hermitian: $\rho^{S\dagger} = \rho^S$
- $\text{Tr} \rho^S = \sum_{\alpha} w_{\alpha} = 1$
- it is positive-semidefinite (so all eigenvalues are $w_{\alpha} \geq 0$)

In the diagonal basis $\rho^S = \sum_{\alpha} w_{\alpha} |w_{\alpha}\rangle\langle w_{\alpha}|$, so the mean value of the operator on S is

$$\langle \psi | O^S | \psi \rangle = \text{Tr}_S \rho^S O^S = \sum_{\alpha} w_{\alpha} \langle w_{\alpha} | O^S | w_{\alpha} \rangle. \quad (1)$$

When more than one target state is used, the density matrix is defined as

$$\rho_{ii'}^S = \sum_l p_l \sum_j \psi_{l,ij} \psi_{l,i'j}, \quad (2)$$

where p_l defines the probability of finding the system in the target state $|\psi_l\rangle$ (not necessarily an eigenstate of the Hamiltonian).

It can be easily shown [2] that the density matrix eigenvalues w_{α} represent the probability of the state $|\psi\rangle$ being in substate $|w_{\alpha}\rangle$. The density matrix leads directly to the optimal states in the system as we demonstrate below.

We define again

$$|\psi\rangle = \sum_{i,j=1}^{M,M'} \psi_{ij} |i\rangle |j\rangle \quad (3)$$

as a state of the $S+E$, having real coefficients for simplicity. Our aim is to obtain a variational wave function $|\hat{\psi}\rangle$ defined in an optimally reduced space, generated by the m vectors of S , $|\alpha\rangle = \sum_{i=1}^m u_{\alpha i} |i\rangle$,

$$|\hat{\psi}\rangle = \sum_{\alpha=1}^m \sum_{j=1}^M a_{\alpha j} |\alpha\rangle |j\rangle \quad (4)$$

such that the difference with the original wave function is minimal with respect to $a_{\alpha j}$

$$||\psi\rangle - |\hat{\psi}\rangle|^2 = 1 - 2 \sum_{\alpha i j} \psi_{ij} a_{\alpha j} u_{\alpha i} + \sum_{\alpha j} a_{\alpha j}^2. \quad (5)$$

This condition leads to

$$\sum_i \psi_{ij} u_{\alpha i} = a_{\alpha j}. \quad (6)$$

Using the definition of the reduced density matrix for S

$$\rho_{ii'}^S = \sum_j \rho_{ij,i'j} = \sum_j \langle j | \langle i | \psi \rangle \langle \psi | i' \rangle | j \rangle = \sum_j \psi_{ij} \psi_{i'j}, \quad (7)$$

and placing (6) into Eq. (5) we obtain

$$1 - \sum_{\alpha i i'} u_{\alpha i} \rho_{ii'}^S u_{\alpha i'} = 1 - \sum_{\alpha=1}^m \omega_{\alpha}, \quad (8)$$

where $u_{\alpha i}$ is the operator that changes basis from $|i\rangle$ to $|\alpha\rangle$, and ω_{α} are the density-matrix eigenvalues. The above expression is minimum for the largest eigenvalues of the density matrix ρ^S , which are all positive or zero and the last term corresponds to the discarded error.

So, summarizing, the best approximation to $|\psi\rangle$ is done by considering the highest eigenvalued (most probable) eigenstates of ρ^S when trimming the basis states. This is the basic mechanism of the DMRG.

2.2 Quantum information analysis

It is useful to consider concepts from quantum information, such as the von Neumann entropy

$$S_{vN} = - \sum_{\alpha} w_{\alpha} \log w_{\alpha} = - \text{Tr } \rho^S \log \rho^S. \quad (9)$$

This quantum entropy gives a quantitative measure of the entanglement between S and E if the target state $|\psi\rangle$ is a pure state. The larger the entropy, the larger the entanglement. If, for example, ρ^S has only one eigenvalue $w_1 = 1$, then there is only one eigenvector in each block and $|\psi\rangle$ is a product state with no entanglement. If the w_{α} decrease rapidly, then it is sufficient to consider only the largest eigenvalues to have a reliable representation of the target state $|\psi\rangle$. The worst case scenario happens when there is no information about the state and all w_{α} have the same value so $S_{vN} = N$, the system size. S_{vN} is the number of qubits (sites with two degrees of freedom) one has to consider to describe the state, so one can estimate that the number of states m one should keep without losing crucial information is $m \sim 2^{S_{vN}}$. The DMRG works best for low quantum entangled systems (it is particularly exact for product states).

This analysis gives us a good insight into the DMRG performance in different systems and dimensions. By using geometric arguments in a $d+1$ -dimensional field theory including a $d-1$ -dimensional hypersurface dividing the system in two, $S+E$, it is shown that the entropy resides essentially at the surface and scales as the area of the hypersurface [16]

$$S_{vN}(L) \propto (L/\lambda)^{d-1}, \quad (10)$$

where λ is an ultraviolet cutoff and L the linear dimension. In one dimension, $d = 1$, a more detailed calculation for gap-less critical systems leads to a logarithmic scaling of the entropy, $S_{vN}(L) = c/3 \ln(L) + \lambda_1$, where c is the central charge of the underlying conformal field theory. A saturated entropy for non critical, gapped systems is obtained [17–19] when the system size exceeds the correlation length. In two dimensions $S_{vN} \sim L$, so one expects a poorer performance of the DMRG.

2.3 Standard algorithm

Now we are ready to introduce the main algorithm in the standard implementation. As said, the DMRG is based on a systematic truncation of the Hilbert space by keeping the most probable states describing a target wave function (e.g. the ground state, or excited states). For this, it is important to define the space in which the Hamiltonian operates (for example, real space, but a description in momentum, orthonormal orbitals or energy space is also feasible). The elements of this space (sites, momenta, orbitals, or energies) are ordered in a one-dimensional way and then it is partitioned into two, not necessarily equal, parts. From now on we will call them sites without loss of generality. In its classical formulation, one begins with a growing or “warm-up” phase starting with a small system, e.g., with N_0 sites, and then gradually increase its size to $N_0+2, N_0+4, \dots$ until the desired length N is reached.

Based on Fig. 1 we define block $[B]$ as a finite chain with l sites having an associated Hilbert space with m states in which operators such as the block Hamiltonian H_B , connecting, and correlation operators are defined and expressed in matrix form. Except for the first iteration, the basis in this block is not explicitly known due to previous basis rotations and reductions which are non-unitary if there has been a reduction of states. An additional site with n states is defined as $[a]$. For a general iteration our system S is formed by blocks $[B]$ and $[a]$: $S = [B] \otimes [a]$. Equivalently, the environment is $E = [a'] \otimes [B']$ (see Fig. 1).

To illustrate the general iteration let us consider the one-dimensional spin $S = 1/2$ Heisenberg Hamiltonian

$$H = \sum_i \mathbf{S}_i \mathbf{S}_{i+1} = S_i^z S_{i+1}^z + \frac{1}{2} (S_i^+ S_{i+1}^- + S_i^- S_{i+1}^+). \quad (11)$$

The general iteration runs as follows:

- (i) Define the Hamiltonian $H_{BB'}$ for the whole system [20] $S+E = [Baa'B']$:

$$[H_{B_1 B_2}]_{ij; i' j'} = [H_{B_1}]_{ii'} \delta_{jj'} + [H_{B_2}]_{jj'} \delta_{ii'} + [S_a^z]_{ii'} [S_{a'}^z]_{jj'} + \frac{1}{2} [S_a^+]_{ii'} [S_{a'}^-]_{jj'} + \frac{1}{2} [S_a^-]_{ii'} [S_{a'}^+]_{jj'} \quad (12)$$

- (ii) Diagonalize $H_{BB'}$ to obtain the ground state $|\psi\rangle$ or other states (target states) using the Lanczos [21] or Davidson [22] algorithms.

- (iii) Calculate the density matrix

$$\rho_{ii'}^S = \sum_j \psi_{ij} \psi_{i'j} \quad (13)$$

on block $S=[Ba]$ defined by states i , tracing over the bath $E=[B'a']$ defined by states j .

- (iv) Diagonalize ρ^S and keep the m states with the largest eigenvalues. The truncation error is $1 - \sum_{\alpha=1}^m \omega_{\alpha}$, which should be kept small, typically much less than 10^{-6} .

- (v) Rotate and change basis of all operators in $[Ba]$ and simultaneously redefine $[Ba] \rightarrow [B]$: for example $H_B = O^\dagger H_{Ba} O$, where O is a rectangular matrix.

- (vi) A new block $[a]$ is added (one site in our case) and the iteration goes back to (i)

Once the desired length N is reached a higher accuracy can be obtained by sweeping to and fro a couple of iterations along the chain without changing N . The block sizes change with the internal variable l as $[Baa'B']$, $N = l+1+1+l'$ where l and $l' = N-l-2$ are the number of sites in B and B' respectively. The density matrix is used to project onto the growing block and stored operators from previous iterations are used for the shrinking block.

It saves time and memory to include symmetries in the DMRG algorithm. For example, for a spin model like the one mentioned above, the total spin z projection S_z is conserved, which is the sum of the spin projections of each constituting block. Total particle number is also a common symmetry that can be easily implemented. When the total quantum number can be

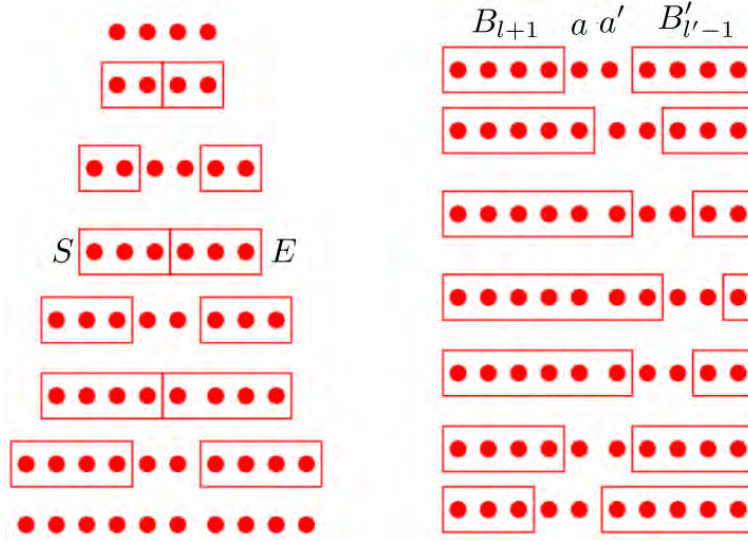


Fig. 1: Iterations of the classical DMRG. Left: warm up growth. Right: Finite-size sweeps. Here system S and environment E are also defined, as well as the blocks used in the iterations.

obtained as an addition of quantum numbers of each block of the system, the density matrix is block diagonal: if the global state $|\psi\rangle = \sum_{i,j} \psi_{ij} |i\rangle |j\rangle$ has N_p particles, $\psi_{ij} = \psi_{ij} \delta_{N_i+N_j, N_p}$, where N_i is the particle number of state $|i\rangle$, the reduced density matrix is

$$\rho_{ii'}^S = \sum_j \psi_{ij} \psi_{i'j}^* = \sum_j \psi_{ij} \psi_{i'j}^* \delta_{N_i+N_j, N_p} \delta_{N_{i'}+N_j, N_p} = \sum_j \psi_{ij} \psi_{i'j}^* \delta_{N_i N_{i'}} ,$$

which is block diagonal with a fixed particle number N_i . The eigenstates of ρ^S will then also have a defined particle number and this means that the renormalization maintains the particle number symmetry. Non abelian or non additive symmetries like $SU(2)$ or total spin are more difficult to implement, however possible in some cases.

3 Calculating dynamical quantities with the DMRG

The density matrix renormalization group can also be used to calculate dynamical properties (mainly at zero temperature) of low-dimensional systems, which are useful to interpret experimental results from, for example, nuclear magnetic resonance (NMR), neutron scattering, optical absorption and photoemission, among others.

The main current approaches for the calculation of spectral functions include the Lanczos method [23–26], the correction vector technique (CV) [24, 27, 28], Fourier transformation of time-dependent excitations [29–32] and Chebyshev polynomials [33, 34].

In this lecture we will focus on the two first ones: the Lanczos dynamics gives complete information of the whole excitation spectrum at the expense of less accuracy for large systems, specially at high energies, while the CV focuses on particular energy values and gives more precise information, being numerically much more expensive, unless the program is parallelized, so several energy values are calculated simultaneously.

We define the following dynamical correlation function at $T = 0$

$$C_A(t-t') = \langle \psi_0 | A^\dagger(t) A(t') | \psi_0 \rangle, \quad (14)$$

where $A^\dagger$ is the Hermitian conjugate of the operator A (other operators are also feasible with slight changes in the formulation), $A(t)$ is the Heisenberg representation of A , and $|\psi_0\rangle$ is the ground state of the system. Its Fourier transform can be written in the Lehmann representation

$$C_A(\omega) = \sum_n |\langle \psi_n | A | \psi_0 \rangle|^2 \delta(\omega - (E_n - E_0)) = -\frac{1}{\pi} \lim_{\eta \rightarrow 0^+} \text{Im } G_A(\omega + i\eta + E_0), \quad (15)$$

where the sum is taken over all the eigenstates $|\psi_n\rangle$ of the Hamiltonian H with energy E_n , E_0 is the ground state energy, and the Green function is defined as

$$G_A(z) = \langle \psi_0 | A^\dagger (z - H)^{-1} A | \psi_0 \rangle, \quad (16)$$

where $z = w + i\eta$ and η is a small shift towards imaginary frequencies, or, equivalently, a finite Lorentzian width of the delta poles.

3.1 Lanczos dynamics

In the Lanczos formalism the function G_A can be written in the form of a continued fraction:

$$G_A(z) = \frac{\langle \psi_0 | A^\dagger A | \psi_0 \rangle}{z - a_0 - \frac{b_1^2}{z - a_1 - \frac{b_2^2}{z - \dots}}}, \quad (17)$$

where the coefficients a_n and b_n can be obtained using the recursion equation [35]

$$|f_{n+1}\rangle = H|f_n\rangle - a_n|f_n\rangle - b_n^2|f_{n-1}\rangle \quad (18)$$

with

$$|f_0\rangle = A|\psi_0\rangle, \quad a_n = \frac{\langle f_n | H | f_n \rangle}{\langle f_n | f_n \rangle}, \quad b_n^2 = \frac{\langle f_n | f_n \rangle}{\langle f_{n-1} | f_{n-1} \rangle}, \text{ and } b_0 = 0. \quad (19)$$

As for finite systems the Green function $G_A(z)$ has a finite number of poles, only a finite number of coefficients (typically less than a few hundreds) a_n and b_n has to be obtained.

For the implementation in the DMRG, one has to take into account several target states using Eq. (2) in order to have a good description of the excitations, for example, the ground state $|\psi_0\rangle$ and the first few $|f_n\rangle$ with $n = 0, 1, \dots$ and $|f_0\rangle = A|\psi_0\rangle$.

3.2 Correction vector dynamics

This method leads to a more precise determination of the spectral functions since it focuses on one particular energy w at a time (or $z = w + i\eta$ if a finite shift is needed). This is achieved by using a correction vector (related to the operator A that can depend on momentum q).

From the Green function Eq. (16), the (complex) correction vector $|x(z)\rangle$ can be defined as

$$|x(z)\rangle = \frac{1}{z - H} A |\psi_0\rangle \quad (20)$$

so the Green function can be calculated as $G(z) = \langle \psi_0 | A^\dagger | x(z) \rangle$. Writing the correction vector in its real and imaginary parts $|x(z)\rangle = |x^r(z)\rangle + i|x^i(z)\rangle$, we obtain

$$\begin{aligned} ((H - w)^2 + \eta^2) |x^i(z)\rangle &= -\eta A |\psi_0\rangle \\ |x^r(z)\rangle &= \frac{1}{\eta} (w - H) |x^i(z)\rangle. \end{aligned} \quad (21)$$

The first equation is solved, for example, using the conjugate gradient method. Here the following target states are kept in the DMRG iterations: the ground state $|\psi_0\rangle$, the first Lanczos vector $A|\psi_0\rangle$ and the correction vector $|x(z)\rangle$. The results lead to reliable excitations for an energy range surrounding this particular point [24].

In [28] a variational formulation of the correction vector technique has been used. From Eq. (21), the following equation is minimized with respect to $|X\rangle$

$$W_{A,\eta}(\omega, X) = \langle X | (H - w)^2 + \eta^2 | X \rangle + \eta \langle \psi_0 | A | X \rangle + \eta \langle X | A | \psi_0 \rangle. \quad (22)$$

For any $\eta \neq 0$ and finite ω this function has a well defined minimum for the quantum state which is the solution of Eq. (21), i.e., $|x^i(z)\rangle$.

4 Using the DMRG as the impurity solver for dynamical mean-field theory

Among the most interesting physical phenomena observed in strongly correlated materials, we can mention high-temperature superconductivity, magnetism, ferroelectricity, and the metal-insulator transition. In spite of the enormous efforts devoted to understanding these phenomena, little progress has been achieved, and this is due to the highly complex character involving the strong correlations mainly of localized electrons. These strong correlations are not correctly treated in methods designed for weakly correlated materials such as density-functional theory (DFT) [36] for which the local density approximation (LDA) [37] and other generalizations are used. Thus, non-perturbative numerical methods are the only reliable approach.

To this end, more than twenty years ago, dynamical mean-field theory (DMFT) was developed [38,39] (see also the chapter by Eva Pavarini). By using it together with LDA it has allowed for band structure calculations of a large variety of correlated materials (see reviews [40,41]) for which DMFT accounts mainly for local interactions [42,43].

DMFT consists of a mapping of the correlated system to an effective interacting quantum impurity problem which has to be solved in a self-consistent way. This is the most computationally expensive step within DMFT and determines its success. Since its development, several impurity solvers have been used, like the iterative perturbation theory (IPT) [44,45], exact diagonalization (ED) [46,47], the Hirsch-Fye quantum Monte Carlo (HFQMC) [48], the continuous time quantum Monte Carlo (CTQMC) [49–53], non-crossing approximations (NCA) [54],

the numerical renormalization group (NRG) [12, 55–57], the rotationally invariant slave-boson mean-field theory (RISB) [58–60] and quantum chemistry-based techniques [61]. However, they all suffer from limitations, for example, the sign problem and the difficulty in reaching low temperatures in the QMC-based algorithms, the difficulty of the NCA in obtaining a reliable solution for the metallic state, the limitation to few lattice sites of the ED, far from the thermodynamic limit, and the reduced high-energy resolution of the NRG technique.

More recently, to overcome some of these difficulties, an impurity solver based on the DMRG technique was proposed [62–66]. With this method one can obtain the density of states directly on the real frequency axis (or with a very small imaginary offset). This, together with the avoidance of the fermionic sign, is the major advantage given by using the DMRG as compared to QMC impurity solvers. More still, no *a priori* approximations are made and the method provides equally reliable solutions for both gapless and gapped phases. The DMRG impurity-solver provides accurate estimates for the distribution of spectral intensities of high frequency features such as the structure of the Hubbard bands, which is of main relevance for the analysis of x-ray photoemission and optical conductivity experiments, among others.

Subsequent related techniques have been proposed, such as using different methods to obtain the dynamical properties within the DMRG [67, 65], or the time-evolution DMRG algorithm (time evolving block decimation, TEBD) [68] for the one- and two-orbital models [69]. Other developments include the kernel polynomial method (Chebyshev expansion for Green functions) [70, 34, 71], a pole decomposition technique within the correction-vector method for the dynamics [72], the block Lanczos approach [73], the application to non-equilibrium DMFT using MPS [74], and other bath geometries [75]. In this work the authors explore other geometries for the impurity bath, showing an increased efficiency for the star environment. In [76], it was shown that the convergence of the DMFT iterative loops on the imaginary energy axis implies a great reduction of computational costs because, mainly, the imaginary-time evolution does not create entanglement. However, the price to be paid is a reduced resolution on the real-frequency axis.

4.1 Implementation for multi-site and multi-orbital problems

Generalizations of the original DMFT can be considered for these cases, which lead to matrix formulations of the DMFT equations (for details see [66], where the operators are defined within the cell which contains N_c orbitals or sites and which is the effective “impurity” to be solved).

For the general formulation let us consider a Hamiltonian which is the sum of a non-interacting term plus local interactions: $H = H^0 + V$ where $V = \sum_i V_i$ and i is the site or cell index (i.e., the cluster containing N_c sites or orbitals). We define the local operators of H^0 as $h_i^0 = \sum_{IJ} t_{IJ} c_{iI\sigma}^\dagger c_{iJ\sigma}$, where $c_{iI\sigma}^\dagger$ creates an electron in cell i and local “orbital” $I = 1, 2, \dots, N_c$ with spin $\sigma = \uparrow, \downarrow$. We also define the local coefficients $T = (t_{IJ})$.

The main approximation of DMFT is to neglect the self-energy between different cells i and j in the lattice, i.e. to consider only the local self-energy, $\Sigma_{ij}(\omega) \approx \Sigma_{\text{cell}}(\omega) \delta_{ij}$, neglecting spatial correlations to a certain degree, albeit with a good treatment of the local dynamical correlations.

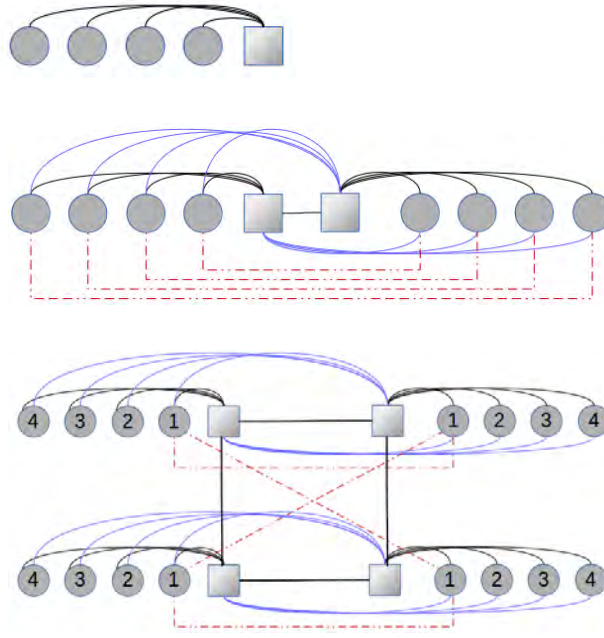


Fig. 2: Schematic representation of Hamiltonian (24) corresponding to the impurity problem for the one, two, and four-site cellular DMFT. See text for details [66].

If we define the non-interacting Green function matrix as $G_0(\omega \mathbf{1} - T)$, the local Green function within the DMFT is now given by [77]

$$G(\omega) = G_0(\omega \mathbf{1} - T - \Sigma(\omega)), \quad (23)$$

which defines the *self-consistency condition* for the $N_c \times N_c$ matrices G and Σ . The lattice problem can now be mapped onto an auxiliary impurity problem that has the same local magnitudes $G(\omega)$ and $\Sigma(\omega)$. This impurity problem should be determined iteratively. Note that G_0 , T , and Σ are $N_c \times N_c$ matrices for the spin-symmetric solution, and $2N_c \times 2N_c$ matrices in the general case. Spatial correlations or the momentum dependence of Σ can be obtained by periodization [78].

The “impurity” Hamiltonian reads

$$H_{\text{imp}} = h_0^0 + V_0 + H_b, \quad (24)$$

where the non-interacting part H_b represents the bath

$$H_b = \sum_{IJq\sigma} \lambda_q^{IJ} b_{Iq\sigma}^\dagger b_{Jq\sigma} + \sum_{IJq} v_q^{IJ} \left[b_{Iq\sigma}^\dagger c_{0J\sigma} + H.c. \right], \quad (25)$$

and $b_{Iq\sigma}^\dagger$ corresponds to the creation operator for the bath-site q , associated to the “orbital” I and spin σ . In Fig. 2 we show a scheme of this effective impurity. Here the circles (squares) represent the non-interacting (interacting impurity) sites. The red lines correspond to the λ_q^{IJ} parameters between bath sites q related to impurities I and J (they are the only hybridization between the baths related to different impurities). The blue lines are the v_q^{IJ} with $I \neq J$ while the black lines are the v_q^{II} . In the bottom scheme we omit some obvious connections for clarity.

The self-consistent iterations of the DMFT can be summarized as follows:

(i) Start with $\Sigma(\omega) = 0$,

(ii) Calculate the Green function

$$G(\omega) = G_0(\omega - T - \Sigma(\omega)), \quad (26)$$

(iii) Obtain the hybridization

$$\Gamma(\omega) = \omega \mathbf{1} - T - \Sigma(\omega) - [G(\omega)]^{-1}, \quad (27)$$

(iv) Find a Hamiltonian representation H_{imp} with hybridization $\tilde{\Gamma}(\omega)$ to approximate $\Gamma(\omega)$. The hybridization $\tilde{\Gamma}(\omega)$ is characterized by the parameters $\mathcal{V}_q = (v_q^{IJ})$ and $\Lambda_q = (\lambda_q^{IJ})$ of H_b through

$$\tilde{\Gamma}(\omega) = \sum_q \mathcal{V}_q \cdot [\omega \mathbf{1} - \Lambda_q]^{-1} \cdot \mathcal{V}_q. \quad (28)$$

(v) Calculate the impurity Green matrix $G_{\text{imp}}(\omega)$ of the Hamiltonian H_{imp} using DMRG,

(vi) Obtain the self-energy

$$\Sigma(\omega) = \omega \mathbf{1} - T - [G_{\text{imp}}(\omega)]^{-1} - \tilde{\Gamma}(\omega). \quad (29)$$

Return to (ii) until convergence.

Step (iv) requires fitting for $\mathcal{V}_q$ and Λ_q (for details see [66]).

As mentioned above, our problem is completely defined through the parameters V_i , G_0 , and T . Notice that G_0 and T are typically well known one-particle quantities for a given lattice problem. Some particular cases are given below. For example, for the Hubbard model, when considering the single-site, one-orbital DMFT, the defining matrices are $V_i = U n_{i\uparrow} n_{i\downarrow}$, $T = -\mu$ and

$$G_0(\omega - T) = \begin{cases} \frac{1}{N} \sum_{\mathbf{k}} [\omega - \varepsilon(\mathbf{k})]^{-1} & \text{Square lattice} \\ 2 [\omega + \sqrt{\omega^2 - 1}]^{-1} & \text{Bethe lattice} \end{cases}$$

where $\varepsilon(\mathbf{k}) = -2t (\cos k_x + \cos k_y) - 4t' \cos k_x \cos k_y$, with $\mathbf{k} = (k_x, k_y)$ the Fourier space of the square lattice with N sites, $N \rightarrow \infty$, and t (t') the (next-)nearest-neighbor hopping integral [79]. For the 2- or 4-site cluster Hubbard model (called the cellular DMFT [80]), $N_c = 2$ or $N_c = 4$, respectively, and the main matrices are

$$V_i = U \sum_{I=1}^{N_c} n_{iI\uparrow} n_{iI\downarrow}, \quad T = \begin{cases} \begin{pmatrix} -\mu & t \\ t & -\mu \end{pmatrix} & \text{c2-DMFT} \\ \begin{pmatrix} -\mu & t & t & t' \\ t & -\mu & t' & t \\ t & t' & -\mu & t \\ t' & t & t & -\mu \end{pmatrix} & \text{c4-DMFT} \end{cases} \quad \text{and}$$

$$G_0(\omega \mathbf{1} - T) = \frac{N_c}{N} \sum_{\tilde{\mathbf{k}}} \left[\omega \mathbf{1} - \tilde{\varepsilon}(\tilde{\mathbf{k}}) \right]^{-1}.$$

Here, T is the non-interacting intracluster matrix and $\tilde{\varepsilon}(\tilde{\mathbf{k}})$ is the intercluster hopping on the superlattice Fourier space $\tilde{\mathbf{k}}$, which is connected to the one-site lattice through

$$\tilde{\varepsilon}(\tilde{\mathbf{k}})_{IJ} = \frac{1}{N_c} \sum_{\mathbf{K}} \exp \left[i(\mathbf{K} + \tilde{\mathbf{k}}) \cdot \mathbf{R}_{IJ} \right] \varepsilon(\mathbf{K} + \tilde{\mathbf{k}}) \quad (30)$$

with $\mathbf{K}$ the intracluster Fourier-space vectors, see Eq. (23) of [81]. The above implementation is done in a real space clustering, the so-called cellular DMFT (CDMFT). An alternative and complementary cluster approach is the Dynamical Cluster Approximation (DCA). For a detailed analysis between the two see [82].

4.1.1 Example: application to the two-orbital Hubbard model

As an application of this method, we studied the non-hybridized two-orbital Hubbard model with different band widths

$$H = \sum_{\langle ij \rangle \alpha \sigma} t_{\alpha} c_{i\alpha\sigma}^{\dagger} c_{j\alpha\sigma} + U \sum_{i\alpha} n_{i\alpha\uparrow} n_{i\alpha\downarrow} + \sum_{i\sigma\sigma'} U_{12} n_{i1\sigma} n_{i2\sigma'}, \quad (31)$$

where $\langle ij \rangle$ are nearest-neighbor sites on a Bethe lattice, $c_{i\alpha\sigma}^{\dagger}$ creates an electron at site i in orbital $\alpha = 1, 2$ with spin σ , and $n_{i\alpha} = n_{i\alpha\uparrow} + n_{i\alpha\downarrow}$. U (U_{12}) is the intra(inter)-orbital Coulomb repulsion between electrons. The nearest neighbor hoppings are $t_1 \geq t_2$, for bands 1 and 2, respectively. We set $t_1 = 0.5$ which defines the unit of energy and we define $\Delta = U - U_{12}$.

We solved this Hamiltonian at half filling (electron-hole symmetric) using the DMFT with the improved impurity solver based on the DMRG described above, which allowed us to obtain the detailed DOS directly on the real axis directly (or with a very small imaginary offset $0.01 < \eta < 0.2$), zero temperature and system sizes of $L = 40 - 60$ sites. Here we show that, thanks to this combined method, we can observe a rich structure in the DOS which had not been seen before with more approximate techniques. We find that a finite density of states at the Fermi energy in one band is correlated with the emergence of well defined quasiparticle states at excited energies $\Delta = U - U_{12}$ in the other band. We have identified these excitations as inter-band holon-doublon bound states [83] (see Fig. (3)).

In summary, the DMFT+DMRG solver produces reliable results for non-local self energies at arbitrary dopings, hybridizations, and interactions, at any energy scale. It also allows for the calculation of large effective “impurities” to study multi-band interacting models and multi-site or multi-momenta clusters. In addition, it also gives reliable results for the case of real impurity problems, such as adsorbed atoms, cold atoms, and interacting nanoscopic systems like quantum dot arrays among others.

This method paves the way towards the inclusion of additional improvements such as symmetries, finite temperature, and more realistic systems by taking into account configurations given by ab-initio methods.

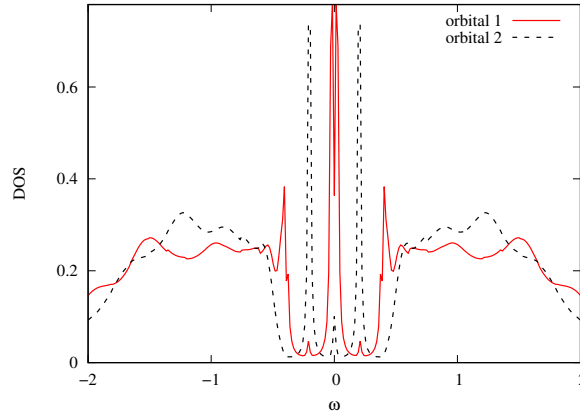


Fig. 3: *Orbital-discriminated DOS for the half-filled two-orbital Hubbard model, (31), for $U = 2.3$, $\Delta = 0.2$ where we observe the existence of in-gap quasiparticle peaks [83].*

5 Conclusions

The DMRG has become one of the most reliable techniques to calculate ground states and dynamical properties of correlated systems. In this lecture we have presented the basic DMRG formalism and given a justification of its performance from a quantum information perspective. This enables the understanding of more recent techniques based on matrix product states and tensor networks. In addition to the calculation of ground state and dynamical properties of models for correlated systems (mainly in low dimensions), we gave an example of how the DMRG can be used to solve the most complex part of the now well established DMFT, i.e, the impurity solver, for the calculation of electronic properties of more realistic models for materials. This technique uses the correction vector to obtain precise Green functions on the real frequency axis directly thus avoiding ill-posed analytic continuation methods from the Matsubara frequencies and fermionic sign problems present in quantum Monte Carlo-based techniques, allowing also for zero temperature calculations. By using a self-consistent bath with low entanglement, it produces precise spectral functions.

Acknowledgements

I would like to acknowledge important contributions of my former PhD student, Yuriel Núñez Fernández, to the optimization of the DMRG as well as to the development and implementation of the DMRG as the impurity solver of the DMFT. I also want to thank my collaborators in this topic, Marcelo Rozenberg, Daniel García, and Pablo Cornaglia for fruitful discussions and joint work. I acknowledge support from projects PICT 2016-0402 from the Argentine ANPCyT and PIP 2015-2017 11220150100538CO (CONICET) and my home institutions, CONICET, CNEA and the University of Cuyo.

References

- [1] S. White, Phys. Rev. Lett. **69**, 2863 (1992)
- [2] S. White, Phys. Rev. B **48**, 10345 (1993)
- [3] I. Peschel, X. Wang, M. Kaulke, and K. Hallberg (eds.): *Density Matrix Renormalization* (Series: Lecture Notes in Physics, Springer, Berlin, 1999)
- [4] K. Hallberg, Adv. Phys. **55**, 477 (2006)
- [5] A. Schollwöck, Rev. Mod. Phys. **77**, 259 (2005); see also R. Noack and S. Manmana, AIP Conf. Proc. **789**, 93 (2005) for a more general review on numerical methods.
- [6] G.K.-L. Chan and S. Sharma, Ann. Rev. Phys. Chem. **62**, 465 (2011)
- [7] S. Wouters and D. van Neck, Eur. Phys. J. D **68**, 272 (2014)
- [8] B. Bauer et al., J. Stat. Mech. P05001 (2011); <http://alps.comp-phys.org>
- [9] <http://quattro.phys.sci.kobe-u.ac.jp/dmrg.html>
- [10] J.C. Bridgeman and C.T. Chubb, J. Phys. A: Math. Theor. **50**, 223001 (2017)
- [11] <http://itensor.org>
- [12] K.G. Wilson, Rev. Mod. Phys. **47**, 773 (1975)
- [13] J. Bray and S. Chui, Phys. Rev. B **19**, 4876 (1979)
- [14] S.R. White and R.M. Noack, Phys. Rev. Lett. **68**, 3487 (1992)
- [15] U. Schollwöck, Ann. Phys. **326**, 96 (2011)
- [16] C. Callan and F. Wilczek, Phys. Lett. B **333**, 55 (1994);
J. Gaiete, Mod. Phys. Lett. A **16**, 1109 (2001)
- [17] J. Latorre, E. Rico and G. Vidal, Quant. Inf. Comp. **6**, 48 (2004)
- [18] G. Vidal, J. Latorre, E. Rico and A. Kitaev, Phys. Rev. Lett. **90**, 227902 (2003)
- [19] V. Korepin, Phys. Rev. Lett. **92**, 96402 (2004)
- [20] In practice it is not necessary to construct the Hamiltonian for the whole system; it is more efficient to handle the blocks separately.
- [21] E. Dagotto, Rev. Mod. Phys. **66**, 763 (1994)
- [22] E.R. Davidson, J. Comput. Phys. **17**, 87 (1975)

- [23] K. Hallberg, Phys. Rev. B **52**, 9827 (1995)
- [24] T. Kühner and S. White, Phys. Rev. B **60**, 335(1999)
- [25] P.E. Dargel, A. Honecker, R. Peters, R.M. Noack, and T. Pruschke, Phys. Rev. B **83**, 161104 (2011)
- [26] P.E. Dargel, A. Wöllert, A. Honecker, I.P. McCulloch, U. Schollwöck, and T. Pruschke, Phys. Rev. B **85**, 205119 (2012)
- [27] S. Ramasesha, Z. Shuai, and J. Brédas, Chem. Phys. Lett. **245**, 224 (1995); S. Ramasesha, S. Pati, H. Krishnamurthy, Z. Shuai, and J.L. Brédas, Phys. Rev. B **54**, 7598 (1996)
- [28] E. Jeckelmann, Phys. Rev. B **66**, 045114 (2002)
- [29] S. White and A. Feiguin, Phys. Rev. Lett. **93**, 076401 (2004)
- [30] U. Schollwöck and S.R. White: *Methods for Time Dependence in DMRG*, in G.G. Batrouni and D. Poilblanc (eds.): *Effective models for low-dimensional strongly correlated systems* (AIP, Melville, New York, 2006)
- [31] J. Sirker and A. Klümper, Phys. Rev. B **71**, 241101 (2005)
- [32] J.J. Garcia-Ripoll, New J. Phys. **8**, 305 (2006)
- [33] A. Holzner, A. Weichselbaum, I.P. McCulloch, U. Schollwöck, and J. von Delft, Phys. Rev. B **83**, 195115 (2011)
- [34] F.A. Wolf, J.A. Justiniano, I.P. McCulloch, and U. Schollwöck, Phys. Rev. B **91**, 115144 (2015)
- [35] E.R. Gagliano and C.A. Balseiro, Phys. Rev. Lett. **59**, 2999 (1987)
- [36] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964)
- [37] R.O. Jones and O. Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989)
- [38] G. Kotliar and D. Vollhardt, Physics Today **57**, 53 (2004)
- [39] A. Georges, G. Kotliar, W. Krauth, and M.J. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996)
- [40] M. Imada and T. Miyake, J. Phys. Soc. Jpn. **79**, 112001 (2010)
- [41] K. Held, Adv. Phys. **56**, 829 (2007)
- [42] V.I. Anisimov, A.I. Poteryaev, M.A. Korotin, A.O. Anokhin and G. Kotliar J. Phys.: Condens. Matter **9**, 7359 (1997)
- [43] A.I. Lichtenstein and M.I. Katsnelson, Phys. Rev. B **57**, 6884 (1998)

- [44] A. Georges and G. Kotliar, Phys. Rev. B **45**, 6479 (1992)
- [45] M.J. Rozenberg, G. Kotliar, and X.Y. Zhang, Phys. Rev. B **49**, 10181 (1994)
- [46] M. Caffarel and W. Krauth, Phys. Rev. Lett. **72**, 1545 (1994)
- [47] Y. Lu, M. Höppner, O. Gunnarsson and M.W. Haverkort, Phys. Rev. B **90**, 085102 (2014)
- [48] J.E. Hirsch and R.M. Fye, Phys. Rev. Lett. **56**, 2521 (1986)
- [49] A.N. Rubtsov, V.V. Savkin, and A.I. Lichtenstein, Phys. Rev. Lett. **72**, 035122 (2005)
- [50] P. Werner, A. Comanac, L. de Medici, M. Troyer, and A.J. Millis, Phys. Rev. Lett. **97**, 076405 (2006)
- [51] H. Park, K. Haule, and G. Kotliar, Phys. Rev. Lett. **101**, 186403 (2008)
- [52] E. Gull, A.J. Millis, A.I. Lichtenstein, A.N. Rubtsov, M. Troyer, and P. Werner, Rev. Mod. Phys. **83**, 349 (2011);
E. Gull, P. Werner, O. Parcollet, and M. Troyer, EPL **82**, 57003 (2008)
- [53] Y. Nomura, S. Sakai, and R. Arita, Phys. Rev. B, **91**, 235107 (2015)
- [54] T. Pruschke, D.L. Cox, and M. Jarrell, Phys. Rev. Lett. **47**, 3553 (1993)
- [55] R. Bulla, Phys. Rev. Lett. **83**, 136 (1999);
R. Bulla, A.C. Hewson, and T. Pruschke, J. Phys.: Cond. Mat. **10**, 8365 (1998)
- [56] K.M. Stadler, Z.P. Yin, J. von Delft, G. Kotliar, and A. Weichselbaum, Phys. Rev. Lett. **115**, 136401 (2015)
- [57] K.M. Stadler, A.K. Mitchell, J. von Delft, and A. Weichselbaum, Phys. Rev. B **93**, 235101 (2016)
- [58] F. Lechermann, A. Georges, G. Kotliar, and O. Parcollet, Phys. Rev. B **76**, 155102 (2007)
- [59] A. Isidori and M. Capone, Phys. Rev. B **80**, 115120 (2009)
- [60] M. Ferrero, P.S. Cornaglia, L. De Leo, O. Parcollet, G. Kotliar, A. Georges, EPL **85**, 57009 (2009)
- [61] D. Zgid and G.K.-L. Chan, J. Chem. Phys. **134**, 094115 (2011);
D. Zgid, E. Gull, and G.K.-L. Chan, Phys. Rev. B **86**, 165128 (2012)
- [62] D.J. García, K. Hallberg, and M.J. Rozenberg, Phys. Rev. Lett. **93**, 246403 (2004)
- [63] D.J. García, E. Miranda, K. Hallberg, M.J. Rozenberg, Phys. Rev. B **75**, 121102(R) (2007)
- [64] Y. Núñez Fernández, D. García, and K. Hallberg, J. Phys.: Conf. Series **568**, 042009 (2014); Y. Núñez Fernández and K. Hallberg, Papers in Physics **9**, 090005 (2017)

- [65] M. Karski, C. Raas, and G. Uhrig, Phys. Rev. B **72**, 113110 (2005)
- [66] Y. Núñez Fernández and K. Hallberg, Front. Phys. **6**, 13 (2018)
- [67] S. Nishimoto, F. Gebhard, and E. Jeckelmann, J. Phys.: Condens. Matter **16**, 7063 (2004)
- [68] G. Vidal, Phys. Rev. Lett. **93**, 040502 (2004); G. Vidal, Phys. Rev. Lett. **91**, 147902 (2003)
- [69] M. Ganahl, M. Aichhorn, P. Thunström, K. Held, H.G. Evertz, and F. Verstraete, Phys. Rev. B **92**, 155132 (2015)
- [70] A. Weisse, G. Wellein, A. Alvermann and H. Fehske, Rev. Mod. Phys. **78**, 275 (2006)
- [71] M. Ganahl, P. Thunström, F. Verstraete, K. Held, and H.G. Evertz, Phys. Rev. B **90**, 045144 (2014)
- [72] R. Peters, Phys. Rev. B **84**, 075139 (2011)
- [73] T. Shirakawa and S. Yunoki, Phys. Rev. B **90**, 195109 (2014)
- [74] F. Wolf, I. McCulloch and U. Schollwöck, Phys. Rev. B **90**, 235131 (2014)
- [75] D. Bauernfeind, M. Zingl, R. Triebl, M. Aichhorn, and H.G. Everts, Phys. Rev. X **7**, 031013 (2017)
- [76] F.A. Wolf, A. Go, I.P. McCulloch, A.J. Millis, U. Schollwöck, Phys. Rev. X **5**, 041032 (2015)
- [77] E. Pavarini, E. Koch, D. Vollhardt, and A. Lichtenstein (eds.): *DMFT at 25: Infinite Dimensions* (Forschungszentrum Jülich, 2014)
<http://www.cond-mat.de/events/correl14>
- [78] T. Stanescu and G. Kotliar, Phys. Rev. B **74**, 125110 (2006)
- [79] S. Sakai, Y. Motome, and M. Imada, Phys. Rev. B **82**, 134505 (2010)
- [80] G. Kotliar, S. Savrasov, G. Pálsson, and G. Biroli, Phys. Rev. Lett. **87**, 186401 (2001)
- [81] T. Maier, M. Jarrell, T. Pruschke, and M. Hettler, Rev. Mod. Phys. **77**, 1027 (2005)
- [82] E. Koch, G. Sangiovanni, and O. Gunnarsson, Phys. Rev. B **78**, 115102 (2008)
- [83] Y. Núñez Fernández, G. Kotliar, and K. Hallberg, Phys. Rev. B **97**, 121113(R) (2018)

10 Dynamical Mean-Field Theory and the Mott Transition

Marcelo Rozenberg

CNRS

Laboratoire de Physique des Solides

Université Paris-Sud, Bât. 510, Orsay 91405, France

Contents

1	Introduction	2
1.1	Strongly correlated systems	2
1.2	Kondo model and Kondo problem	4
2	Dynamical mean-field theory: a primer	6
2.1	Green functions in a nutshell	6
2.2	The DMFT self-consistency equations	8
2.3	DMFT on the Bethe lattice	10
2.4	Quantum impurity problem solvers	13
2.5	Long-range order	15
3	The Mott-Hubbard transition in DMFT	16
3.1	V_2O_3 a strongly correlated material with a metal-insulator transition	17
3.2	The Mott-Hubbard transition	18
3.3	Band-structure evolution across the metal-insulator transition	19
3.4	Coexistence of solutions and the first-order transition line	21
3.5	Endless directions	22
4	Hands-on exercise (with IPT code): the Mott-Hubbard transition	29

1 Introduction

The discovery of the high-temperature cuprate superconductors in the late 80's triggered a strong interest in the physics of transition-metal oxides. It was soon realized that understanding these systems posed a significant theoretical challenge, namely, to describe electronic systems where the independent-electron approximation fails. This became known as the problem of strongly correlated electron systems, and to a large extent remains a challenge. Nevertheless, some significant progress has been accomplished. In this lecture we shall be concerned with a particularly successful approach, namely, dynamical mean-field theory (DMFT), which was developed in the 90's and has allowed to gain new insights into the problem of strong correlations. Specifically, it provided a significant advance in our understanding and description of one of the classic problems in the field, the Mott metal-insulator transition. For a detailed account on how DMFT was developed, the interested reader is referred to the review article [1]. The goal of the present lecture is to introduce DMFT and its application to the problem of the Mott-Hubbard transition in a pedagogical manner, putting emphasis on the new concepts that it brought to light. The lecture is aimed at final-year undergraduates, beginning graduates, or anybody looking for an accessible presentation to the concepts of DMFT, *including experimentalists*. The lecture is supplemented with a computational code, which allows the interested reader to solve the basic DMFT equations. We also propose a set of problems that will guide the reader in the discovery of the physics of the Mott-Hubbard metal-insulator transition.

We shall begin by illustrating, from an experimental point of view, the manifestations of strong correlation phenomena with special attention to that of the Mott transition. We shall then describe in simple terms the DMFT approach by drawing an analogy with the classic mean-field theory of spin models. We then move on to consider the solution of the prototype model of strongly correlated systems, the Hubbard model, which is a minimal model to capture the metal-insulator transition. We shall discuss the transition as a function of interaction strength, temperature, and doping in both, the paramagnetic and the antiferromagnetic phase. We shall describe some basic experimental data on a material that is widely considered to exhibit an actual Mott transition and discuss the connection to theoretical results of the Hubbard model within DMFT.

1.1 Strongly correlated systems

How do we know that we are dealing with a strongly correlated system? This question is important, because the models and their solutions should illustrate precisely those aspects. There are a few physical phenomena which we may consider to be key. The first that we can mention is the presence of complex phase diagrams (Fig. 1). Ordinary materials, are either metals or insulators, or even semiconductors if their gaps are small with respect to room temperature. Common examples are gold, diamond (gap 5.5 eV), and silicon (gap 0.67 eV), respectively. They are relatively easy to understand already by looking whether the outermost electronic orbital shell is partially filled or not. In the case of carbon, the $2p$ orbital has two electrons and this permits different structural arrangements that lift the orbital degeneracy. Thus if the degeneracy is fully

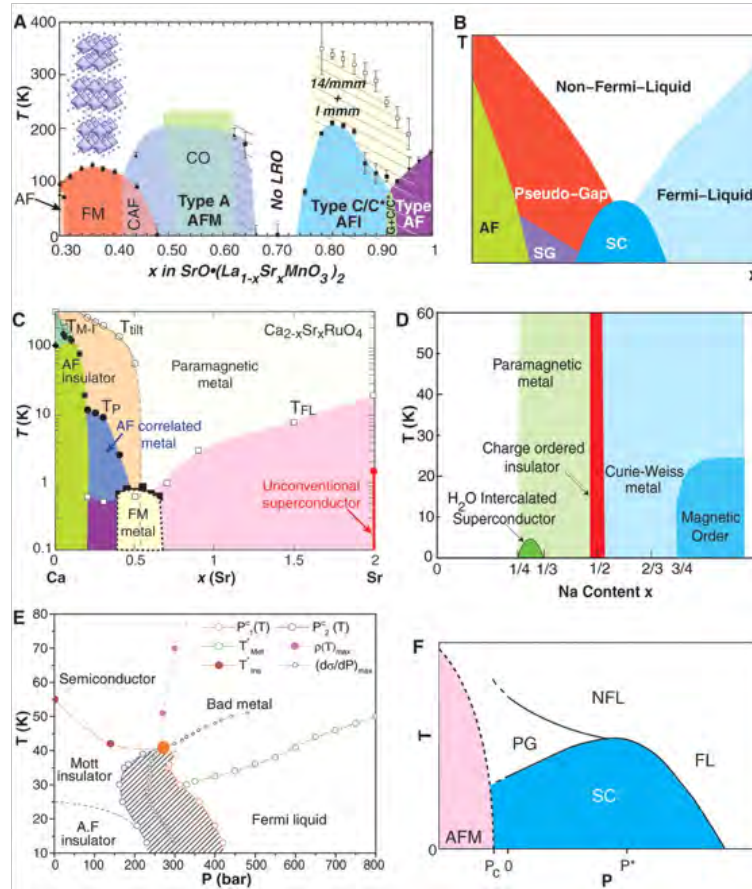


Fig. 1: A complex phase diagram is characteristic of systems with strong electronic correlations. Examples from manganites (A), cuprates (B), ruthenates (C), cobaltates (D), 2-d κ -organics (E), and heavy fermions (F). From [2].

lifted, the $2p$ band is full as in diamond, but if not, it is metallic as in graphite. In the case of gold, the outermost shell is the partially filled $6s^1$, which leads to a metallic structure. These simple materials remain in their stable phase upon heating from low temperatures, without significant changes in their electronic structure. Only at very large temperatures, well above room temperature, the crystalline structure may eventually give up due to phonon excitations. Strongly correlated systems are different. They exhibit dramatic changes in their electronic properties, even at temperatures smaller than room temperature. Examples of these phenomena are the metal insulator transition in V_2O_3 and in the family of nickelates $XNiO_3$ ($X = La, Sm, Pr$, etc.) the magneto-resistance of manganites $La_{1-y}X_yMnO_3$ ($X = Sr, Ca$, etc.) and the superconductivity in cuprates such as $La_{2-y}Sr_yCuO_4$, $Bi_2Sr_2Ca_1Cu_2O_{8+y}$, $YBa_2Cu_3O_{7-y}$, $HgBa_2Ca_2Cu_3O_8$, among many others [2]. Iridates, such as Sr_2IrO_4 [3, 4] are currently receiving a great deal of attention for their potential “topological” properties. And we may also mention more exotic structures, such as the molecular crystals of “buckyballs” A_3C_{60} ($A = K, Rb, Cs$, etc.) that may exhibit superconductivity at $\sim 35K$ [5, 6]. These changes in their electronic transport properties are also correlated with anomalous spectroscopic properties, which involve transfer of spectral intensity that takes place over energy scales of the order of an eV. This becomes significant when we realize that $1 \text{ eV} \sim 11\,000 \text{ K}$. So the question is how, by heating up a material to $\sim 100 \text{ K}$, we may observe changes on energy scales 100 times larger.

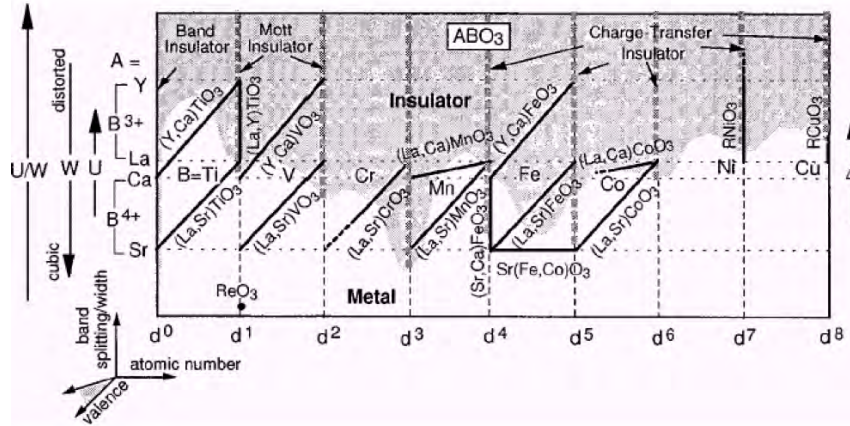


Fig. 2: Schematic phase diagram of transition-metal oxides as a function of the partial filling of the 3d-orbital band and the intensity of Coulomb correlations. Mott insulators are found at integer fillings of the d-shell. From [7].

The most amusing playground for strongly correlated physics, from the point of view of materials, has been that of transition-metal oxides. In particular those transition metals that occupy the third row of the periodic table, filling the 3d orbital shell. As a function of the filling we find a large variety of oxide materials, which are expected to be metals from density-functional theory (DFT) calculations, but are found to be insulators, as illustrated in the Fig. 2. Moreover, those unexpected insulators lead to anomalous metallic states upon chemical doping. A survey of those systems has been condensed into an excellent review by Imada, Fujimori, and Tokura [8]. A practical, but certainly non-rigorous definition of strongly correlated systems could be given: They are those materials whose electronic state and band-structure fail to be described by DFT.

1.2 Kondo model and Kondo problem

One of the oldest problems in strongly correlated materials, and certainly one of the, conceptually, most important, is that of the observed minimum of the resistivity in metals with magnetic impurities, which led to the formulation of the Kondo model. The physical phenomenon consists on the observation of a minimum in the resistivity of an ordinary metal with a small amount of magnetic impurities, such as gold with Mn impurities, as schematically depicted in Fig. 3. The problem was theoretically addressed by Kondo, who considered a Hamiltonian of an ordinary metal of bandwidth W , interacting with an embedded single magnetic impurity with a spin interaction J . Kondo showed that diagrammatic perturbation theory broke down at a low temperature, where logarithmic divergences developed. This became the Kondo problem, which led to very important developments. We may mention a wonderful paper by Anderson, known as “Poor’s man scaling,” which was an important step in the right direction. Eventually, Wilson invented the numerical renormalization group (NRG) in the 70’s, providing an exact numerical solution and a conceptual breakthrough. The problem was analytically solved by Andrei and Tsvelik in the 80’s using the Bethe Ansatz, a highly technical mathematical methodology. An important concept that emerged from the solution of the Kondo problem was that of the

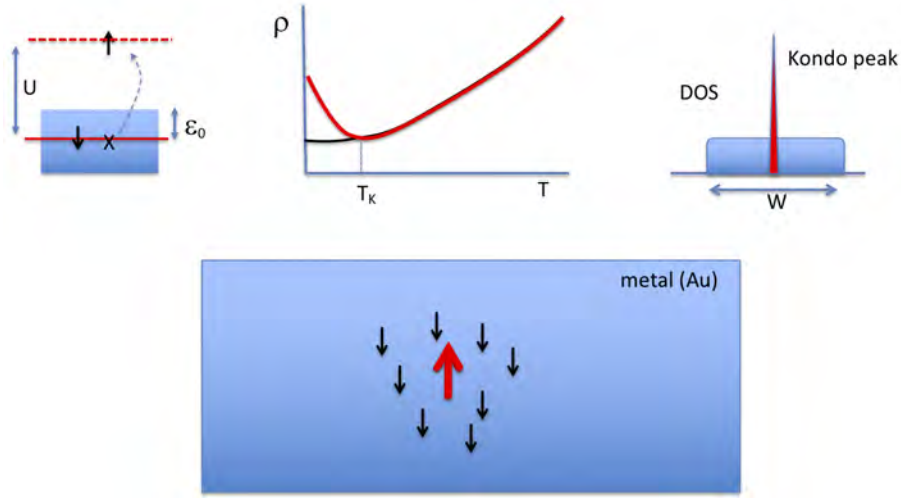


Fig. 3: Schematic illustration of the Kondo effect. Top: (left) The magnetic moment of an impurity site emerges when the second electron is hindered to doubly-occupy a site below the Fermi sea due to the interaction energy cost U (dotted energy level). The physical manifestation of the Kondo effect are: (middle) a logarithmic increase of the resistivity at low T (red line), with respect to the non-magnetic impurity scattering (black line); and (right) a sharp resonance in the DOS at the Fermi energy (the Kondo peak). Bottom: Illustration of the (dynamic) screening of the magnetic impurity by the conduction electrons forming a many-body singlet state.

Kondo resonance and the Kondo temperature. The former is a peak in the local density of states (at the site of the magnetic impurity), which corresponds to a many-body state where the metallic electrons dynamically screen the magnetic moment of the impurity, forming a singlet state. This phenomenon occurs below the Kondo temperature, which is exponentially small: $T_K \sim W e^{-W/J}$. The solution of this problem already illustrates the characteristics of strong correlations we have mentioned before: There is a change in the electronic conduction (Kondo minimum) and in the spectral properties (Kondo resonance), which all occur at a low temperature (Kondo temperature) well below the bare energies scales of the model (W and $J \sim \text{eV}$, while $T_K \sim 10 \text{ K}$).

The Kondo model can be generalized into the single impurity Anderson model (SIAM), where the magnetic impurity is represented by an atomic site with energy $\varepsilon_0 \leq 0$ (beneath the surface of the Fermi sea) and a local Coulomb repulsion U . For large values of U , the double occupation of the site is penalized so the orbital occupied by only one electron describes a magnetic impurity. The atomic site is hybridized with the conduction band of the metallic host via an amplitude V . This permits the conduction electrons to briefly (doubly) occupy the impurity site, screening its spin. Because of the high energetic cost U , one of the electrons of the impurity returns to the metal, which may produce a “spin-flip” of the impurity spin. These processes lead to the formation of a non-magnetic many-body state involving both the impurity and the conduction electron degrees of freedom. Similarly as in the Kondo model, this occurs below a low temperature scale. The relation between the two models is $J \sim V^2/U$, for the case where $\varepsilon_0 = 0$ and U is large. An instructive problem to solve is to consider a minimal SIAM as a two-site Hamiltonian problem, where one “impurity” site has energy ε_0 and a correlation term

$Un_{0\uparrow}n_{0\downarrow}$, the other “conduction band” site has energy zero, and they are hybridized by a hopping amplitude V . The density of states (DOS) of the sites can be computed, along with all interesting observables, such as the magnetic correlation functions and the magnetic moments, with relatively small numerical effort, even at finite T . Already the DOS will show a striking temperature dependence, with transfers of spectral weight across large energy scales, the emergence of a precursor of the Kondo peak, and the magnetic screening of the impurity spin. As we shall see below, the importance of the Kondo and the SIAM is not only conceptual: they form the very heart of the DMFT method.

2 Dynamical mean-field theory: a primer

The best way to introduce dynamical mean-field Theory (DMFT) is to draw an analogy with the familiar mean-field theory of the Ising model, which is a text book case of statistical physics. However, before doing that we need to give a brief introduction to Green functions (GF), since these mathematical objects are central to the formulation of DMFT. Unlike the Ising spins, GF are frequency (or time) dependent objects, hence the “dynamical” aspect of the DMFT. We shall avoid mathematical rigor and focus just on the aspects of the GF that we need to carry on the discussion. We shall avoid using vectors, also in the sake of keeping the notation light. The meaning should be always clear from the context. There are excellent text books on the topic of GFs, a classic one is that by G. Mahan [9].

2.1 Green functions in a nutshell

Physically, the GF are mathematical objects that characterize the propagation of particles through the lattice. Therefore they have site and time coordinates, or equivalently, lattice momentum and frequency (k, ω) . In this case, the physical interpretation is that the GF describes the process of adding (removing) a particle with energy $\omega > 0$ ($\omega < 0$) and momentum k . The GF are complex functions that are defined on the whole complex plane $G(k, z)$, with $z \in \mathbb{C}$, where $\omega = \text{Re}(z)$. They are analytic, so they are defined on both, the real and imaginary axis. In practice we use both. The real frequency axis GF provide functions that can be compared with experiments. For instance, the imaginary part of the local GF (i.e., at position $x=0$) as a function of real frequency ω is the density of states (DOS) $\rho(\omega)$, which is measured by photoemission and scanning tunneling spectroscopy experiments,

$$\rho(\omega) = -\frac{1}{\pi} \text{Im} G_{\text{loc}}(\omega) = -\frac{1}{\pi} \text{Im} \sum_k G(k, \omega). \quad (1)$$

In contrast to the continuous variable ω , on the imaginary frequency axis the GF is defined on a set of discrete frequencies, called (fermionic) Matsubara frequencies $\omega_n = (2n+1)\pi/\beta$, with $n \in \mathbb{Z}$ and $\beta \equiv 1/T$ is the inverse temperature. The interest of using GFs on the imaginary axis is that they are often easier to compute than their real axis counterparts. There is an important

price to pay, however, which is the need to analytically continue the $G(k, i\omega_n)$ to obtain the $G(k, \omega)$ so comparisons can be made with experiments.

The simplest GF is that of a single orbital state, isolated, and of energy ε_0 ,

$$G(\omega) = \frac{1}{\omega + \varepsilon_0 - i\eta}. \quad (2)$$

The imaginary part (i.e., the DOS) has a delta function from the simple pole at ε_0 . For a tight binding Hamiltonian, with a band that disperses as ε_k , the GF becomes a function of k and ω

$$G(k, \omega) = \frac{1}{\omega - \varepsilon_k + i\eta}, \quad (3)$$

whose imaginary part has delta-peaks at the poles of $G(k, \omega)$ that provide the electronic band energy dispersion ε_k , while $i\eta$ provides a small width to the peaks. It is called the *spectral function*, $A(k, \omega) = -\text{Im } G(k, \omega)/\pi$ and is, in principle, measured in ARPES experiments.

The lifetime of the excitations is given by the inverse of the frequency-width of the peaks appearing in the spectral functions $A(k, \omega)$. In a non-interacting system, as in a tight binding Hamiltonian with dispersion ε_k , the single particle states are eigenstates and are stationary, so they have an infinite lifetime. Accordingly, the poles of the GF become infinitely narrow delta-peaks $\delta(k - k_0, \omega - \varepsilon_{k_0})$ in the spectral function.

Interactions, such as on-site Coulomb repulsion, that in the Hubbard model have the same form, $U n_{\uparrow} n_{\downarrow}$, as in the SIAM affect the non-interacting band structure. They may change the energy dispersion and can also change the lifetime of the excitations. In extreme cases, they can qualitatively change the energy dispersion or “electronic structure.” We shall see a concrete instance in the Mott-Hubbard metal-insulator transition. Mathematically, the effects of the interactions is encoded in the calculation of another complex function that shares the same analytic properties as the GF. It is called the self-energy $\Sigma(k, \omega)$. Thus solving the many-body problem of an interacting model amounts to obtaining the self-energy. Let’s write down the definitions for the concrete case of a Hubbard model (HM)

$$H = H_0 + U \sum_i n_{i\uparrow} n_{i\downarrow} \quad \text{with} \quad H_0 = -t \sum_{\langle i,j \rangle \sigma} c_{i\sigma}^\dagger c_{j\sigma}, \quad (4)$$

where $\langle i, j \rangle$ denote nearest neighbors sites and $n_{i\sigma} = c_{i\sigma}^\dagger c_{i\sigma}$. Thus, we have for the GF

$$G_0(k, \omega) = \frac{1}{\omega - \varepsilon_k + i\eta} \quad \text{and} \quad G(k, \omega) = \frac{1}{\omega - \varepsilon_k - \Sigma(k, \omega)}, \quad (5)$$

where G_0 is called the non-interacting GF. We see in the expression of the GF how the Σ function can modify the electronic dispersion of H_0 : $\text{Re } \Sigma$ changes the energy of the excitations, while $\text{Im } \Sigma$ changes their lifetime $\tau_L = 1/\text{Im } \Sigma(\omega=0)$.

Two more definitions will be useful. The notion of quasiparticle residue Z and renormalized mass m^* . They both serve to parametrize the effect of interactions for the low energy band-structure. We can write it as the sum of two contributions

$$G(k, \omega) \approx \frac{Z}{\omega - Z \varepsilon_k} + (1 - Z) G_{\text{inc}}(k, \omega). \quad (6)$$

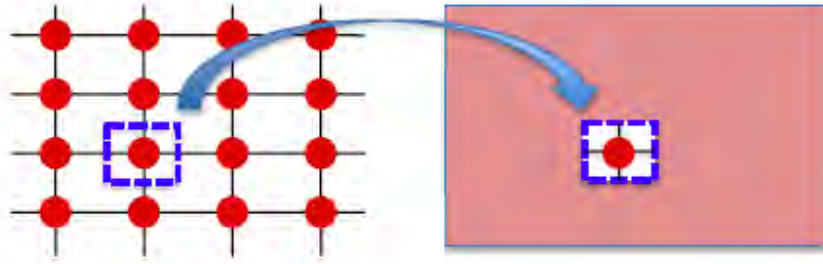


Fig. 4: One site from the lattice is embedded into an effective electronic bath. The bath is self-consistently determined to best represent the lattice environment of the site. This single-site quantum-impurity problem remains a non-trivial many-body problem.

The quasiparticle residue is $0 \leq Z \leq 1$, and represents the part of the DOS which remains with a well defined energy dispersive structure. The excitations are modified as $Z\varepsilon_k$, which implies that the electronic band becomes flatter, i.e., has a higher mass. If $\varepsilon_k = -t \cos(ka) \approx k^2/2m$, we see that $Z\varepsilon_k$ gives an enhanced effective mass $m^* = m/Z > m$. Since the DOS is normalized to 1, the spectral weight which is not in the quasiparticle part at low energy has to appear at higher energy. This contribution does not have a very well defined dispersion (due to short lifetimes from a large $\text{Im } \Sigma$) and thus we have the factor $1-Z$ in front of the second contribution, which we call *incoherent* G_{inc} .

We have gone fast on these definitions. There are excellent text books for the interested reader to learn more details [9].

2.2 The DMFT self-consistency equations

We use the functional integral formalism to introduce the main DMFT equations. Again, there are excellent text books on that formalism too [10]. The method is based on writing down the action of a Hamiltonian model. We shall skip all the details of the formalism and simply write down directly the most important expressions, which should be clear enough. For simplicity we shall focus on the Hubbard model (HM) defined above in Eq. (4). The action of the HM reads

$$S = \int_0^\beta d\tau \left(\sum_{i\sigma} c_{i\sigma}^\dagger \partial_\tau c_{i\sigma} - t \sum_{\langle i,j \rangle \sigma} c_{i\sigma}^\dagger c_{j\sigma} - \mu \sum_{i\sigma} n_{i\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow} \right). \quad (7)$$

The model is defined on a given lattice. The next step is, as in standard mean-field theory, to single out a site and try to replace the original lattice problem by an effective *quantum impurity problem* (QIP) embedded in a medium that is determined so to best represent the original environment of the site. This is pictorially represented in Fig. 4. More concretely, we split the action into that of the single lattice site $i = 0$ at the “origin,” S_0 , the rest of the lattice, $S^{(0)}$, and the coupling between them, ΔS : $S = S_0 + \Delta S + S^{(0)}$

$$S_0 = \int_0^\beta d\tau \left(c_{0\sigma}^\dagger (\partial_\tau - \mu) c_{0\sigma} + U n_{0\uparrow} n_{0\downarrow} \right) \quad \text{and} \quad \Delta S = \int_0^\beta d\tau \left(-t \sum_{\langle i,0 \rangle \sigma} c_{i\sigma}^\dagger c_{0\sigma} + c_{0\sigma}^\dagger c_{i\sigma} \right). \quad (8)$$

By means of standard field-theory methods we integrate out the rest of the lattice and write the full effective action on the lattice site at the origin alone as

$$S_{\text{eff}} = \int_0^\beta d\tau \left(c_{0\sigma}^\dagger (\partial_\tau - \mu) c_{0\sigma} - \sum_{ij\sigma} t_{0i} t_{j0} G_{ij}^{(0)} + \dots + U \sum_i n_{i\uparrow} n_{i\downarrow} \right). \quad (9)$$

$G_{ij}^{(0)}$ denotes the exact propagator of the lattice with the 0-site excluded and (...) stand for higher order terms in the hopping t . Up to now there are no approximations, but the problem remains too hard to deal with. DMFT corresponds to taking the limit of large dimensionality, large lattice connectivity or large lattice coordination. An evident problem is that the number of neighbor sites to 0 grows to infinity (i.e., the band-width would grow to ∞). Metzner and Vollhardt [11] realized that the reasonable way to cure this problem is by rescaling the hopping amplitude $t \rightarrow t/\sqrt{d}$, where d is the number of spatial dimensions. The simplest way to see this is noting that the typical value of the kinetic energy for a random k vector is $E_{\text{kin}} = \sum_i^d -t \cos(ka) \propto (\sqrt{d}t)^2$ (using the central limit theorem). A key consequence of this scaling is that the (...) in Eq. (9) vanish (as they are higher order in t), which is a great simplification. Thus, we recognize from S_{eff} the quantum impurity problem that we were looking for

$$S_{\text{QIP}} = \sum_{n\sigma} c_{0\sigma}^\dagger \mathcal{G}_0^{-1}(i\omega_n) c_{0\sigma} + \beta U n_{0\uparrow} n_{0\downarrow}, \quad (10)$$

with the “non-interacting” GF of the QIP defined as

$$\mathcal{G}_0^{-1} = i\omega_n + \mu - t^2 \sum_{(ij)} G_{ij}^{(0)}(i\omega_n), \quad (11)$$

where (ij) denotes the sites neighboring 0 and we went from imaginary time τ to the Matsubara frequency ω_n representation. Notice that the $\mathcal{G}_0$ is the bare propagator of the QIP and should not be confused with G_0 , which is the bare local propagator of the lattice. The last term represents the environment of the impurity, which still needs to be determined. Its physical interpretation is that an electron at the impurity site has an amplitude t to hop out to a neighbor site i , then it propagates through the rest of the lattice from i to j with $G_{ij}^{(0)}$, and returns from site j back to the impurity site with a second hop t . The 0 index in $\mathcal{G}_0$ indicates that it is the non-interacting GF of the impurity problem, but this object is, in general, different from the original lattice local non-interacting GF.

$G_{ij}^{(0)}$ is a fully interacting GF whose solution is, in principle, at least as hard as the original problem. So we need to do something about it. Diagrammatically, we can write the cavity GF

$$G_{ij}^{(0)} = G_{ij} - \frac{G_{i0} G_{0j}}{G_{00}}, \quad (12)$$

where the second term subtracts from the first all the diagrams that go back to the origin, and its denominator takes care of the double counting of local diagrams. Notice that $G^{(0)}$ has now been written in terms of the lattice GF. If we *assume* a k -independent self-energy we can express the lattice GF in real space by summing over specifying the geometry of the lattice and Fourier

transforming. Then, inserting Eq. (12) into (11) the sum over spatial indices can be performed and one obtains the self-consistency equation [1]

$$\mathcal{G}_0^{-1}(i\omega_n) - \Sigma(i\omega_n) = \sum_k \frac{1}{i\omega_n + \mu - \varepsilon_k - \Sigma(i\omega_n)} = G_{\text{loc}}(i\omega_n). \quad (13)$$

This expression is valid for all lattices. For instance, on a hyper-cubic lattice, which is the generalization of the square and cubic lattices to high dimensions, the integral over k can be done as a sum over the hyper-cubic lattice single-particle energies ε

$$G_{\text{loc}}(i\omega) = \int_{-\infty}^{+\infty} \frac{1}{i\omega_n + \mu - \varepsilon - \Sigma(i\omega_n)} \text{DOS}(\varepsilon) d\varepsilon, \quad (14)$$

where

$$\text{DOS}(\varepsilon) = \frac{e^{-\varepsilon^2/2t^2}}{t\sqrt{2\pi}} \quad (15)$$

is a Gaussian function (again using the central limit theorem).

A few comments are in order now. From Eq. (13) and (10) we see that $\Sigma(i\omega_n)$ is the self-energy of the QIP. It is obtained as the solution of the many-body single-site problem, i.e., the QIP, which depends on $\mathcal{G}_0$. Thus, we can write $\Sigma = \Sigma[\mathcal{G}_0]$ so that the self-consistent nature of Eq. (13) becomes evident. The key feature that links this equation to the original lattice problem is that it can be shown that *at the self-consistent point*, the QIP self-energy $\Sigma(i\omega_n)$ *coincides* with the exact self-energy of the lattice $\Sigma(i\omega_n)$ [1]. Crucially, in the limit of large dimensionality or lattice connectivity and with the re-scaling of the hopping made above, the lattice self-energy *is* k -independent, which validates the assumption made to obtain Eq.13 [1]. Hence, at the self-consistent point, we also recognize on the right hand side of Eq. (13) the *local* GF of the lattice problem $G_{\text{loc}}(i\omega_n)$, which we set out to solve.

So the issue is now reduced to obtaining the self-energy, given the impurity $\mathcal{G}_0$. We see that given a guess for $\mathcal{G}_0$, we solve the many-body QIP to obtain a guess for Σ . We input that into the r.h.s. of Eq. (13) to obtain a guess for G_{loc} . Then, from $G_{\text{loc}} + \Sigma$ we get a *new* guess for $\mathcal{G}_0$. This has to be iterated until self-consistency is attained. Then the problem is solved as we obtain the DMFT solution for the *lattice* GF as

$$G(k, i\omega_n) = \frac{1}{i\omega_n - \varepsilon_k - \Sigma(i\omega_n)}. \quad (16)$$

This equation, with a k -independent Σ is exact for lattices in infinite spatial dimensions or infinite connectivity. However, the procedure may be adopted for lattices in any dimension and that case the DMFT and the k -independence of Σ become an approximation, which is at the root of the realistic DMFT approach for materials that we shall describe later.

2.3 DMFT on the Bethe lattice

Another illuminating light can be cast on the DMFT self-consistency condition by considering the Bethe lattice. There are two main features that make this lattice a very popular choice for

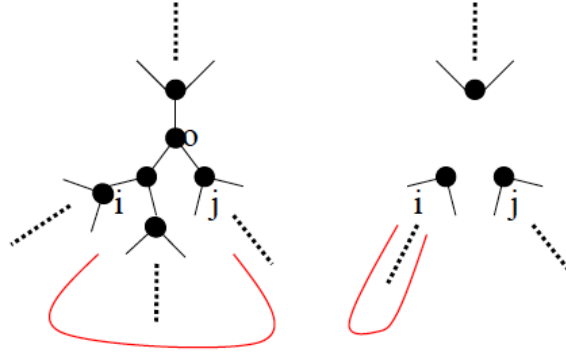


Fig. 5: The Bethe lattice (for connectivity $d = 3$). The red line indicates the “cavity” propagator. When the origin (0) is taken out, the electrons have to leave and return to the origin via the same neighbor site (i), rendering the cavity propagator site-diagonal, cf. Eq. (12).

DMFT studies. The first one is that the $\text{DOS}(\omega)$ is a semi-circle, which has a finite band-width and band-edges similar to 3D cubic lattices. The second is that the self-consistency equations are easier to derive. The Bethe lattice of coordination z is a “branching tree”, where from each node emanates a number z of branches. In Fig. 13 we show the case $z=3$. In this type of lattice the cavity $G_{ij}^{(0)}$ is easy to obtain. In the left panel of the figure we depict with a red line a propagation from site i to j , both nearest neighbors of the 0-site. The cavity propagator has to be obtained with the 0-site removed. The right panel shows that when we do that, an electron that has hopped from 0 to i can only return back to 0 hopping from i . In other words, $G_{ij}^{(0)} = G_{ii}^{(0)} \delta_{ij}$. Now, if we consider the limit of large lattice coordination, $z \rightarrow \infty$, we see from the right panel of Fig. 5 that the $G_{ii}^{(0)}$ is identical to G_{ii} , which by definition is G_{loc} . Thus, from the cavity Eq. (12)

$$G_{ij}^{(0)} = G_{ii}^{(0)} \delta_{ij} = G_{ii} \delta_{ij} = G_{\text{loc}} \delta_{ij}. \quad (17)$$

Replacing into Eq. (11), we can perform the sum to get

$$\mathcal{G}_0^{-1} = i\omega_n + \mu - t^2 G_{\text{loc}}(i\omega_n), \quad (18)$$

where we used that the hopping is rescaled, as mentioned before, by $t \rightarrow t/\sqrt{z}$. The self-consistency equation for the Bethe lattice has a very compact and intuitive form, and avoids the need of the ε integral of the hyper-cubic case.

From the point of view of the QIP problem we observe that the quantum impurity is embedded in a medium that is nothing but t^2 times the local GF. The problem is dealt with similarly as before: given a guess for the $\mathcal{G}_0$, the many-body problem of the 0-site is solved. An interacting GF for the impurity is obtained and a new guess for $\mathcal{G}_0$ is simply computed from Eq. (18). The iteration proceeds until convergence is attained and at that point, as before, the $G_{\text{loc}}(i\omega_n)$ becomes the local GF of the original lattice problem.

To show that the self-consistency condition obeys the general form of Eq. (13), we first compute the local *non-interacting* Bethe lattice $\text{DOS}(\varepsilon)$. Since $U=0$, we have $G_{\text{loc}}=\mathcal{G}_0$. Thus, replacing into Eq. (18) and solving the quadratic equation, we can get the $G_{\text{loc}}(i\omega_n)$ and from the

imaginary part of the analytic continuation to the real-frequency axis

$$\text{DOS}(\varepsilon) = \frac{2}{\pi D^2} \sqrt{1 - \left(\frac{\varepsilon}{D}\right)^2} \quad \text{for } |\varepsilon| < D, \quad (19)$$

where we introduced the half-bandwidth $D = 2t$. Inserting this $\text{DOS}(\varepsilon)$ into the general Eq. (14) and using the Dyson equation for the quantum impurity, $G_{\text{loc}}^{-1} = \mathcal{G}_0^{-1} - \Sigma$, one re-derives the Bethe lattice self-consistency condition Eq. (18).

Experience has shown that for simple model Hamiltonians there are no qualitative differences in the DMFT solutions of different lattices, therefore, the simplicity of the Bethe lattice often justifies the choice.

An important feature of the solution is that the lattice self-energy Σ , which coincides with that of the QIP at self-consistency, does not depend on momentum. This is evidently not a characteristic of a Σ in general, so we may ask if this makes sense. Or, in other words, when should we expect Σ to be independent, or weakly dependent, of the momentum? The answer is for physical problems where the stronger interactions are local and when the lattice coordination is large. In fact, the latter follows from a mathematical statement, which is that the DMFT solution becomes exact in the limit of large spatial dimensions [11, 1]. In practice, one may then expect that for cubic, bcc, and fcc lattices with coordinations 6, 8, and $12 \gg 1$, the approach should be reliable. On the other hand, in regard to locality of interactions, we may expect them to dominate the physics when the orbitals are small with respect to the interatomic distances of the material (or more simply to the lattice spacing). This is the case for two types of materials, such as the $3d$ transition-metal oxides and the heavy fermions. Among the first we have the strongly correlated materials that we listed in the Introduction (Sec. 1). Heavy fermions are typically inter-metallic compounds, such as CeCu_6 , CeAl_3 , UBe_{13} , and UPt_3 , that have an ordinary metal and one with active f -electrons, such as the actinides.

How about the case for a material that has a relatively low coordination but does have strong local interactions? That is the case of an important class of materials, the high- T_c cuprate and the iron-based superconductors. Both these systems have layered structures. The answer may depend on the person that this question is asked to. But more fairly, one should say that the relevance of DMFT may depend on the type of physical question that one is asking. We shall see some examples later on that shall illustrate this point.

We can summarize the DMFT method and its self-consistent nature by drawing an analogy with the familiar mean-field theory of an Ising model. This is schematically shown in Fig. 6. On the left panel we have the main DMFT ingredients: a lattice model and Hamiltonian; the mapping to the quantum impurity problem and the effective action of the single-site of the lattice; the restoration of the spatial translation invariance of the lattice by enforcing a self-consistency constraint; and the requirement for a rescaling of the original hopping parameter so that all terms in the model remain of finite energy when one takes the limit of high dimensionality. As we can see in the right panel of the figure, all those ingredients have a counterpart in the Ising model MFT. One can observe that in the latter the local magnetization m (or h_{eff}) is an a priori unknown that needs to be determined, similarly as the $\mathcal{G}_0$. Another feature is that the



$$H = - \sum_{i,j,\sigma} t_{ij} c_{i\sigma}^\dagger c_{j\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow}$$

$$H = \sum_{i,j} J_{ij} S_i S_j$$

$$S_{\text{eff}}[\mathcal{G}_0] = - \int \int d\tau d\tau' c_{0\sigma}^\dagger \underline{\underline{\mathcal{G}_0^{-1}}} c_{0\sigma} + U \int d\tau n_{0\uparrow} n_{0\downarrow} \quad H_{\text{eff}} = \left(\sum_i J_{0i} S_i \right) S_0 = z J \underline{\underline{m}} S_0 = h_{\text{eff}} S_0$$

$$\underline{\underline{\mathcal{G}_0^{-1}}} = i\omega_n + \mu - t^2 G(i\omega_n)$$

$$\underline{\underline{m}} = \langle S_0 \rangle = \tanh(\beta z J \underline{\underline{m}})$$

$$t_{ij} \sim 1/\sqrt{z}$$

$$J_{ij} \sim 1/z$$

Fig. 6: Analogy between Hubbard model DFMT and Ising model MFT. We highlight the similar role of the “origin” site 0 and the a priori unknown “cavity” or “Weiss field” function $\underline{\underline{\mathcal{G}_0^{-1}}}$ and the mean magnetization $\underline{\underline{m}}$. Also, both methods become exact in the limit of large dimensions (or connectivity z) after the required “rescaling” of the hopping in DMFT and the magnetic interaction in the MFT. Notice that in the Hubbard model the super-exchange $J \sim 4t^2/U$ also becomes rescaled by $1/z$.

numerical difficulty to solve the model is dramatically reduced by mapping to a single site. Solving a single Ising-spin is trivial, however, solving the QIP still remains a difficult many-body problem. A variety of techniques have been developed over the years to obtain reliable numerical solutions.

2.4 Quantum impurity problem solvers

An important technical point that we should mention is the practical solution of the QIP, which remains a non-trivial many-body problem [12]. From the start we should say that despite almost 30 years of work, where a variety of methods have been proposed [13], there is no single ideal one. We shall briefly comment on the most important techniques. We recall that the goal is to solve an arbitrary single-impurity Anderson model (SIAM), where the interacting atomic site is hybridized to an environment, or “bath” that is specified by a $\text{DOS}(\omega)$. In the standard SIAM the bath represents a metal, but in the present case the bath is a function that evolves under the iterative procedure. For a Bethe lattice it coincides with the local GF as we discussed above. The main techniques are the following:

Quantum Monte Carlo: This is a finite temperature method that is performed on the imaginary time axis, so it produces solutions to the model on the Matsubara frequency axis. This has the drawback that it requires the additional step of analytic continuation, which presents significant technical problems regarding the reliability and the precision of the spectra. This

can be mitigated improving the accuracy of the MC calculation. It is perhaps the most powerful method. It was originally implemented via a Trotter expansion of the action and a discrete Hubbard-Stratonovich transformation [14, 1]. More recently, a continuous time formulation was developed, based on a statistical sum of diagrams [15]. Its main advantages are that it is numerically exact (in the statistical sense), that its scaling to multi-orbital models is not bad, and that is easily parallelizable. Among its main drawbacks are the need of analytic continuation that we mentioned, the increased numerical cost to lower the simulation temperature, and the so called “minus sign” problem that prevents the solution of certain multi-orbital problems, especially in the case of cluster methods.

Exact Diagonalization: In this method one adopts a bath of non-interacting fictitious atoms that are coupled to the impurity site. The bath is thus defined by the atomic energies and the coupling amplitudes. Given a set of values for these parameters, a SIAM Hamiltonian is exactly diagonalized by standard techniques and the GF is obtained. The solution is used, via the self-consistency equation (Eq. (13) and 18) to compute the new bath, which is fit to obtain a new set of parameters for the fictitious atoms [1]. The main advantages of this method is that it can be formulated at zero or finite temperature, that it does not pose the problem of analytic continuation, and that its accuracy can be systematically improved. Its drawbacks are that it is numerically costly, especially for multi-orbital models and going to finite temperatures (requires full diagonalization of the Hamiltonian). This is due to the poor scaling of the size of the Hilbert space, which severely limits the number of sites in the bath, typically to about 10, which makes the pole structure of the GF quite discrete. This problem can be overcome by representing the bath with a linear chain and using the DMRG method for the solution [16, 17]. One can implement baths with up to 100 atoms. However, the scaling to multi-orbital models is still poor. We may also mention the solution of the SIAM using Wilson’s NRG method [18]. This approach also allows to implement large atomic baths and provides excellent accuracy at low frequency and zero temperature. Its main shortcoming is, as for DMFT-DMRG, the poor scaling for multi-orbital models moreover it is not particularly advantageous for the study of insulating states.

Iterative Perturbation Theory: The IPT method has both, remarkable advantages and limitations. It is based on a perturbative evaluation to the second order in U/t of the self-energy. The method is very simple and fast. It provided extremely valuable insights on the Mott transition. Its value relies on a fortunate fact, namely, it provides an asymptotically correct solution in the large coupling limit. This is by no means obvious and, apparently, it just works by a lucky stroke. Perturbation theory is by construction good at small coupling, and by no means should be expected to work at large U/t . However, it is not hard to demonstrate it. It is most simply done in the case of a Bethe lattice. We just need to know that, for $0 < \tau < \beta$, $\frac{1}{\beta} \sum_n e^{-i\omega_n \tau} \frac{1}{i\omega_n} = -\frac{1}{2}$, if ω_n is a fermionic Matsubara frequency. From Eq. (18) in the atomic limit ($t = 0$), we see that at half filling $\mathcal{G}_0(i\omega_n) = \frac{1}{i\omega_n}$. Thus, for $0 < \tau < \beta$, we have $\mathcal{G}_0(\tau) = -\frac{1}{2}$ and $\mathcal{G}_0(-\tau) = \frac{1}{2}$. The 2nd order diagram of $\Sigma(i\omega_n)$ therefore takes the value

$$\Sigma^{(2)}(i\omega_n) = -U^2 \int_0^\beta d\tau e^{i\omega_n \tau} [\mathcal{G}_0(\tau)]^2 \mathcal{G}_0(-\tau) = \left(\frac{U}{2}\right)^2 \int_0^\beta d\tau e^{i\omega_n \tau} \mathcal{G}_0(\tau) = \frac{U^2}{4} \frac{1}{i\omega_n}. \quad (20)$$

Using this result and Eq. (14) and (19), we obtain for the G_{loc} in the large- U limit

$$G_{\text{loc}}(i\omega_n) \approx \frac{2}{i\omega_n - (U/2)^2/i\omega_n + \sqrt{(i\omega_n - (U/2)^2/i\omega_n)^2 - D^2}} \quad (21)$$

and from its imaginary part on the real axis we get the DOS

$$\text{DOS}(\omega) \approx \frac{2}{\pi D^2} \sqrt{(\omega - (U/2)^2/\omega)^2 - D^2} \quad \text{with} \quad |\omega - (U/2)^2/\omega| < D \quad (22)$$

which corresponds to two bands of width $\approx D$, split by a gap $\approx U$, where we recall $D = 2t$ is the half-bandwidth of the original non-interacting lattice model, so this solution correctly captures the “atomic” limit of $D \ll U$.

How do we know that it works in between where $U \sim D$? It is because the IPT solution can be benchmarked with the two exact methods that we described before. Comparisons have shown that, quite remarkably, the IPT solution reproduces most of the physical features of the Mott-Hubbard transition at both zero and finite temperature, including the first-order metal-insulator transition that ends in a finite temperature critical point [19, 1]. The surprise gets even bigger: Besides the Hubbard model, the second-order “recipe” for Σ also qualitatively works for the solution of the periodic Anderson model [20] and for the dimer Hubbard model [21]. Unfortunately, the IPT only works at the particle-hole symmetric point. Upon doping the systems, pathological behavior occurs, such as negative compressibility.

In Sec. 4 we provide a link to the IPT source code and propose simple exercises to guide the “hands-on” discovery of the Mott-Hubbard transition.

2.5 Long-range order

DMFT can also be used to explore simple types of magnetic (charge, position, orbital, etc.) long-range ordering, such as ferromagnetic and Néel (i.e. checkerboard) order. In the first case, the equations remain the same, one just needs to consider that the bath may be different for spin up and down. In the case of Néel order, one has to explicitly take into account the two sub-lattices, say A and B. In that case the self-consistency equations read

$$\begin{aligned} \mathcal{G}_{0\sigma}^A(i\omega_n) &= \frac{1}{i\omega_n + \mu - t^2 \mathcal{G}_{\sigma}^B(i\omega_n)} \\ \mathcal{G}_{0\sigma}^B(i\omega_n) &= \frac{1}{i\omega_n + \mu - t^2 \mathcal{G}_{\sigma}^A(i\omega_n)} \end{aligned} \quad (23)$$

and using the symmetry properties of Néel order between sub-lattices, we get just one equation

$$\mathcal{G}_{0\sigma}(i\omega_n) = \frac{1}{i\omega_n + \mu - t^2 \mathcal{G}_{-\sigma}(i\omega_n)}. \quad (24)$$

Note that $G_{\sigma}(i\omega_n)$ is equal to $-G_{-\sigma}^*(i\omega_n)$ in the particle-hole symmetric case, but in general it is not.

3 The Mott-Hubbard transition in DMFT

Before describing the solution of the Hubbard model within DMFT and its relation to the Mott transition, we shall describe some experimental background to motivate this study. The Mott transition is a central problem of strongly correlated systems, and has been occupying a center stage since the discovery of the high T_c cuprate superconductors in the 80's [22], followed by the manganites in the 90's [2], and so on [8]. The interest has been essentially non-stop, with the most recent instance being the fascinating discovery of superconductivity in twisted bi-layer graphene [23]. Once again, this validates the notion that we already discussed in the introduction, namely, that interesting physics always emerges close to a Mott transition. Hence, the relevance of this physical concept.

The Mott transition is a metal-insulator transition (MIT), and the concept goes well before the cuprates, to an argument made by Mott in the 40's [24]. He argued that by considering the dependence of the kinetic and potential energy as a function of the electron density in a solid, one should expect a discontinuous phase transition. In simplest terms the argument is that $E_{\text{kin}} \sim k_F^2/m \sim 1/a^2 \sim n^{2/3}$, while the $E_{\text{pot}} \sim e^2/a \sim n^{1/3}$, so that at low n the E_{pot} dominates, while at large n the kinetic energy does. Hence, as a function of the density, a first order transition should occur between an insulator and a metal. While this argument advances the notion of competition of electronic delocalization versus Coulomb repulsion, which are the ingredients of the Hubbard model, Mott's argument does not immediately apply. What came to be known as the Mott-Hubbard MIT is a phenomenon that occurs at "half-filling," that is, when a band has an occupation of one electron per site (remember that a band has room for two electrons due to the spin), thus, at electronic density $n = 1$. If a band is half-filled, it is partially filled and should have plenty of states just above the Fermi energy. So it should be a metal. Thus, the Mott insulator state is an insulator state that is realized in a half-filled band due to strong Coulomb interactions. Intuitively, if Coulomb repulsion dominates, it will cost a lot of energy to bring two electrons onto the same atomic site. Hence if $n = 1$, the best one can do is to avoid the *double occupation* of any site, which can be achieved by localizing each electron onto a respective atomic site. They would then be locked in their positions, since to move they would have to jump to a neighboring site and that is too costly. Thus, we may think of the Mott state as a global Coulomb blockade.

We should also emphasize that the notion of a Mott transition does not involve a change in the symmetry, such as antiferromagnetic ordering or a lattice dimerization. These two phenomena, associated to the names of Slater and Peierls, can open a gap in the band but do not require strong interactions. Thus, we call them weak-coupling mechanisms, since they emerge from perturbation theory, such as Hartree-Fock. In Fig. 7 we schematically illustrate this point. The ordering effectively doubles the unit cell, thus halving the Brillouin zone (BZ) and doubling the bands that open a gap at the border of the BZ [25].

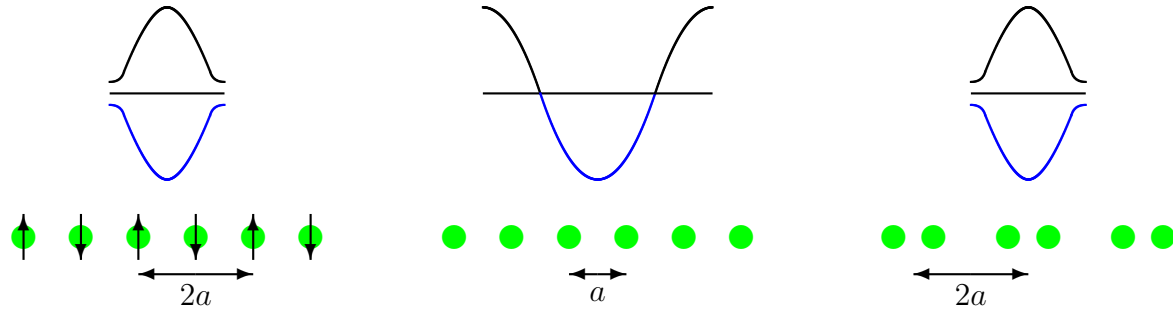


Fig. 7: Schematic representation of the weak-coupling gap opening from the effective lattice parameter doubling $2a$ due to magnetic (left) or lattice (right) symmetry breaking. These mechanisms are associated to Slater and Peierls respectively.

3.1 V_2O_3 a strongly correlated material with a metal-insulator transition

Now we can ask ourselves the question, is the Mott transition realized in nature? Or, is it just an idealized concept? The answer is yes, to both. The Mott transition that we just described is obviously a very idealized and simplified situation: exact half-filling, no change in the symmetry, no disorder, no temperature effect, etc. etc. Yet, and quite remarkably, there seems to be a realization of the Mott-transition concept in the compound V_2O_3 . In Fig. 8 we show the phase diagram of this famous compound. We observe three phases, an antiferromagnetic insulator (AFI) at low temperature, and two paramagnetic phases at intermediate temperatures, one insulator (PMI) and one metallic (PMM) that are separated by a first-order line, which ends in a critical point (CP). The important feature is that the MIT occurs with no change in lattice symmetry. It can be driven by temperature or by pressure (hydrostatic or chemical). The small Cr and Ti substitution is considered to slightly change the lattice spacing, hence the bandwidth, but not the number of carriers. In fact, the no change in the lattice symmetry is easily understood from the fact that if one starts, say in the PMI next to the first-order line, and then heats-up the system just above the CP, then applies pressure just past the CP, and cool down again, one arrives to the qualitatively different PMM phase, all through a continuous and smooth process. This feature shows that the MIT in V_2O_3 is qualitatively similar to the familiar liquid-vapor transition in water.

The study of the MIT in V_2O_3 continues to be a matter of strong scientific interest, attention and debates with many interesting findings that continuously challenge our physical understanding of this compound. Among the most recent and exciting discoveries is that strong electric pulsing may collapse the Mott insulating state and that the phenomenon may be exploited to implement *artificial neurons* [26–28]. From a conceptual point of view understanding this new and unexpected *neuromorphic functionality* [29] poses a significant challenge, namely to describe the Mott transition out-of-equilibrium. This is a topic of great current interest [30], which is also motivated by the fast development of so called “pump-probe” experiments.

After this brief introduction to the Mott phenomenon and its relevance, we now move on to describe the Mott-Hubbard transition within DMFT.

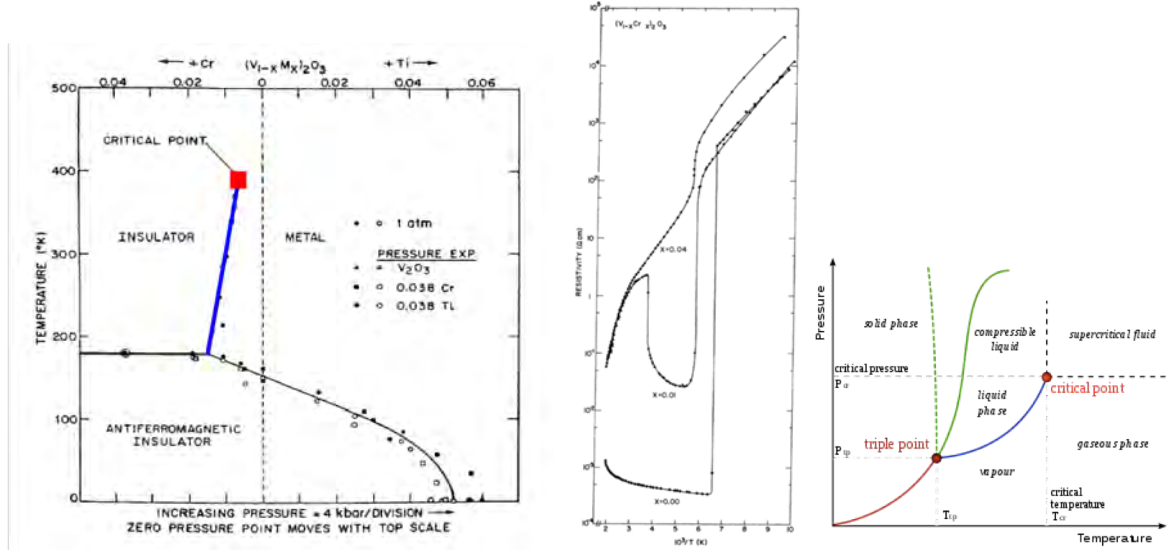


Fig. 8: Left: Phase diagram of V_2O_3 . Negative and positive chemical pressure can be applied with a few % of Cr and Ti substitution. The blue line indicates the first-order line between the metal and the insulator within the paramagnetic phase. The orange square indicates the second-order critical end-point. Center: Resistivity of V_2O_3 across the meta-insulator transitions. Right: A similar first-order line and critical end-point are present in the familiar water-vapor phase transition.

3.2 The Mott-Hubbard transition

As mentioned before, the Mott insulator is realized at half-filling, i.e., one electron per site in a mono-atomic lattice. From experiments, we observe that the transition can be driven by changing the bandwidth, i.e., applying pressure, and by heating. Thus, in the framework of a Hubbard model, which can be considered as a minimal model to capture the physics, one may explore the behavior at half-filling and as a function of the ratio of interaction to bandwidth (U/W), temperature (T), and doping (δ) away from half-filling, i.e., $\delta = n - 1$. For simplicity, we shall consider the case of a Bethe lattice with a semicircular DOS. We shall adopt as unit of energy of the problem the half-bandwidth $D = W/2$ that we set equal to one, unless indicated. As we discussed before, to study the Mott transition as observed in Cr-doped V_2O_3 , we need to restrict ourselves to paramagnetic (PM) states. However, the solution of the Hubbard model on a bipartite lattice, such as Bethe or the (hyper-)cubic, has strong “Néel nesting” that favors an AFI state at low T . Thus, we shall ignore for the moment antiferromagnetic solutions and only be concerned with paramagnetic ones, and we shall explore whether Coulomb repulsion can lead to the break down of a metallic state in the half-filled band of a tight-binding model. In Fig. 9 we show the beginning of the answer to this question within DMFT. We show the evolution of the DOS of a half-filled tight-binding model on a Bethe lattice at $T = 0$ as the interaction U is increased.

The main feature of the solution is the existence of a MIT as a function of increasing interaction strength U . We observe that the $DOS(\omega = 0)$ is finite at low U , but becomes zero when an insulating gap opens at large U . We can look at the nature of this evolution in more detail. We

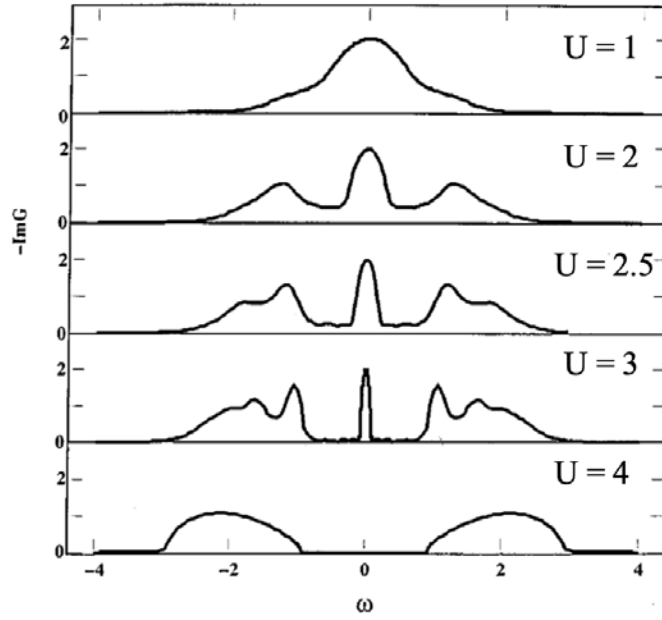


Fig. 9: $DOS(\omega)$ of the HM at $T = 0$ for increasing values of U . The quasiparticle peak narrows as U increases until it collapses at the critical value U_{c2} . From [31].

observe that there is a peak developing at low frequency which becomes increasingly narrow. This is called the quasi-particle peak and its origin can be connected to the Kondo physics we discussed in the introduction. The narrowing of the peak corresponds to an increase of the effective mass as the MIT is approached. A detailed numerical study shows that the effective mass goes as $m^* \sim 1/Z \sim (U_{c2} - U)$, where Z is the quasiparticle residue, i.e., the spectral intensity of the peak, that we defined before, and $U_{c2} \approx 3D$ is the critical value of the interaction where the MIT occurs. The other feature that we observe is the growth of two large peaks with the spectral weight that is lost from the central peak (i.e. $1 - Z$) at frequencies $\pm U/2$. After the transition only these two peaks are left and are separated by a charge gap $\Delta \approx U - 2D$. They are called the Hubbard bands.

3.3 Band-structure evolution across the metal-insulator transition

It is interesting to go back to the lattice to observe what is the nature of the electronic states that conform these peaks. In the case of a Bethe lattice the notion of momentum space is not obvious, so instead of labelling the single-particle states by their momentum quantum number we shall use their single particle energy ε . So the dispersion relation of the interacting energies that is usually denoted by $E = E(k)$ with $k \in BZ$ becomes $E = E(\varepsilon)$ with $\varepsilon \in [-D, D]$. So, Eq. (16) becomes

$$G(\varepsilon, i\omega_n) = \frac{1}{i\omega_n - \varepsilon - \Sigma(i\omega_n)} \quad (25)$$

In the non-interacting case $U = 0$ and $\Sigma = 0$. Then, the GF has poles at ε and the non-interacting dispersion is simply linear with $E(\varepsilon) = \varepsilon$ and $\varepsilon \in [-D, D]$. One intuitive way to

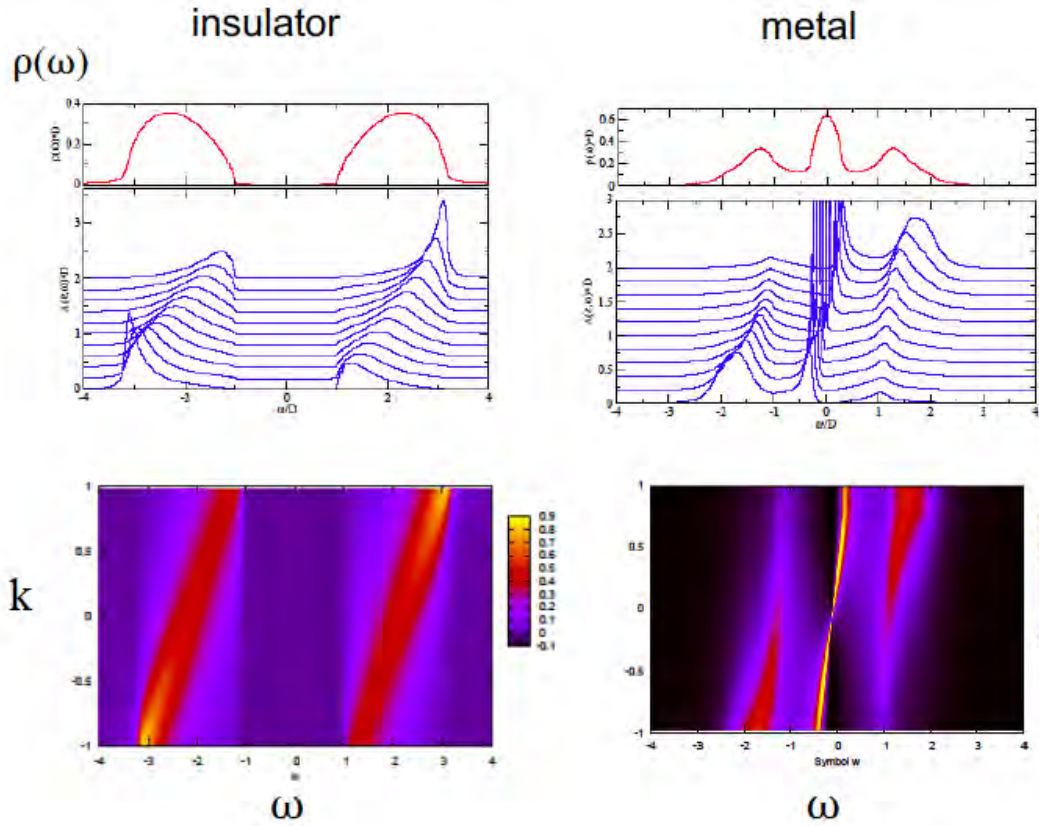


Fig. 10: Top: DOS of the HM at $T = 0$ for the Mott insulator $U > U_{c2}$ (left) and the correlated metal $U < U_{c2}$ (right). Middle: $A(\varepsilon, \omega)$ for increasing values of $\varepsilon \in [-D, D]$. Bottom: Idem in a color intensity plot (the label k plays an analogous role as ε denoting the quantum number of the non-interacting single-particle states (see text)).

think of this band structure is that it is qualitatively similar to that of a 1D tight binding model, i.e., a linearized $-\cos(k)$. Thus, the state $\varepsilon = -D$, with $E(-D) = -D$ is the bottom of the band, i.e., often the Γ -point; the state $\varepsilon = +D$ with $E(+D) = +D$ is the top of the band and the edge of the BZ; and the state $\varepsilon = 0$ with $E(0) = 0$ is the Fermi energy. From Eq. (25) we clearly see that the self-energy function encodes the interaction effects, as it modifies the pole structure of the non-interacting GF and also broadens the poles giving a finite lifetime to the electronic states. In Fig. 10 we show the spectral functions $A(\varepsilon, \omega) = -\text{Im} G(\varepsilon, \omega)/\pi$ of the interacting model for the strongly correlated metal and for the insulator. The $\omega < 0$ spectral functions are measured in ARPES photoemission experiments.

We observe various features which are worth pointing out. In the case of the metal we observe that the pole structure at low energy remains a collection of sharp resonances with a linear dispersion. This indicates that the metallic states are well defined quasi-particle states (i.e., have long life-times) and that their mass is enhanced by the effect of interactions. The effective mass is larger because the band is flatter, as it disperses linearly between $[-ZD, +ZD]$, where we recall that the quasiparticle residue $Z < 1$. Thus, heavy mass aside, the states are qualitatively similar to Bloch waves and we call these states *coherent*. We also observe the Hubbard bands developing at higher frequency. While they show dispersive features, the nature of the propagation that they describe is very different. The electronic states are broad in energy,

which indicates that the excitations are rather short-lived and the propagation of the electrons is *incoherent*. In other words, they describe rather localized particle-like states that are qualitatively different from momentum eigenstates. This correlated metallic state is perhaps the most profound physical insight that emerged from DMFT. We see that the solution of the Hubbard model within DMFT is a concrete realization of a quantum many-body electronic state which simultaneously shares both, wave-like and particle-like features [32].

The Mott-Hubbard insulator has only incoherent Hubbard bands with a dispersion that resembles the non-interacting dispersion, but split by the Coulomb repulsion U , thus approximately follows $E(\varepsilon) = \pm \frac{U}{2}\varepsilon$. We note that the lifetime has a non-trivial variation across the BZ, with a more quasiparticle-like character at the bottom and top of the lower and upper Hubbard bands respectively, and becomes more incoherent close to the gap edges.

3.4 Coexistence of solutions and the first-order transition line

Another remarkable feature of the MIT is the coexistence of solutions. In Fig. 9 we showed the evolution of the DOS(ω) as the interaction U is increased, which displays a MIT at a critical value U_{c2} . However, this is not the only transition. If one starts from the insulator at large U and reduces the interaction one observes that the two Hubbard bands get closer and the gap $\Delta \approx U - 2D$ shrinks. The remarkable feature is that this insulating solution continues to exist for $U < U_{c2}$. The solution eventually breaks down at a value $U_{c1} \approx 2D$, where the gap closes. Thus, for $U \in [U_{c1}, U_{c2}]$ two qualitatively different solutions one metallic the other insulating coexist. This feature can be considered analogous to the coexistence of solutions in the MFT of the ferromagnetic Ising model, with all-up and all-down.

There are many consequences that follow from this feature. If keeping the occupation fixed at $n = 1$, at particle-hole symmetry, and increasing the temperature, the solutions will not disappear. They smoothly evolve, giving rise to a coexistence region in the U - T plane. The evolution of the two solutions with increasing T is qualitatively different.

The metallic one has a low frequency quasiparticle peak. Its width can be related to the Kondo energy scale of the associated impurity problem. This sets a dynamically generated new low-energy scale in the system, much smaller than D (i.e. t) and U . As T is increased, the Kondo resonance can no longer be sustained and the dynamical singlet state that the impurity forms with the bath is broken. The energy scale of the quasiparticle peak is $\sim (U_{c2} - U)$, thus we may expect that the correlated metallic solution will break down along a line $U_{c2}(T_{MIT})$, with T_{MIT} being proportional to $(U_{c2}(0) - U)$. This expectation is indeed realized as shown in Fig. 11. Near U_{c2} there is a significant magnetic moment due to the penalizing effect of U on the probability of double occupation. Moreover, above the line $U_{c2}(T)$, the temperature is too large for the moment to be Kondo screened and we are left with an incoherent collection of disordered magnetic moments at each lattice site.

On the other hand, as T is increased in the insulating solution, the gap may get thermally filled without any significant effect. Thus we may expect that the coexistence region in the U - T plane has a triangular shape, which is actually the case as shown in Fig. 11. The figure also indicates

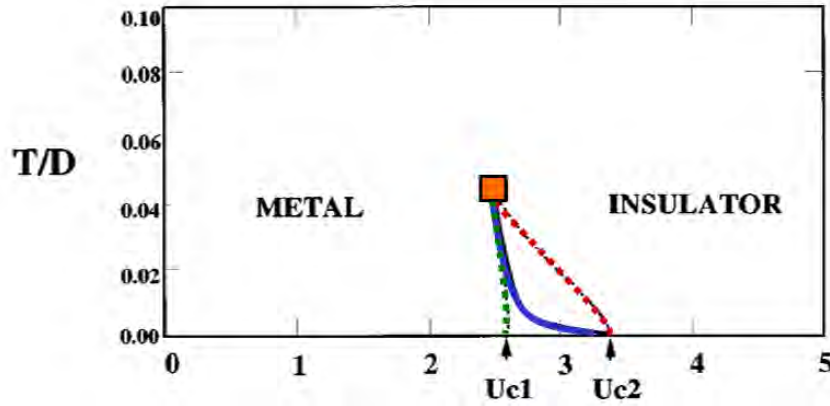


Fig. 11: Phase diagram in the U - T plane. The dotted lines show the region where the paramagnetic metallic and insulating solutions coexist. The red dotted line is $U_{c2}(T)$ and the green dotted line is $U_{c1}(T)$. The blue line denotes the first-order transition where the free energies of the two solutions cross. The orange square denotes the finite- T critical end-point.

the line where the free energies of the two solutions cross, which denotes a first order metal-insulator transition in this model. The fact that the metal is more stable at low T , due to the additional energy gain of the Kondo state, implies also that the physical transition is from a metal to an insulator upon heating. This is qualitatively the case in the MIT within the paramagnetic phase of V_2O_3 that we discussed before. We should perhaps remark here that there is another vanadate, VO_2 , that also displays a transition driven by temperature between two paramagnetic states. However, in that case and contrary to V_2O_3 , the transition is from an insulator to a metal upon heating. Thus the two transitions are qualitatively different. Nevertheless, one may also understand the transition in VO_2 as a Mott transition with a two-site quantum impurity, where the insulating ground-state wins as the two moments screen each other into a local singlet. Such an insulator-metal transition has been discussed recently [33,21,34].

3.5 Endless directions

We have described the core of the Mott transition physics that was unveiled by the introduction of the DMFT approach formulated as a mapping of the (infinite-dimensional) lattice problem onto a self-consistent quantum impurity [35–37,31]. From that starting point an endless number of problems and extensions have been explored and continue to be developed. We shall briefly mention some of them here.

3.5.1 Doping driven Mott transition

In this lecture we have investigated the MIT at half-filling keeping one electron per site. The system is at particle-hole (p-h) symmetry, hence the $DOS(\omega)$ are always even functions. In this situation we have seen that if the interaction U is strong enough and the temperature T low enough the system is in a Mott insulator state. We can destabilize this insulator state by *doping* the system, i.e., by changing the particle occupation by δ . This can be done by changing the chemical potential away from the p-h symmetry at $\mu = U/2$. In the simple single band

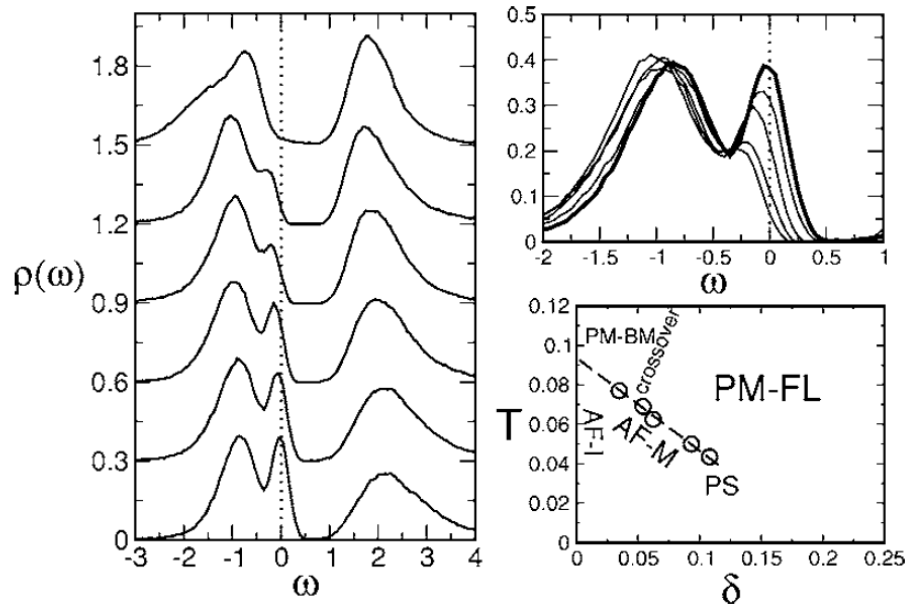


Fig. 12: Left: Paramagnetic $DOS(\omega)$ of a lightly hole-doped Hubbard model at $U = 3.125$, $T = 0.1$, and increasing $\delta = 0.003, 0.0076, 0.0114, 0.022, 0.038$, and 0.055 , from top to bottom. Top right: Detail of the evolution of the quasiparticle peak in the previous results. Bottom right: Phase diagram as a function of δ and T for $U > U_{c2}$. From [38].

Hubbard model that we consider it is equivalent to dope particles or holes, the resulting GFs are related by the change $\omega \rightarrow -\omega$. The effect of doping is to create a correlated metallic state. It shares several features of the $n = 1$ correlated metal that we have already described. It has a narrow quasiparticle peak at $\omega = 0$ that is flanked by the two Hubbard bands. The spectral intensity of the quasiparticle peak is in this case controlled by the doping, with $Z \approx 1/\delta$. Thus the renormalized bandwidth is $\sim \delta D$. This is again a small energy scale and increasing the temperature will destroy the quasiparticle peak. The way this takes place is qualitatively different from the p-h symmetric case. As shown in Fig. 12 one observes that the quasiparticle peak becomes very asymmetric with respect to the origin. This signals a departure from the Fermi liquid state and is associated to the notion of *resilient quasiparticles* and of *bad metal* states [38]. In a bad metal state the system has spectral weight at the Fermi energy but the quasiparticles have very short lifetimes or, equivalently, very large scattering rates that lead to a resistivity in excess of the Ioffe-Regel limit. In other words the mean-free path becomes shorter than the lattice spacing. This feature is often observed in strongly correlated systems including the metallic phase of the vanadates that we mentioned before and the high- T_c cuprate superconductors.

By a continuity argument one should also expect that the coexistence region of solutions must extend into the non p-h symmetric case for $\mu \neq U/2$. This feature has been investigated in [39] where the main consequence was the finding of a divergence in the electronic compressibility. This electronic anomaly can be considered as a precursor for charge density waves, phase separation, and lattice structural changes [40].

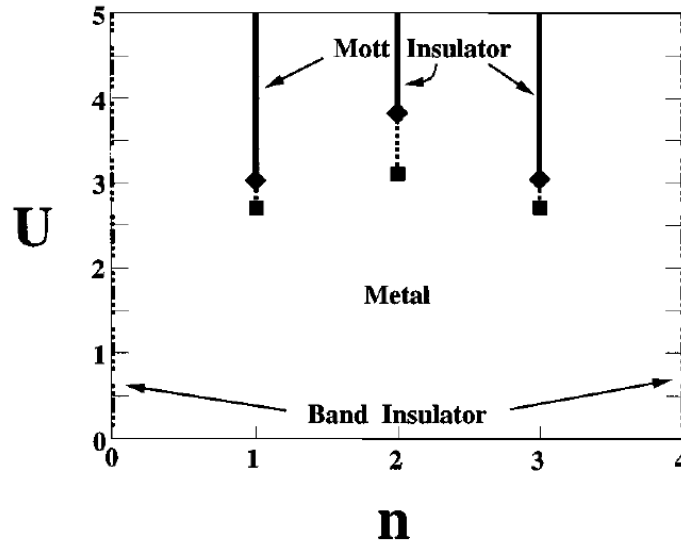


Fig. 13: Phase diagram of the 2-orbital degenerate band Hubbard model. See Fig. 2 for comparison. From [43].

3.5.2 Mott-Hubbard transition in the presence of long-range order

So far we have been mostly concerned with the paramagnetic state. In Sec. 2.5 we described how DMFT can be extended to consider phases with LRO, such as Néel antiferromagnetism. In fact, the lowest energy solution of the Hubbard model at half-filling and $T = 0$ in bipartite lattices, such as Bethe or the hyper-cubic, is an antiferromagnetic insulator (AFI). This state is the most stable below an ordering temperature T_N , which depends on the value of U . At small U the AFI state can be considered as a Slater AF, with a T_N and a gap Δ that are both small, and grow exponentially with U . This state is rather well captured by Hartree-Fock MFT. In contrast, at high values of U , the electrons are Mott localized and the AFI should be considered as a Heisenberg AF where the ordering follows from the super-exchange interaction $J = 4t^2/U$. In this case, the gap is large and $\sim U$, while the $T_N \sim J$, so it decreases with increasing U . This situation is realized in high- T_c superconductors, which have a large gap $\sim$ eV and a T_N one or two orders of magnitude smaller.

An interesting issue is to explore the behavior of the model when one dopes away from the half-filled AF Mott insulator. Despite a significant amount of work done in DMFT, there are few studies that consider this question in the Hubbard model and the detailed evolution remains rather poorly known [38]. From those studies the physical picture that emerges is that of a heavy-mass renormalized quasiparticle band at low frequencies, which is split in two due to the effective doubling of the lattice periodicity. These coherent bands are flanked by Hubbard bands, which are separated by an energy $\sim U$ and have different spectral intensities for the up and down spin projections. Interestingly, these features are qualitatively similar to those obtained in solutions of Cluster DMFT calculations, which unlike “standard” DMFT incorporate spatial spin fluctuations [41, 42].

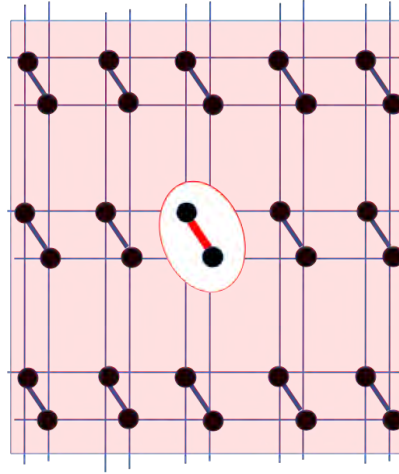


Fig. 14: *Schematic representation of the embedding of a dimer (cluster) in the DHM. This model can be considered the simplest instance of CDMFT.*

3.5.3 Multi-orbital models

DMFT can be naturally extended to consider atomic sites with multiple orbitals that lead to multiple bands. This was initially done for the simplest case of two orbitals with identical hopping amplitudes that leads to two degenerate bands [43]. In Fig. 13 we show the phase diagram in the U - n plane. The interesting feature is that for an N -orbital Hubbard model, Mott transitions are found for fractional dopings $n/2N$ with $n = 1, \dots, 2N-1$, which extends the notion of half-filling to the multi-orbital case. The intuitive way to think about this is that Mott states are found when there is an integer filling of electrons at a lattice site. The energy to add an extra electron to the lattice, would be the Coulomb charging energy $\sim U$. Interestingly, the correlated metal and insulator states are qualitatively similar to those at half-filling, presenting Hubbard bands and heavy quasiparticle bands. Moreover, Mott transitions also have regions of coexistence and therefore the MITs have first order character.

One novelty that the multi-orbital models incorporate is the Hund interaction. This is responsible for the FM alignment of electrons occupying the same atomic site, creating large local magnetic moments as in the colossal magnetoresistive manganites [2]. The large magnetic moment is more difficult to screen, leading to a decrease of the Kondo temperature of the associated impurity model. This has as a consequence the emergence of correlated metallic states with low coherence temperatures and bad metallic features [44, 45]. These systems are known as Hund's metals and their study is relevant for correlated materials like the iron based superconductors [46].

3.5.4 Cluster DMFT

Since DMFT is exact in the limit of ∞ - d , where the lattice problem is mapped to a single impurity by taking one site and embedding it in a self-consistent environment, a natural extension is to consider the embedding of a small portion of the lattice, or a cluster. Such approaches go by the name of cluster-DMFT (CDMFT) [47] and dynamical cluster approximation (DCA) [48].

In the first case one defines a cluster of atoms in real space and embeds it in an effective self-consistent medium, pretty much as we have described before (Sec. 2.2). This is schematically illustrated for the case of a lattice of dimers in Fig. 14. The dimer Hubbard model (DHM) can be exactly solved within CDMFT method [49, 33], and has recently been considered for the interpretation of experiments in VO_2 . However, when applied to general lattices there are certain technical difficulties to restore translational invariance and different approximate schemes have been proposed.

The DCA method is formulated in reciprocal space. It is based on computing a coarse grained self-energy of a finite cluster, i.e., in a space of discrete momentum K , which is then used to obtain estimates of the actual infinite-lattice self-energy (where the momentum k is continuous). Thus the method is fully formulated in momentum space, including the generalization of $\mathcal{G}_0(i\omega_n)$ to a $\mathcal{G}_0(K, i\omega_n)$. So the issue of restoring translational invariance does not emerge. However, there is a price to pay, which is the discontinuity of the self-energy that is defined of coarse-grain “patches” of the BZ.

Both extensions of DMFT enable the exploration of momentum dependence. One of the main results that these approaches provided is the notion of momentum-space differentiation. Namely, the possibility that a Mott gap may open only in certain regions of the Fermi surface. This provides an interpretation to the intriguing observation of “Fermi arcs” in the cuprates [50].

3.5.5 Realistic DMFT or LDA+DMFT

DMFT has also been extended to incorporate real material-specific information. This methodology goes by the names of DFT+DMFT, LDA+DMFT, or Realistic-DMFT [47]. Schematically, the approach retains the same mapping onto an QIP and its self-consistent solution, but the material-specific electronic structure dispersion replaces the ε_k in the k -summations, Eq. (13). Since LDA solution to the DFT equations is a self-consistent method itself, this opens the door to a variety of possible schemes. We shall not dwell further on this topic since there is the dedicated lecture by E. Pavarini.

3.5.6 Out-of-equilibrium Mott transition

An exciting new frontier is the investigation of the correlated systems driven away from equilibrium conditions. This is relevant for recently developed experimental techniques such as “pump-probe” spectroscopies that give access to the time-resolved evolution of a strongly interacting quantum material. Particularly interesting to us is the possibility of driving a Mott insulator out of equilibrium, which has been considered in a variety of experimental studies. Here we cite some particularly interesting ones related to vanadates [51, 52] that indicate that a sharp insulator-metal transition (IMT) can be induced by strong enough light irradiation. Particularly interesting is that these experiments seem to have provided strong evidence of the meta-stable states (i.e. the state coexistence) that has been predicted by DMFT studies of the Hubbard and dimer Hubbard models. The latter case is particularly interesting is where a photoemission study of the Mott insulator VO_2 was conducted [53]. The main observations where

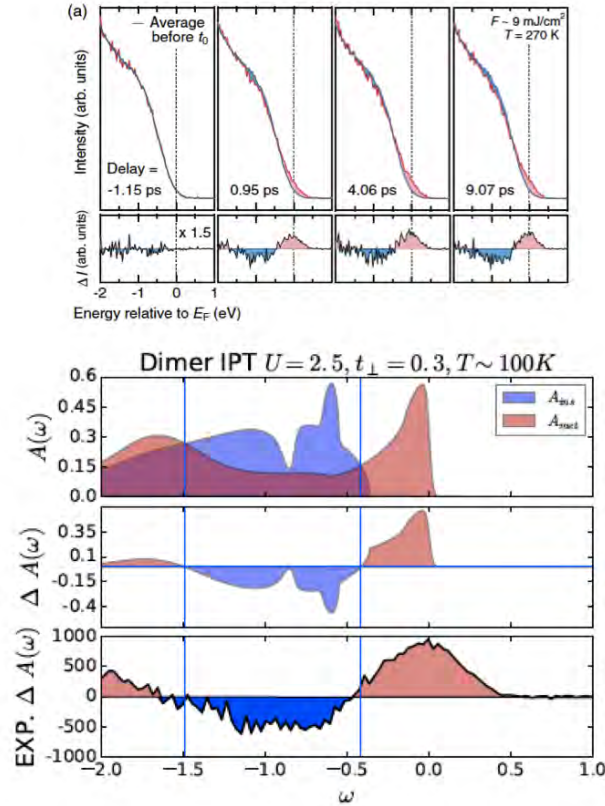


Fig. 15: Top: Photoemission spectra of the VO_2 insulator “before and after” the pump pulse. Bottom: Difference between after- and before-pump spectra showing the transfer of spectral intensity across the pump-driven Mott transition. Mid-top: Occupied part ($\omega < 0$) of the $\text{DOS}(\omega)$ of the insulator (blue) and metal (red) coexistent solutions of the DHM. Mid-bottom: Difference between the model $\text{DOS}(\omega)$. Bottom: Detail of experimental photoemission difference data.

the existence of a sharp light-fluence threshold for the IMT, that the resulting metallic state was very long lived (> 10 ps) and that the photoemission spectrum was different from the high- T metallic state. We found that this intriguing metallic state could be the realization of a *monoclinic* metal in VO_2 , which emerges from a DMFT study of the Mott insulating state of a dimer Hubbard model [33, 21, 34]. In Fig. 15 we show the experimental variation of the spectrum across the IMT along with the results from the theoretical study [54]

The time-resolved behavior across the Mott-Hubbard transition can be studied by extending the DMFT approach to the out-of-equilibrium situation by adopting the Kadanoff-Baym and Keldysh GFs formalism. These extension of DMFT received a great deal of attention in recent years [55, 30]. Despite significant progress and new insights the detailed solution of the problem in the coupling regime where there are coexistent solutions still remains a challenge. We shall not devote much more here and point to the lecture on this topic by J. Freericks.

3.5.7 Mottronics for Artificial Intelligence

One of the latest and perhaps most original and exciting developments regarding the Mott transition is the possibility of using the Mott insulators to fabricate artificial electronic neurons for spiking neural networks [28]. This is particularly timely given the current explosion and inter-

est in artificial intelligence (AI). The algorithms of AI are often based on neural networks (NN) such as the classic Hopfield model, popular in the 70's, to the more modern convolutional NN. In general, NN have two types of units, a non-linear input-output device, the neuron, and devices that interconnects the neurons and modulate the intensity of their coupling, the synapses. Models of NN can be implemented in software or in hardware. In the first case, a notable example is the AlphaGo code that has defeated the world champion of Go [56]. However, running this algorithm requires a powerful supercomputer that consumes several kW. In the second case, powerful chips are built using state-of-the-art electronics that can implement a million neurons, such as TrueNorth [57]. They are energetically efficient but their main limitation is that they require almost 10^{10} transistors to implement a million neurons (synapses require relatively fewer transistors than neurons). While this accomplishment is remarkable, these chips are still several orders of magnitude below the 10^9 neurons in a cat's brain. This situation opens the way for a disruptive technology, which may implement artificial neurons using far fewer components.

Recently, we have shown that a Mott insulator may accomplish this task [58, 29]. The key finding was that Mott insulators under electric pulsing realize a *neuromorphic functionality*, which consists in behaving analogously to the leaky-integrate-and-fire (LIF) model of spiking neurons [59]. The LIF model is a classic and basic model of biological neurons. It describes the integration of electric input that arrives at a neuron through its dendrites, the leakage during the time in-between arriving input spikes, and the fire of an action potential when the integrated input reaches a threshold. A Mott insulator under electric pulsing may behave similarly. The key feature is that in narrow gap Mott insulators, such as GaTa_4Se_8 [60] or V_2O_3 , when a strong voltage is applied, creating a field of the order of kV/cm, a collapse of the resistance is observed after a certain delay time $\tau_d \sim \text{tens of } \mu\text{s}$ [61]. Let us now consider applying instead of a constant voltage a train of pulses, where the duration of each pulse τ_p is smaller than τ_d . It is easy to understand that if the time between the pulses τ_w is very long, then each pulse is an independent perturbation that will not produce the resistive collapse of the Mott state. On the other extreme, if τ_w goes to zero, then the pulses will simply accumulate and produce the collapse after n_{sw} pulses, where $n_{sw}^{min} = \tau_d/\tau_p$. Thus, n_{sw} is a function of τ_w that increases from n_{sw}^{min} to ∞ . This behavior was originally predicted by a phenomenological model of resistive breakdown in Mott insulators and experimentally observed [61], as illustrated in Fig. 16.

The phenomenological model consisted of a resistor network, where the key assumption for the resistive units was the existence of two resistive states. One more stable with high resistance and a metastable one with low resistance. This assumption was motivated by the coexistence of solutions of the DMFT studies of Hubbard models that we described in this lecture. Interestingly, the equations that describe the resistor network model can be shown to be analogous to that of the LIF model of neurons [29], where the role of spikes is played by the applied pulses. The “firing” of an action potential corresponds to the current spike through the Mott insulator as its resistance collapses.

Interestingly, following a different line of work a group at Hewlett-Packard has proposed an implementation of another classic biological neuron model, the Hodgkin-Huxley model [59], using NbO_2 , which is also a Mott insulator material [62].

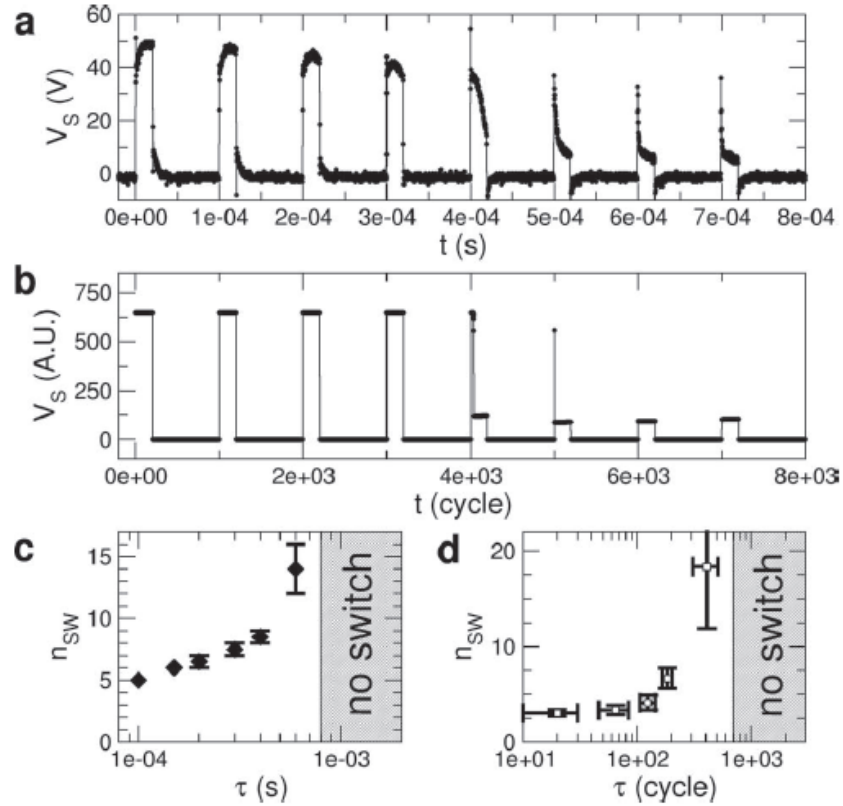


Fig. 16: *Top: Collapse of the Mott insulator state in GaTa_4Se_8 observed after 5 pulses by the collapse of the resistance (i.e., the voltage drop on the sample). Middle: Resistor network model simulation that qualitatively captures the behavior of the Mott insulator under strong electric pulsing. Bottom: Systematic behavior of n_{sw} as a function of the time between pulses τ_w (left is experimental data and right model simulations). From [61].*

Thus the neuromorphic functionalities of Mott insulators, owing to the unique non-linear behavior of their I - V characteristics, are emerging as a new and exciting road towards bringing Mott materials to the realm of future electronics — or rather *Mottronics*.

4 Hands-on exercise (with IPT code): The Mott-Hubbard transition

Many of the plots in this lecture illustrating the Mott-Hubbard transition were obtained by solving the DMFT equations using an impurity solver based on iterative perturbation theory (Sect. 2.4) [35, 31]. This approximate method has the advantage of being simple and providing qualitatively good solutions across the transition. The interested reader is invited to download the IPT codes and go through the proposed exercises that serve as a guide for a hands-on exploration of the Mott-Hubbard metal-insulator transition in DMFT. Codes are available for free at <http://mycore.core-cloud.net/index.php/s/oAz0lIWuBM90Gqt>.

References

- [1] A. Georges, G. Kotliar, W. Krauth, and M. Rozenberg, *Rev. Mod. Phys.* **68**, 13 (1996)
- [2] E. Dagotto, *Science* **309**, 257 (2005)
- [3] B.J. Kim *et al.*, *Science* **323**, 1329 (2009)
- [4] V. Brouet *et al.*, *Phys. Rev. B* **92**, 081117(R) (2015)
- [5] H. Alloul, *EPJ Web of Conferences* **23**, 00015 (2012)
- [6] P. Wzietek *et al.*, *Phys. Rev. Lett.* **112**, 066401 (2014)
- [7] A. Fujimori, *J. Phys. Chem. Solids* **53**, 1595 (1992)
- [8] M. Imada, A. Fujimori, and Y. Tokura, *Rev. Mod. Phys.* **70**, 1039 (1998)
- [9] G. Mahan: *Many-Particle Physics*, (Springer, New York, 2000)
- [10] J.W. Negele, and H. Orland: *Quantum Many-particle Systems* (Perseus, 1998)
- [11] W. Metzner and D. Vollhardt, *Phys. Rev. Lett.* **62**, 324 (1989)
- [12] E. Gull *et al.*, *Rev. Mod. Phys.* **83**, 349 (2011)
- [13] O. Parcollet, M. Ferrero, T. Ayrar, H. Hafermann, I. Krivenko, L. Messio, and P. Seth, *Comp. Phys. Comm.* **196**, 398 (2015)
- [14] J.E. Hirsch and R.M. Fye, *Phys. Rev. Lett.* **56**, 2521 (1986)
- [15] E. Gull, A.J. Millis, A.I. Lichtenstein, A.N. Rubtsov, M. Troyer, and P. Werner, *Rev. Mod. Phys.* **83**, 349 (2011)
- [16] D.J. García, K. Hallberg, and M.J. Rozenberg, *Phys. Rev. Lett.* **93**, 246403 (2004)
- [17] D.J. García, E. Miranda, K. Hallberg, and M.J. Rozenberg, *Phys. Rev. B* **75**, 121102(R) (2007)
- [18] R. Bulla, T.A. Costi, and T. Pruschke, *Rev. Mod. Phys.* **80**, 395 (2008)
- [19] G. Kotliar, E. Lange, and M.J. Rozenberg, *Phys. Rev. Lett.* **84**, 5180 (2000)
- [20] M.J. Rozenberg, G. Kotliar, and H. Kajueter, *Phys. Rev. B* **54**, 8452 (1996)
- [21] O. Nájera, M. Civelli, V. Dobrosavljević, and M.J. Rozenberg, *Phys. Rev. B* **97**, 045108 (2018)
- [22] B. Keimer, S.A. Kivelson, M.R. Norman, S. Uchida, and J. Zaanen, *Nature* **518**, 179 (2015)

- [23] Y. Cao *et al.*, Nature **556**, 80 (2018)
- [24] N. Mott: *Metal-Insulator Transitions* (Taylor & Francis, 1997)
- [25] N.W. Ashcroft and N. Mermin: *Solid State Physics* (Brooks, 1976)
- [26] I.H. Inoue and M.J. Rozenberg, Adv. Funct. Mat. **18**, 2289 (2008)
- [27] E. Janod *et al.*, Adv. Func. Mat. **25**, 6287 (2016)
- [28] J. del Valle, G. Ramírez, M.J. Rozenberg, and I.K. Schuller, Jour. Appl. Phys. **124**, 211101 (2018)
- [29] P. Stoliar *et al.*, Adv. Funct. Mat. **27**, 1604740 (2017)
- [30] H. Aoki, N. Tsuji, M. Eckstein, M. Kollar, T. Oka, and P. Werner, Rev. Mod. Phys. **86**, 779 (2014)
- [31] X.Y. Zhang, M.J. Rozenberg, and G. Kotliar, Phys. Rev. Lett. **70**, 1666 (1993)
- [32] M. Takizawa *et al.*, Phys. Rev. B. **80**, 235104 (2009)
- [33] O. Nájera, M. Civelli, V. Dobrosavljević, and M.J. Rozenberg, Phys. Rev. B **95**, 035113 (2017)
- [34] H.T. Stinson *et al.*, Nat. Commun. **9**, 3604 (2018)
- [35] A. Georges and G. Kotliar, Phys. Rev. B **45**, 6479 (1992)
- [36] M. Jarrell, Phys. Rev. Lett. **69**, 168 (1992)
- [37] M.J. Rozenberg, X.Y. Zhang, and G. Kotliar, Phys. Rev. Lett. **69**, 1236 (1992)
- [38] A. Camjayi, R. Chitra, and M.J. Rozenberg, Phys. Rev. B Rapid Comm. **73**, 041103 (2006)
- [39] G. Kotliar, S. Muthry, and M.J. Rozenberg, Phys. Rev. Lett. **89**, 046401 (2002)
- [40] S.R. Hassan, A. Georges, and H.R. Krishnamurthy, Phys. Rev. Lett. **94**, 036402 (2005)
- [41] L. Fratino *et al.*, Phys. Rev. B **96**, 241109(R) (2017)
- [42] S. Sakai *et al.*, Phys. Rev. Lett. **111**, 107001 (2013)
- [43] M. Rozenberg, Phys. Rev. B **55**, 4855(R) (1997)
- [44] A. Georges, L. de' Medici, and J. Mrarlje, Annu. Rev. Condens. Matter Phys. **4**, 137 (2013)

- [45] L. de' Medici: *Hund's Metals Explained*, in E. Pavarini, E. Koch, R. Scalettar, and R. Martin (eds.): *The Physics of Correlated Insulators, Metals, and Superconductors Modeling and Simulation Vol. 7* (Forschungszentrum Jülich, 2017) <http://www.cond-mat.de/events/correl17>
- [46] Z.P. Yin, K. Haule, and G. Kotliar, *Nat. Mater.* **10**, 932 (2011)
- [47] G. Kotliar *et al.*, *Rev. Mod. Phys.* **78**, 865 (2006)
- [48] T. Maier, M. Jarrell, T. Pruschke, and M.H. Hettler, *Rev. Mod. Phys.* **77**, 1027 (2005)
- [49] G. Moeller, V. Dobrosavljević, and A.E. Ruckenstein, *Phys. Rev. B* **59**, 6846 (1999)
- [50] M. Civelli *et al.*, *Phys. Rev. Lett.* **95**, 106402 (2005)
- [51] G. Lantz *et al.*, *Nat. Commun.* **8**, 13917 (2017)
- [52] V.R. Morrison *et al.*, *Science* **346**, 445 (2014)
- [53] R. Yoshida *et al.*, *Phys. Rev. B* **89**, 205114 (2014)
- [54] O. Nájera: *Study of the dimer Hubbard Model within Dynamical Mean Field Theory and its application to VO₂* (Ph.D. Thesis Université Paris-Saclay 2017) <https://tel.archives-ouvertes.fr/tel-01690699>
- [55] J.K. Freericks, V.M. Turkowski, and V. Zlatić, *Phys. Rev. Lett.* **97**, 266408 (2006)
- [56] D. Silver *et al.*, *Nature* **529**, 484 (2016)
- [57] P.A. Merolla *et al.*, *Science* **345**, 668 (2014)
- [58] US Patent 15307269: *Single-component artificial neuron based on Mott insulators, network of artificial neurons and corresponding manufacturing method*, Inventors: L. Cario, B. Corraze, P. Stoliar, V. Tranchant, E. Janod, M.-P. Belsand, and M.J. Rozenberg
- [59] D.W. Gerstner, W.M. Kistler, R. Naud, and L. Paninski: *Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition* (Cambridge University Press, 2014)
- [60] A. Camjayi *et al.*, *Phys. Rev. Lett.* **113**, 086404 (2014)
- [61] P. Stoliar *et al.*, *Adv. Mater.* **25**, 3222 (2013)
- [62] M.D. Pickett *et al.*, *Nat. Mater.* **12**, 114 (2013)

11 Dynamical Mean-Field Theory for Materials

Eva Pavarini

Institute for Advanced Simulation

Forschungszentrum Jülich

Contents

1	Introduction	2
2	From DMFT to LDA+DMFT	5
2.1	DMFT for a toy model: The Hubbard dimer	5
2.2	Non-local Coulomb interaction	12
2.3	Quantum-impurity solvers: Continuous-time quantum Monte Carlo	13
2.4	Hartree-Fock versus DMFT approximation	17
2.5	DMFT for the one-band Hubbard model	22
2.6	DMFT for multi-orbital models	25
3	Building materials-specific many-body models	27
3.1	Model construction	27
3.2	Localization of the basis and range of the Coulomb interaction	30
3.3	Hubbard Hamiltonians for t_{2g} and e_g systems	31
3.4	Spin-orbit interaction and effects of the basis choice	32
3.5	Non-spherical Coulomb terms and double-counting correction	35
4	Conclusion	37
A	Eigenstates of two-site models	38
A.1	Hubbard dimer	38
A.2	Anderson molecule	40
B	Lehmann representation of the local Green function	42

1 Introduction

In 1965, when the two founding papers [1] of density-functional theory (DFT) had just been published in the Physical Review, very few could recognize the revolution in the making. Efficient techniques for materials science applications were missing and computers were not that powerful. And yet in 50 years, thanks to remarkable ideas, novel algorithms and steady advance in computational power, brilliant minds brought DFT to its present splendor, making it the *standard model* of condensed matter physics. In 1998, in his Nobel lecture [2], Walter Kohn, described the most important contributions of DFT to science with these words

The first is in the area of fundamental understanding. Theoretical chemists and physicists, following the path of the Schrödinger equation, have become accustomed to think in a truncated Hilbert space of single particle orbitals. The spectacular advances achieved in this way attest to the fruitfulness of this perspective. However, when very high accuracy is required, so many Slater determinants are required (in some calculations up to $\sim 10^9$!) that comprehension becomes difficult.

DFT changed the focus from the N -electron ground-state wavefunction $\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N)$ to a three-dimensional variable, the electronic ground-state density $n(\mathbf{r})$, or other directly measurable quantities, such as response functions. These, within a given approximation of the DFT exchange-correlation functional, could be calculated *from first principles*, i.e., using as input the type of atoms involved and, at most, their positions. In practice, $n(\mathbf{r})$ is obtained by mapping the actual many-body problem onto an auxiliary single-electron Hamiltonian with the same ground-state electron density, the Kohn-Sham Hamiltonian. The associated Kohn-Sham eigenvalues are thus in principle merely Lagrange multipliers. Remarkably, however, the big success of DFT came, in part, from bold applications of the theory beyond its actual realm of validity. To general surprise, the Kohn-Sham eigenvalues turned out to be in many cases excellent approximations to the actual elementary excitations of a given material. Early on it became clear, however, that this Ansatz fails qualitatively for a whole category of systems, those in which local Coulomb repulsion effects are large, also known as *strongly-correlated materials*. Paradoxically, simple models describing generic features of the microscopic mechanism are much more effective than DFT-based material-specific calculations in describing strong-correlation phenomena. This happens, e.g., for the Kondo effect, heavy-fermion behavior, or the metal-insulator transition. Thus criticisms arose. Particularly outspoken in this contest was another Nobel laureate, P.W. Anderson, who emphasized the emergent nature of a true many-body phenomenon [3]. Emergent states typically elude a simple mapping to an effective non-interacting system. For a while, two apparently not compatible philosophies thus coexisted. The first-principles school identified in the material dependence the essential ingredient for understanding the real world, and tried to correct the failures of the practical implementations of DFT by corrections, often ad hoc, of the exchange-correlation potential. Instead, the many-body models school identified canonical models which explain specific emergent phenomena, dismissing the materials dependence as non-relevant, non-generic detail. Only in recent years these two world-views started to merge, and it became apparent that both sides were right and wrong at

the same time. While the local Coulomb repulsion is indeed key, materials aspects turn out to be essential for understanding real correlated materials. In hindsight, we can now put this debate in a different perspective. The electronic many-body problem, in the non-relativistic limit and in the Born-Oppenheimer approximation, is described by the Hamiltonian

$$\hat{H}_e = -\frac{1}{2} \sum_i \nabla_i^2 - \sum_i \sum_\alpha \frac{Z_\alpha}{|\mathbf{r}_i - \mathbf{R}_\alpha|} + \sum_{i>j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{\alpha>\alpha'} \frac{Z_\alpha Z_{\alpha'}}{|\mathbf{R}_\alpha - \mathbf{R}_{\alpha'}|}, \quad (1)$$

where $\{\mathbf{r}_i\}$ are electron coordinates, $\{\mathbf{R}_\alpha\}$ nuclear coordinates and Z_α the nuclear charges. Using a complete one-electron basis, for example the basis $\{\phi_a(\mathbf{r})\}$, where $\{a\}$ are the quantum numbers, we can write this Hamiltonian in second quantization as

$$\hat{H}_e = - \underbrace{\sum_{ab} t_{ab} c_a^\dagger c_b}_{\hat{H}_0} + \frac{1}{2} \underbrace{\sum_{aa'bb'} U_{aa'bb'} c_a^\dagger c_{a'}^\dagger c_{b'} c_b}_{\hat{H}_U}.$$

Here the hopping integrals are given by

$$t_{ab} = - \int d\mathbf{r} \bar{\phi}_a(\mathbf{r}) \left(-\frac{1}{2} \nabla^2 - \underbrace{\sum_\alpha \frac{Z_\alpha}{|\mathbf{r} - \mathbf{R}_\alpha|}}_{v_{\text{en}}(\mathbf{r})} \right) \phi_b(\mathbf{r}),$$

while the elements of the Coulomb tensor are

$$U_{aa'bb'} = \int d\mathbf{r}_2 \int d\mathbf{r}_1 \bar{\phi}_a(\mathbf{r}_1) \bar{\phi}_{a'}(\mathbf{r}_2) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \phi_{b'}(\mathbf{r}_2) \phi_b(\mathbf{r}_1).$$

In principle, all complete one-electron bases are equivalent. In practice, since, in the general case, we cannot solve the N -electron problem exactly, some bases are better than others. One possible choice for the basis are the Kohn-Sham orbitals, $\{\phi_a^{\text{KS}}(\mathbf{r})\}$, obtained, e.g., in the local density approximation (LDA) or its simple extensions.¹ In this case, it is useful to replace the electron-nuclei interaction $v_{\text{en}}(\mathbf{r})$ with the DFT potential $v_{\text{R}}(\mathbf{r})$, which includes in addition the Hartree term $v_{\text{H}}(\mathbf{r})$ and the (approximate) exchange-correlation potential $v_{\text{xc}}(\mathbf{r})$

$$v_{\text{R}}(\mathbf{r}) = v_{\text{en}}(\mathbf{r}) + \underbrace{\int d\mathbf{r}' \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}}_{v_{\text{H}}(\mathbf{r})} + v_{\text{xc}}(\mathbf{r}),$$

so that

$$\tilde{t}_{ab} = - \int d\mathbf{r} \bar{\phi}_a^{\text{KS}}(\mathbf{r}) \left(-\frac{1}{2} \nabla^2 + v_{\text{R}}(\mathbf{r}) \right) \phi_b^{\text{KS}}(\mathbf{r}). \quad (2)$$

To avoid double counting (DC), we have however to subtract from $\hat{H}_U$ the term $\hat{H}_{\text{DC}}$, which describes the Coulomb terms already included in the hopping integrals

$$\hat{H}_e = - \underbrace{\sum_{ab} \tilde{t}_{ab} c_a^\dagger c_b}_{\hat{H}_0 = \hat{H}_e^{\text{LDA}}} + \frac{1}{2} \underbrace{\sum_{aa'bb'} \tilde{U}_{aa'bb'} c_a^\dagger c_{a'}^\dagger c_{b'} c_b}_{\Delta \hat{H}_U} - \hat{H}_{\text{DC}}.$$

¹For the purpose of many-body calculations the differences between LDA, GGA or their simple extensions are in practice negligible; for simplicity, in the rest of the lecture, we thus adopt LDA as representative functional.

For weakly-correlated systems, in the Kohn-Sham basis, the effects included in $\Delta\hat{H}_U$ can, in first approximation, either be neglected or treated as a perturbation. This implies that $\hat{H}_e^{\text{LDA}} \sim \hat{H}_{\text{eff}}$, where $\hat{H}_{\text{eff}}$ is the effective model which provides a good description of the system (at least) at low energy, and which describes emergent effective “elementary particles” and their interactions. Hypothetically, one could imagine that $\hat{H}_{\text{eff}}$ is obtained via a canonical transformation, so that $\hat{H}_{\text{eff}} \sim \hat{S}^{-1} \hat{H}_e \hat{S}$, although the exact form of the operator $\hat{S}$ is unknown.

A defining feature of strong-correlation effects is that they cannot be described via a single-electron Hamiltonian, however. A model of form $\hat{H}_e^{\text{LDA}}$ does not describe correctly the Mott metal-insulator transition, no matter what the specific values of the parameters $\tilde{t}_{ab}$ are.² Thus for strongly-correlated systems the low-energy effective model must have a different form. For Mott systems a canonical Hamiltonian is the Hubbard model

$$\hat{H} = - \sum_{\sigma} \sum_{ii'} t^{i,i'} c_{i\sigma}^{\dagger} c_{i'\sigma} + U \sum_i \hat{n}_{i\uparrow} \hat{n}_{i\downarrow}, \quad (3)$$

which includes, in addition to a single-electron term, the on-site Coulomb repulsion. This Hamiltonian captures the essence of the Mott transition. At half filling, for $U = 0$ it describes a paramagnetic metal, and for $t^{i,i'}(1-\delta_{i,i'})=0$ an insulating set of paramagnetic atoms. Unfortunately, differently from Hamiltonians of type $\hat{H}_e^{\text{LDA}}$, Hubbard-like models cannot be solved exactly in the general case. Remarkably, till 30 years ago, no method for describing the complete phase diagram of (3) in one coherent framework, including the paramagnetic insulating phase, was actually known. This changed between 1989 and 1992, when the dynamical mean-field theory (DMFT) was developed [4–7]. The key idea of DMFT consists in mapping the Hubbard model onto a self-consistent *auxiliary quantum-impurity problem*, which can be solved exactly. The mapping is based on the *local dynamical self-energy approximation*, very good for realistic three-dimensional lattices—and becoming exact in the infinite coordination limit [4, 5].

DMFT was initially applied to simple cases, due to limitations in model building, computational power and numerical methods for solving the auxiliary impurity problem (the quantum impurity solvers). In the last twenty years remarkable progress lifted many of these limitations. First, reliable schemes to build realistic low-energy materials-specific Hubbard-like models have been devised, in particular using Kohn-Sham localized Wannier functions. This is remarkable, given that we do not know the exact operator $\hat{S}$ which gives the effective low-energy Hamiltonian, and thus a truly systematic derivation is not possible. Second, key advances in quantum impurity solvers and increasingly more powerful supercomputers made it possible to study always more complex many-body Hamiltonians. The approach which emerged, consisting in solving within DMFT materials-specific many-body Hamiltonian constructed via LDA, is known as the LDA+DMFT method [8–10]. This technique (and its extensions) is now the state-of-the-art for describing strongly-correlated materials. In this lecture I will outline the basic ideas on which the method is based, its successes and its limitations.

²One can obtain an insulator by reducing the symmetry, e.g, by increasing the size of the primitive cell. This Slater-type insulator has however different properties than a Mott-type insulator.

2 From DMFT to LDA+DMFT

In this section we introduce the basics of DMFT. We start from a case for which we can perform analytic calculations, the two-site Hubbard Hamiltonian. This is a toy model, useful to illustrate how the method works, but for which, as we will see, DMFT is not a good approximation. Indeed, the Hubbard dimer is the worst case for DMFT, since the coordination number is the lowest possible. Next we extend the formalism to the one-band and then to the multi-orbital Hubbard Hamiltonian. For three-dimensional lattices the coordination number is typically large and thus DMFT is an excellent approximation. In Sec. 3 we describe modern schemes to construct materials-specific many-body models. They are based on Kohn-Sham Wannier orbitals, calculated, e.g, using the LDA functional. The solution of such models via DMFT defines the LDA+DMFT method.

2.1 DMFT for a toy model: The Hubbard dimer

The two-site Hubbard model is given by

$$\hat{H} = \varepsilon_d \sum_{i\sigma} \hat{n}_{i\sigma} - t \sum_{\sigma} \left(c_{1\sigma}^\dagger c_{2\sigma} + c_{2\sigma}^\dagger c_{1\sigma} \right) + U \sum_i \hat{n}_{i\uparrow} \hat{n}_{i\downarrow},$$

with $i = 1, 2$. The ground state for $N = 2$ electrons (half filling) is the singlet³

$$|G\rangle_H = \frac{a_2(t, U)}{\sqrt{2}} \left(c_{1\uparrow}^\dagger c_{2\downarrow}^\dagger - c_{1\downarrow}^\dagger c_{2\uparrow}^\dagger \right) |0\rangle + \frac{a_1(t, U)}{\sqrt{2}} \left(c_{1\uparrow}^\dagger c_{1\downarrow}^\dagger + c_{2\uparrow}^\dagger c_{2\downarrow}^\dagger \right) |0\rangle \quad (4)$$

with

$$a_1^2(t, U) = \frac{1}{\Delta(t, U)} \frac{\Delta(t, U) - U}{2}, \quad a_2^2(t, U) = \frac{4t^2}{\Delta(t, U)} \frac{2}{\Delta(t, U) - U},$$

and

$$\Delta(t, U) = \sqrt{U^2 + 16t^2}.$$

The energy of this state is

$$E_0(2) = 2\varepsilon_d + \frac{1}{2}(U - \Delta(t, U)).$$

In the $T \rightarrow 0$ limit, using the Lehmann representation (see Appendix B), one can show that the local Matsubara Green function for spin σ takes then the form

$$G_{i,i}^\sigma(i\nu_n) = \frac{1}{4} \left(\frac{1 + w(t, U)}{i\nu_n - (E_0(2) - \varepsilon_d + t - \mu)} + \frac{1 - w(t, U)}{i\nu_n - (E_0(2) - \varepsilon_d - t - \mu)} \right. \\ \left. + \frac{1 - w(t, U)}{i\nu_n - (-E_0(2) + U + 3\varepsilon_d + t - \mu)} + \frac{1 + w(t, U)}{i\nu_n - (-E_0(2) + U + 3\varepsilon_d - t - \mu)} \right),$$

³Eigenstates and eigenvalues of the Hubbard dimer for arbitrary filling can be found in Appendix A.1.

where $\nu_n = \pi(2n+1)/\beta$ are fermionic Matsubara frequencies, $\mu = \varepsilon_d + U/2$ is the chemical potential, and the weight is $w(t, U) = 2a_1(t, U)a_2(t, U)$. The local Green function can be rewritten as the average of the Green function for the bonding ($k = 0$) and the anti-bonding state ($k = \pi$), i.e.,

$$G_{i,i}^\sigma(i\nu_n) = \frac{1}{2} \left(\underbrace{\frac{1}{i\nu_n + \mu - \varepsilon_d + t - \Sigma^\sigma(0, i\nu_n)}}_{G^\sigma(0, i\nu_n)} + \underbrace{\frac{1}{i\nu_n + \mu - \varepsilon_d - t - \Sigma^\sigma(\pi, i\nu_n)}}_{G^\sigma(\pi, i\nu_n)} \right).$$

The self-energy is given by

$$\Sigma^\sigma(k, i\nu_n) = \frac{U}{2} + \frac{U^2}{4} \frac{1}{i\nu_n + \mu - \varepsilon_d - \frac{U}{2} - e^{ik} 3t}.$$

The self-energies $\Sigma^\sigma(0, i\nu_n)$ and $\Sigma^\sigma(\pi, i\nu_n)$ differ due to the phase $e^{ik} = \pm 1$ in their denominators. The local self-energy is, by definition, the average of the two

$$\Sigma_l^\sigma(i\nu_n) = \frac{1}{2} \left(\Sigma^\sigma(\pi, i\nu_n) + \Sigma^\sigma(0, i\nu_n) \right) = \frac{U}{2} + \frac{U^2}{4} \frac{i\nu_n + \mu - \varepsilon_d - \frac{U}{2}}{(i\nu_n + \mu - \varepsilon_d - \frac{U}{2})^2 - (3t)^2}.$$

The difference

$$\Delta \Sigma_l^\sigma(i\nu_n) = \frac{1}{2} \left(\Sigma^\sigma(\pi, i\nu_n) - \Sigma^\sigma(0, i\nu_n) \right) = \frac{U^2}{4} \frac{3t}{(i\nu_n + \mu - \varepsilon_d - \frac{U}{2})^2 - (3t)^2},$$

thus measures the importance of non-local effects; it would be zero if the self-energy was independent of k . Next we define the hybridization function

$$F^\sigma(i\nu_n) = \frac{(t + \Delta \Sigma_l^\sigma(i\nu_n))^2}{i\nu_n + \mu - \varepsilon_d - \Sigma_l^\sigma(i\nu_n)}$$

which for $U = 0$ becomes

$$F_0^\sigma(i\nu_n) = \frac{t^2}{i\nu_n}.$$

By using these definitions, we can rewrite the local Green function as

$$G_{i,i}^\sigma(i\nu_n) = \frac{1}{i\nu_n + \mu - \varepsilon_d - F^\sigma(i\nu_n) - \Sigma_l^\sigma(i\nu_n)}. \quad (5)$$

It is important to point out that, as one may see from the formulas above, the local Green function and the local self-energy satisfy the following *local Dyson equation*

$$\Sigma_l^\sigma(i\nu_n) = \frac{1}{\mathfrak{G}_{i,i}^\sigma(i\nu_n)} - \frac{1}{G_{i,i}^\sigma(i\nu_n)},$$

where $\mathfrak{G}_{i,i}^\sigma(i\nu_n)$ is given by

$$\mathfrak{G}_{i,i}^\sigma(i\nu_n) = \frac{1}{i\nu_n + \mu - \varepsilon_d - F^\sigma(i\nu_n)}.$$

Thus, one could think of mapping the Hubbard dimer into an auxiliary quantum-impurity model, chosen such that, within certain approximations, the impurity Green function is as close as possible to the local Green function of the original problem. How can we do this? Let us adopt as auxiliary model the Anderson molecule

$$\hat{H}^A = \varepsilon_s \sum_{\sigma} \hat{n}_{s\sigma} - t \sum_{\sigma} \left(c_{d\sigma}^{\dagger} c_{s\sigma} + c_{s\sigma}^{\dagger} c_{d\sigma} \right) + \varepsilon_d \sum_{\sigma} \hat{n}_{d\sigma} + U \hat{n}_{d\uparrow} \hat{n}_{d\downarrow}. \quad (6)$$

The first constraint would be that Hamiltonian (6) has a ground state with the same occupations of the 2-site Hubbard model, i.e., at half filling, $n_d = n_s = 1$. Such a *self-consistency condition* is satisfied if $\varepsilon_s = \mu = \varepsilon_d + U/2$. This can be understood by comparing the Hamiltonian matrices of the two models in the Hilbert space with $N = 2$ electrons. To this end, we first order the two-electron states of the Hubbard dimer as

$$\begin{aligned} |1\rangle &= c_{1\uparrow}^{\dagger} c_{2\uparrow}^{\dagger} |0\rangle, & |4\rangle &= \frac{1}{\sqrt{2}} (c_{1\uparrow}^{\dagger} c_{2\downarrow}^{\dagger} - c_{1\downarrow}^{\dagger} c_{2\uparrow}^{\dagger}) |0\rangle, \\ |2\rangle &= c_{1\downarrow}^{\dagger} c_{2\downarrow}^{\dagger} |0\rangle, & |5\rangle &= c_{1\uparrow}^{\dagger} c_{1\downarrow}^{\dagger} |0\rangle, \\ |3\rangle &= \frac{1}{\sqrt{2}} (c_{1\uparrow}^{\dagger} c_{2\downarrow}^{\dagger} + c_{1\downarrow}^{\dagger} c_{2\uparrow}^{\dagger}) |0\rangle, & |6\rangle &= c_{2\uparrow}^{\dagger} c_{2\downarrow}^{\dagger} |0\rangle. \end{aligned}$$

In this basis the Hamiltonian of the Hubbard dimer has the matrix form

$$\hat{H}_2(\varepsilon_d, U, t) = \begin{pmatrix} 2\varepsilon_d & 0 & 0 & 0 & 0 & 0 \\ 0 & 2\varepsilon_d & 0 & 0 & 0 & 0 \\ 0 & 0 & 2\varepsilon_d & 0 & 0 & 0 \\ 0 & 0 & 0 & 2\varepsilon_d & -\sqrt{2}t & -\sqrt{2}t \\ 0 & 0 & 0 & -\sqrt{2}t & 2\varepsilon_d + U & 0 \\ 0 & 0 & 0 & -\sqrt{2}t & 0 & 2\varepsilon_d + U \end{pmatrix}.$$

The ground state, the singlet given in Eq. (4), can be obtained by diagonalizing the lower 3×3 block. For the Anderson molecule, ordering the basis in the same way ($1 \rightarrow d, 2 \rightarrow s$), this Hamiltonian becomes

$$\hat{H}_2^A(\varepsilon_d, U, t; \varepsilon_s) = \begin{pmatrix} \varepsilon_d + \varepsilon_s & 0 & 0 & 0 & 0 & 0 \\ 0 & \varepsilon_d + \varepsilon_s & 0 & 0 & 0 & 0 \\ 0 & 0 & \varepsilon_d + \varepsilon_s & 0 & 0 & 0 \\ 0 & 0 & 0 & \varepsilon_d + \varepsilon_s & -\sqrt{2}t & -\sqrt{2}t \\ 0 & 0 & 0 & -\sqrt{2}t & 2\varepsilon_d + U & 0 \\ 0 & 0 & 0 & -\sqrt{2}t & 0 & 2\varepsilon_s \end{pmatrix}.$$

Comparing the lower 3×3 block of $\hat{H}_2^A(\varepsilon_d, U, t; \varepsilon_s)$ with the corresponding block of $\hat{H}_2(\varepsilon_d, U, t)$ we can see that, unless $\varepsilon_s = \mu = \varepsilon_d + U/2$, the two doubly occupied states $|5\rangle$ and $|6\rangle$ have different energies and thus the two sites $i = 1, 2$ are differently occupied in the ground state.

By setting $\varepsilon_s = \mu$ we find that $\hat{H}_2^A(\varepsilon_d, U, t; \mu) = \hat{H}_2(\varepsilon_d + \frac{U}{4}, \frac{U}{2}, t)$. The $N = 2$ ground state of $\hat{H}_2^A(\varepsilon_d, U, t; \mu)$ has thus the form of the ground-state for the Hubbard dimer

$$|G\rangle_A = \frac{a_2(t, U/2)}{\sqrt{2}} \left(c_{d\uparrow}^\dagger c_{s\downarrow}^\dagger - c_{d\downarrow}^\dagger c_{s\uparrow}^\dagger \right) |0\rangle + \frac{a_1(t, U/2)}{\sqrt{2}} \left(c_{d\uparrow}^\dagger c_{d\downarrow}^\dagger + c_{s\uparrow}^\dagger c_{s\downarrow}^\dagger \right) |0\rangle,$$

and the condition $n_s = n_d = 1$ is satisfied. Since $\varepsilon_s \neq \varepsilon_d$, however, the eigenstates of $\hat{H}^A$ for one electron ($N = 1$) or one hole ($N = 3$) are not the bonding and antibonding states.⁴ The impurity Green function is then given by

$$G_{d,d}^\sigma(i\nu_n) = \frac{1}{4} \left(\frac{1 + w'(t, U)}{i\nu_n - (E_0(2) - E_-(1) - \mu)} + \frac{1 - w'(t, U)}{i\nu_n - (E_0(2) - E_+(1) - \mu)} \right. \\ \left. + \frac{1 + w'(t, U)}{i\nu_n - (E_-(3) - E_0(2) - \mu)} + \frac{1 - w'(t, U)}{i\nu_n - (E_+(3) - E_0(2) - \mu)} \right),$$

where

$$E_0(2) - E_\pm(1) - \mu = - \left(E_\pm(3) - E_0(2) - \mu \right) = -\frac{1}{4} \left(2\Delta(t, U/2) \pm \Delta(t, U) \right),$$

and

$$w'(t, U) = \frac{1}{2} \frac{32t^2 - U^2}{\Delta(t, U)\Delta(t, U/2)}.$$

After some rearrangement we obtain a much simpler expression

$$G_{d,d}^\sigma(i\nu_n) = \frac{1}{i\nu_n + \mu - \varepsilon_d - \mathcal{F}_0^\sigma(i\nu_n) - \Sigma_A^\sigma(i\nu_n)}.$$

The impurity self-energy equals the local self-energy of the Hubbard dimer

$$\Sigma_A^\sigma(i\nu_n) = \frac{U}{2} + \frac{U^2}{4} \frac{i\nu_n}{(i\nu_n)^2 - (3t)^2}.$$

The hybridization function is given by

$$\mathcal{F}_0^\sigma(i\nu_n) = \frac{t^2}{i\nu_n}.$$

For $U = 0$, $G_{d,d}^\sigma(i\nu_n)$ equals the non-interacting impurity Green function

$$G_{d,d}^{0\sigma}(i\nu_n) = \frac{1}{i\nu_n + \mu - \varepsilon_d - \mathcal{F}_0^\sigma(i\nu_n)}.$$

The impurity Green function thus satisfies the *impurity Dyson equation*

$$\Sigma_A^\sigma(i\nu_n) = \frac{1}{G_{d,d}^{0\sigma}(i\nu_n)} - \frac{1}{G_{d,d}^\sigma(i\nu_n)}.$$

⁴The complete list of eigenvalues and eigenvectors of the Anderson molecule for $\varepsilon_s = \varepsilon_d + U/2$ and arbitrary electron number N can be found in Appendix A.2.

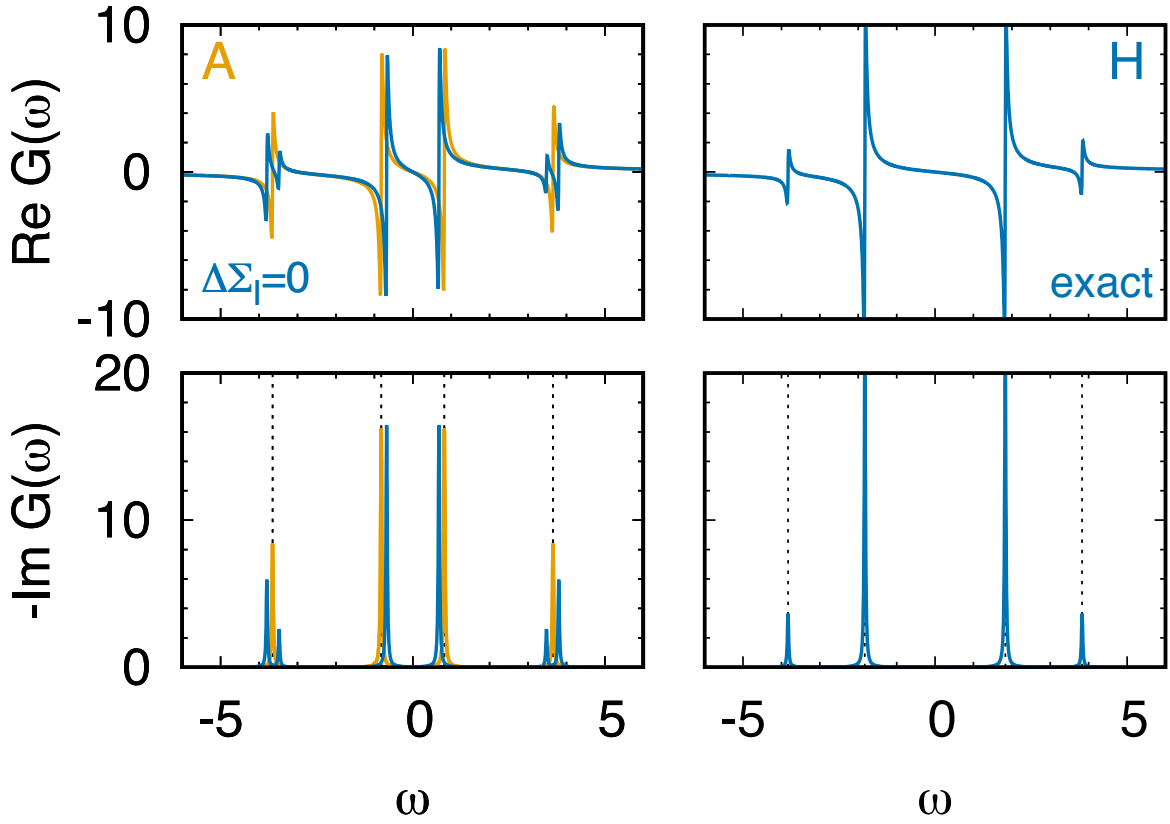


Fig. 1: Green functions of the Hubbard dimer ($t = 1, U = 4$) and the Anderson molecule ($\varepsilon_s = \varepsilon_d + U/2$) in the zero temperature limit. Left panels, blue: Hubbard dimer with local self-energy only, i.e., with $\Delta\Sigma_l^\sigma(\omega) = 0$. Left panels, orange: Anderson molecule. Right panels: Exact Green function of the Hubbard dimer. Dashed lines: Poles of the Green function of the Anderson molecule (left) or Hubbard dimer (right).

In Fig. 1 we show the impurity Green function of the Anderson molecule (orange, left panels) and the local Green function of the 2-site Hubbard model, in the local self-energy approximation (blue, left panels) and exact (blue, right panels). Comparing left and right panels we can see that setting $\Delta\Sigma_l^\sigma(\omega) = 0$ yields large errors. The left panels demonstrate, however, that the spectral function of the Anderson molecule is quite similar to the one of the Hubbard dimer with $\Delta\Sigma_l^\sigma(\omega) = 0$. The small remaining deviations come from the fact that, for the Hubbard dimer, in the impurity Dyson equation, the non-interacting impurity Green function is replaced by $\mathfrak{G}_{i,i}^\sigma(i\nu_n)$ in the local self-energy approximation, i.e., with the *bath* Green function

$$\mathcal{G}_{i,i}^\sigma(i\nu_n) = \frac{1}{i\nu_n + \mu - \varepsilon_d - \mathcal{F}_l^\sigma(i\nu_n)},$$

where

$$\mathcal{F}_l^\sigma(i\nu_n) = \frac{t^2}{i\nu_n + \mu - \varepsilon_d - \Sigma_A^\sigma(i\nu_n)}.$$

We are now in the position of explaining how DMFT works for the Hamiltonian of the Hubbard dimer, choosing the Anderson molecule Hamiltonian (6) as the auxiliary quantum-impurity model. The procedure can be split in the following steps

1. Build the initial quantum impurity model with $G_{d,d}^{0\sigma}(i\nu_n) = G_{i,i}^{0\sigma}(i\nu_n)$. The initial bath is thus defined by energy $\varepsilon_s = \varepsilon_d$ and hopping t .
2. Calculate the local Green function $G_{d,d}^\sigma(i\nu_n)$ for the auxiliary model.
3. Use the local Dyson equation to calculate the impurity self-energy

$$\Sigma_A^\sigma(i\nu_n) = \frac{1}{G_{d,d}^{0\sigma}(i\nu_n)} - \frac{1}{G_{d,d}^\sigma(i\nu_n)}.$$

4. Calculate the local Green function of the Hubbard dimer setting the self-energy equal to the one of the quantum-impurity model

$$G_{i,i}^\sigma(i\nu_n) \sim \frac{1}{2} \left(\frac{1}{i\nu_n + \mu - \varepsilon_d + t - \Sigma_A^\sigma(i\nu_n)} + \frac{1}{i\nu_n + \mu - \varepsilon_d - t - \Sigma_A^\sigma(i\nu_n)} \right).$$

5. Calculate a new bath Green function $\mathcal{G}_{i,i}^\sigma(i\nu_n)$ from the local Dyson equation

$$\mathcal{G}_{i,i}^\sigma(i\nu_n) = \frac{1}{\Sigma_A^\sigma(i\nu_n) + 1/G_{i,i}^\sigma(i\nu_n)}.$$

6. Build a new $G_{d,d}^{0\sigma}(i\nu_n)$ from $\mathcal{G}_{i,i}^\sigma(i\nu_n)$.
7. Restart from the second step.
8. Iterate till self-consistency, i.e., here till $n_d^\sigma = n_i^\sigma$ and $\Sigma_A^\sigma(i\nu_n)$ does not change any more.

The Anderson molecule satisfies the self-consistency requirements for $\varepsilon_s = \mu$. The remaining difference between $G_{d,d}^\sigma(i\nu_n)$, the impurity Green function, and $G_{i,i}^\sigma(i\nu_n)$, the local Green function of the Hubbard dimer in the local self-energy approximation, arises from the difference in the associated hybridization functions

$$\Delta\mathcal{F}_l(i\nu_n) = \mathcal{F}_l^\sigma(i\nu_n) - \mathcal{F}_0^\sigma(i\nu_n) = t^2 p^2 \left(-\frac{2}{i\nu_n} + \frac{1}{i\nu_n - \varepsilon_a} + \frac{1}{i\nu_n + \varepsilon_a} \right)$$

where $p^2 = U^2/8\varepsilon_a^2$ and $\varepsilon_a = \sqrt{9t^2 + U^2}/4$. If we use the Anderson molecule as quantum-impurity model we neglect $\Delta\mathcal{F}_l(i\nu_n)$; the error made is small, as shown in the left panels of Fig. 1. To further improve we would have to modify the auxiliary model adding more bath sites. Remaining with the Anderson molecule, let us compare in more detail its spectral function with the exact spectral function of the Hubbard dimer. Fig. 2 shows that the evolution as a function of U is different for the two Hamiltonians. Anticipating the discussion of later sections, if we compare to the spectral function of the actual lattice Hubbard model, we could say that the Anderson molecule partially captures the behavior of the central “quasi-particle” or “Kondo”

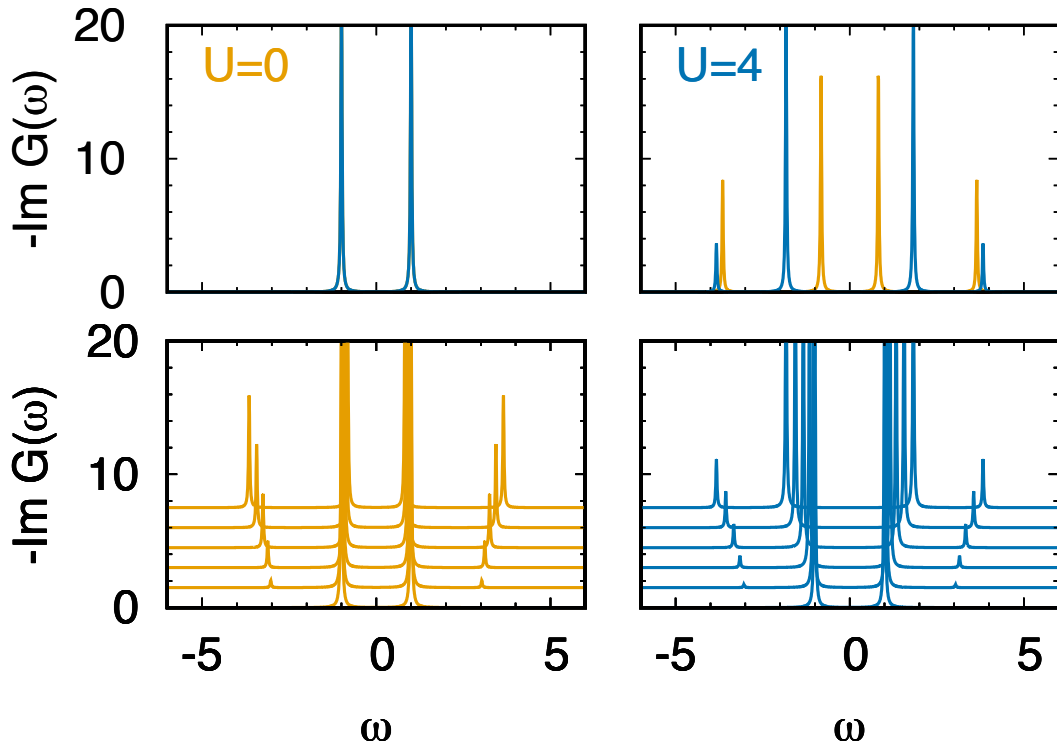


Fig. 2: *Imaginary part of the Green function of the Anderson molecule (orange) and Hubbard dimer (blue) in the zero temperature limit. For the Hubbard dimer the exact Green functions are used, as in the right panels of Fig. 1. Parameters: $t = 1$, $\varepsilon_s = \mu$. Top: $U = 0$ (left) and $U = 4t$ (right). Bottom: Evolution with increasing U from 0 to $4t$ in equal steps.*

peak with increasing U , although the Kondo effect itself is unrealistically described; as a matter of fact, the Kondo energy gain (the “Kondo temperature”) is perturbative ($\propto t^2/U$) in the case of the Anderson molecule, while it is exponentially small for a Kondo impurity in a metallic bath. On the other hand, the Hubbard dimer captures well the Hubbard bands and the gap in the large- U limit. The example of the Anderson molecule also points to the possible shortcomings of DMFT calculations for the lattice Hubbard model (3) in which the quantum-impurity model is solved via exact diagonalization, however using a single bath site or very few; this might perhaps be sufficient in the limit of large gap,⁵ but is bound to eventually fail approaching the metallic regime. Indeed, this failure is one of the reasons why the solution of the Kondo problem required the development of—at the time new—non-perturbative techniques such as the numerical renormalization group. Finally, the example of the Hubbard dimer shows that DMFT is not a good approximation for molecular complexes with two (or few) correlated sites. This is because in such systems the coordination number is the lowest possible, the worst case for dynamical mean-field theory. In three dimensional crystals, instead, the coordination number is typically large enough to make DMFT an excellent approximation.

⁵For a discussion of bath parametrization in exact diagonalization and the actual convergence with the number of bath sites for the lattice Hubbard model see Ref. [11].

2.2 Non-local Coulomb interaction

In Sec. 2.1 we have seen that the local Coulomb interaction gives rise, *alone*, to non-local self-energy terms, which can be very important. What is, instead, the effect of the non-local part of the Coulomb interaction? For a Hubbard dimer, extending the Coulomb interaction to first neighbors leads to the Hamiltonian

$$\begin{aligned} \hat{H} = & \varepsilon_d \sum_{i\sigma} \hat{n}_{i\sigma} - t \sum_{\sigma} \left(c_{1\sigma}^\dagger c_{2\sigma} + c_{2\sigma}^\dagger c_{1\sigma} \right) + U \sum_{i=1,2} \hat{n}_{i\uparrow} \hat{n}_{i\downarrow} \\ & + \sum_{\sigma\sigma'} (V - 2J_V - J_V \delta_{\sigma\sigma'}) \hat{n}_{1\sigma} \hat{n}_{2\sigma'} - J_V \sum_{i \neq i'} \left(c_{i\uparrow}^\dagger c_{i'\downarrow} c_{i'\downarrow}^\dagger c_{i\uparrow} + c_{i'\uparrow}^\dagger c_{i\downarrow} c_{i\downarrow}^\dagger c_{i'\uparrow} \right), \end{aligned}$$

where the parameters in the last two terms are the intersite direct (V) and exchange (J_V) Coulomb interaction. For two electrons the Hamiltonian, in a matrix form, becomes

$$\hat{H}_2^{\text{NL}} = \begin{pmatrix} 2\varepsilon_d + V - 3J_V & 0 & 0 & 0 & 0 & 0 \\ 0 & 2\varepsilon_d + V - 3J_V & 0 & 0 & 0 & 0 \\ 0 & 0 & 2\varepsilon_d + V - 3J_V & 0 & 0 & 0 \\ 0 & 0 & 0 & 2\varepsilon_d + V - J_V & -\sqrt{2}t & -\sqrt{2}t \\ 0 & 0 & 0 & -\sqrt{2}t & 2\varepsilon_d + U & -J_V \\ 0 & 0 & 0 & -\sqrt{2}t & -J_V & 2\varepsilon_d + U \end{pmatrix}.$$

Since $J_V > 0$, the effect of J_V is to lower the energy of triplet states with respect singlet states. This might change the nature of the ground state. If, however, J_V is sufficiently small, the ground state remains a singlet. Setting for simplicity $J_V = 0$, we can notice that $\hat{H}_2^{\text{NL}}$ equals $\hat{H}_2(\varepsilon'_d, U', t)$, the corresponding $N=2$ -electron Hamiltonian of the $J_V=V=0$ Hubbard dimer, with parameters $\varepsilon'_d = \varepsilon_d + V/2$ and $U' = U - V$. The $N=2$ ground state is thus still given by Eq. (4), provided, however, that we replace U with U' in the coefficients. Eventually, in the limit case $U=V$, $\hat{H}_2^{\text{NL}}$ equals the corresponding Hamiltonian of an effective non-correlated dimer. What happens away from half filling? For $N=1$ electrons, eigenvectors and eigenvalues are the same as in the $V=0$ case; for $N=3$ electrons all energies are shifted by $2V$. This leads to further differences in the local Green function with respect to the $V=0$ case—in addition to those arising from replacing U with U' ; to some extent, these additional changes can be interpreted as a hopping enhancement from t to $t + V/2$. Putting all these results together, we could thus say that, in first approximation, the (positive) intersite coupling V effectively reduces the strength of correlations.

In conclusion, *strong-correlation* effects typically appear when the *local term of the electron-electron repulsion dominates*, i.e., when it is much larger than long-range terms. Instead, a hypothetical system in which the Coulomb interaction strength is independent on the distance between sites (here $U=V$) is likely to be already well described via an *effective weakly correlated model*. Of course, in real materials, the effects of long-range Coulomb repulsion can be much more complicated than in the two-site model just discussed, but the general considerations made here remain true even in realistic cases.

2.3 Quantum-impurity solvers: Continuous-time quantum Monte Carlo

For the case of the Anderson molecule exact diagonalization is the simplest quantum impurity solver and the one that provides most insights. Methods based on quantum Monte Carlo (QMC) sampling are often, however, the only option for realistic multi-orbital and/or multi-site models. Here we explain how to obtain the impurity Green function of the Anderson molecule via strong-coupling continuous-time QMC [12], a very successful QMC-based quantum impurity solver. In this approach, the first step consists in splitting the Hamiltonian into bath ($\hat{H}_{\text{bath}}$), hybridization ($\hat{H}_{\text{hyb}}$), and local ($\hat{H}_{\text{loc}}$) terms

$$\hat{H}^A = \underbrace{\varepsilon_s \sum_{\sigma} \hat{n}_{s\sigma}}_{\hat{H}_{\text{bath}}} - t \underbrace{\sum_{\sigma} \left(c_{d\sigma}^{\dagger} c_{s\sigma} + c_{s\sigma}^{\dagger} c_{d\sigma} \right)}_{\hat{H}_{\text{hyb}}} + \underbrace{\varepsilon_d \sum_{\sigma} \hat{n}_{d\sigma} + U \hat{n}_{d\uparrow} \hat{n}_{d\downarrow}}_{\hat{H}_{\text{loc}}}.$$

Next, we write the partition function Z as a perturbation series in the hybridization. To this end, we define $\hat{H}_0 = \hat{H}_{\text{bath}} + \hat{H}_{\text{loc}}$ and rewrite the partition function as

$$Z = \text{Tr} \left(e^{-\beta(\hat{H}_0 - \mu\hat{N})} \hat{V}(\beta) \right)$$

where the operator $\hat{V}(\beta)$ is given by

$$\hat{V}(\beta) = e^{\beta(\hat{H}_0 - \mu\hat{N})} e^{-\beta(\hat{H}_0 + \hat{H}_{\text{hyb}} - \mu\hat{N})} = \sum_m \underbrace{\int_0^{\beta} d\tau_1 \cdots \int_{\tau_{m-1}}^{\beta} d\tau_m}_{\int d\boldsymbol{\tau}^m} \underbrace{(-1)^m \prod_{l=m}^1 \hat{H}_{\text{hyb}}(\tau_l)}_{\hat{O}^m(\boldsymbol{\tau})},$$

and

$$\hat{H}_{\text{hyb}}(\tau_l) = e^{\tau_l(\hat{H}_0 - \mu\hat{N})} \hat{H}_{\text{hyb}} e^{-\tau_l(\hat{H}_0 - \mu\hat{N})} = -t \sum_{\sigma} \left(c_{d\sigma_l}^{\dagger}(\tau_l) c_{s\sigma_l}(\tau_l) + c_{s\sigma_l}^{\dagger}(\tau_l) c_{d\sigma_l}(\tau_l) \right).$$

In this expansion, the only terms that contribute to the trace are even order ones ($m = 2k$) and they are products of d (and s) creator-annihilator couples. We can thus rewrite

$$\int d\boldsymbol{\tau}^{2k} \longrightarrow \int d\boldsymbol{\tau}^k \int d\bar{\boldsymbol{\tau}}^k \qquad \hat{O}^{2k}(\boldsymbol{\tau}) \longrightarrow \sum_{\boldsymbol{\sigma}, \bar{\boldsymbol{\sigma}}} \hat{O}_{\boldsymbol{\sigma}, \bar{\boldsymbol{\sigma}}}^{2k}(\boldsymbol{\tau}, \bar{\boldsymbol{\tau}})$$

where

$$\hat{O}_{\boldsymbol{\sigma}, \bar{\boldsymbol{\sigma}}}^{2k}(\boldsymbol{\tau}, \bar{\boldsymbol{\tau}}) = (t)^{2k} \prod_{i=1}^k \left(c_{d\bar{\sigma}_i}^{\dagger}(\bar{\tau}_i) c_{s\bar{\sigma}_i}(\bar{\tau}_i) c_{s\sigma_i}^{\dagger}(\tau_i) c_{d\sigma_i}(\tau_i) \right).$$

The vector $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \dots, \sigma_k)$ gives the spins $\{\sigma_i\}$ associated with the k impurity destructors at imaginary times $\{\tau_i\}$, while $\bar{\boldsymbol{\sigma}} = (\bar{\sigma}_1, \bar{\sigma}_2, \dots, \bar{\sigma}_k)$ gives the spins $\{\bar{\sigma}_i\}$ associated with the k impurity creators at imaginary times $\{\bar{\tau}_i\}$. It follows that the local and bath traces can be

decoupled and the partition function can be rewritten as

$$\begin{aligned}\frac{Z}{Z_{\text{bath}}} &= \sum_k \int^k d\boldsymbol{\tau} \int^k d\bar{\boldsymbol{\tau}} \sum_{\boldsymbol{\sigma}, \bar{\boldsymbol{\sigma}}} d_{\bar{\boldsymbol{\sigma}}, \boldsymbol{\sigma}}^k(\boldsymbol{\tau}, \bar{\boldsymbol{\tau}}) t_{\boldsymbol{\sigma}, \bar{\boldsymbol{\sigma}}}^k(\boldsymbol{\tau}, \bar{\boldsymbol{\tau}}) \\ d_{\bar{\boldsymbol{\sigma}}, \boldsymbol{\sigma}}^k(\boldsymbol{\tau}, \bar{\boldsymbol{\tau}}) &= (t)^{2k} \text{Tr}_{\text{bath}} \left(e^{-\beta(\hat{H}_{\text{bath}} - \mu \hat{N}_s)} \mathcal{T} \prod_{i=k}^1 c_{s\sigma_i}^\dagger(\tau_i) c_{s\bar{\sigma}_i}(\bar{\tau}_i) \right) / Z_{\text{bath}} \\ t_{\boldsymbol{\sigma}, \bar{\boldsymbol{\sigma}}}^k(\boldsymbol{\tau}, \bar{\boldsymbol{\tau}}) &= \text{Tr}_{\text{loc}} \left(e^{-\beta(\hat{H}_{\text{loc}} - \mu \hat{N}_d)} \mathcal{T} \prod_{i=k}^1 c_{d\sigma_i}(\tau_i) c_{d\bar{\sigma}_i}^\dagger(\bar{\tau}_i) \right),\end{aligned}$$

where

$$c_{d\sigma}(\tau) = e^{\tau(\hat{H}_{\text{loc}} - \mu \hat{N}_d)} c_{d\sigma} e^{-\tau(\hat{H}_{\text{loc}} - \mu \hat{N}_d)}, \quad c_{s\sigma}(\tau) = e^{\tau(\hat{H}_{\text{bath}} - \mu \hat{N}_s)} c_{s\sigma} e^{-\tau(\hat{H}_{\text{bath}} - \mu \hat{N}_s)}.$$

The trace involving only bath operators is simple to calculate, since $\hat{H}_{\text{bath}}$ describes an independent-electron problem, for which Wick theorem holds. It is given by the determinant

$$d_{\bar{\boldsymbol{\sigma}}, \boldsymbol{\sigma}}^k(\boldsymbol{\tau}, \bar{\boldsymbol{\tau}}) = \det \left(F_{\bar{\boldsymbol{\sigma}}, \boldsymbol{\sigma}}^k(\boldsymbol{\tau}, \bar{\boldsymbol{\tau}}) \right)$$

of the $k \times k$ hybridization-function matrix, with elements

$$\left(F_{\bar{\boldsymbol{\sigma}}, \boldsymbol{\sigma}}^k(\boldsymbol{\tau}, \bar{\boldsymbol{\tau}}) \right)_{i', i} = F_{\bar{\sigma}_{i'}, \sigma_i}(\bar{\tau}_{i'} - \tau_i)$$

and

$$F_{\bar{\sigma}, \sigma}(\bar{\tau} - \tau) = \delta_{\bar{\sigma}, \sigma} \frac{t^2}{1 + e^{-\beta(\varepsilon_s - \mu)}} \times \begin{cases} -e^{-\tau(\varepsilon_s - \mu)} & \tau > 0, \\ +e^{-(\beta + \tau)(\varepsilon_s - \mu)} & \tau < 0 \end{cases}$$

This is the imaginary time Fourier transform of the hybridization function introduced previously

$$F_{\bar{\sigma}, \sigma}(i\nu_n) = \frac{t^2}{i\nu_n - (\varepsilon_s - \mu)} \delta_{\bar{\sigma}, \sigma}.$$

The calculation of the local trace is in general more complicated. In the case discussed here, the local Hamiltonian does not flip spins. Thus only terms with an equal number of creation and annihilation operators *per spin* contribute to the local trace, and we can express the partition function in expansion orders per spin, k_σ . This yields

$$\frac{Z}{Z_{\text{bath}}} = \left(\prod_\sigma \sum_{k_\sigma=0}^\infty \int^{k_\sigma} d\boldsymbol{\tau}_\sigma \int^{k_\sigma} d\bar{\boldsymbol{\tau}}_\sigma \right) d_{\bar{\boldsymbol{\sigma}}, \boldsymbol{\sigma}}^k(\boldsymbol{\tau}, \bar{\boldsymbol{\tau}}) t_{\boldsymbol{\sigma}, \bar{\boldsymbol{\sigma}}}^k(\boldsymbol{\tau}, \bar{\boldsymbol{\tau}})$$

where the vectors $\boldsymbol{\sigma} = (\sigma_\uparrow, \sigma_\downarrow)$ and $\bar{\boldsymbol{\sigma}} = (\bar{\sigma}_\uparrow, \bar{\sigma}_\downarrow)$ have $(k_\uparrow, k_\downarrow)$ components, and for each k_σ component $\sigma_i = \bar{\sigma}_i = \sigma$. Thus

$$t_{\boldsymbol{\sigma}, \bar{\boldsymbol{\sigma}}}^k(\boldsymbol{\tau}, \bar{\boldsymbol{\tau}}) = \text{Tr}_{\text{loc}} \left(e^{-\beta(\hat{H}_{\text{loc}} - \mu \hat{N}_d)} \mathcal{T} \prod_\sigma \prod_{i=k_\sigma}^1 c_{d\sigma_i}(\tau_{\sigma_i}) c_{d\bar{\sigma}_i}^\dagger(\bar{\tau}_{\bar{\sigma}_i}) \right).$$

The latter can be calculated analytically. To do this, first we parametrize all configurations for a given spin via a timeline $[0, \beta)$ plus a number of creator/annihilator couples which define

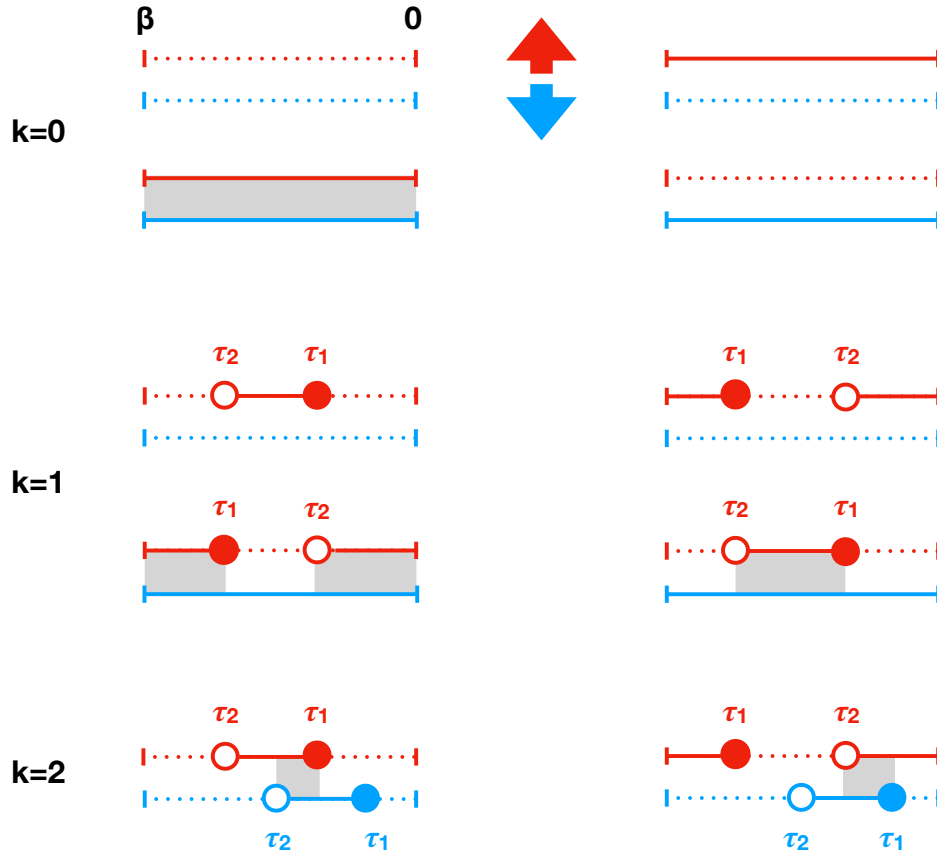


Fig. 3: Representative configurations contributing to the local trace at zero, first and second order. The timelines for spin up are red and those for spin down are blue. The filled circles correspond to the insertion of a creator (time τ_1), and the empty circles to the insertion of a destructor (time τ_2). Dotted lines represent the vacuum state for a given spin, full lines the occupied state. The grey boxes indicate the regions in which $l_{\uparrow,\downarrow} \neq 0$.

segments on the timeline. At zero order two possible configurations exist per spin, an empty timeline, which corresponds to the vacuum state $|0\rangle$, and a full timeline, which corresponds to the state $c_{d\sigma}^\dagger|0\rangle$. A given configuration yields at order k

$$t_{\sigma,\bar{\sigma}}^k(\boldsymbol{\tau}, \bar{\boldsymbol{\tau}}) = \left(\prod_{\sigma} s_{\sigma}^{k_{\sigma}} \right) e^{-\sum_{\sigma\sigma'} ((\varepsilon_d - \mu)\delta_{\sigma\sigma'} + \frac{U}{2}(1 - \delta_{\sigma,\sigma'})) l_{\sigma,\sigma'}}$$

where $l_{\sigma,\sigma'}$ is the length of the overlap of the τ segments for spin σ and σ' , while $s_{\sigma} = \text{sign}(\tau_{\sigma_1} - \bar{\tau}_{\sigma_1})$ is the fermionic sign. Possible configurations at order $k = 0, 1, 2$ are shown in Fig. 3. Next, the partition function is represented as the sum over all configurations $\{c\}$, i.e., in short

$$Z = \sum_c w_c = \sum_c |w_c| \text{sign } w_c.$$

In a compact form, we can write $w_c = d\boldsymbol{\tau}_c d_c t_c$ where $d\boldsymbol{\tau}_c = \prod_{\sigma} \prod_i^{k_{\sigma}} d\tau_{\sigma_i} d\bar{\tau}_{\bar{\sigma}_i}$, and d_c and t_c are the bath and local traces for the configuration c . This expression of the partition function

shows that we can interpret $|w_c|$ as the sampling weight of configuration c . A generic observable $\hat{O}$ can then be obtained as the Monte Carlo average on a finite number of configurations N_c

$$\langle \hat{O} \rangle = \frac{\sum_c \langle \hat{O} \rangle_c |w_c| \text{sign } w_c}{\sum_c |w_c| \text{sign } w_c} = \frac{\sum_c \text{sign } w_c \langle \hat{O} \rangle_c |w_c| / \sum_c |w_c|}{\sum_c \text{sign } w_c |w_c| / \sum_c |w_c|} \approx \frac{\frac{1}{N_c} \sum_c \langle \hat{O} \rangle_c \text{sign } w_c}{\frac{1}{N_c} \sum_c \text{sign } w_c}.$$

The term $\frac{1}{N_c} \sum_c \text{sign } w_c$ in the denominator is the average fermionic sign. When this is small, much longer runs are required to obtain data of the same quality; eventually the computational time can become so long that the calculation is in practice impossible—in these cases we have a sign problem. In practice, the QMC simulation starts from a random configuration c . Next we propose an update $c \rightarrow c'$. Within the Metropolis algorithm, the acceptance ratio is

$$R_{c \rightarrow c'} = \min \left(1, \frac{p_{c' \rightarrow c} |w_{c'}|}{p_{c \rightarrow c'} |w_c|} \right)$$

where $p_{c \rightarrow c'}$ is the proposal probability for the update $c \rightarrow c'$. In the approach described here, known as *segment solver*, the basic updates are addition and removal of segments, antisegments (segments winding over the borders of the timeline, see Fig. 3), or complete lines. As example, let us consider the insertion of a segment for spin σ . A segment is made by a creator and a destructor. The creator is inserted at time τ_{in} ; the move is rejected if τ_{in} is in a region where a segment exists. If created, the segment can have at most length l_{max} , given by the distance between τ_{in} and the time at which the next creator is. Hence

$$p_{c \rightarrow c'} = \frac{d\bar{\tau}}{\beta} \frac{d\tau}{l_{\text{max}}}$$

The proposal probability of the reverse move (removing a segment) is instead given by the inverse of the number of existing segments

$$p_{c' \rightarrow c} = \frac{1}{k_\sigma + 1}$$

The acceptance ratio for the insertion of a segment becomes then

$$R_{c \rightarrow c'} = \min \left(1, \frac{\beta l_{\text{max}}}{k_\sigma + 1} \left| \frac{d_{c'} t_{c'}}{d_c t_c} \right| \right).$$

For the impurity Green function, here the most important observable, the direct average yields

$$\langle \hat{O} \rangle_c = \langle G_{d,d}^\sigma \rangle_c = \sum_{\sigma'} \sum_{i=1}^{k_\sigma} \sum_{j=1}^{k_\sigma} \Delta(\tau, \tau_{\sigma'j} - \bar{\tau}_{\sigma'j}) (M^{k'_\sigma})_{\sigma'j, \sigma'i} \delta_{\sigma_{\sigma j}, \sigma} \delta_{\bar{\sigma}_{\sigma i}, \sigma}$$

where $M^k = (F^k)^{-1}$ is the inverse of the hybridization matrix and

$$\Delta(\tau, \tau') = -\frac{1}{\beta} \begin{cases} \delta(\tau - \tau') & \tau' > 0 \\ -\delta(\tau - (\tau' + \beta)) & \tau' < 0 \end{cases}.$$

2.4 Hartree-Fock versus DMFT approximation

Let us now compare the exact solution of the Hubbard dimer with the result of the Hartree-Fock (HF) approximation, which consists in replacing

$$\hat{H}_U = U \sum_i \hat{n}_{i\uparrow} \hat{n}_{i\downarrow} \rightarrow \hat{H}_U^{\text{HF}} = U \sum_i (\hat{n}_{i\uparrow} \bar{n}_{i\downarrow} + \hat{n}_{i\downarrow} \bar{n}_{i\uparrow} - \bar{n}_{i\uparrow} \bar{n}_{i\downarrow}), \quad (7)$$

where $\bar{n}_{i\sigma}$ is the HF expectation value of the operator $\hat{n}_{i\sigma}$. It is convenient to define

$$\begin{aligned} n_i &= \bar{n}_{i\uparrow} + \bar{n}_{i\downarrow} & n &= \frac{1}{2}(n_1 + n_2) & \delta n &= \frac{1}{2}(n_1 - n_2) \\ m_i &= \frac{1}{2}(\bar{n}_{i\uparrow} - \bar{n}_{i\downarrow}) & m_+ &= \frac{1}{2}(m_1 + m_2) & m_- &= \frac{1}{2}(m_1 - m_2) \end{aligned}$$

Inverting these relations, in the absence of charge disproportionation ($\delta n = 0$), we find

$$\bar{n}_{i\uparrow} = (m_+ + (-1)^{i-1} m_-) + n/2 \quad \bar{n}_{i\downarrow} = -(m_+ + (-1)^{i-1} m_-) + n/2$$

The Hartree-Fock version of the Hubbard dimer Hamiltonian equals the non-interacting Hamiltonian plus a shift of the on-site level. This shift depends on the site and the spin

$$\begin{aligned} \hat{H}_{\text{HF}} &= \sum_{i\sigma} (\varepsilon_d + \Delta_{i\sigma}) \hat{n}_{i\sigma} - t \sum_{\sigma} (c_{1\sigma}^\dagger c_{2\sigma} + c_{2\sigma}^\dagger c_{1\sigma}) - \Delta_0 \\ \Delta_0 &= 2U \left(\frac{n^2}{4} - m_+^2 - m_-^2 \right) \\ \Delta_{i\sigma} &= U \left(\frac{1}{2} n - \sigma (m_+ + (-1)^{i-1} m_-) \right), \end{aligned}$$

where $\sigma = +1$ for spin up and $\sigma = -1$ for spin down. Thus we can write immediately the local Green function matrix for site i . It is convenient to use the site basis, hence, to calculate the matrix $G_{i,i'}^\sigma(i\nu_n)$. Then we have

$$G_{i,i}^\sigma(i\nu_n) = \left(\begin{array}{cc} i\nu_n - (\varepsilon_d - \mu + \Sigma_{1,1}^\sigma(i\nu_n)) & t \\ t & i\nu_n - (\varepsilon_d - \mu + \Sigma_{2,2}^\sigma(i\nu_n)) \end{array} \right)_{i,i}^{-1}$$

where we introduced the diagonal self-energy matrix

$$\Sigma_{i,i'}^\sigma(i\nu_n) = \Delta_{i\sigma} \delta_{i,i'}.$$

This shows that the self-energy is not dependent on the frequency, i.e., Hartree-Fock is a *static* mean-field approach. The value of the parameters m_+ and m_- have to be found solving the *self-consistent equations*

$$\bar{n}_{i\sigma} = \langle \hat{n}_{i\sigma} \rangle = \frac{1}{\beta} \sum_n e^{-i\nu_n 0^-} G_{i,i}^\sigma(i\nu_n) = G_{i,i}^\sigma(0^-). \quad (8)$$

Since the exact ground state of the Hubbard dimer is a singlet, let us consider first the anti-ferromagnetic Hartree Fock solution ($m_+=0$ and $m_- \neq 0$). In this case, the Hartree-Fock self-energy depends on the site and the interaction couples k states. This becomes clear rewriting the Green-function matrix in the basis of the bonding ($k = 0$) and anti-bonding ($k = \pi$) creation/annihilation operators

$$G^\sigma(i\nu_n) = \frac{1}{2} \begin{pmatrix} i\nu_n - (\varepsilon_d - t - \mu + \frac{1}{2} \sum_i \Sigma_{i\sigma}(i\nu_n)) & \frac{1}{2} \sum_i (-1)^{i-1} \Sigma_{i\sigma}(i\nu_n) \\ \frac{1}{2} \sum_i (-1)^{i-1} \Sigma_{i\sigma}(i\nu_n) & i\nu_n - (\varepsilon_d + t - \mu + \frac{1}{2} \sum_i \Sigma_{i\sigma}(i\nu_n)) \end{pmatrix}^{-1}.$$

The diagonal terms are identical, hence

$$\Sigma_{0,0}^\sigma(i\nu_n) = \Sigma_{\pi,\pi}^\sigma(i\nu_n) = \frac{1}{2} (\Sigma_{1,1}^\sigma(i\nu_n) + \Sigma_{2,2}^\sigma(i\nu_n)).$$

The off-diagonal terms $\Sigma_{0,\pi}^\sigma(i\nu_n)$ and $\Sigma_{\pi,0}^\sigma(i\nu_n)$ are not zero, however. This tells us that, by introducing the antiferromagnetic HF correction, we lower the symmetry of the system. Let us now calculate explicitly the eigenstates for different fillings. It is sufficient to diagonalize $\hat{H}_1$, the Hamiltonian in the 1-electron sector; the many-electron states can be obtained by filling the one-electron states respecting the Pauli principle. The Hamiltonian $\hat{H}_1$ can be written as $\hat{H}_1 = \hat{H}'_1 + \varepsilon_d \hat{N} - \Delta_0$, and, in the antiferromagnetic case we then have

$$\hat{H}'_1 = \begin{pmatrix} U(\frac{1}{2}n - m_-) & -t & 0 & 0 \\ -t & U(\frac{1}{2}n + m_-) & 0 & 0 \\ 0 & 0 & U(\frac{1}{2}n + m_-) & -t \\ 0 & 0 & -t & U(\frac{1}{2}n - m_-) \end{pmatrix}.$$

This leads to the (normalized) states

$ 1\rangle_l$	$E_l(1)$
$ 1\rangle_3 = a_2 1, 1/2, \uparrow\rangle_1 - a_1 1, 1/2, \uparrow\rangle_2$	$\varepsilon_0(1) + \Delta_1(t, U)$
$ 1\rangle_2 = a_1 1, 1/2, \downarrow\rangle_1 - a_2 1, 1/2, \downarrow\rangle_2$	$\varepsilon_0(1) + \Delta_1(t, U)$
$ 1\rangle_1 = a_1 1, 1/2, \uparrow\rangle_1 + a_2 1, 1/2, \uparrow\rangle_2$	$\varepsilon_0(1) - \Delta_1(t, U)$
$ 1\rangle_0 = a_2 1, 1/2, \downarrow\rangle_1 + a_1 1, 1/2, \downarrow\rangle_2$	$\varepsilon_0(1) - \Delta_1(t, U)$

where $\varepsilon_0(1) = \varepsilon_d + U(1/2 + 2m_-^2 - n^2/2)$ and $a_1^2 = \frac{1}{2} \left(1 + \frac{Um_-}{\Delta_1(t, U)} \right)$, while $\Delta_1(t, U) = \sqrt{(m_-U)^2 + t^2}$. At half filling, if we assume that only the ground doublet is occupied, solving the self-consistent equations (8) yields

$$m_- = 0 \quad \text{or} \quad m_- = \frac{1}{2} \sqrt{1 - \frac{4t^2}{U^2}}.$$

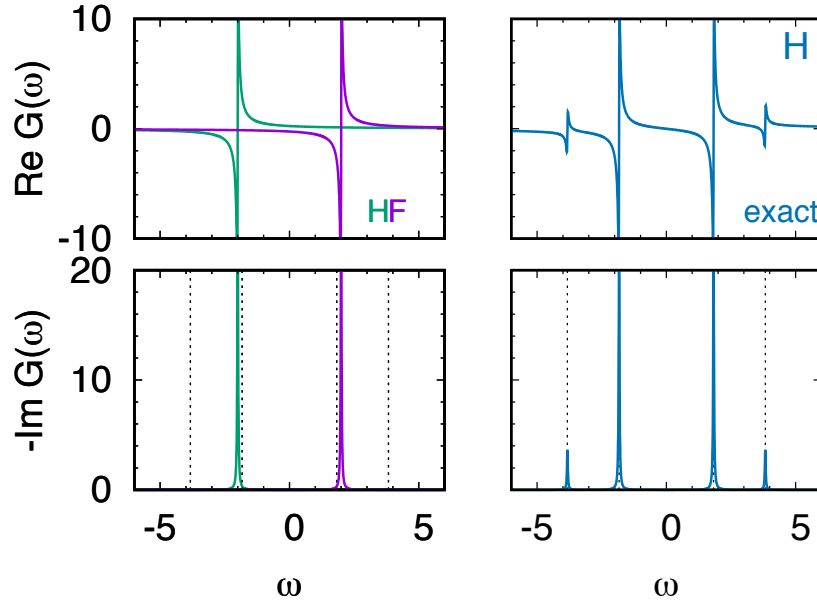


Fig. 4: Local Green function $G^\sigma(\omega)$ of the Hubbard dimer ($t = 1, U = 4$) at half filling and in the zero temperature limit. Left: Antiferromagnetic Hartree-Fock (HF) approximation; green and violet distinguish $G^\uparrow(\omega)$ and $G^\downarrow(\omega)$ at a given site. Right: Exact local Green function, for which $G^\uparrow(\omega) = G^\downarrow(\omega)$. Dashed lines: Poles of the exact local Green function.

As a result, for the non-trivial solution ($m_- \neq 0$), the gap at half filling is $E_g^{\text{HF}} = 2\Delta_1(t, U) = U$. In Fig. 4, Hartree-Fock and exact spectral function are compared for a specific parameters choice. Looking at the size of the gap only, one could naively infer that, for the case shown, Hartree-Fock is an excellent approximation, better than DMFT (see Fig. 2 top-right panel). This would be, however, the wrong conclusion. Although, due to the small coordination number, in this specific case, DMFT indeed sizably underestimates the gap, it nevertheless captures the essential nature of the ground state. Instead, while HF, in this specific case,⁶ gives an almost exact gap, it does it via a qualitatively incorrect description (antiferromagnetic ground state). Increasing the coordination number, DMFT approaches the exact solution, while the Hartree-Fock approximation, which is based on a wrong picture, does not. Let us analyze in more detail the differences between the Hartree-Fock approximation and the exact solution. The Hartree-Fock Hamiltonian for two electrons is given by $\hat{H}_2 = \hat{H}'_2 + \varepsilon_d \hat{N} - \Delta_0$, and

$$\hat{H}'_2 = \begin{pmatrix} U & 0 & 0 & -2Um_- & 0 & 0 \\ 0 & U(1 - 2m_+) & 0 & 0 & 0 & 0 \\ 0 & 0 & U(1 + 2m_+) & 0 & 0 & 0 \\ -2Um_- & 0 & 0 & U & -\sqrt{2}t & -\sqrt{2}t \\ 0 & 0 & 0 & -\sqrt{2}t & U & 0 \\ 0 & 0 & 0 & -\sqrt{2}t & 0 & U \end{pmatrix}$$

⁶The Hartree-Fock approximation often overestimates the size of the gap, however.

For $m_+=0$ and $m_- \neq 0$ (antiferromagnetic solution) the eigenvalues and eigenvectors are

$ 2\rangle_l$	$E_l(2)$
$ 2\rangle_5 = \frac{1}{\sqrt{2}} \left(2, 0, 0\rangle_0 + a_2 2, 1, 0\rangle - \frac{a_1}{\sqrt{2}} (2, 0, 0\rangle_1 + 2, 0, 0\rangle_2) \right)$	$\varepsilon_0(2) + 2\Delta_1(t, U)$
$ 2\rangle_4 = \frac{1}{\sqrt{2}} (2, 0, 0\rangle_1 - 2, 0, 0\rangle_2)$	$\varepsilon_0(2)$
$ 2\rangle_3 = 2, 1, 1\rangle$	$\varepsilon_0(2)$
$ 2\rangle_2 = 2, 1, -1\rangle$	$\varepsilon_0(2)$
$ 2\rangle_1 = a_1 2, 1, 0\rangle + a_2 \frac{1}{\sqrt{2}} (2, 0, 0\rangle_1 + 2, 0, 0\rangle_2)$	$\varepsilon_0(2)$
$ 2\rangle_0 = \frac{1}{\sqrt{2}} \left(2, 0, 0\rangle_0 - a_2 2, 1, 0\rangle + \frac{a_1}{\sqrt{2}} (2, 0, 0\rangle_1 + 2, 0, 0\rangle_2) \right)$	$\varepsilon_0(2) - 2\Delta_1(t, U)$

where $\varepsilon_0(2) = 2\varepsilon_d + U(1 + 2m_-^2 - n^2/2)$, and $a_1^2 = t^2/\Delta_1^2(t, U)$. The antiferromagnetic Hartree-Fock ground state has an overlap with the correct ground state, however incorrectly mixes triplet and singlet states, thus breaking the rotational symmetry of the model. For this reason, its energy, in the large U limit, is $2\varepsilon_d - 2t^2/U$ and not $2\varepsilon_d - 4t^2/U$ as in the exact case. For the ferromagnetic solution of the Hartree-Fock equations ($m_-=0$ and $m_+ \neq 0$) the eigenvalues and eigenvectors are instead

$ 2\rangle_l$	$E_l(2)$
$ 2\rangle_5 = 2, 1, -1\rangle$	$\varepsilon_0^+(2) + 2Um_+$
$ 2\rangle_4 = \frac{1}{\sqrt{2}} \left(2, 0, 0\rangle_0 - \frac{1}{\sqrt{2}} (2, 0, 0\rangle_1 + 2, 0, 0\rangle_2) \right)$	$\varepsilon_0^+(2) + 2t$
$ 2\rangle_3 = \frac{1}{\sqrt{2}} (2, 0, 0\rangle_1 - 2, 0, 0\rangle_2)$	$\varepsilon_0^+(2)$
$ 2\rangle_2 = 2, 1, 0\rangle$	$\varepsilon_0^+(2)$
$ 2\rangle_1 = \frac{1}{\sqrt{2}} \left(2, 0, 0\rangle_0 + \frac{1}{\sqrt{2}} (2, 0, 0\rangle_1 + 2, 0, 0\rangle_2) \right)$	$\varepsilon_0^+(2) - 2t$
$ 2\rangle_0 = 2, 1, 1\rangle$	$\varepsilon_0^+(2) - 2Um_+$

where $\varepsilon_0^+(2) = 2\varepsilon_d + U(1 + 2m_+^2 - n^2/2)$. The ferromagnetic Hartree-Fock correction thus yields an incorrect sequence of levels; the ground state for large U/t , indicated as $|2\rangle_0$ in the table, has no overlap with the exact ground state of the Hubbard dimer. It is, instead, one of the states of the first excited triplet. An important observation is that, despite the errors, the energy difference between ferro- and antiferro-magnetic ground state is

$$E_{AF} - E_F \sim -\frac{2t^2}{U},$$

which is indeed the correct value in the small t/U limit. It does not correspond, however, to the actual singlet-triplet excitation energy, $\Gamma \sim 4t^2/U$.

We can now directly compare the Hartree-Fock and the dynamical mean-field approximation for the two-site Hubbard model. Both approaches are based on the solution of self-consistent mean-field-type equations, and therefore they are both mean-field methods. In Hartree-Fock the self-energy is frequency-independent (static), while in DMFT depends on the frequency (dynamical). In Hartree-Fock the self-energy can be site-dependent, as we have seen in the antiferromagnetic case. In DMFT, allowing for a site-dependent self-energy leads to cluster or cluster-like extensions of the method. If the same unit cell is used in DMFT and Hartree-Fock, we can identify another relation between the two methods. In the infinite-frequency limit, the DMFT self-energy equals the Hartree-Fock self-energy. This can be shown analytically in a simple way. For the case of the two-site Hubbard model, the expansion of the local lattice Green function at high frequency in the local-self-energy approximation is

$$G_{i,i}^{\sigma}(i\nu_n) = \frac{1}{i\nu_n} - \frac{(\mu - \varepsilon_d - \Sigma_l^{\sigma}(\infty))}{(i\nu_n)^2} + O\left(\frac{1}{(i\nu_n)^3}\right).$$

Instead, the expansion of the impurity Green-function can be written as

$$\begin{aligned} G_{d,d}^{\sigma}(i\nu_n) &= \int_0^{\beta} d\tau e^{i\nu_n\tau} G_{d,d}^{\sigma}(\tau) \\ &= \frac{1}{i\nu_n} \int_0^{\beta} d\tau \frac{d}{d\tau} e^{i\nu_n\tau} G_{d,d}^{\sigma}(\tau) = \frac{1}{i\nu_n} \left(-G_{d,d}^{\sigma}(\beta^-) - G_{d,d}^{\sigma}(0^+) - \int_0^{\beta} d\tau e^{i\nu_n\tau} \frac{dG_{d,d}^{\sigma}(\tau)}{d\tau} \right) \\ &= \frac{1}{i\nu_n} + \frac{1}{(i\nu_n)^2} \left(\left. \frac{dG_{d,d}^{\sigma}(\tau)}{d\tau} \right|_{\beta^-} + \left. \frac{dG_{d,d}^{\sigma}(\tau)}{d\tau} \right|_{0^+} + \int_0^{\beta} d\tau e^{i\nu_n\tau} \frac{d^2 G_{d,d}^{\sigma}(\tau)}{d\tau^2} \right) \\ &= \frac{1}{i\nu_n} + \frac{\left. \frac{dG_{d,d}^{\sigma}(\tau)}{d\tau} \right|_{\beta^-} + \left. \frac{dG_{d,d}^{\sigma}(\tau)}{d\tau} \right|_{0^+}}{(i\nu_n)^2} + O\left(\frac{1}{(i\nu_n)^3}\right) \end{aligned}$$

where

$$\left. \frac{dG_{d,d}^{\sigma}(\tau)}{d\tau} \right|_{\beta^-} + \left. \frac{dG_{d,d}^{\sigma}(\tau)}{d\tau} \right|_{0^+} = -\left\langle \left\{ [\hat{H}_A - \mu\hat{N}, c_{d\sigma}], c_{d\sigma}^{\dagger} \right\} \right\rangle = -\mu + \varepsilon_d + U\langle \hat{n}_{-\sigma} \rangle$$

From this result we can conclude that, assuming self-consistency has been reached,

$$\Sigma_l^{\sigma}(\infty) = U\langle \hat{n}_{-\sigma} \rangle.$$

This is exactly the Hartree-Fock expression of the self-energy that we found earlier, and indeed equals the infinite-frequency limit of the DMFT self-energy we previously calculated; however, the occupations $\langle \hat{n}_{-\sigma} \rangle$ in DMFT and Hartree-Fock calculations are typically not the same. In the case of the dimer, DMFT yields $\langle \hat{n}_{-\sigma} \rangle = 1/2$; this would correspond in Hartree-Fock to a trivial solution, in which the self-energy merely shifts all the energy levels by the same amount, and has therefore no relevant effects.

2.5 DMFT for the one-band Hubbard model

The Hubbard Hamiltonian (3) is in principle the simplest model for the description of the Mott metal-insulator transition. In the tight-binding approximation it becomes

$$\hat{H} = \varepsilon_d \sum_{\sigma i} \hat{n}_{i\sigma} - t \sum_{\sigma \langle ii' \rangle} c_{i\sigma}^\dagger c_{i'\sigma} + U \sum_i \hat{n}_{i\uparrow} \hat{n}_{i\downarrow}, \quad (9)$$

where $\langle ii' \rangle$ is a sum over first neighbors. As discussed in the introduction, for $U = 0$, at half-filling, this Hamiltonian describes a metallic band. For $t = 0$ it describes an insulating collection of disconnected atoms. Somewhere in between, at a critical value of t/U , a metal to insulator transition must occur. In this section we will discuss the DMFT solution of (9) and the picture of the metal-insulator transition emerging from it. The first step consists in mapping the original many-body Hamiltonian into an effective quantum-impurity model, such as the Anderson Hamiltonian

$$\hat{H}^A = \underbrace{\sum_{\mathbf{k}\sigma} \varepsilon_{\mathbf{k}}^s \hat{n}_{\mathbf{k}\sigma}}_{\hat{H}_{\text{bath}}} + \underbrace{\sum_{\mathbf{k}\sigma} \left(V_{\mathbf{k}}^s c_{\mathbf{k}\sigma}^\dagger c_{d\sigma} + \text{h.c.} \right)}_{\hat{H}_{\text{hyb}}} + \underbrace{\varepsilon_d \sum_{\sigma} \hat{n}_{d\sigma} + U \hat{n}_{d\uparrow} \hat{n}_{d\downarrow}}_{\hat{H}_{\text{imp}}}.$$

In this model the on-site Coulomb repulsion U appears only in the impurity Hamiltonian, $\hat{H}_{\text{imp}}$, while the terms $\hat{H}_{\text{bath}}$ and $\hat{H}_{\text{hyb}}$, describe, respectively, the bath and the bath-impurity hybridization. In the next step, the quantum-impurity model is solved. Differently from the case of the Anderson molecule, this cannot be done analytically. It requires non-perturbative numerical methods, such as exact diagonalization, the numerical renormalization group, or QMC. Here we describe the DMFT self-consistency loop for a QMC quantum-impurity solver. Solving the quantum-impurity model yields the impurity Green function $G_{d,d}^\sigma(i\nu_n)$. From the impurity Dyson equation we can calculate the impurity self-energy

$$\Sigma_A^\sigma(i\nu_n) = \left(G_{d,d}^{0\sigma}(i\nu_n) \right)^{-1} - \left(G_{d,d}^\sigma(i\nu_n) \right)^{-1}.$$

Next, we adopt the local approximation, i.e., we assume that the self-energy of the Hubbard model equals the impurity self-energy. Then, the local Green function is given by

$$G_{i_c,i_c}^\sigma(i\nu_n) = \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}} \frac{1}{i\nu_n + \mu - \varepsilon_{\mathbf{k}} - \Sigma_A^\sigma(i\nu_n)},$$

where $N_{\mathbf{k}}$ is the number of $\mathbf{k}$ points. The local Dyson equation is used once more, this time to calculate the bath Green function $\mathcal{G}^\sigma(i\nu_n)$, which in turn defines a new quantum-impurity model. This procedure is repeated until self-consistency is reached, i.e., the number of electrons is correct and the self-energy does not change anymore (within a given numerical accuracy). In this situation we have

$$G_{i_c,i_c}^\sigma(i\nu_n) \sim G_{d,d}^\sigma(i\nu_n).$$

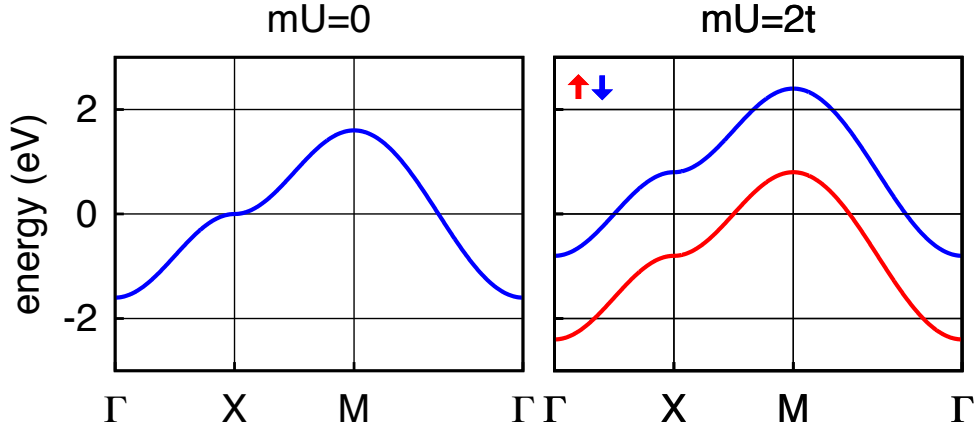


Fig. 5: The metal-insulator transition in ferromagnetic Hartree-Fock. The calculation is for a square lattice tight-binding model with dispersion $\varepsilon_{\mathbf{k}} = -2t(\cos k_x + \cos k_y)$.

It is important to underline that self-consistency is key to the success of DMFT in describing the metal-to-insulator transition. This can, perhaps, be best understood looking once more at the effects of self-consistency in a simpler approach, the static mean-field Hartree-Fock method.⁷ If we chose the same primitive cell as in DMFT, the Hartree-Fock self-energy matrix is

$$\Sigma_{i,i'}^{\sigma}(i\nu_n) = U \left(\frac{n}{2} - \sigma m \right) \delta_{i,i'},$$

where $\sigma = +1$ for spin up and $\sigma = -1$ for spin down and $m = m_+ = (n_{\uparrow} - n_{\downarrow})/2$, with $n_{\sigma} = n_{i\sigma}$. The approximation is then identical to replacing the Hubbard Hamiltonian with

$$\hat{H}_{\text{HF}} = \sum_{\mathbf{k}\sigma} \left[\varepsilon_{\mathbf{k}} + U \left(\frac{1}{2} - \sigma m \right) \right] \hat{n}_{\mathbf{k}\sigma}. \quad (10)$$

This shows that $h_{\text{eff}} = 2Um$ plays the role of an effective magnetic field (Weiss field). The self-consistency criterion is

$$\bar{n}_{\sigma} = \bar{n}_{i\sigma} = \langle \hat{n}_{i\sigma} \rangle_{\text{HF}},$$

where the expectation value $\langle \hat{n}_{i\sigma} \rangle_{\text{HF}}$ is calculated using the Hamiltonian $\hat{H}_{\text{HF}}$, which in turn depends on $\bar{n}_{\sigma}$ via m . This gives the self-consistency equation

$$m = \frac{1}{2} \frac{1}{N_{\mathbf{k}}} \sum_{\mathbf{k}\sigma} \frac{\sigma e^{-\beta(\varepsilon_{\mathbf{k}} + U(\frac{1}{2} - \sigma m) - \mu)}}{1 + e^{-\beta(\varepsilon_{\mathbf{k}} + U(\frac{1}{2} - \sigma m) - \mu)}}.$$

If we set $m = 0$ the equation is satisfied; for such a trivial solution the static mean-field correction in Eq. (10) merely redefines the chemical potential and has therefore no effect. For sufficiently large U , however, a non-trivial solution ($m \neq 0$) can be found. If $m \neq 0$ the spin up and spin down bands split, and eventually a gap can open. This is shown in Fig. 5. The static mean-field correction in Eq. (10) equals the contribution of the Hartree diagram to the self-energy, $\Sigma_{\text{H}}^{\sigma}(i\nu_n) = U\bar{n}_{-\sigma}$. In many-body perturbation theory, however, $\bar{n}_{\sigma} = 1/2$, i.e., $m = 0$.

⁷Keeping in mind that many self-consistent solutions obtained with the Hartree-Fock method are spurious.

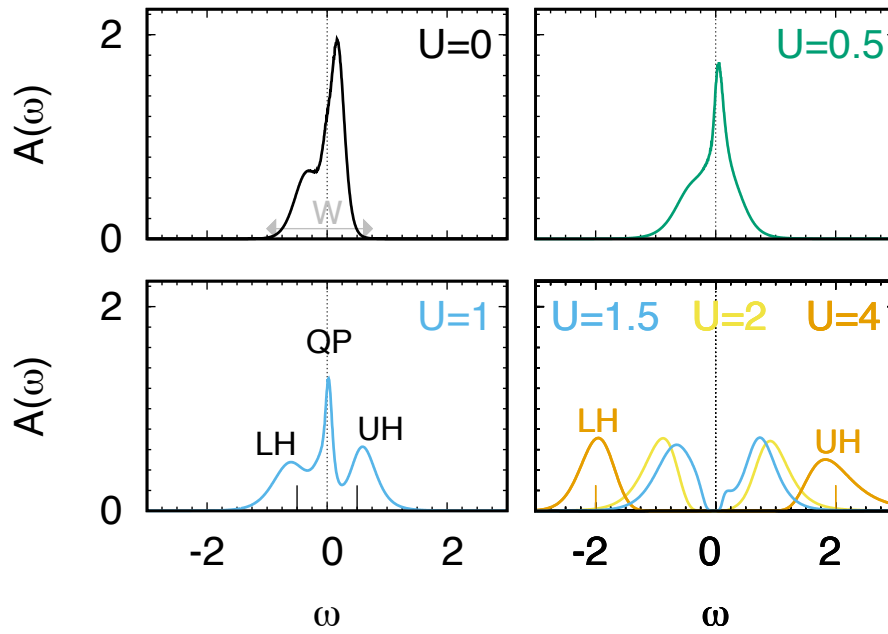


Fig. 6: *VOMoO₄: LDA+DMFT spectral function at finite temperature for $0 \leq U \leq 4$. Energies are in eV and spectral functions in states/spin/eV. The calculations have been done using a continuous-time hybridization-expansion QMC solver [13]. A detailed LDA+DMFT study of the electronic and magnetic properties *VOMoO₄* can be found in Ref. [14].*

In the self-consistent static mean-field approximation, instead, m can differ from zero, and a phenomenon not described by the mere Hartree diagram can be captured, ferromagnetism in a correlated metal. If mU is larger than the bandwidth, the system can even become an insulator. In DMFT the role of the Weiss field is played by the bath Green function $\mathcal{G}_{i,i}^\sigma(i\nu_n)$. The emerging picture of the Mott transition is described in Fig. 6 for a representative single-band material. In the $U = 0$ limit, the spectral function $A_0(\omega)$ is metallic at half filling (top left panel). For finite U , if we set $\Sigma_A^\sigma(\omega) = 0$ as initial guess, the DMFT self-consistency loop starts with $A(\omega) = A_0(\omega)$. For small U/t , the converged spectral function $A(\omega)$ is still similar to $A_0(\omega)$. This can be seen comparing the $U = 0.5$ and $U = 0$ panels in Fig. 6. Further increasing U/t , sizable spectral weight is transferred from the zero-energy quasi-particle peak to the lower (LH) and upper (UH) Hubbard bands, centered at $\omega \sim \pm U/2$. This can be observed in the $U = 1$ panel of Fig. 6. The system is still metallic, but with strongly renormalized masses and short lifetimes, reflected in the narrow quasi-particle (QP) peak. Finally, for U larger than a critical value ($U \geq 1.5$ in the figure) a gap opens and the system becomes a Mott insulator. When this happens the self-energy diverges at low frequency, where

$$\Sigma_A^\sigma(\omega + i0^+) \sim \frac{U}{2} + \frac{A}{\omega + i0^+}.$$

In the large U/t limit the gap increases linearly with the Coulomb repulsion, i.e., $E_g(1) \sim U - W$, where W is the bandwidth.

2.6 DMFT for multi-orbital models

The multi-orbital Hubbard-like Hamiltonian has the form

$$\begin{aligned}\hat{H} &= \hat{H}_0 + \hat{H}_U \\ \hat{H}_0 &= - \sum_{ii'} \sum_{\sigma} \sum_{mm'} t_{m\sigma, m'\sigma'}^{i, i'} c_{im\sigma}^\dagger c_{i'm'\sigma'} \\ \hat{H}_U &= \frac{1}{2} \sum_i \sum_{\sigma\sigma'} \sum_{mm'} \sum_{pp'} U_{mpm'p'} c_{im\sigma}^\dagger c_{ip\sigma'}^\dagger c_{ip'\sigma'} c_{im'\sigma},\end{aligned}$$

where m, m' and p, p' are different orbitals and the Coulomb tensor is local. The DMFT approach can be extended to solve models of this form, mapping them to multi-orbital quantum-impurity models. The main changes with respect to the formalism introduced in the previous section are then the following

$$\begin{aligned}\varepsilon_{\mathbf{k}} &\rightarrow (H_{\mathbf{k}})_{m\sigma, m'\sigma'} & (i\nu_n + \mu) &\rightarrow (i\nu_n + \mu) \hat{1}_{m\sigma, m'\sigma'} \\ t_{m\sigma, m'\sigma'}^{i, i'} &\rightarrow t_{m\sigma, m'\sigma'}^{i, i'} & \varepsilon_d &\rightarrow \varepsilon_{m\sigma, m'\sigma'}^{i, i'} = -t_{m\sigma, m'\sigma'}^{i, i}\end{aligned}$$

where $\hat{1}$ is the identity matrix. As a consequence, the local Green function, the bath Green function, the hybridization function and the self-energy also become matrices

$$\mathcal{G}^\sigma(i\nu_n) \rightarrow \mathcal{G}_{m, m'}^{\sigma, \sigma'}(i\nu_n) \quad G^\sigma(i\nu_n) \rightarrow G_{m, m'}^{\sigma, \sigma'}(i\nu_n) \quad \Sigma^\sigma(i\nu_n) \rightarrow \Sigma_{m, m'}^{\sigma, \sigma'}(i\nu_n).$$

The corresponding generalization of the self-consistency loop is shown schematically in Fig. 7. Although the extension of DMFT to Hubbard models with many orbitals might appear straightforward, in practice it is not. The bottleneck is the solution of the generalized multi-orbital quantum-impurity problem. The most flexible solvers available so far are all based on QMC. Despite being flexible, QMC-based approaches have limitations. These can be classified in two types. First, with increasing the number of degrees of freedom, calculations become very quickly computationally too expensive—how quickly depends on the specific QMC algorithm used and the actual implementation. Thus, going beyond a rather small number of orbitals and reaching the zero-temperature limit is unfeasible in practice. The second type of limitation is more severe. Increasing the number of degrees of freedom leads, eventually, to the infamous sign problem; when this happens, QMC calculations cannot be performed at all. In order to deal with limitations of the first type, it is crucial to restrict QMC calculations to the essential degrees of freedom; furthermore, we should exploit symmetries, develop fast algorithms and use the power of massively parallel supercomputers to reduce the actual computational time. For the second type of problems not a lot can be done; nevertheless, it has been shown that a severe sign problem might appear earlier with some basis choices than with others [13]. Although eventually we cannot escape it, this suggests that the model set up can be used as a tool to expand the moderate sign-problem zone. For what concerns symmetries, in the paramagnetic case and in absence of spin-orbit interaction or external fields, an obvious symmetry to exploit

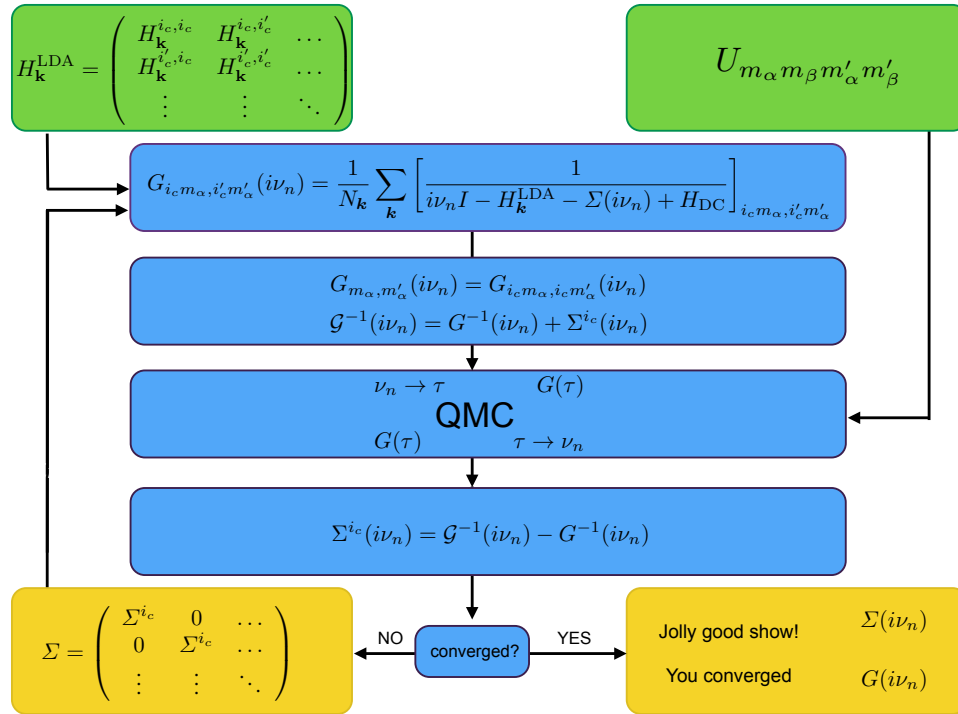


Fig. 7: LDA+DMFT self-consistency loop. The one-electron Hamiltonian is built in the basis of Bloch states obtained from localized Wannier functions, for example in the local-density approximation (LDA); this gives $H_{\mathbf{k}}^{\text{LDA}}$. The set $\{i_c\}$ labels the equivalent correlated sites inside the unit cell. The local Green-function matrix is at first calculated using an initial guess for the self-energy matrix. The bath Green-function matrix is then obtained via the Dyson equation and used to construct an effective quantum-impurity model. The latter is solved via a quantum-impurity solver, here quantum Monte Carlo (QMC). This yields the impurity Green-function matrix. Through the Dyson equation the self-energy is then obtained, and the procedure is repeated until self-consistency is reached.

is the rotational invariance of spins, from which follows

$$A_{m,m'}^{\sigma,\sigma'}(i\nu_n) = \delta_{\sigma,\sigma'} A_{m,m'}(i\nu_n),$$

where $A = \mathcal{G}, G, \Sigma$. In addition, if we use a basis of real functions, the local Green-function matrices are real and symmetric in imaginary time τ , hence

$$A_{m,m'}^{\sigma,\sigma'}(i\nu_n) = \delta_{\sigma,\sigma'} A_{m,m'}(i\nu_n) = \delta_{\sigma,\sigma'} A_{m',m}(i\nu_n).$$

Finally, often the unit cell contains several equivalent correlated sites, indicated as $\{i_c\}$ in Fig. 7. In order to avoid expensive cluster calculations, we can use space-group symmetries to construct the matrices $\mathcal{G}, G, \Sigma$ at a given site i'_c from the corresponding matrices at an equivalent site, e.g., $i_c = 1$. Space-group symmetries also tell us if some matrix elements are zero. For example, for a model with only t_{2g} (or only e_g) states, in cubic symmetry, in the paramagnetic case and in absence of spin-orbit interaction or external fields, we have

$$A_{m,m'}^{\sigma,\sigma'}(i\nu_n) = \delta_{\sigma,\sigma'} A_{m,m}(i\nu_n) \delta_{m,m'}.$$

3 Building materials-specific many-body models

3.1 Model construction

How do we build realistic Hubbard-like models for correlated materials? The state-of-the art approach relies on constructing, for a given system, *materials-specific* Kohn-Sham Wannier functions $\phi_{im\sigma}^{\text{KS}}(\mathbf{r})$. These can be obtained via electronic structure calculations based on density-functional theory [8–10], e.g., in the LDA approximation.⁸ After we have built the complete one-electron basis, the first steps in model-building are those already described in the introduction. We recall here the essential points and then discuss the next stage. The many-body Hamiltonian can be expressed as $\hat{H} = \hat{H}_0 + \hat{H}_U - \hat{H}_{\text{DC}}$, with

$$\begin{aligned}\hat{H}_0 &= \hat{H}^{\text{LDA}} = - \sum_{\sigma} \sum_{ii'} \sum_{mm'} t_{m,m'}^{i,i'} c_{im\sigma}^{\dagger} c_{i'm'\sigma}, \\ \hat{H}_U &= \frac{1}{2} \sum_{ii'jj'} \sum_{\sigma\sigma'} \sum_{mm'pp'} U_{mp\ m'p'}^{ij\ i'j'} c_{im\sigma}^{\dagger} c_{jp\sigma'}^{\dagger} c_{j'p'\sigma'} c_{i'm'\sigma}.\end{aligned}$$

The double-counting correction $\hat{H}_{\text{DC}}$ arises from the fact that the hopping integrals are calculated replacing the electron-nuclei interaction $v_{\text{en}}(\mathbf{r})$ with the self-consistent DFT reference potential

$$v_{\text{R}}(\mathbf{r}) = v_{\text{en}}(\mathbf{r}) + \underbrace{\int d\mathbf{r}' \frac{1}{|\mathbf{r} - \mathbf{r}'|}}_{v_{\text{H}}(\mathbf{r})} + v_{\text{xc}}(\mathbf{r}),$$

which includes the long-range Hartree term $v_{\text{H}}(\mathbf{r})$ and the exchange-correlation contribution $v_{\text{xc}}(\mathbf{r})$. To avoid to count these terms twice, we thus subtract from $\hat{H}_U$ the effects already included in $\hat{H}_0$

$$\hat{H}_U \rightarrow \Delta\hat{H}_U = \hat{H}_U - \hat{H}_{\text{DC}}.$$

Unfortunately we do not know which important correlation effects are indeed included in $\hat{H}_0$ via $v_{\text{R}}(\mathbf{r})$, and therefore the exact expression of $\Delta\hat{H}_U$ is also unknown. The remarkable successes of the LDA suggest, however, that in many materials the LDA is overall a good approximation, and therefore, in those systems at least, the term $\Delta\hat{H}_U$ can be completely neglected. What about strongly-correlated materials? Even in correlated systems, most likely the LDA works rather well for the delocalized electrons or in describing the average or the long-range Coulomb effects. Thus one can think of separating the electrons into *uncorrelated* and *correlated*; only for the latter we do take the correction $\Delta\hat{H}_U$ into account explicitly, assuming furthermore that $\Delta\hat{H}_U$ is local or almost local [8], since we know that it is the local term which is responsible for most non-trivial many-body effects. Typically, correlated electrons are those that partially retain their atomic character, e.g., those that originate from localized *d* and *f* shells; for convenience, here we assume that in a given system they stem from a single atomic shell *l* (e.g., *d* for

⁸Using GGA or similar functionals in place of LDA yields minor differences in the many-body Hamiltonian; on the other hand, using LDA+*U* or similar approximations yields Hartree-Fock-like effects that would have to be subtracted via the double-counting correction.

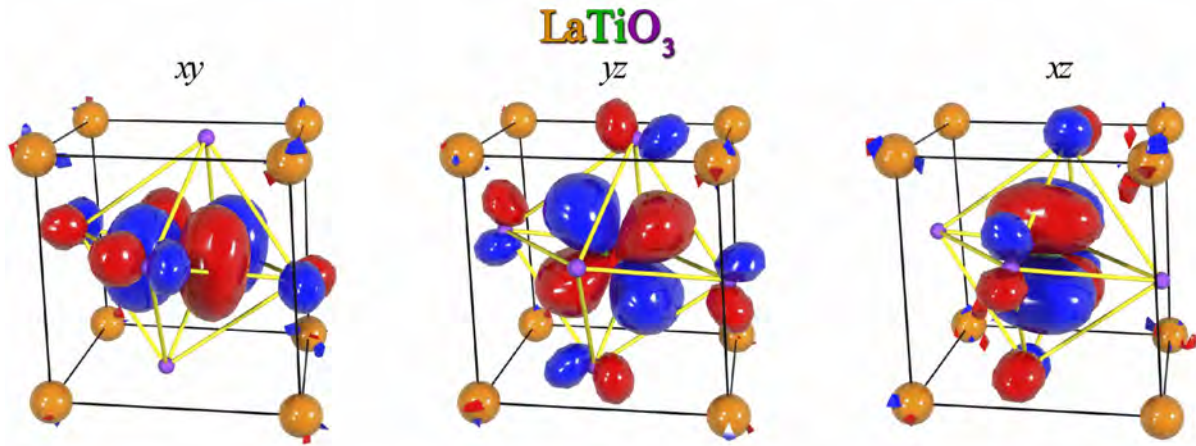


Fig. 8: NMTO Wannier-like orbitals for t_{2g} states in LaTiO_3 obtained via massive downfolding to the t_{2g} bands. The t_{2g} -like orbitals have O p tails at the neighboring O sites reflecting the distortions of the lattice. The figure has been taken from Ref. [15].

transition-metal oxides or f for heavy-fermion systems) and label their states with the atomic quantum numbers l and $m = -l, \dots, l$ of that shell. Thus

$$U_{mpm'p'}^{iji'j'} \sim \begin{cases} U_{mpm'p'}^l & iji'j' = iiii \quad \wedge \quad mp, m'p' \in l \\ 0 & iji'j' \neq iiii \quad \vee \quad mp, m'p' \notin l. \end{cases}$$

Within this approximation $\Delta\hat{H}_U$ is replaced by $\Delta\hat{H}_U^l = \hat{H}_U^l - \hat{H}_{\text{DC}}^l$, where $\hat{H}_{\text{DC}}^l$ is, e.g., given by the static mean-field contribution of $\hat{H}_U^l$. There is a drawback in this procedure, however. By splitting electrons into correlated and uncorrelated we implicitly assume that the main effect of the latter is the renormalization or *screening* of parameters for the former, in particular of the Coulomb interaction. The computation of screening effects remains, unfortunately, a challenge. The calculation of exact screening would require the solution of the original many-body problem, taking all degrees of freedom into account, an impossible task. Commonly-used approximate schemes are the constrained LDA approximation (cLDA) and the constrained random-phase approximation (RPA) [8–10]. Both methods give reasonable estimates of screened Coulomb parameters for DMFT calculations. Typically cRPA calculations include more screening channels and are performed for less localized bases than cLDA calculations; thus cRPA parameters turn out to be often smaller than cLDA ones. To some extent, the difference can be taken as an estimate of the error bar. After we have selected the electrons for which we think is necessary to include explicitly the Hubbard correction, in order to build the final Hamiltonian for DMFT calculations, it is often convenient to integrate out or *downfold*, in part or completely, the weakly correlated states. There are different degrees of downfolding. The two opposite extreme limits are (i) *no downfolding*, i.e., keep explicitly in the Hamiltonian all weakly-correlated states (ii) *massive downfolding*, i.e., downfold all weakly correlated states. If we perform massive downfolding, e.g., downfold to the d (or e_g or t_{2g}) bands at the Fermi level, the Hamiltonian relevant for DMFT takes a simpler form. The LDA part is limited

to the selected orbitals or bands, which, in the ideal case, are decoupled from the rest

$$\hat{H}^{\text{LDA}} = - \sum_{\sigma} \sum_{ii'} \sum_{m_{\alpha} m'_{\alpha}} t_{m_{\alpha}, m'_{\alpha}}^{i, i'} c_{im_{\alpha}\sigma}^{\dagger} c_{i'm'_{\alpha}\sigma}.$$

The local *screened* Coulomb interaction for this set of orbitals is the on-site tensor

$$\hat{H}_U^l = \frac{1}{2} \sum_i \sum_{\sigma\sigma'} \sum_{m_{\alpha} m'_{\alpha}} \sum_{m_{\beta} m'_{\beta}} U_{m_{\alpha} m_{\beta} m'_{\alpha} m'_{\beta}} c_{im_{\alpha}\sigma}^{\dagger} c_{im_{\beta}\sigma'}^{\dagger} c_{im'_{\beta}\sigma'} c_{im'_{\alpha}\sigma}.$$

It is important to point out that the level of downfolding does not modify the hardness of the quantum-impurity problem. If, for example, in studying a transition-metal oxide, we plan to treat only $3d$ bands as correlated, it does not matter if we perform calculations with a Hamiltonian containing also, e.g., O p states, or we rather downfold all states but the $3d$ and work with a set of Wannier basis spanning the $3d$ -like bands only. The number of correlated orbitals in the quantum-impurity problem is the same.⁹

One advantage of massive downfolding is that the double-counting correction typically becomes a shift of the chemical potential, and it is therefore not necessary to calculate it explicitly. A second important advantage is that the interpretation of the final results is simpler. Instead, a disadvantage is that the basis functions are less localized, and therefore the approximation of the Coulomb interaction to a local operator might be less justified, and in some cases it might be necessary to include non-local Coulomb terms. The effect of downfolding on the localization of Wannier functions is illustrated for example in Fig. 8. Finally, another disadvantage of massive downfolding is that the energy window in which the model is valid is more narrow.

Considering all advantages and disadvantages, what is then the best way of performing DMFT calculations? There is no universal answer to this question; it depends on the problem we are trying to solve and the system we are studying. Independently of the degree of downfolding we choose, it is important to point out that a clear advantage of Wannier functions in general is that they carry information about the lattice, bonding, chemistry and distortions. This can be seen once more in Fig. 8, where orbitals are tilted and deformed by the actual structure and chemistry of the compound. Indeed, one might naively think of using an “universal” basis, for example atomic functions, the same for all systems, and thus calculating the hopping integrals using simply the electron-nuclei interaction $v_{en}(\mathbf{r})$. Apart the complications arising from the lack of orthogonality, such a basis has no built-in materials-specific information, except lattice positions. It is therefore a worse starting point to describe the electronic structure, even in the absence of correlations; larger basis sets are required to reach the same accuracy. From the point of view of LDA+DMFT, an advantage of an universal basis would be that it is free from double-counting corrections; on the other hand, however, exactly because we do not use the LDA potential and LDA orbitals to calculate the hopping integrals, we also cannot count on the successes of LDA in the description of average and long-range Coulomb effects. For these reasons *ab-initio* Wannier functions remain so far the basis of choice. They can be built via the Nth-Order Muffin-Tin Orbital (NMTO) method [15], the maximal-localization scheme [16] or projectors. Fig. 8 shows examples of NMTO-based Wannier functions.

⁹The choice might influence how severe the QMC sign problem is, however.

3.2 Localization of the basis and range of the Coulomb interaction

No matter what construction procedure is used, a common characteristic of *ab-initio* Wannier functions is that they are site-centered and localized.¹⁰ A question naturally arises: How crucial is it to use localized functions as one-electron basis? This is an important point, since we have seen that strong-correlation effects arise in systems in which the on-site Coulomb interaction is much larger than longer-range terms. Let us consider therefore two opposite extreme limits. The first is the case in which the basis functions are independent of the lattice position (i.e., they are totally delocalized). For such a basis choice the Coulomb interaction parameters would be the same for every couple of lattice sites, no matter how distant. Thus a Hubbard-like model would be hard to justify. In the second extreme case, we adopt a hypothetical basis so localized that $\psi_{im\sigma}(\mathbf{r})\overline{\psi_{i'm'\sigma'}(\mathbf{r})} \sim \delta_{i,i'}\delta(\mathbf{r} - \mathbf{T}_i)$. Even for such a basis choice, the unscreened Coulomb interaction is *not* local. It is given by

$$U_{mp\ m'p'}^{ij\ i'j'} \propto \frac{\delta_{i,i'}\delta_{j,j'}}{|\mathbf{T}_i - \mathbf{T}_j|},$$

hence it decays slowly with distance, although the (divergent) on-site term dominates. More generally, we can conclude that by increasing the localization of the basis we enhance the importance of the on-site Coulomb repulsion with respect to long-range terms; this better justifies Hubbard-like models—although we have to remember that most of the long-range part of the Coulomb interaction is in any case subtracted via the double-counting correction $\hat{H}_{DC}$. The extreme case of the $\delta(\mathbf{r} - \mathbf{T}_i)$ functions also illustrates, however, how far we can go. A major problem with the extremely localized basis discussed above is that it would make it impossible to properly describe bonding, since the hopping integrals would be zero. Although such a basis is, of course, never used to build many-body models, there is a tempting approximation that has similar flaws. If one uses DFT-based electronic-structure techniques that tile the space in interstitial and non-overlapping atomic spheres (e.g., the LAPW method), it is tempting to use as basis for correlated electrons the atomic functions *defined inside the atomic spheres*. These functions are, by construction, much more localized than Wannier orbitals (even when no down-folding is performed in the Wannier construction). However, they *do not form a complete basis set* in the space of square-integrable functions. This is obvious because such a basis does not even span the LDA bands; to reproduce the bands we need, in addition, functions defined in the interstitial region. This is illustrated in Fig. 9 for a simple example of two quantum well potentials.¹¹ We therefore cannot use it to write the many-body Hamiltonian in the usual form $\hat{H}_0 + \hat{H}_U$. In conclusion, a basis which, as *ab-initio* Wannier functions, is complete and indeed spans the bands, is better justified, although we somewhat lose in localization.

¹⁰Differences in localizations between the various construction procedures are actually small for the purpose of many-body calculations, provided that the same bands are spanned in the same way.

¹¹Another, but less severe, problem of atomic sphere truncations is that the results will depend on the sphere size, in particular when atomic spheres are small.

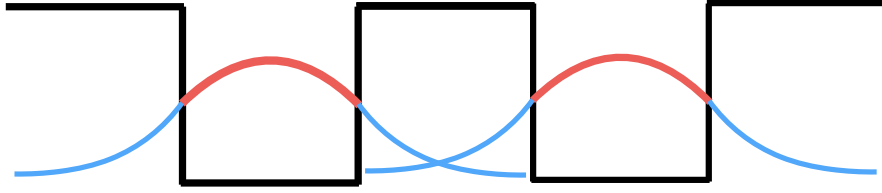


Fig. 9: *The problem of two quantum wells. The figure shows (schematically) for each well the wavefunction of a bound state. If we consider only the part of the wavefunction inside its own well (red in the figure), the differential overlap (and hence the hopping integral) between functions centered on different wells would be zero.*

3.3 Hubbard Hamiltonians for t_{2g} and e_g systems

Several strongly-correlated transition-metal oxides can be described via minimal materials-specific Hubbard-like models which involve only t_{2g} or only e_g bands. A representative system of this kind is the layered perovskite Sr_2RuO_4 with the $4d\ t_{2g}^4$ electronic configuration. Its crystal structure is shown in Fig. 10 (left side), together with the associated LDA bands crossing the Fermi level (top-right panel). Due to the layered structure the xz and yz bands are quasi one-dimensional and the xy band is quasi two-dimensional. Thus, the t_{2g} bands give rise, in first approximation, to a Fermi surface made of four crossing lines (from the xz, yz bands) and a circle (from the xy band), shown schematically in the bottom-right panel of Fig. 10. Experimentally, Sr_2RuO_4 is a correlated metal down to 1.5 K; below this temperature it becomes an anomalous superconductor. Other representative cases of t_{2g} systems are the Mott insulating perovskites LaTiO_3 and YTiO_3 , with the electronic configuration $3d\ t_{2g}^1$. A paradigmatic e_g system is instead the orbitally ordered insulator KCuF_3 , with the electronic configuration $t_{2g}^6 e_g^3$. For all these materials, if we massively downfold all LDA bands but the t_{2g} (or the e_g), the resulting 3-band (or 2-band) generalized Hubbard model takes the following form

$$\begin{aligned} \hat{H} = & \sum_{i\sigma} \sum_{mm'} \varepsilon_{m,m'} c_{im\sigma}^\dagger c_{im'\sigma} - \sum_{\sigma} \sum_{i \neq i'} \sum_{mm'} t_{m,m'}^{i,i'} c_{im\sigma}^\dagger c_{i'm'\sigma} \\ & + U \sum_{i,m} \hat{n}_{im\uparrow} \hat{n}_{im\downarrow} + \frac{1}{2} \sum_{\substack{i\sigma\sigma' \\ m \neq m'}} (U - 2J - J\delta_{\sigma,\sigma'}) \hat{n}_{im\sigma} \hat{n}_{im'\sigma'} \\ & - J \sum_{i,m \neq m'} \left(c_{im\uparrow}^\dagger c_{im\downarrow}^\dagger c_{im'\uparrow} c_{im'\downarrow} + c_{im\uparrow}^\dagger c_{im\downarrow} c_{im'\downarrow}^\dagger c_{im'\uparrow} \right), \end{aligned}$$

where $m, m' = xy, yz, xz$ for t_{2g} and $m, m' = 3z^2 - r^2, x^2 - y^2$ for e_g . The parameters U and J are the direct and exchange screened Coulomb integrals. The Coulomb interaction $\hat{H}_U$ is here assumed to have full $O(3)$ rotational symmetry, as in the atomic limit.¹² The first two terms of

¹²A derivation of the Coulomb interaction tensor for the free atom can be found in my chapter in Ref. [8]. There the differences in the values of U and J for the t_{2g} and e_g orbitals is also discussed.

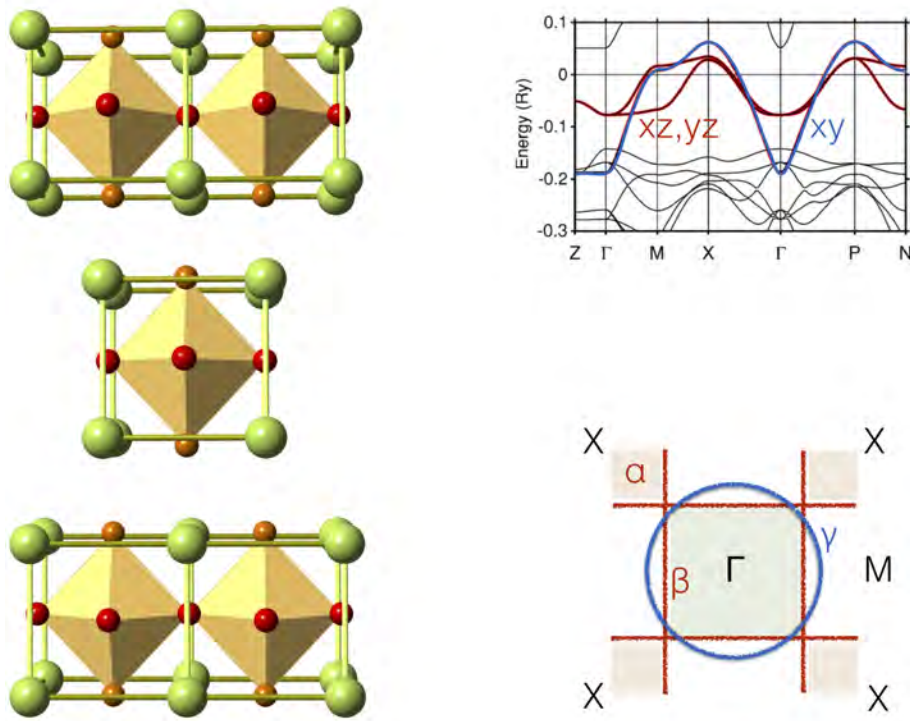


Fig. 10: Left: Crystal structure of the tetragonal layered perovskite Sr_2RuO_4 . Right: Low-energy LDA band structure (top) and schematic representation of the associated Fermi surface (bottom). The band structure was calculated using the Nth-Order Muffin-Tin Orbital (NMTO) method. The figure is rearranged from Ref. [17].

$\hat{H}_U$ are the so-called density-density terms, and the last two are the pair-hopping and spin-flip interactions. We dropped the double-counting correction $\hat{H}_{\text{DC}}$, which in this case is a mere shift of the chemical potential. The energies $\varepsilon_{m,m'}$ are the elements of the crystal-field matrix. In the case of cubic symmetry, the crystal-field matrix, the self-energy, the Green function and the spectral function are all diagonal in orbital space. For low-symmetry systems, however, this is not true. It can be seen in Fig. 11, which shows the diagonal and off-diagonal elements of the spectral-function matrix for the orthorhombic Mott insulator YTiO_3 .

3.4 Spin-orbit interaction and effects of the basis choice

In many interesting systems the spin-orbit interaction $\hat{H}_{\text{SO}}$ plays an important role. In the atomic limit, for the d shells the spin-orbit interaction is

$$\hat{H}_{\text{SO}} = \sum_{\mu} \lambda_{\mu} \sum_{mm'} \sum_{\sigma\sigma'} \epsilon_{m\sigma,m'\sigma'}^{\mu} c_{m\sigma}^{\dagger} c_{m'\sigma'}, \quad \epsilon_{m\sigma,m'\sigma'}^{\mu} = \langle m\sigma | l_{\mu} s_{\mu} | m'\sigma' \rangle,$$

where $\mu = x, y, z$, and λ_{μ} are the spin-orbit couplings, with $\lambda_{\mu} = \lambda$ in $O(3)$ symmetry, and

$$\lambda \sim g\mu_B^2 \left\langle \frac{1}{r} \frac{d}{dr} v_R(r) \right\rangle.$$

For t_{2g} systems the matrix with elements $\hat{\epsilon}_{m\sigma,m'\sigma'} = \sum_{\mu} \lambda_{\mu} \epsilon_{m\sigma,m'\sigma'}^{\mu}$ is given by

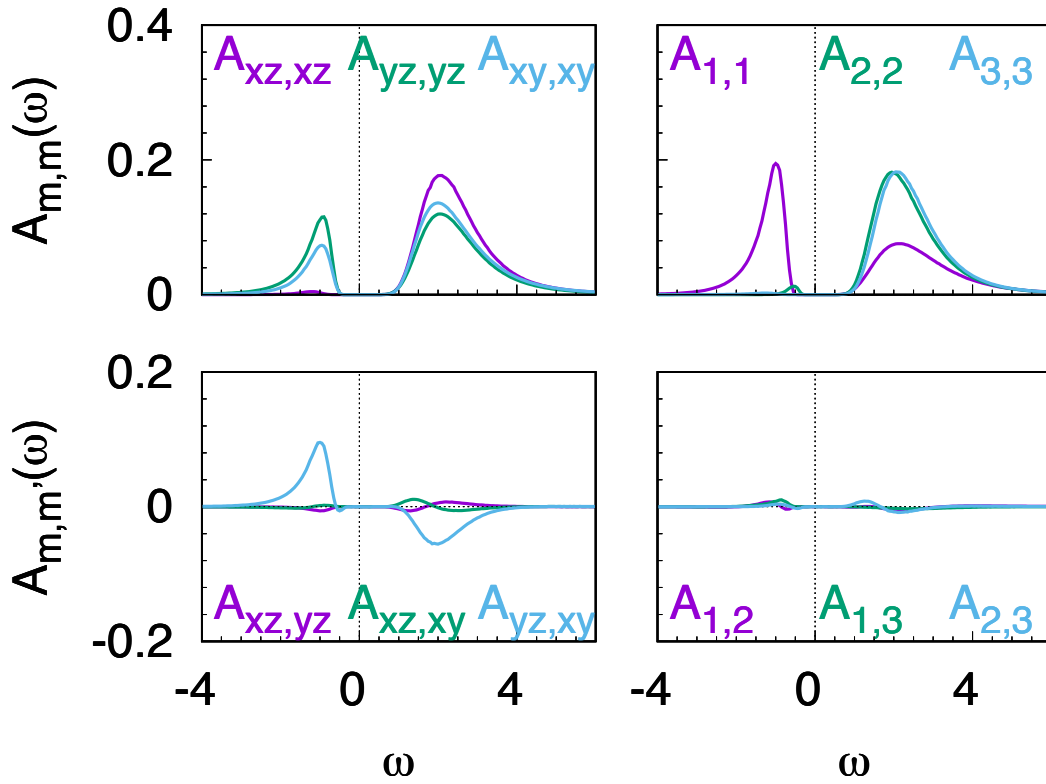


Fig. 11: The LDA+DMFT spectral function matrix of the orbitally-ordered t_{2g}^1 system YTiO_3 , in the (xz, yz, xy) basis (left panels) and in the crystal-field basis (right panels) [15, 22].

$$\hat{\varepsilon} = \begin{pmatrix} \varepsilon_{xy} & 0 & 0 & 0 & \frac{\lambda_y}{2} & -\frac{i\lambda_x}{2} \\ 0 & \varepsilon_{yz} & \frac{i\lambda_z}{2} & -\frac{\lambda_y}{2} & 0 & 0 \\ 0 & -\frac{i\lambda_z}{2} & \varepsilon_{xz} & \frac{i\lambda_x}{2} & 0 & 0 \\ 0 & -\frac{\lambda_y}{2} & -\frac{i\lambda_x}{2} & \varepsilon_{xy} & 0 & 0 \\ \frac{\lambda_y}{2} & 0 & 0 & 0 & \varepsilon_{yz} & -\frac{i\lambda_z}{2} \\ \frac{i\lambda_x}{2} & 0 & 0 & 0 & \frac{i\lambda_z}{2} & \varepsilon_{xz} \end{pmatrix}.$$

Although $\hat{H}_{\text{SO}}$ looks like an innocent one-body term, it turns out that, for materials, simulations including this term are more difficult. This has two reasons: (i) QMC calculations involve Green function matrices of larger size, e.g., 6×6 as in the case just discussed, hence they are from the start computationally more demanding; (ii) QMC calculations are often hampered by a much stronger sign problem; even when it can be tamed, this often happens at the price of further increasing the computational time. Thus, specific basis choices and approximations are used. A possible approach consists in working in the basis that diagonalizes the non-interacting local Green function or the non-interacting local Hamiltonian; such a choice is known to reduce the sign problem, as was first shown in Ref. [13] for the case without spin-orbit interaction. For a system with tetragonal symmetry, the states that diagonalize the local Green function belong either to the Γ_6 or to the Γ_7 irreducible representations, both 2-dimensional. There are two

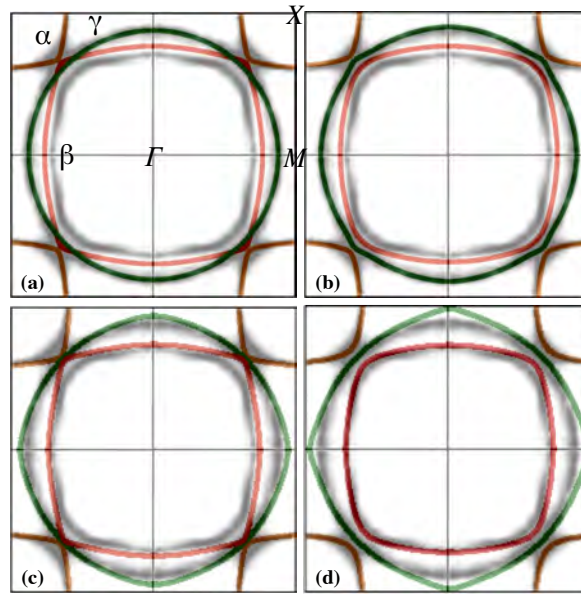


Fig. 12: Fermi surface of Sr_2RuO_4 calculated with LDA (a), LDA+SO (b), LDA+DMFT (c) and LDA+SO+DMFT (d) [18]. The grey maps are experimental results from Ref. [21].

(coupled) Γ_7 representations, defining the space $\Gamma'_7 \oplus \Gamma''_7$. The analytic expression of these states can be found in Refs. [17–19]. The transformation to the $\Gamma_6 \oplus \Gamma'_7 \oplus \Gamma''_7$ basis is, of course, in principle, a mere basis change. Approximations are made, however, if all off-diagonal elements of the Green function are set to zero or the Coulomb tensor is truncated to further tame the sign problem or to reduce the computational time.

It has to be noticed that the Coulomb interaction of the t_{2g} Hubbard Hamiltonian is *only invariant under orthogonal transformations* of the basis. Thus if we change basis to the Γ_6 - Γ'_7 - Γ''_7 representation, the form of the interaction tensor changes. The exact expression of the Coulomb tensor in the angular momentum basis (i.e., the one relevant when the t_{2g} crystal-field splitting is zero) can be found in Ref. [20]. For the Coulomb interaction tensor in the basis of spherical harmonics you can instead look the derivation in Ref. [8].

In the presence of crystal-field splitting, if the spin-orbit interaction does not dominate, it is often preferable to perform the calculations in the t_{2g} basis. To this end, it is key to make QMC codes very efficient in order to reduce as much as possible statistical errors and increase the average sign. Exact LDA+SO+DMFT calculations in the t_{2g} basis have been successfully performed, e.g., for Sr_2RuO_4 , using an interaction-expansion continuous-time quantum Monte Carlo solver, and an orbital-dependent phase which makes the Green function matrix real [18, 19]. This approach allowed us to study, for example, the effects of the spin-orbit interaction on the Fermi surface without approximations. The results are shown in Fig. 12 in comparison with experimental data; we will discuss them in the next section.

3.5 Non-spherical Coulomb terms and double-counting correction

While for a free atom the Coulomb tensor is spherical (symmetry $O(3)$), in a material the screened Coulomb tensor has, in general, the symmetry of the lattice. Taking into account non-spherical Coulomb terms is, in general hard, both because they make QMC calculations more difficult and can worsen the sign problem, and because in their presence the double-counting correction has to be explicitly accounted for, even when massive downfolding is used. For these reasons they are typically neglected. Recently it was shown that they can play, however, a very important role for the Fermi surface [18]. Let us therefore discuss how the double-counting correction can be treated with and without such terms, following the approach of Ref. [18]. One of the classical approximations for the double-counting correction is the so called “around mean-field” approximation. The idea is that LDA describes well the average Coulomb term, in the absence of orbital polarization. This is equivalent to using as double-counting correction the Hartree term of the Coulomb interaction tensor, i.e., the operator

$$\begin{aligned}\hat{H}_U^{\text{DC}} = & U \sum_m (\hat{n}_{m\uparrow} \bar{n}_{m\downarrow} + \bar{n}_{m\uparrow} \hat{n}_{m\downarrow}) + (U - 2J) \sum_{m \neq m'} (\hat{n}_{m\uparrow} \bar{n}_{m'\downarrow} + \bar{n}_{m\uparrow} \hat{n}_{m'\downarrow}) \\ & + (U - 3J) \sum_{\sigma} \sum_{m > m'} (\hat{n}_{m\sigma} \bar{n}_{m'\sigma} + \bar{n}_{m\sigma} \hat{n}_{m'\sigma}) - \mu \hat{N}_d \\ & - U \sum_m \bar{n}_{m\uparrow} \bar{n}_{m\downarrow} + (U - 2J) \sum_{m \neq m'} \bar{n}_{m\uparrow} \bar{n}_{m'\downarrow} + (U - 3J) \sum_{\sigma} \sum_{m > m'} \bar{n}_{m\sigma} \bar{n}_{m'\sigma}\end{aligned}$$

where $\bar{n}_{m\sigma} = n/d$, if n is the number of the correlated electrons per site and d the orbital degeneracy. Within this approximation we have, after collecting all terms,

$$\begin{aligned}\hat{H}_U^{\text{DC}} = & (\delta\mu - \mu) \hat{N}_d - \frac{n^2}{d} (U(2d - 1) - 5(d - 1)) \\ \delta\mu = & \frac{n}{d} (U(2d - 1) - 5J(d - 1)).\end{aligned}$$

If we perform massive downfolding to the correlated bands, as previously mentioned, this is merely a shift of the chemical potential and can therefore be neglected. Let us now consider the case in which the Coulomb interaction has an additional term that does not change the average U but has tetragonal symmetry

$$\Delta \hat{H}_U = \frac{\Delta U}{3} (2\hat{n}_{xy\uparrow} \hat{n}_{xy\downarrow} - \hat{n}_{xz\uparrow} \hat{n}_{xz\downarrow} - \hat{n}_{yz\uparrow} \hat{n}_{yz\downarrow})$$

We can now use the around mean-field approximation for this term as well. We find

$$\begin{aligned}\Delta \hat{H}_U^{\text{DC}} = & \frac{n}{6} \frac{\Delta U}{3} \sum_{\sigma} (2\hat{n}_{xy\sigma} - \hat{n}_{xz\sigma} - \hat{n}_{yz\sigma}) = \frac{n}{6} \Delta U \sum_{\sigma} \hat{n}_{xy\sigma} - \delta\mu' \hat{N} \\ \delta\mu' = & \frac{n}{6} \frac{\Delta U}{3}.\end{aligned}$$

This term, in addition to a shift of the chemical potential, yields an effective change of the crystal-field splitting ε_{CF} , and has therefore to be accounted for explicitly.

How does ΔU change the Fermi surface of Sr_2RuO_4 ? The Fermi surface is determined by the poles of the Green function at zero frequency. These depend on the non-interacting Hamiltonian and the self-energy matrix at zero frequency. In the Fermi-liquid regime, and within the DMFT approximation, the effect of the self-energy is merely to modify the on-site part of the Hamiltonian, i.e., the crystal-field splitting and the spin-orbit couplings

$$\varepsilon_{\text{CF}} \rightarrow \varepsilon_{\text{CF}} + \Delta\varepsilon_{\text{CF}}(0),$$

$$\lambda_\mu \rightarrow \lambda_\mu + \Delta\lambda_\mu(0).$$

Both $\Delta\varepsilon_{\text{CF}}(0)$ and $\Delta\lambda_\mu(0)$ are positive for Sr_2RuO_4 , and lead to an almost doubling of the LDA parameters. The positive $\Delta\varepsilon_{\text{CF}}(0)$ shrinks the β sheet (xz/yz bands) and enlarges the γ (xy band) sheet. This can be understood from the schematic Fermi surface and the LDA band structure in Fig. 10. Enhancing the crystal-field splitting corresponds to slightly moving the xy band downwards and the xz/yz bands upwards with respect to the Fermi level. The enhancement of the spin-orbit couplings has a large Hartree-Fock component [19], since the spin-orbit interaction yields a small but finite off-diagonal occupation matrix. For an $O(3)$ -symmetric Coulomb tensor, the Hartree-Fock enhancement of the spin-orbit coupling is thus

$$\begin{aligned} \frac{\Delta\lambda_z}{2} &= i(U - 3J) n_{xz,yz}^{\uparrow\uparrow} \\ \frac{\Delta\lambda_y}{2} &= - (U - 3J) n_{xy,yz}^{\uparrow\downarrow}, \\ \frac{\Delta\lambda_x}{2} &= -i(U - 3J) n_{xz,xy}^{\uparrow\downarrow}, \end{aligned}$$

where $n_{m,\sigma}^{\sigma\sigma'}$ are the off-diagonal elements of the density matrix. The Coulomb-enhanced spin-orbit coupling improves the agreement with the experimental Fermi surface at the degeneracy points (e.g., along the Γ - X direction). The agreement with ARPES data, however, further deteriorates for the γ sheet. This can be seen in Fig. 12, in which the LDA and LDA+DMFT Fermi surface are shown on top of ARPES data from Ref. [21].

Including correlation effects has thus two opposite effects: on the one hand, the agreement with experiments improves (with respect to LDA) for the β sheet; on the other hand, it deteriorates for the γ sheet. This can be seen comparing either panels (a) and (c) or panels (b) and (d) in Fig. 12. Introducing tetragonal Coulomb terms, and in particular the term ΔU , however, reduces the crystal-field enhancement to

$$\varepsilon_{\text{CF}} \rightarrow \varepsilon_{\text{CF}} + \Delta\varepsilon'_{\text{CF}}(0),$$

where $\Delta\varepsilon'_{\text{CF}}(0)$ becomes almost zero for cRPA-based estimates of ΔU . This leads to an almost perfect Fermi surface, as shown in Fig. 13. Non-spherical Coulomb terms turn out to be more important for properties that reflect the point symmetry of the lattice, like the Fermi surface, than for properties that average over orbitals, like the total spectral function [19].

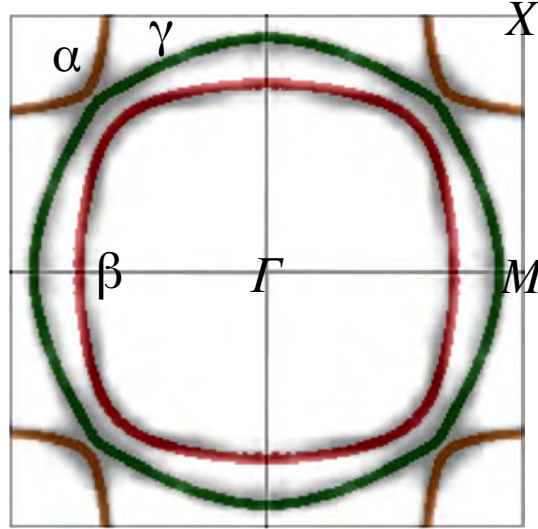


Fig. 13: The LDA+SO+DMFT Fermi surface of Sr_2RuO_4 calculated including the effects of the non-spherical Coulomb term ΔU .

4 Conclusion

The LDA+DMFT approach and its extension has proved very successful for describing correlated materials. It has shown us that materials details do matter, contrarily to what often was assumed in the past; for example a crystal field much smaller than the bandwidth can favor the Mott metal-insulator transition [22]. The method is also becoming very versatile. It is now possible, e.g., to study multi-orbital models including the full Coulomb vertex and spin-orbit interaction, and to calculate response functions or bosonic excitations. Still, many challenges remain. Models with more than a few orbitals or sites and out of equilibrium phenomena remain, e.g., very hard to study. Furthermore, since the birth of DFT, Moore’s law constantly helped as much as new algorithms in extending the frontier; that has come now slowly to a halt. Quantum computers and artificial intelligence, or other new, not yet foreseeable, technological advances might help us in the future. The present deceleration, however, urges us to think to what we should strive for. It is perhaps easier to discuss first what we *do not* need. We do not need a magical calculating machine that gives us, either via an “exact” first principles scheme or via artificial intelligence, answers with no explanations. Reproducing the data is, of course, the aim of theory, but not the only goal. As scientists we need to know why. The danger of giving this question up is evident if we look back at the historical controversy between copernican and ptolemaic system [23]. At the time, the ptolemaic system was the one which agreed better with experiments. Indeed, it had been fine-tuned along the years, via a series of ad hoc assumptions. The copernican system was initially not performing so well, because it was based on circular orbits. But, at the end, it is the copernican system that made a contribution to science, by identifying the correct picture. A theory should reproduce the generic aspects of a phenomenon, but providing at the same time, as Walter Kohn underlined, *understanding*, via a coherent description and the essential details. This should be our aim.

Appendices

A Eigenstates of two-site models

A.1 Hubbard dimer

The Hamiltonian of the Hubbard dimer is given by

$$\hat{H} = \varepsilon_d \sum_{\sigma} \sum_{i=1,2} n_{i\sigma} - t \sum_{\sigma} \left(c_{1\sigma}^{\dagger} c_{2\sigma} + c_{2\sigma}^{\dagger} c_{1\sigma} \right) + U \sum_{i=1,2} \hat{n}_{i\uparrow} \hat{n}_{i\downarrow}.$$

It commutes with the number of electron operator $\hat{N}$, with the total spin $\hat{S}$ and with $\hat{S}_z$. Thus we can express the many-body states in the atomic limit as

$ N, S, S_z\rangle$		N	S	$E(N, S)$	
$ 0, 0, 0\rangle$	$=$	$ 0\rangle$	0	0	
$ 1, 1/2, \sigma\rangle_1$	$=$	$c_{1\sigma}^\dagger 0\rangle$	1	1/2	ε_d
$ 1, 1/2, \sigma\rangle_2$	$=$	$c_{2\sigma}^\dagger 0\rangle$	1	1/2	ε_d
$ 2, 1, 1\rangle$	$=$	$c_{2\uparrow}^\dagger c_{1\uparrow}^\dagger 0\rangle$	2	1	$2\varepsilon_d$
$ 2, 1, -1\rangle$	$=$	$c_{2\downarrow}^\dagger c_{1\downarrow}^\dagger 0\rangle$	2	1	$2\varepsilon_d$
$ 2, 1, 0\rangle$	$=$	$\frac{1}{\sqrt{2}}\left(c_{1\uparrow}^\dagger c_{2\downarrow}^\dagger + c_{1\downarrow}^\dagger c_{2\uparrow}^\dagger\right) 0\rangle$	2	1	$2\varepsilon_d$
$ 2, 0, 0\rangle_0$	$=$	$\frac{1}{\sqrt{2}}\left(c_{1\uparrow}^\dagger c_{2\downarrow}^\dagger - c_{1\downarrow}^\dagger c_{2\uparrow}^\dagger\right) 0\rangle$	2	0	$2\varepsilon_d$
$ 2, 0, 0\rangle_1$	$=$	$c_{1\uparrow}^\dagger c_{1\downarrow}^\dagger 0\rangle$	2	0	$2\varepsilon_d + U$
$ 2, 0, 0\rangle_2$	$=$	$c_{2\uparrow}^\dagger c_{2\downarrow}^\dagger 0\rangle$	2	0	$2\varepsilon_d + U$
$ 3, 1/2, \sigma\rangle_1$	$=$	$c_{1\sigma}^\dagger c_{2\uparrow}^\dagger c_{2\downarrow}^\dagger 0\rangle$	3	1/2	$3\varepsilon_d + U$
$ 3, 1/2, \sigma\rangle_2$	$=$	$c_{2\sigma}^\dagger c_{1\uparrow}^\dagger c_{1\downarrow}^\dagger 0\rangle$	3	1/2	$3\varepsilon_d + U$
$ 4, 0, 0\rangle$	$=$	$c_{1\uparrow}^\dagger c_{1\downarrow}^\dagger c_{2\uparrow}^\dagger c_{2\downarrow}^\dagger 0\rangle$	4	0	$4\varepsilon_d + 2U$

Let us order the $N = 1$ states as in the table above, first the spin up and then spin down block. For finite t the Hamiltonian matrix for $N = 1$ electrons takes then the form

$$\hat{H}_1 = \begin{pmatrix} \varepsilon_d & -t & 0 & 0 \\ -t & \varepsilon_d & 0 & 0 \\ 0 & 0 & \varepsilon_d & -t \\ 0 & 0 & -t & \varepsilon_d \end{pmatrix}.$$

This matrix can be easily diagonalized and yields the *bonding* (−) and *antibonding* (+) states

$ 1, S, S_z\rangle_\alpha$	$E_\alpha(1, S)$	$d_\alpha(1, S)$
$ 1, 1/2, \sigma\rangle_+ = \frac{1}{\sqrt{2}}(1, 1/2, \sigma\rangle_1 - 1, 1/2, \sigma\rangle_2)$	$\varepsilon_d + t$	2
$ 1, 1/2, \sigma\rangle_- = \frac{1}{\sqrt{2}}(1, 1/2, \sigma\rangle_1 + 1, 1/2, \sigma\rangle_2)$	$\varepsilon_d - t$	2

where $d_\alpha(N)$ is the spin degeneracy of the α manifold.

For $N = 2$ electrons (half filling), the hopping integrals only couple the three $S = 0$ states, and therefore the Hamiltonian matrix is given by

$$\hat{H}_2 = \begin{pmatrix} 2\varepsilon_d & 0 & 0 & 0 & 0 & 0 \\ 0 & 2\varepsilon_d & 0 & 0 & 0 & 0 \\ 0 & 0 & 2\varepsilon_d & 0 & 0 & 0 \\ 0 & 0 & 0 & 2\varepsilon_d & -\sqrt{2}t & -\sqrt{2}t \\ 0 & 0 & 0 & -\sqrt{2}t & 2\varepsilon_d + U & 0 \\ 0 & 0 & 0 & -\sqrt{2}t & 0 & 2\varepsilon_d + U \end{pmatrix}.$$

The eigenvalues and the corresponding eigenvectors are

$ 2, S, S_z\rangle_\alpha$	$E_\alpha(2, S)$	$d_\alpha(2, S)$
$ 2, 0, 0\rangle_+ = a_1 2, 0, 0\rangle_0 - \frac{a_2}{\sqrt{2}}(2, 0, 0\rangle_1 + 2, 0, 0\rangle_2)$	$2\varepsilon_d + \frac{1}{2}(U + \Delta(t, U))$	1
$ 2, 0, 0\rangle_o = \frac{1}{\sqrt{2}}(2, 0, 0\rangle_1 - 2, 0, 0\rangle_2)$	$2\varepsilon_d + U$	1
$ 2, 1, m\rangle_o = 2, 1, m\rangle$	$2\varepsilon_d$	3
$ 2, 0, 0\rangle_- = a_2 2, 0, 0\rangle_0 + \frac{a_1}{\sqrt{2}}(2, 0, 0\rangle_1 + 2, 0, 0\rangle_2)$	$2\varepsilon_d + \frac{1}{2}(U - \Delta(t, U))$	1

where

$$\Delta(t, U) = \sqrt{U^2 + 16t^2},$$

and

$$a_1^2 = a_1^2(t, U) = \frac{1}{\Delta(t, U)} \frac{\Delta(t, U) - U}{2} \quad a_2^2 = a_2^2(t, U) = \frac{4t^2}{\Delta(t, U)} \frac{2}{(\Delta(t, U) - U)},$$

so that $a_1 a_2 = 2t/\Delta(t, U)$. For $U = 0$ we have $a_1 = a_2 = 1/\sqrt{2}$, and the two states $|2, 0, 0\rangle_-$ and $|2, 0, 0\rangle_+$ become, respectively, the state with two electrons in the bonding orbital and the state with two electrons in the antibonding orbital; they have energy $E_\pm(2, 0) = 2\varepsilon_d \pm 2t$; the remaining states have energy $2\varepsilon_d$ and are non-bonding. For $t > 0$, the ground state is unique and it is always the singlet $|2, 0, 0\rangle_-$; in the large U limit its energy is

$$E_-(2, 0) \sim 2\varepsilon_d - 4t^2/U.$$

In this limit the energy difference between the first excited state, a triplet state, and the singlet ground state is thus equal to the Heisenberg antiferromagnetic coupling

$$E_o(2, 1) - E_-(2, 0) \sim 4t^2/U = \Gamma.$$

Finally, for $N = 3$ electrons, eigenstates and eigenvectors are

$ 3, S, S_z\rangle_\alpha$	$E_\alpha(3)$	$d_\alpha(3, S)$
$ 3, 1/2, \sigma\rangle_+ = \frac{1}{\sqrt{2}}(1, 1/2, \sigma\rangle_1 + 1, 1/2, \sigma\rangle_2)$	$3\varepsilon_d + U + t$	2
$ 3, 1/2, \sigma\rangle_- = \frac{1}{\sqrt{2}}(1, 1/2, \sigma\rangle_1 - 1, 1/2, \sigma\rangle_2)$	$3\varepsilon_d + U - t$	2

If we exchange holes and electrons, the $N = 3$ case is identical to the $N = 1$ electron case. This is due to the particle-hole symmetry of the model.

A.2 Anderson molecule

The Hamiltonian of the Anderson molecule is given by

$$\hat{H} = \varepsilon_s \sum_{\sigma} \hat{n}_{2\sigma} - t \sum_{\sigma} (c_{1\sigma}^\dagger c_{2\sigma} + c_{2\sigma}^\dagger c_{1\sigma}) + \varepsilon_d \sum_{\sigma} \hat{n}_{1\sigma} + U \hat{n}_{1\uparrow} \hat{n}_{1\downarrow}.$$

In the atomic limit, its eigenstates states can be classified as

$ N, S, S_z\rangle$		N	S	$E(N, S)$
$ 0, 0, 0\rangle =$	$ 0\rangle$	0	0	0
$ 1, 1/2, \sigma\rangle_1 =$	$c_{1\sigma}^\dagger 0\rangle$	1	1/2	ε_d
$ 1, 1/2, \sigma\rangle_2 =$	$c_{2\sigma}^\dagger 0\rangle$	1	1/2	ε_s
$ 2, 1, 1\rangle =$	$c_{2\uparrow}^\dagger c_{1\uparrow}^\dagger 0\rangle$	2	1	$\varepsilon_d + \varepsilon_s$
$ 2, 1, -1\rangle =$	$c_{2\downarrow}^\dagger c_{1\downarrow}^\dagger 0\rangle$	2	1	$\varepsilon_d + \varepsilon_s$
$ 2, 1, 0\rangle =$	$\frac{1}{\sqrt{2}} (c_{1\uparrow}^\dagger c_{2\downarrow}^\dagger + c_{1\downarrow}^\dagger c_{2\uparrow}^\dagger) 0\rangle$	2	1	$\varepsilon_d + \varepsilon_s$
$ 2, 0, 0\rangle_0 =$	$\frac{1}{\sqrt{2}} (c_{1\uparrow}^\dagger c_{2\downarrow}^\dagger - c_{1\downarrow}^\dagger c_{2\uparrow}^\dagger) 0\rangle$	2	0	$\varepsilon_d + \varepsilon_s$
$ 2, 0, 0\rangle_1 =$	$c_{1\uparrow}^\dagger c_{1\downarrow}^\dagger 0\rangle$	2	0	$2\varepsilon_d + U$
$ 2, 0, 0\rangle_2 =$	$c_{2\uparrow}^\dagger c_{2\downarrow}^\dagger 0\rangle$	2	0	$2\varepsilon_s$
$ 3, 1/2, \sigma\rangle_1 =$	$c_{1\sigma}^\dagger c_{2\uparrow}^\dagger c_{2\downarrow}^\dagger 0\rangle$	3	1/2	$\varepsilon_d + 2\varepsilon_s$
$ 3, 1/2, \sigma\rangle_2 =$	$c_{2\sigma}^\dagger c_{1\uparrow}^\dagger c_{1\downarrow}^\dagger 0\rangle$	3	1/2	$2\varepsilon_d + \varepsilon_s + U$
$ 4, 0, 0\rangle =$	$c_{1\uparrow}^\dagger c_{1\downarrow}^\dagger c_{2\uparrow}^\dagger c_{2\downarrow}^\dagger 0\rangle$	4	0	$2\varepsilon_d + 2\varepsilon_s + U$

For $N = 1$ electrons the Hamiltonian can be written in the matrix form

$$\hat{H}_1 = \begin{pmatrix} \varepsilon_d & -t & 0 & 0 \\ -t & \varepsilon_s & 0 & 0 \\ 0 & 0 & \varepsilon_d & -t \\ 0 & 0 & -t & \varepsilon_s \end{pmatrix}.$$

The eigenstates are thus

$ 1, S, S_z\rangle_\alpha$	$E_\alpha(1, S)$	$d_\alpha(1, S)$
$ 1, 1/2, \sigma\rangle_+ = \alpha_1 1, 1/2, \sigma\rangle_1 - \alpha_2 1, 1/2, \sigma\rangle_2$	$\frac{1}{2}(\varepsilon_d + \varepsilon_s + \sqrt{(\varepsilon_d - \varepsilon_s)^2 + 4t^2})$	2
$ 1, 1/2, \sigma\rangle_- = \alpha_2 1, 1/2, \sigma\rangle_1 + \alpha_1 1, 1/2, \sigma\rangle_2$	$\frac{1}{2}(\varepsilon_d + \varepsilon_s - \sqrt{(\varepsilon_d - \varepsilon_s)^2 + 4t^2})$	2

where $d_\alpha(N)$ is the spin degeneracy of the α manifold. For $\varepsilon_s = \varepsilon_d + U/2$ the eigenvalues are

$$E_\pm(1, S) = \varepsilon_d + \frac{1}{4}(U \pm \Delta(U, t)),$$

while the coefficients are $\alpha_1 = a_1(t, U)$ and $\alpha_2 = a_2(t, U)$.

For $N=2$ electrons, the hopping integrals only couple the $S=0$ states. The Hamiltonian is

$$\hat{H}_2 = \begin{pmatrix} \varepsilon_d + \varepsilon_s & 0 & 0 & 0 & 0 & 0 \\ 0 & \varepsilon_d + \varepsilon_s & 0 & 0 & 0 & 0 \\ 0 & 0 & \varepsilon_d + \varepsilon_s & 0 & 0 & 0 \\ 0 & 0 & 0 & \varepsilon_d + \varepsilon_s & -\sqrt{2}t & -\sqrt{2}t \\ 0 & 0 & 0 & -\sqrt{2}t & 2\varepsilon_d + U & 0 \\ 0 & 0 & 0 & -\sqrt{2}t & 0 & 2\varepsilon_s \end{pmatrix}$$

For $\varepsilon_s = \varepsilon_d + U/2$ the eigenvalues and the corresponding eigenvectors are

$ 2, S, S_z\rangle_\alpha$	$E_\alpha(2, S)$	$d_\alpha(2, S)$
$ 2, 0, 0\rangle_+ = b_1 2, 0, 0\rangle_0 - \frac{b_2}{\sqrt{2}}(2, 0, 0\rangle_1 + 2, 0, 0\rangle_2)$	$2\varepsilon_d + \frac{U}{2} + \frac{1}{4}(U + 2\Delta(t, \frac{U}{2}))$	1
$ 2, 0, 0\rangle_o = \frac{1}{\sqrt{2}}(2, 0, 0\rangle_1 - 2, 0, 0\rangle_2)$	$2\varepsilon_d + U$	1
$ 2, 1, m\rangle_o = 2, 1, m\rangle$	$2\varepsilon_d + \frac{U}{2}$	3
$ 2, 0, 0\rangle_- = b_2 2, 0, 0\rangle_0 + \frac{b_1}{\sqrt{2}}(2, 0, 0\rangle_1 + 2, 0, 0\rangle_2)$	$2\varepsilon_d + \frac{U}{2} + \frac{1}{4}(U - 2\Delta(t, \frac{U}{2}))$	1

where $b_1 = a_1(t, U/2)$ and $b_2 = a_2(t, U/2)$. These states have the same form as in the case of the Hubbard dimer; the ground state energy and the weight of doubly occupied states in $|2, 0, 0\rangle_-$ differ, however. Finally, for $N = 3$ electrons, the eigenstates are

$ 3, S, S_z\rangle_\alpha$	$E_\alpha(3, S)$	$d_\alpha(3, S)$
$ 3, 1/2, \sigma\rangle_+ = \alpha_2 1, 1/2, \sigma\rangle_1 + \alpha_1 1, 1/2, \sigma\rangle_2$	$3\varepsilon_d + U + \frac{1}{4}(U + \Delta(t, U))$	2
$ 3, 1/2, \sigma\rangle_- = \alpha_1 1, 1/2, \sigma\rangle_1 - \alpha_2 1, 1/2, \sigma\rangle_2$	$3\varepsilon_d + U + \frac{1}{4}(U - \Delta(t, U))$	2

B Lehmann representation of the local Green function

For a single-orbital model, the local Matsubara Green function for a given site i is defined as

$$G_{i,i}^{\sigma}(i\nu_n) = - \int_0^{\beta} d\tau e^{i\nu_n\tau} \langle \mathcal{T} c_{i\sigma}(\tau) c_{i\sigma}^{\dagger}(0) \rangle,$$

where $\mathcal{T}$ is the time-ordering operator, $\beta = 1/k_B T$, and ν_n a fermionic Matsubara frequency. Let us assume to know all eigenstates $|N_l\rangle$ and their energy $E_l(N)$, for arbitrary number of electrons N . Thus, formally

$$G_{i,i}^{\sigma}(i\nu_n) = - \frac{1}{Z} \sum_{Nl} \int_0^{\beta} d\tau e^{i\nu_n\tau} e^{-\Delta E_l(N)\beta} \langle N_l | c_{i\sigma}(\tau) c_{i\sigma}^{\dagger}(0) | N_l \rangle,$$

where $Z = \sum_{Nl} e^{-\Delta E_l(N)\beta}$ is the partition function, $\Delta E_l(N) = E_l(N) - \mu N$ with μ the chemical potential, and $c_{i\sigma}^{\dagger}(0) = c_{i\sigma}^{\dagger}$. We now insert a complete set of states, obtaining

$$\begin{aligned} G_{i,i}^{\sigma}(i\nu_n) &= - \frac{1}{Z} \sum_{l'N'} \int_0^{\beta} d\tau e^{i\nu_n\tau} e^{-\Delta E_l(N)\beta} \langle N_l | c_{i\sigma}(\tau) | N'_{l'} \rangle \langle N'_{l'} | c_{i\sigma}^{\dagger} | N_l \rangle \\ &= - \frac{1}{Z} \sum_{l'N'} \int_0^{\beta} d\tau e^{-\Delta E_l(N)\beta} e^{(i\nu_n + \Delta E_l(N) - \Delta E_{l'}(N'))\tau} |\langle N'_{l'} | c_{i\sigma}^{\dagger} | N_l \rangle|^2 \\ &= \frac{1}{Z} \sum_{l'N'} \frac{e^{-\Delta E_{l'}(N')\beta} + e^{-\Delta E_l(N)\beta}}{i\nu_n + \Delta E_l(N) - \Delta E_{l'}(N')} |\langle N'_{l'} | c_{i\sigma}^{\dagger} | N_l \rangle|^2. \end{aligned}$$

Due to the weight $|\langle N'_{l'} | c_{i\sigma}^{\dagger}(0) | N_l \rangle|^2$ only the terms for which $N' = N+1$ contribute. Thus, after exchanging the labels $l'N' \leftrightarrow lN$ in the first addend, we obtain the Lehmann representation

$$G_{i,i}^{\sigma}(i\nu_n) = \sum_{l'N} \frac{e^{-\beta\Delta E_l(N)}}{Z} \left(\frac{|\langle (N-1)_{l'} | c_{i\sigma} | N_l \rangle|^2}{i\nu_n - \Delta E_l(N) + \Delta E_{l'}(N-1)} + \frac{|\langle (N+1)_{l'} | c_{i\sigma}^{\dagger} | N_l \rangle|^2}{i\nu_n - \Delta E_{l'}(N+1) + \Delta E_l(N)} \right).$$

Let us consider as example the atomic limit of the Hubbard model at half filling. In this case all sites are decoupled; there are four eigenstates per site, the vacuum $|0\rangle$, with $\Delta E(0) = 0$, the doublet $|1_{\sigma}\rangle = c_{i\sigma}^{\dagger}|0\rangle$, with $\Delta E_{\sigma}(1) = -U/2$, and the doubly-occupied singlet $|2\rangle = c_{i\uparrow}^{\dagger}c_{i\downarrow}^{\dagger}|0\rangle$, with $\Delta E(2) = 0$. Furthermore, $Z = 2(1 + e^{\beta U/2})$ and

$$|\langle (N-1)_{l'} | c_{i\sigma} | N_l \rangle|^2 = \begin{cases} 1 & \text{if } |N_l\rangle = |2\rangle \vee |1_{\sigma}\rangle \\ 0 & \text{otherwise} \end{cases} \quad |\langle (N+1)_{l'} | c_{i\sigma}^{\dagger} | N_l \rangle|^2 = \begin{cases} 1 & \text{if } |N_l\rangle = |0\rangle \vee |1_{-\sigma}\rangle \\ 0 & \text{otherwise} \end{cases}$$

Thus, after summing up the four non-zero contributions, we find

$$G_{i,i}^{\sigma}(\nu_n) = \frac{1}{2} \left(\frac{1}{i\nu_n + U/2} + \frac{1}{i\nu_n - U/2} \right).$$

References

- [1] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964);
W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965)
- [2] W. Kohn, Rev. Mod. Phys. **71**, 1253 (1999)
- [3] P.W. Anderson: *More and different – Notes from a thoughtful curmudgeon*
(World Scientific, Singapore, 2011)
- [4] W. Metzner and D. Vollhardt, Phys. Rev. Lett. **62**, 324 (1989)
- [5] E. Müller-Hartmann, Z. Phys. B **74**, 507 (1989);
Z. Phys. B **76**, 211 (1989); Int. J. Mod. Phys. B **3**, 2169 (1989)
- [6] A. Georges and G. Kotliar, Phys. Rev. B **45**, 6479 (1992)
- [7] M. Jarrell, Phys. Rev. Lett. **69**, 168 (1992)
- [8] E. Pavarini, E. Koch, D. Vollhardt, A. Lichtenstein (Eds.):
The LDA+DMFT approach to strongly-correlated materials,
Reihe Modeling and Simulation, Vol. 1 (Forschungszentrum Jülich, 2011)
<http://www.cond-mat.de/events/correl11>
- [9] E. Pavarini, E. Koch, A. Lichtenstein, D. Vollhardt (Eds.):
DMFT at 25: Infinite Dimensions,
Reihe Modeling and Simulation, Vol. 4 (Forschungszentrum Jülich, 2014)
<http://www.cond-mat.de/events/correl14>
- [10] E. Pavarini, E. Koch, A. Lichtenstein, D. Vollhardt (Eds.):
DMFT: From Infinite Dimensions to Real Materials,
Reihe Modeling and Simulation, Vol. 8 (Forschungszentrum Jülich, 2018)
<http://www.cond-mat.de/events/correl18>
- [11] E. Koch, G. Sangiovanni, and O. Gunnarsson, Phys. Rev. B **78**, 115102 (2008)
- [12] E. Gull, A. J. Millis, A. I. Lichtenstein, A. N. Rubtsov, M. Troyer, and P. Werner,
Rev. Mod. Phys. **83**, 349 (2011)
- [13] A. Flesch, E. Gorelov, E. Koch, E. Pavarini, Phys. Rev. B **87**, 195141 (2013)
- [14] A. Kiani and E. Pavarini, Phys. Rev. B **94**, 075112 (2016)
- [15] E. Pavarini, A. Yamasaki, J. Nuss, and O.K. Andersen,
New. J. Phys **7**, 188 (2005)
- [16] N. Marzari, A.A. Mostofi, J.R. Yates, I. Souza, and D. Vanderbilt,
Rev. Mod. Phys. **84**, 1419 (2012)

-
- [17] G. Zhang and E. Pavarini, Phys. Status Solidi RRL, 1800211 (2018)
 - [18] G. Zhang, E. Gorelov, E. Sarvestani, and E. Pavarini, Phys. Rev. Lett. **116**, 106402 (2016)
 - [19] E. Sarvestani, G. Zhang, E. Gorelov, and E. Pavarini, Phys. Rev. B **97**, 085141 (2018)
 - [20] G. Zhang and E. Pavarini, Phys. Rev. B **99**, 125102 (2019)
 - [21] A. Damascelli, D.H. Lu, K.M. Shen, N.P. Armitage, F. Ronning, D.L. Feng, C. Kim, Z.-X. Shen, T. Kimura, Y. Tokura, Z.Q. Mao, and Y. Maeno, Phys. Rev. Lett. **85**, 5194 (2000)
 - [22] E. Pavarini, S. Biermann, A. Poteryaev, A.I. Lichtenstein, A. Georges, and O.K. Andersen, Phys. Rev. Lett. **92**, 176403 (2004)
 - [23] S. Singh: *Big Bang. The most important scientific discovery of all time and why you need to know about it* (Harper Perennial, London, 2005)

12 Analytical Properties of Self-Energy and Luttinger-Ward Functional

Robert Eder

Institut für Festkörperphysik

Karlsruhe Institute of Technology

Contents

1	Introduction	2
2	Green function, self-energy, and their analytical properties	2
3	Proof of the theorem by Luttinger and Ward	9
3.1	Statement of the theorem	9
3.2	The case $\lambda = 0$	10
3.3	Calculation of $\partial\Omega/\partial\lambda$	11
3.4	Definition and properties of the Luttinger-Ward functional	12
3.5	Calculation of $\partial\tilde{\Omega}/\partial\lambda$	18
4	Conserving approximations	20
5	Conclusion	24
A	A theorem on Fourier transforms	25
B	A theorem on determinants	26

1 Introduction

In 1961 Luttinger and Ward (LW) published a seminal paper [1] which became the foundation of many important developments in the quantum theory of many particle systems. In particular, they gave an explicit expression for the grand canonical potential Ω of an interacting Fermion system. A key step thereby was the construction of the Luttinger-Ward functional, a functional of the Green function which essentially describes the deviation of Ω from a non-interacting system. The expression for Ω became the basis for the derivation of the famous Luttinger theorem [2], which states that interactions between electrons do not change the volume of the Fermi surface. Subsequently, Baym and Kadanoff [3] investigated the question, under which conditions approximate response functions for systems of interacting particles comply with certain conservation laws, i.e., what is the criterion for the construction of conserving approximations. Baym showed [4] that the Luttinger-Ward functional thereby plays a key role in that a self-energy derived from an approximate Luttinger-Ward functional always gives rise to a conserving approximation. It is the purpose of the present notes to give an introduction to Green functions and the self-energy, derive the LW expression for Ω and briefly discuss the ideas of Baym regarding conserving approximations.

2 Green function, self-energy, and their analytical properties

In this section we discuss Green functions, their properties and use. Thereby we will also refer to the representation of Green functions in terms of Feynman diagrams but we will not give a derivation of these. Excellent introductions to this subject can be found in various textbooks [5–7]. In the present notes we try to be consistent with Fetter-Walecka (FW) [6].

We consider a system of interacting fermions and assume that there is some complete basis of single-electron states. Each state is labeled by a set of quantum numbers, α , we denote the number of different sets α as n_α . Introducing Fermionic creation/annihilation operators $c_\alpha^\dagger/c_\alpha$ for electrons in these states, the Hamiltonian, assumed to be time-independent, can be written as $H = H_0 + H_1$ with

$$H_0 = \sum_{\alpha,\beta} t_{\alpha,\beta} c_\alpha^\dagger c_\beta, \quad (1)$$

$$H_1 = \frac{1}{2} \sum_{\alpha,\beta,\gamma,\delta} V_{\alpha,\beta,\delta,\gamma} c_\alpha^\dagger c_\beta^\dagger c_\gamma c_\delta. \quad (2)$$

Note the factor of $1/2$ and the ‘inverted’ order of indices on the interaction matrix element V in (2) which follows from the prescription for second quantization [5–7]. H commutes with the operator $\hat{N}$ of particle number, which means that eigenstates of H have a fixed particle number. In all that follows we consider a grand canonical ensemble with inverse temperature $\beta = 1/k_B T$ and chemical potential μ . Introducing $K = H - \mu N$ the thermal average of any operator $\hat{O}$ is

$$\langle \hat{O} \rangle_{th} = \frac{1}{Z} \text{Tr} \left(e^{-\beta K} \hat{O} \right)$$

with the grand partition function

$$Z = \text{Tr} (e^{-\beta K}). \quad (3)$$

For any operator $\hat{O}$ the imaginary time Heisenberg operator is $\hat{O}(\tau) = e^{\tau K/\hbar} \hat{O} e^{-\tau K/\hbar}$. Now let $\hat{A}$ and $\hat{B}$ be any two operators and $n_B = [\hat{N}, \hat{B}]$, $n_A = [\hat{N}, \hat{A}]$. The *imaginary time Green function* then is $G_{A,B}(\tau, \tau') = -\langle T[\hat{A}(\tau)\hat{B}(\tau')] \rangle_{th}$ where T is the time-ordering operator. It is easy to see that $G_{A,B}$ is a function of $\tau - \tau'$ only and setting $\tau' = 0$ we find the more explicit expression (with $\xi_B = (-1)^{n_B}$)

$$G_{A,B}(\tau) = -\Theta(\tau) \langle \hat{A}(\tau) \hat{B} \rangle_{th} - \xi_B \Theta(-\tau) \langle \hat{B} \hat{A}(\tau) \rangle_{th} \quad (4)$$

$$= \frac{1}{Z} \left(-\Theta(\tau) \sum_{i,j} e^{-\beta K_i} e^{\frac{\tau}{\hbar}(K_i - K_j)} \langle i | \hat{A} | j \rangle \langle j | \hat{B} | i \rangle \right. \\ \left. - \xi_B \Theta(-\tau) \sum_{i,j} e^{-\beta K_i} e^{\frac{\tau}{\hbar}(K_j - K_i)} \langle i | \hat{B} | j \rangle \langle j | \hat{A} | i \rangle \right), \quad (5)$$

Here $|i\rangle$ are the exact eigenstates of H and $K_i = E_i - \mu N_i$ the corresponding eigenvalues of K with energy E_i and particle number N_i . It is obvious from (5) that G can be different from zero only if $n_A = -n_B$. The τ -dependence of both terms in (5) is $e^{(-\beta + \frac{|\tau|}{\hbar})K_i} e^{-\frac{|\tau|}{\hbar}K_j}$. Since the K_i are bounded from below (namely by the K for the ground state with the given μ) but unbounded from above in the thermodynamical limit, G is well-defined only for $\tau \in [-\beta\hbar, \beta\hbar]$ [8]. Using the cyclic property of the trace one can show that for $\tau \in [-\beta\hbar, 0]$ one has $G(\tau + \beta\hbar) = \xi_B G(\tau)$. Accordingly, G can be expanded in a Fourier series

$$G(\tau) = \frac{1}{\beta\hbar} \sum_{\nu=-\infty}^{\infty} e^{-i\omega_\nu \tau} G(i\omega_\nu), \\ G(i\omega_\nu) = \int_0^{\beta\hbar} d\tau e^{i\omega_\nu \tau} G(\tau), \quad (6)$$

with $\omega_\nu = \frac{\nu\pi}{\beta\hbar}$ and integer ν . The ω_ν are called Matsubara frequencies and for even n_B (odd n_B) only even ν (odd ν) contribute in the Fourier expansion. Using (5) one finds

$$G_{A,B}(i\omega_\nu) = \frac{1}{Z} \sum_{i,j} \frac{e^{-\beta K_i} - \xi_B e^{-\beta K_j}}{i\omega_\nu + \frac{1}{\hbar}(K_i - K_j)} \langle i | \hat{A} | j \rangle \langle j | \hat{B} | i \rangle. \quad (7)$$

Next, we consider the retarded real-time Green function. The real-time Heisenberg operator is $\hat{O}(\tau) = e^{itK/\hbar} \hat{O} e^{-itK/\hbar}$ and the retarded real-time Green function is

$$G_{A,B}^R(t) = -i\Theta(t) \left(\langle \hat{A}(t) \hat{B} \rangle_{th} - \xi_B \langle \hat{B} \hat{A}(t) \rangle_{th} \right). \quad (8)$$

It is straightforward to write down the expression of $G_{A,B}^R(t)$ corresponding to (5), and using the theorem in Appendix A one finds its Fourier transform

$$G_{A,B}^R(\omega) = \frac{1}{Z} \lim_{\eta \rightarrow 0^+} \sum_{i,j} \frac{e^{-\beta K_i} - \xi_B e^{-\beta K_j}}{\omega + i\eta + \frac{1}{\hbar}(K_i - K_j)} \langle i | \hat{A} | j \rangle \langle j | \hat{B} | i \rangle. \quad (9)$$

Comparing with (7) it is obvious that $G_{A,B}^R(\omega)$ can be obtained from $G_{A,B}(i\omega_\nu)$ by replacing $i\omega_\nu \rightarrow \omega + i0^+$. In other words, there is one function $G_{A,B}(z)$ of the complex variable z – usually called ‘the’ Green function – which gives $G_{A,B}(i\omega_\nu)$ when it is evaluated for the Matsubara frequencies, and $G_{A,B}^R(\omega)$ when it is evaluated along a line infinitesimally above the real axis. Equation (7) is the *Lehmann representation* of the Green function. The existence of ‘the’ Green function is the very reason why the imaginary time Green function is so useful. In principle, the quantity of physical interest is the real-time Green function. For example one can show that if a system is acted upon by a time dependent perturbation of the form $H_p = f(t)\hat{B}$ (with $\hat{B}$ some Hermitean operator and $f(t)$ a real function), the change of the expectation value of some operator $\hat{A}$ which is linear in $f(t)$ is given by (see FW (32.2))

$$\delta \langle \hat{A} \rangle(t) = \frac{1}{\hbar} \int_{-\infty}^{\infty} dt' G_{AB}^R(t-t') f(t').$$

On the other hand, the imaginary-time Green function can be evaluated approximately by using the powerful technique of expansion in Feynman diagrams [5–7], which is *not* possible for the real-time Green functions. The standard way to obtain the real-time Green function, which is used over and over again in the literature, is to first obtain an approximate $G_{A,B}(i\omega_\nu)$ by doing an expansion in Feynman diagrams and then continue this analytically to a line infinitesimally above the real axis to obtain the real-time Green function of physical interest.

We now specialize to the *single-particle Green function* which corresponds to $\hat{A} = c_\alpha$, $\hat{B} = c_\beta^\dagger$ (so that $\xi_B = -1$). It may be viewed as a matrix of dimension $n_\alpha \times n_\alpha$, denoted by $\mathbf{G}$, and we can write it

$$G_{\alpha,\beta}(z) = \frac{1}{Z} \sum_{i,j} e^{-\beta K_i} \left[\frac{\langle i|c_\beta^\dagger|j\rangle \langle j|c_\alpha|i\rangle}{z - (K_i - K_j)/\hbar} + \frac{\langle i|c_\alpha|j\rangle \langle j|c_\beta^\dagger|i\rangle}{z - (K_j - K_i)/\hbar} \right], \quad (10)$$

$$= \int_{-\infty}^{\infty} d\omega \frac{\rho_{\alpha,\beta}^{(-)}(\omega)}{z - \omega} + \int_{-\infty}^{\infty} d\omega \frac{\rho_{\alpha,\beta}^{(+)}(\omega)}{z - \omega} \quad (11)$$

$$\rho_{\alpha,\beta}^{(-)}(\omega) = \frac{1}{Z} \sum_{i,j} e^{-\beta K_i} \langle i|c_\beta^\dagger|j\rangle \langle j|c_\alpha|i\rangle \delta(\omega - (K_i - K_j)/\hbar), \quad (12)$$

$$\rho_{\alpha,\beta}^{(+)}(\omega) = \frac{1}{Z} \sum_{i,j} e^{-\beta K_i} \langle i|c_\alpha|j\rangle \langle j|c_\beta^\dagger|i\rangle \delta(\omega - (K_j - K_i)/\hbar). \quad (13)$$

Since $K_j - K_i$ is real, $\mathbf{G}(z)$ has a number of poles on the real axis. For both, $\rho^{(-)}$ and $\rho^{(+)}$, one has $\rho_{\alpha,\beta}^*(\omega) = \rho_{\beta,\alpha}(\omega)$ i.e., the ω -dependent matrix $\boldsymbol{\rho}(\omega)$ is Hermitean. Moreover, for any vector $\mathbf{v}$ of length n_α we have

$$\sum_{\alpha,\beta} v_\alpha^* \rho_{\alpha,\beta}^{(-)}(\omega) v_\beta = \frac{1}{Z} \sum_{i,j} e^{-\beta K_i} |\langle j|c_\mathbf{v}|i\rangle|^2 \delta(\omega - (K_i - K_j)/\hbar) > 0$$

with $c_\mathbf{v} = \sum v_\alpha c_\alpha$ so that $\boldsymbol{\rho}$ are positive definite. It follows from the Hermiticity of $\boldsymbol{\rho}$ that $[\mathbf{G}(z)]^+ = \mathbf{G}(z^*)$ which shows that, for complex z , $\mathbf{G}(z)$ is *not* Hermitean.

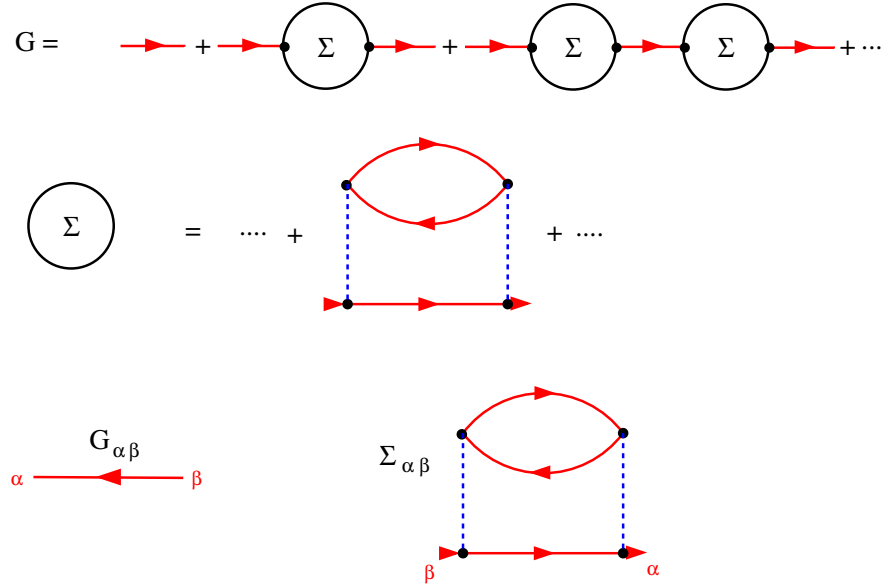


Fig. 1: Top: Graphical representation of the Dyson equation. Middle: Self-energy diagrams have two open ends. Bottom: The convention for the representation of the Green function implies the labeling of the open ends of the self-energy.

We now assume that the range of ω where the elements of $\rho^{(\pm)}$ are different from zero is finite, which means that the change in energy upon adding or removing an electron is bounded. Then we consider the limit $|z| \rightarrow \infty$ and expand

$$\frac{1}{z \pm (K_j - K_i)/\hbar} \rightarrow \frac{1}{z} \mp \frac{K_j - K_i}{\hbar z^2} + O\left(\frac{1}{z^3}\right).$$

Inserting this into (10) and using $(K_j - K_i)\langle j|c_\alpha|i\rangle = \langle j|[K, c_\alpha]|i\rangle$ we find

$$G_{\alpha,\beta}(z) \rightarrow \frac{\delta_{\alpha,\beta}}{z} + \frac{\left\langle \left\{ c_\beta^\dagger, [c_\alpha, K] \right\} \right\rangle_{th}}{\hbar z^2} + O\left(\frac{1}{z^3}\right).$$

Using the Hamiltonian (1) and (2) one finds

$$\left\langle \left\{ c_\beta^\dagger, [c_\alpha, K] \right\} \right\rangle_{th} = t_{\alpha,\beta} - \mu \delta_{\alpha,\beta} + \sum_{\gamma,\delta} (V_{\alpha,\gamma,\beta,\delta} - V_{\alpha,\gamma,\delta,\beta}) \langle c_\gamma^\dagger c_\delta \rangle_{th}. \quad (14)$$

The term involving V ‘looks like’ the Hartree-Fock potential $V_{\alpha,\beta}^{(HF)}$, however, whereas for the true Hartree-Fock potential the thermal average has to be taken using the Hartree-Fock wave functions and energies, the thermal average in (14) has to be taken using the fully interacting eigenstates and energies. Keeping this subtle difference in mind we still call the third term the Hartree-Fock potential $V_{\alpha,\beta}^{(HF)}$ so that

$$\mathbf{G}(z) \rightarrow \frac{1}{z} + \frac{\mathbf{t} - \mu + \mathbf{V}^{(HF)}}{\hbar z^2} + O\left(\frac{1}{z^3}\right). \quad (15)$$

As already mentioned the imaginary-time Green function can be expanded in Feynman diagrams and the self-energy $\Sigma(i\omega_\nu)$ be introduced in the standard way see Figure 1. The self-energy can be expanded in diagrams which have two ‘entry points’ an incoming and an outgoing one. Following FW [6] we represent the Green function $G_{\alpha\beta}$ and also the noninteracting Green function $G_{\alpha\beta}^0$ by a directed line with an arrow running $\beta \rightarrow \alpha$ (the reason is, that it is the *creation operator* which has the index β , see (10)). In the Dyson equation the matrix indices of the Green function and the self-energy must take the form of consecutive matrix products, e.g., $G_{\delta\alpha}^0 \Sigma_{\alpha\beta} G_{\beta\gamma}^0$ – otherwise the summation of the geometric series would not be possible. It follows that the element $\Sigma_{\alpha\beta}$ must have the label α on the outgoing entry and the label β on the incoming one, see Figure 1. The diagrammatic expansion shows [5–7] that the Green function obeys the Dyson equation

$$\begin{aligned} \left(i\omega_\nu - \frac{1}{\hbar} (\mathbf{t} - \mu) - \Sigma(i\omega_\nu) \right) \mathbf{G}(i\omega_\nu) &= 1 \\ \left(-\partial_\tau - \frac{1}{\hbar} (\mathbf{t} - \mu) \right) \mathbf{G}(\tau) - \int_0^{\beta\hbar} \Sigma(\tau - \tau') \mathbf{G}(\tau') d\tau' &= \delta(\tau), \end{aligned} \quad (16)$$

where the second equation is the Fourier-transform of the first and FW (25.21) was used. The inverse of the Green function thus is

$$\mathbf{G}^{-1}(z) = z - \frac{1}{\hbar} (\mathbf{t} - \mu) - \Sigma(z)$$

On the other hand, from (15) we find

$$\mathbf{G}^{-1}(z) \rightarrow z - \frac{\mathbf{t} - \mu}{\hbar} - \frac{\mathbf{V}^{(HF)}}{\hbar} + O\left(\frac{1}{z}\right) \Rightarrow \Sigma(z) \rightarrow \frac{\mathbf{V}^{(HF)}}{\hbar} + O\left(\frac{1}{z}\right).$$

Accordingly, the quantity $\bar{\Sigma} = \Sigma - \mathbf{V}^{(HF)}$ vanishes as $1/z$ for large $|z|$.

Next, let $\mathbf{v}$ be any complex vector of length n_α and consider

$$f(z) = \sum_{\alpha,\beta} v_\alpha^* G_{\alpha,\beta}(z) v_\beta = \frac{1}{Z} \sum_{i,j} \frac{e^{-\beta K_i} + e^{-\beta K_j}}{z + (K_i - K_j)/\hbar} |\langle j | c_{\mathbf{v}}^\dagger | i \rangle|^2, \quad (17)$$

where $c_{\mathbf{v}}^\dagger = \sum v_\alpha c_\alpha^\dagger$ and we have used (7). Putting $z = x + iy$ the imaginary part reads

$$\mathcal{I}f(z) = -\frac{y}{Z} \sum_{i,j} \frac{e^{-\beta K_i} + e^{-\beta K_j}}{(x + (K_i - K_j)/\hbar)^2 + y^2} |\langle j | c_{\mathbf{v}}^\dagger | i \rangle|^2.$$

This expression cannot be zero unless $y = 0$. It follows that for z away from the real axis all eigenvalues of $\mathbf{G}(z)$ are different from zero [9], otherwise we might choose $\mathbf{v}$ to be either the right- or left-hand eigenvector for eigenvalue 0 and obtain $f(z) = 0$ (note that for complex z $\mathbf{G}(z)$ is not Hermitean which means that right- and left-hand eigenvectors will be different). In fact, this is the very condition that $\mathbf{G}^{-1}(z)$ does exist and it follows that for all complex z the determinant of $\mathbf{G}(z)$ is different from zero. Using Cramer’s rule we find the elements of the inverse Green function

$$G_{\alpha,\beta}^{-1}(z) = \frac{(-1)^{\alpha+\beta} \det \mathbf{M}_{\alpha,\beta}(z)}{\det \mathbf{G}(z)},$$

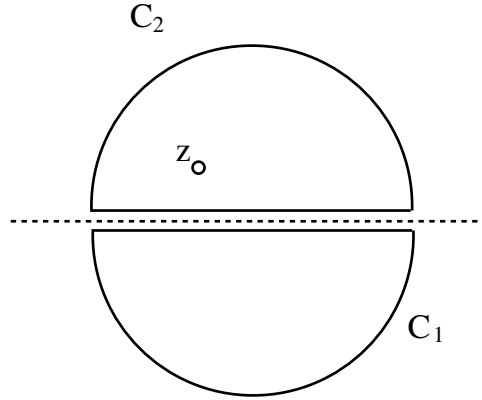


Fig. 2: Integration contours for the proof of the spectral representation of Σ . The dashed line is the real z' axis.

where $M_{\alpha,\beta}(z)$ is the respective minor of $G(z)$, i.e., the matrix $G(z)$ with row α and column β discarded. Since for z away from the real axis all elements of $G(z)$ are finite, see Eq. 10, and the determinant of $G(z)$ different from zero it follows that away from the real axis all elements of $G^{-1}(z)$ are analytical functions of z . Starting from $G^{-1}(z)G(z) = 1$, taking the Hermitean conjugate and using $[G(z)]^+ = G(z^*)$ we find $[G^{-1}(z)]^+ = G^{-1}(z^*)$. Since both, $(G^0)^{-1}(z) = z - (t - \mu)/\hbar$ and V^{HF} , obey this relation, it follows that $\bar{\Sigma}(z)$ alone obeys this relation as well, i.e., $\bar{\Sigma}(z^*) = \bar{\Sigma}^+(z)$. For real ω , we define the real matrices $K(\omega)$ and $J(\omega)$ by

$$\bar{\Sigma}(\omega + i0^+) = K(\omega) + iJ(\omega) \Rightarrow \bar{\Sigma}(\omega - i0^+) = K^T(\omega) - iJ^T(\omega). \quad (18)$$

Next, we define

$$\begin{aligned} \bar{\Sigma}^{(+)}(z) &= \frac{1}{2} (\bar{\Sigma}(z) + \bar{\Sigma}^T(z)) \Rightarrow \bar{\Sigma}^{(+)}(\omega \pm i0^+) = K^{(+)}(\omega) \pm iJ^{(+)}(\omega), \\ \bar{\Sigma}^{(-)}(z) &= \frac{i}{2} (\bar{\Sigma}(z) - \bar{\Sigma}^T(z)) \Rightarrow \bar{\Sigma}^{(-)}(\omega \pm i0^+) = -J^{(-)}(\omega) \pm iK^{(-)}(\omega), \end{aligned} \quad (19)$$

where $K^{(\pm)} = \frac{1}{2}(K \pm K^T)$ and $J^{(\pm)} = \frac{1}{2}(J \pm J^T)$. Now consider the integration contours in Figure 2 which consist of lines infinitesimally above and below the real axis and semicircles at infinity. Since $\bar{\Sigma}$ is analytic away from the real axis we have for z in the upper half-plane

$$\oint_{C_1} dz' \frac{\bar{\Sigma}^{(\pm)}(z')}{z' - z} = 0 \Rightarrow \int_{-\infty}^{\infty} d\omega \frac{K^{(\pm)}(\omega)}{\omega - z} = i \int_{-\infty}^{\infty} d\omega \frac{J^{(\pm)}(\omega)}{\omega - z}.$$

The second equation follows, because the integrand is $\propto 1/z'^2$ for large $|z'|$ so that the arc does not contribute and infinitesimally below the real axis we can use (18). Next we can use Cauchy's theorem and write

$$\bar{\Sigma}^{(\pm)}(z) = \frac{1}{2\pi i} \oint_{C_2} dz' \frac{\bar{\Sigma}^{(\pm)}(z')}{z' - z} \Rightarrow \bar{\Sigma}^{(\pm)}(z) = \pm \frac{1}{\pi} \int_{-\infty}^{\infty} d\omega \frac{J^{(\pm)}(\omega)}{\omega - z}.$$

Reverting (19) and recalling the definition of $\bar{\Sigma}$ we finally arrive at the spectral representation of the self-energy, as derived by Luttinger [10]

$$\Sigma(z) = \mathbf{V}^{(HF)} + \frac{1}{\pi} \int_{-\infty}^{\infty} d\omega \frac{\boldsymbol{\sigma}(\omega)}{\omega - z} \quad (20)$$

with $\boldsymbol{\sigma}(\omega) = \mathbf{J}^{(+)}(\omega) + i\mathbf{J}^{(-)}(\omega)$, which shows that $\boldsymbol{\sigma}(\omega)$ is Hermitean. It should be noted that in deriving (20) we have made use only of the Dyson equation and certain rigorous analytical properties of the Green function. This is therefore a completely rigorous result. It should also be noted that we have shown that for any z off the real axis $\mathbf{G}^{-1}(z)$ is well-defined. This shows that there is a unique mapping $\mathbf{G}(z) \rightarrow \Sigma(z)$.

We have introduced the self-energy using the diagrammatic expansion of the Green function but it can also be defined in another way. Using $-\hbar\partial_{\tau}\hat{O} = [\hat{O}, K]$ and $\partial_{\tau}\Theta(\pm\tau) = \pm\delta(\tau)$ we find the equation of motion of the imaginary time Green function:

$$-\hbar\partial_{\tau}G_{\alpha,\beta}(\tau) = \hbar\delta_{\alpha,\beta}\delta(\tau) + \sum_{\nu} (t_{\alpha,\nu} - \mu\delta_{\alpha,\nu}) G_{\nu,\beta}(\tau) - \sum_{\nu,\kappa,\lambda} V_{\alpha,\nu,\kappa,\lambda} \langle T[(c_{\nu}^{\dagger}c_{\lambda}c_{\kappa})(\tau)c_{\beta}^{\dagger}(0)] \rangle_{th}$$

The time-ordered product in the last term can be written as $G_2(\lambda\tau, \kappa\tau, \beta 0, \nu\tau^+)$ where

$$G_2(\lambda\tau_1, \kappa\tau_2, \beta\tau_3, \nu\tau_4) = (-1)^2 \langle T[c_{\lambda}(\tau_1)c_{\kappa}(\tau_2)c_{\nu}^{\dagger}(\tau_4)c_{\beta}^{\dagger}(\tau_3)] \rangle_{th}$$

is the two-particle imaginary time Green function (the factor of $(-1)^2$ has to be replaced by $(-i)^2$ for the real-time Green function—it is always the square of the prefactor of the single-particle Green function). Comparing with the Dyson equation (16) it is obvious that G_2 and Σ are related as

$$-\sum_{\nu,\kappa,\lambda} V_{\alpha,\nu,\kappa,\lambda} G_2(\lambda\tau, \kappa\tau, \beta 0, \nu\tau^+) = \hbar \int_0^{\hbar\beta} d\tau' \Sigma_{\alpha,\gamma}(\tau-\tau') G_{\gamma,\beta}(\tau'). \quad (21)$$

Frequently an approximate Green function is found by expressing G_2 as a functional of G . For example, the Hartree-Fock approximation corresponds to replacing

$$G_2(\lambda\tau_1, \kappa\tau_2, \beta\tau_3, \nu\tau_4) \rightarrow G(\lambda\tau_1, \beta\tau_3) G(\kappa\tau_2, \nu\tau_4) - G(\lambda\tau_1, \nu\tau_4) G(\kappa\tau_2, \beta\tau_3).$$

Inserting this into (21) one obtains

$$\begin{aligned} -\sum_{\nu,\kappa,\lambda} V_{\alpha,\nu,\kappa,\lambda} G_2(\lambda\tau, \kappa\tau, \beta 0, \nu\tau^+) &= \sum_{\nu,\kappa,\lambda} (V_{\alpha,\nu,\lambda,\kappa} - V_{\alpha,\nu,\kappa,\lambda}) \langle c_{\nu}^{\dagger}c_{\kappa} \rangle G_{\lambda,\beta}(\tau) \\ \Rightarrow \Sigma_{\alpha,\nu}(\tau) &= \frac{1}{\hbar} V_{\alpha,\nu}^{(HF)} \delta(\tau) \end{aligned}$$

which is what one would have expected for the self-energy in Hartree-Fock approximation.

3 Proof of the theorem by Luttinger and Ward

3.1 Statement of the theorem

In the present section we consider a solid, described by an LCAO-like Hamiltonian. Then we have $\alpha = (i, n, \nu, \sigma)$ where $i \in \{1, \dots, N\}$ denotes the unit cell, $n \in \{1, \dots, n_{Atom}\}$ the number of atoms in the basis, $\nu \in \{s, p_x, p_y, p_z, d_{xy} \dots\}$ the type of orbital and σ the z -component of spin. The number of all orbitals in the unit cell is n_{orb} . The Fourier transform of the Fermion operators is

$$c_{\mathbf{k},\beta}^\dagger = \frac{1}{\sqrt{N}} \sum_i e^{i\mathbf{k} \cdot (\mathbf{R}_i + \mathbf{r}_n)} c_{i,n,\nu,\sigma}^\dagger,$$

where we have introduced the orbital index $\beta = (n, \nu, \sigma)$. Since this second compound index always comes together with either a momentum $\mathbf{k}$ or a cell index i , no misunderstanding is possible. The Hamiltonian now can be written as

$$H_0 = \sum_{\mathbf{k}} \sum_{\alpha,\beta} t_{\alpha,\beta}(\mathbf{k}) c_{\mathbf{k},\alpha}^\dagger c_{\mathbf{k},\beta}, \quad (22)$$

$$H_1 = \frac{1}{2} \sum_{\mathbf{k}, \mathbf{k}', \mathbf{q}} \sum_{\alpha,\beta,\gamma,\delta} V_{\alpha,\beta,\delta,\gamma}(\mathbf{k}, \mathbf{k}', \mathbf{q}) c_{\mathbf{k}+\mathbf{q},\alpha}^\dagger c_{\mathbf{k}'-\mathbf{q},\beta}^\dagger c_{\mathbf{k}',\gamma} c_{\mathbf{k},\delta}. \quad (23)$$

Equation (22) defines the $2n_{orb} \times 2n_{orb}$ matrix $\mathbf{t}(\mathbf{k})$, whose eigenvalues $E_n(\mathbf{k})$ give the noninteracting band structure. This formulation allows H_0 to describe magnetic systems or include spin-orbit coupling.

The grand canonical potential $\Omega(T, \mu)$ contains all thermodynamical information about a system at fixed temperature T and chemical potential μ . It is defined as the logarithm of the grand partition function

$$\Omega = -\frac{1}{\beta} \ln(Z) \quad \text{with} \quad Z = \sum_i e^{-\beta(E_i - \mu N_i)},$$

where the sum is over all eigenstates of the system with energy E_i and particle number N_i . The latter can indeed be evaluated for noninteracting particles and in this way one obtains for example the grand canonical potential of noninteracting Bloch electrons

$$\Omega = -\frac{1}{\beta} \sum_{n=1}^{2n_{orb}} \sum_{\mathbf{k}} \ln(1 + e^{-\beta(E_n(\mathbf{k}) - \mu)}). \quad (24)$$

As shown in textbooks of statistical mechanics, expression (24) allows to derive the complete thermodynamics of metals. However, it is in general not possible to evaluate the grand partition function for a system of *interacting particles* of macroscopic size.

Luttinger and Ward, however, derived a relation for the grand canonical potential of interacting Fermions [1]. More precisely, they considered the following quantity

$$\Omega' = -\lim_{\eta \rightarrow 0^+} \frac{1}{\beta} \sum_{\mathbf{k}, \nu} e^{i\omega_\nu \eta} \left(\ln \det(-\mathbf{G}^{-1}(\mathbf{k}, i\omega_\nu)) + \text{Tr}(\mathbf{G}(\mathbf{k}, i\omega_\nu) \Sigma(\mathbf{k}, i\omega_\nu)) \right) + \Phi[\mathbf{G}]. \quad (25)$$

Here $\sum_\nu$ denotes summation over the Fermionic Matsubara frequencies and $\Phi[\mathbf{G}]$ is the so-called Luttinger-Ward functional which is defined as a sum over closed, linked Feynman-diagrams (the precise definition will be discussed below). The important point here is that a closed Feynman diagram is simply a number, so that $\Phi[\mathbf{G}]$ indeed assigns a (real) number to each possible Green function $\mathbf{G}$. As regards the logarithm of the determinant in (25) we recall that the determinant of a matrix is given by the product of its eigenvalues (the matrix need not be Hermitean for this to be true) so that the logarithm of the determinant is the sum of the logarithms of the (complex) eigenvalues of $-\mathbf{G}^{-1}$.

In the following we want to show that in fact $\Omega' = \Omega$, the true grand canonical potential, and thereby follow the original proof by Luttinger and Ward. The basic idea is to multiply the interaction part of the Hamiltonian, (2), by a scale factor, $H_1 \rightarrow \lambda H_1$, then show $\Omega' = \Omega$ for $\lambda = 0$, i.e., the noninteracting limit, and next show that $\partial_\lambda \Omega' = \partial_\lambda \Omega$. Obviously, this proves the identity of the two expressions for any λ .

3.2 The case $\lambda = 0$

In this limit $\Sigma = 0$ and $\Phi[\mathbf{G}] = 0$ (the latter property follows because all interaction lines in all diagrams are zero) so that only the first term in (25) remains and

$$\begin{aligned} \mathbf{G}^{-1}(\mathbf{k}, \omega) &= \omega + (\mu - \mathbf{t}(\mathbf{k}))/\hbar, \\ \ln \det(-\mathbf{G}^{-1}(\mathbf{k}, \omega)) &= \sum_{n=1}^{2n_{orb}} \ln \left(-\omega - (\mu - E_n(\mathbf{k}))/\hbar \right). \end{aligned} \quad (26)$$

We now replace the sum over Matsubara frequencies by a contour integration, which is a standard trick used in field theory (see, e.g., section 25 of FW) and obtain

$$-\frac{1}{\beta} \sum_\nu e^{i\omega_\nu \eta} \ln \det(-\mathbf{G}^{-1}(\mathbf{k}, i\omega_\nu)) = \frac{\hbar}{2\pi i} \oint_{\mathcal{C}} d\omega f(\omega) e^{\omega \eta} \ln \det(-\mathbf{G}^{-1}(\mathbf{k}, \omega)) \quad (27)$$

where

$$f(\omega) = \frac{1}{e^{\beta \hbar \omega} + 1},$$

is the Fermi function and the contour $\mathcal{C}$ encircles the imaginary axis in counterclockwise fashion, see Figure 3(a). Next we note that the integrals along the two clover-shaped contours in Figure 3(b) are zero, *provided* the integrand is analytic in the interior of the two curves. Since the Fermi function has all of its poles along the imaginary axis, which is outside of the curves in Figure 3(b), we only need to consider possible singularities of $\ln \det(-\mathbf{G}^{-1}(\mathbf{k}, \omega))$. In principle, the complex logarithm has a branch-cut along the negative real axis which could be problematic. However, a quick glance at (26) shows, that as long as ω has a nonvanishing imaginary part, the argument of the logarithm can never be purely real. Singularities of the logarithm thus occur only on the real axis, which also is exterior to the contours in Figure 3(b). The integral along the contours in Figure 3(b) therefore is indeed zero. Next, Jordan's lemma

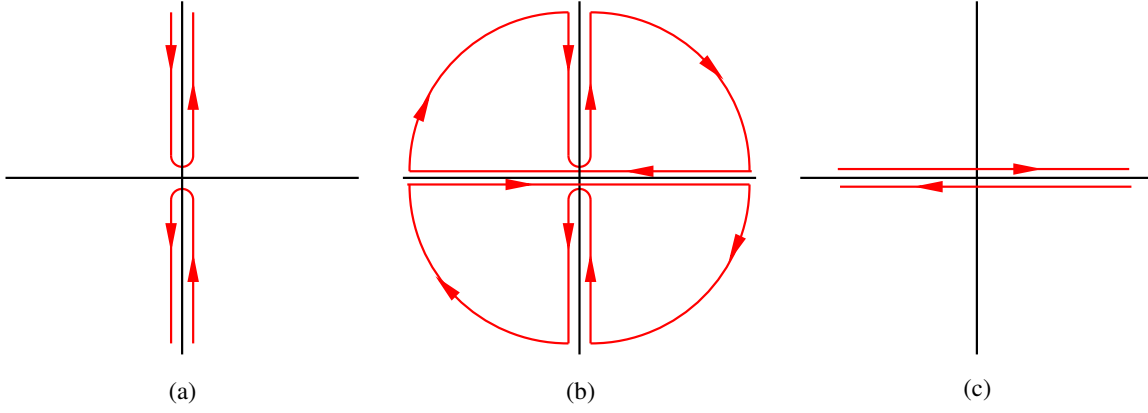


Fig. 3: (a) Integration contour $\mathcal{C}$ used in (27). Since the integrals along the two contours in (b) are zero and the contributions from the circular arcs vanish, the integral along the contour in (a) is equal to that over the contour $\mathcal{C}'$ in (c).

can be invoked to establish that the integral along the large semicircles vanishes. Here the Fermi function $f(\omega)$ guarantees that the contribution from the semicircle with $\text{Re}(\omega) > 0$ vanishes, whereas the factor $e^{\omega\eta}$ does the same for the semicircle with $\text{Re}(\omega) < 0$. It follows that the integral along the contour $\mathcal{C}$ in Figure 3a is equal to that along the contour $\mathcal{C}'$ in Figure 3c (note the inverted direction of the curves in Figure 3c as compared to Figure 3b). Next, we insert

$$f(\omega) = -\frac{1}{\beta\hbar} \frac{d}{d\omega} \ln(1 + e^{-\beta\hbar\omega}) \quad (28)$$

and integrate by parts. Thereby the Fermi function and the factor $e^{\eta\omega}$ again make sure that the contributions from $\text{Re}(\omega) = \pm\infty$ vanish and we obtain

$$\begin{aligned} & \frac{1}{\beta} \frac{1}{2\pi i} \oint_{\mathcal{C}'} d\omega \ln(1 + e^{-\beta\hbar\omega}) \frac{d}{d\omega} \left(e^{\eta\omega} \sum_{n=1}^{2n_{orb}} \ln(-\omega + (\mu - E_n(\mathbf{k}))/\hbar) \right) \\ &= \frac{1}{\beta} \frac{1}{2\pi i} \oint_{\mathcal{C}'} d\omega \ln(1 + e^{-\beta\hbar\omega}) e^{\eta\omega} \sum_{n=1}^{2n_{orb}} \frac{\hbar}{\hbar\omega + \mu - E_n(\mathbf{k})} + \mathcal{O}(\eta). \end{aligned}$$

Now we substitute $\hbar\omega \rightarrow z$ and use the theorem of residues (remembering that $\mathcal{C}'$ encircles the poles of the Green function on the real axis in *clockwise* fashion) and after taking the limit $\eta \rightarrow 0$ obtain the expression (24), which completes the first step of the proof.

3.3 Calculation of $\partial\Omega/\partial\lambda$

To obtain the derivative of the true grand potential Ω with respect to λ we start from the formula

$$\begin{aligned} \lambda \frac{\partial}{\partial\lambda} \Omega(\lambda) &= -\frac{1}{\beta} \lambda \frac{\partial}{\partial\lambda} \ln(\text{Tr}(e^{-\beta(H_0 + \lambda H_1) - \mu N})) \\ &= \frac{1}{Z} \text{Tr}(\lambda H_1 e^{-\beta(H_0 + \lambda H_1) - \mu N}) = \langle \lambda H_1 \rangle_\lambda \end{aligned}$$

where $\langle \dots \rangle_\lambda$ denotes the thermal average *calculated at interaction strength* λ . The last quantity thus is the expectation value of the interaction Hamiltonian for interaction strength λ . It can be

computed by making use of the equation of motion of the Green function which is a procedure found in many textbooks, see e.g. Equation (23.14) of FW. One has

$$\langle \lambda H_1 \rangle_\lambda = -\frac{1}{2} \lim_{\tau \rightarrow 0^-} \sum_{\mathbf{k}} \text{Tr} \left(\hbar \frac{\partial}{\partial \tau} - \mu + \mathbf{t}(\mathbf{k}) \right) \mathbf{G}_\lambda(\mathbf{k}, \tau),$$

where the subscript λ on the Green function implies that this is the exact Green function for interaction strength λ . Next we recall the Dyson equation (16), which holds for any λ

$$\left(-\partial_\tau + \frac{\mu}{\hbar} - \frac{1}{\hbar} \mathbf{t}(\mathbf{k}) \right) \mathbf{G}(\mathbf{k}, \tau) - \int_0^{\beta\hbar} \Sigma(\mathbf{k}, \tau - \tau') \mathbf{G}(\mathbf{k}, \tau') d\tau' = \delta(\tau).$$

Since $\delta(\tau < 0) = 0$ we have $\lim_{\tau \rightarrow 0^-} \delta(\tau) = 0$ and obtain

$$\begin{aligned} \lambda \frac{\partial}{\partial \lambda} \Omega(\lambda) &= \frac{\hbar}{2} \lim_{\tau \rightarrow 0^-} \sum_{\mathbf{k}} \int_0^{\beta\hbar} d\tau' \text{Tr} \left(\Sigma_\lambda(\mathbf{k}, \tau - \tau') \mathbf{G}_\lambda(\mathbf{k}, \tau') \right) \\ &= \frac{1}{2\beta} \sum_{\mathbf{k}, \nu} \text{Tr} \Sigma_\lambda(\mathbf{k}, i\omega_\nu) \mathbf{G}_\lambda(\mathbf{k}, i\omega_\nu). \end{aligned} \quad (29)$$

3.4 Definition and properties of the Luttinger-Ward functional

As already mentioned $\Phi[\mathbf{G}]$ is defined as a sum over infinitely many Feynman diagrams with certain properties. The diagrams which contribute are *closed*, which means they have no external lines. They are moreover *connected*, which means that they cannot be decomposed into sub-diagrams which are not connected by lines. And finally, only *skeleton diagrams* are taken into account in the Luttinger-Ward functional. A skeleton diagram is a diagram where no Green function line contains a self-energy insertion. In other words, it is impossible to draw a box around any part of the diagram, so that only two Green function lines cross the box.

At this point we need to discuss an important property of the skeleton diagrams. Let us consider a self-energy diagram. It contains one Green function line from the entry-point to the exit-point, and a number of Green function loops. Starting from the entry-point we may follow the Green function line and draw a circle around each self-energy insertion that we encounter until we reach the exit point. This procedure will eliminate a number of loops, that means enclose them in a self-energy insertion. Then, we continue along the first interaction line which is not eliminated until we reach a Fermion loop that is not yet eliminated. We follow the Green function line along this loop and again draw a circle around each self-energy insertion. We proceed to the next interaction line that has not yet been eliminated and so on. We end up with a diagram in which all self-energy insertions are inside circles. Replacing the circles by straight lines, we obviously obtain a skeleton-diagram for the self-energy. It is easy to see that the skeleton diagram to which a given self-energy diagram is reduced is unique. All self-energy diagrams thus can be grouped into classes such that all members of one class can be reduced to the same skeleton diagram. Conversely, all members of one class can be obtained by starting out from the skeleton-diagram and inserting the full Green function for each Green function line in the

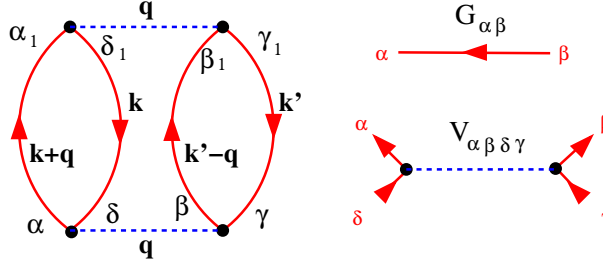


Fig. 4: Left: A diagram contributing to the Luttinger-Ward functional. Right: the elements of the diagram.

diagram, which we write as

$$\Sigma(\mathbf{k}, \omega) = \sum_n \Sigma^{(s,n)}(\mathbf{k}, \omega). \quad (30)$$

Here $\Sigma^{(s,n)}$ denotes the set of all n^{th} order skeleton diagrams (i.e. diagrams with n interaction lines) with the Green function lines replaced by the full Green function.

Having defined the diagrams contributing to $\Phi[\mathbf{G}]$ each diagram is now translated into a multiple sum according to the standard Feynman rules for the imaginary-time Green function in momentum space (see section 25 of FW). However, there is one crucial difference: whereas in a standard Feynman diagram a Green function line corresponds to a factor $\mathbf{G}^0(\mathbf{k}, \omega)$ (the non-interacting Green function), in the Luttinger-Ward functional we replace $\mathbf{G}^0(\mathbf{k}, \omega) \rightarrow \mathbf{G}(\mathbf{k}, \omega)$ where $\mathbf{G}(\mathbf{k}, \omega)$ is the argument of the functional $\Phi[\mathbf{G}]$. As an example, the expression corresponding to the diagram in Figure (4) is

$$\left(\frac{-1}{\beta \hbar^2 N}\right)^2 (-1)^2 \sum_{\mathbf{k}, \mathbf{k}', \mathbf{q}} \sum_{\alpha, \beta, \gamma, \delta} \sum_{\alpha_1, \beta_1, \gamma_1, \delta_1} \sum_{\nu, \nu', \mu} V_{\alpha, \beta, \delta, \gamma}(\mathbf{k}, \mathbf{k}', \mathbf{q}) V_{\delta_1, \gamma_1, \alpha_1, \beta_1}(\mathbf{k}+\mathbf{q}, \mathbf{k}'-\mathbf{q}, -\mathbf{q}) \\ \times G_{\alpha_1, \alpha}(\mathbf{k}+\mathbf{q}, i\omega_\nu + \omega_\mu) G_{\delta, \delta_1}(\mathbf{k}, i\omega_\nu) G_{\beta_1, \beta}(\mathbf{k}'-\mathbf{q}, \omega_{\nu'} - \omega_\mu) G_{\gamma, \gamma_1}(\mathbf{k}', \omega_{\nu'}). \quad (31)$$

The Luttinger-Ward functional $\Phi[\mathbf{G}]$ thus consists of an infinite sum of multiple sums which involve the interaction matrix elements V of the Hamiltonian (23) and the function $\mathbf{G}$ for which the functional is to be evaluated.

Let us briefly discuss the scaling with system size, N . By the Feynman rules an n^{th} order diagram has the prefactor $(1/N)^n$. On the other hand, there are n interaction lines, and $2n$ Green function lines, so that there are $3n$ momenta. The n interaction lines give rise to $2n$ momentum conservation conditions, one for each end of a line. However, in a *closed* diagram one of these momentum conservation conditions is fulfilled trivially so that there remain $n+1$ momenta to be summed over (see the above example). Each sum runs over N momenta so that the total diagram is of order N , as it has to be because Ω is an extensive quantity.

In addition to the factors originating from the Feynman rules, each diagram is multiplied by $-1/(\beta S)$ where the positive integer S is the *symmetry factor* of the diagram. A very detailed discussion of these symmetry factors is given in section 2.3 of Negele-Orland [7]. The definition is as follows: first, the diagram is drawn such that all interaction lines are in x -direction. The n interaction lines of a diagram are labeled by integers $i \in \{1 \dots n\}$ and the

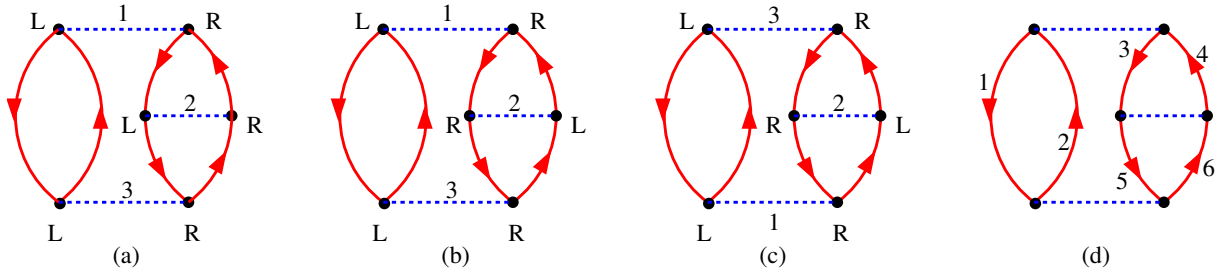


Fig. 5: Determination of the symmetry factor S for a diagram.

ends of each interaction line are labeled by R and L (for ‘right end’ and ‘left end’), see Figure 5(a). Any Green function line in the diagram now can be labeled by the ends of the interaction lines where it departs and where it ends: $(i, S_1) \rightarrow (j, S_2)$ with $i, j \in \{1 \dots n\}$ and $S_1, S_2 \in \{R, L\}$. Obviously, the diagram is characterized completely by the $2n$ ‘directed quadruples’ $(i, S_1) \rightarrow (j, S_2)$. Then, we consider the following operations on the diagrams: a) any permutation of the indices i , b) exchange of the labels R and L on an arbitrary number of interaction lines, c) any combination of a permutation followed by label exchanges. Such an operation obviously changes the quadruples which characterize the connectivity of the diagram: $[(i, S_1) \rightarrow (j, S_2)] \rightarrow [(i', S'_1) \rightarrow (j', S'_2)]$. The symmetry factor of a diagram then is the number of symmetry operations—including identity—where the new labels $(i', S'_1) \rightarrow (j', S'_2)$ are a permutation of the old ones, $(i, S_1) \rightarrow (j, S_2)$ (Negele-Orland then call the transformed diagram a deformation of the first one). As an example, consider the diagram in Figure 5(a). Label exchange on, say, interaction line 2 leads to the diagram shown in 5(b) which, however, is not a deformation of the original diagram. This can be seen by considering, e.g., the line connecting the R -end of 1 and the R -end of 2. In 5(a) this line would have the label $(2, R) \rightarrow (1, R)$, whereas it would be $(1, R) \rightarrow (2, R)$ in 5(b). This means that the direction of momentum flow along the line would be reversed. On the other hand, the permutation of the labels 1 and 3 followed by label exchange on interaction line 2 leads to the diagram 5(c) which indeed is a deformation of the original diagram. In Figure 5(d) the Green function lines are numbered by $1 \rightarrow 6$ and Table 1 gives the quadruples corresponding to these lines in Figures 5(a) and 5(c). Obviously the two sets of quadruples are a permutation of each other.

Line	5(a)	5(c)
1	$(1, L) \rightarrow (3, L)$	$(3, L) \rightarrow (1, L)$
2	$(3, L) \rightarrow (1, L)$	$(1, L) \rightarrow (3, L)$
3	$(1, R) \rightarrow (2, L)$	$(3, R) \rightarrow (2, R)$
4	$(2, R) \rightarrow (1, R)$	$(2, L) \rightarrow (3, R)$
5	$(2, L) \rightarrow (3, R)$	$(2, R) \rightarrow (1, R)$
6	$(3, R) \rightarrow (2, R)$	$(1, R) \rightarrow (2, L)$

Table 1: Quadruples describing the connectivity of the diagrams Figure 5a and Figure 5c. The numbers of the Green function lines are given in Figure 5d.

It turns out that this is the only symmetry operation which leaves the diagram invariant, so that, taking into account the identity operation, the diagram has $S = 2$. Since a symmetry operation corresponds to a permutation of the quadruples $(i, S_1) \rightarrow (j, S_2)$ which characterize the individual Green function lines in a diagram it defines a mapping between these lines whereby each line is mapped onto the one which gets its label. For example, from Table 1 one reads off the corresponding mapping for the operation connecting 5a and 5c:

$$\begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 6 & 5 & 4 & 3 \end{array}$$

If two Green function lines i and j are mapped onto each other, the lines are equivalent in the sense that the diagram could be deformed such that the deformed diagram is precisely the same as the original one but line j now taking the place of line i and vice versa.

Let us now assume that a diagram has the symmetry factor S . This means that all Green function lines can be grouped into disjunct classes such that the lines belonging to one class are mapped onto each other by one of the S symmetry operations. For example, the diagram in 5 has the classes $(1, 2)$, $(3, 6)$, and $(4, 5)$. Since a diagram with n interaction lines has $2n$ Green function lines the number of classes is $2n/S$ which will be of importance later on.

Next, we want to see the meaning of this definition. In fact, the Luttinger-Ward functional is the *generating functional* of the self-energy, or, more precisely,

$$\frac{\partial \Phi}{\partial G_{\alpha\beta}(\mathbf{k}, i\omega_\nu)} = \frac{1}{\beta} \Sigma_{\beta,\alpha}(\mathbf{k}, i\omega_\nu). \quad (32)$$

To see this, consider an infinitesimal change $G_{\alpha\beta}(\mathbf{k}, i\omega_\nu) \rightarrow G_{\alpha\beta}(\mathbf{k}, i\omega_\nu) + \delta G_{\alpha\beta}(\mathbf{k}, i\omega_\nu)$ as in Figure 6. The initial diagrams correspond to multiple sums over products of Green functions where all internal frequencies, momenta, and orbital indices are summed over, subject to the condition of energy/momentum conservation at each interaction vertex, see (31). The first order change then also can be viewed as a sum of diagrams but with a single missing line—this corresponds to the variation δG which has been factored out. Another way to state this is to say that differentiating with respect to an element of $\mathbf{G}$ amounts to successively ‘open’ each of the lines in the initial closed diagram and sum the remaining diagrams. These remaining diagrams obviously ‘look like’ self-energy diagrams in that they have two entry points. We now need to show, however, that the diagrams not only ‘look like’ possible contributions to the skeleton diagram expansion of the self-energy, but that they come with exactly the right numerical prefactors. At this point, the additional prefactors of $-1/\beta S$ turn out to be crucial.

We first note that the momentum and frequency which flow into/out of the diagram are fixed by the momentum and frequency of δG . As regards the orbital indices, we recall that $G_{\alpha\beta}$ corresponds to a directed line $\beta \rightarrow \alpha$. The resulting self-energy-like diagrams therefore all have the matrix index α on their incoming entry and β on their outgoing entry, and comparing with Figure 4 we see that this assignment of indices corresponds to $\Sigma_{\beta\alpha}$. Moreover, all internal momenta, frequencies, and matrix indices in the remaining diagrams are summed over—subject to the condition of frequency and momentum conservation at the interaction lines—as would

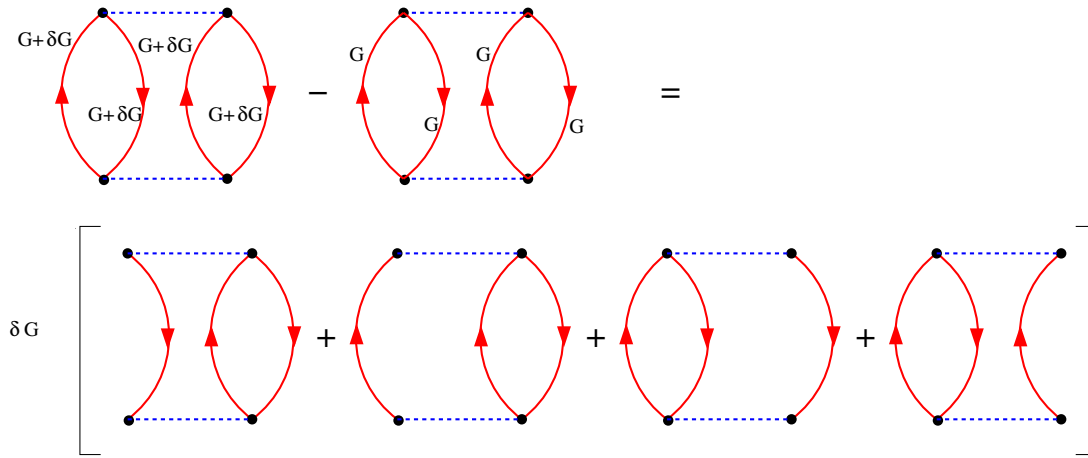


Fig. 6: Variation of G implies ‘opening’ the lines of a Feynman diagram.

be the case in the true self-energy diagrams. Second, the order n of a diagram—that means the number of interaction lines—is not changed by opening a Green function line, so that the prefactor $-1/(\beta\hbar^2 N)^n$ of the closed diagram is also the correct prefactor for the resulting self-energy diagram. Third, opening a Green function line reduces the number of closed Fermion loops by 1 and the factor (-1) in $(-\frac{1}{\beta S})$ takes care of this. And, lastly, we need to discuss the symmetry factor S . Let us consider a diagram with n interaction lines, which accordingly has $2n$ Green function lines and moreover assume that the diagram has the symmetry factor S . As we saw above, the $2n$ Green function lines can be divided into classes of S members which are mapped into each other by the symmetry operations and the number of these classes is $2n/S$. A symmetry operation maps a Green function line i onto an equivalent one j , so that it is possible to deform the diagram such that it looks exactly the same as the original one, but with line j in place of line i . This means, however, that ‘opening’ the line i also gives *exactly* the same self-energy diagram as opening line j . Accordingly, from the single closed diagram of degree n with symmetry factor S we obtain $2n/S$ different skeleton diagrams for the self-energy, and each is produced S times, see also Figure 7. This factor of S , however, precisely cancels the prefactor $1/S$. It follows, that each skeleton-diagram for the self-energy is produced with same prefactor $1/\beta$. Differentiating $\Phi[G]$ with respect to $G_{\alpha\beta\gamma}(\mathbf{k}, i\omega_\nu)$ thus gives $1/\beta$ times the sum of all skeleton diagrams for $\Sigma_{\beta\alpha}(\mathbf{k}, i\omega_\nu)$, with the noninteracting Green function replaced by the full one, and this is exactly $\Sigma_{\beta\alpha}(\mathbf{k}, i\omega_\nu)$ itself, see (30), so that (32) is proved.

We have just seen, that all skeleton-diagrams for the self-energy can be obtained by differentiating the Luttinger-Ward functional with respect to G , whereby the differentiation corresponds to ‘opening’ one line in a closed diagram. We then may ask if this operation can be reversed, namely if the Luttinger-Ward functional can be obtained by starting from the skeleton-diagram expansion of the self-energy and ‘close’ the diagrams by ‘reconnecting’ the entry-points of the self-energy by a Green function. More precisely, we consider

$$\frac{1}{\beta} \sum_{\nu, \mathbf{k}} \sum_{\alpha, \beta} G_{\alpha, \beta}(\mathbf{k}, i\omega_\nu) \Sigma_{\beta, \alpha}^{(s, n)}(\mathbf{k}, i\omega_\nu). \quad (33)$$

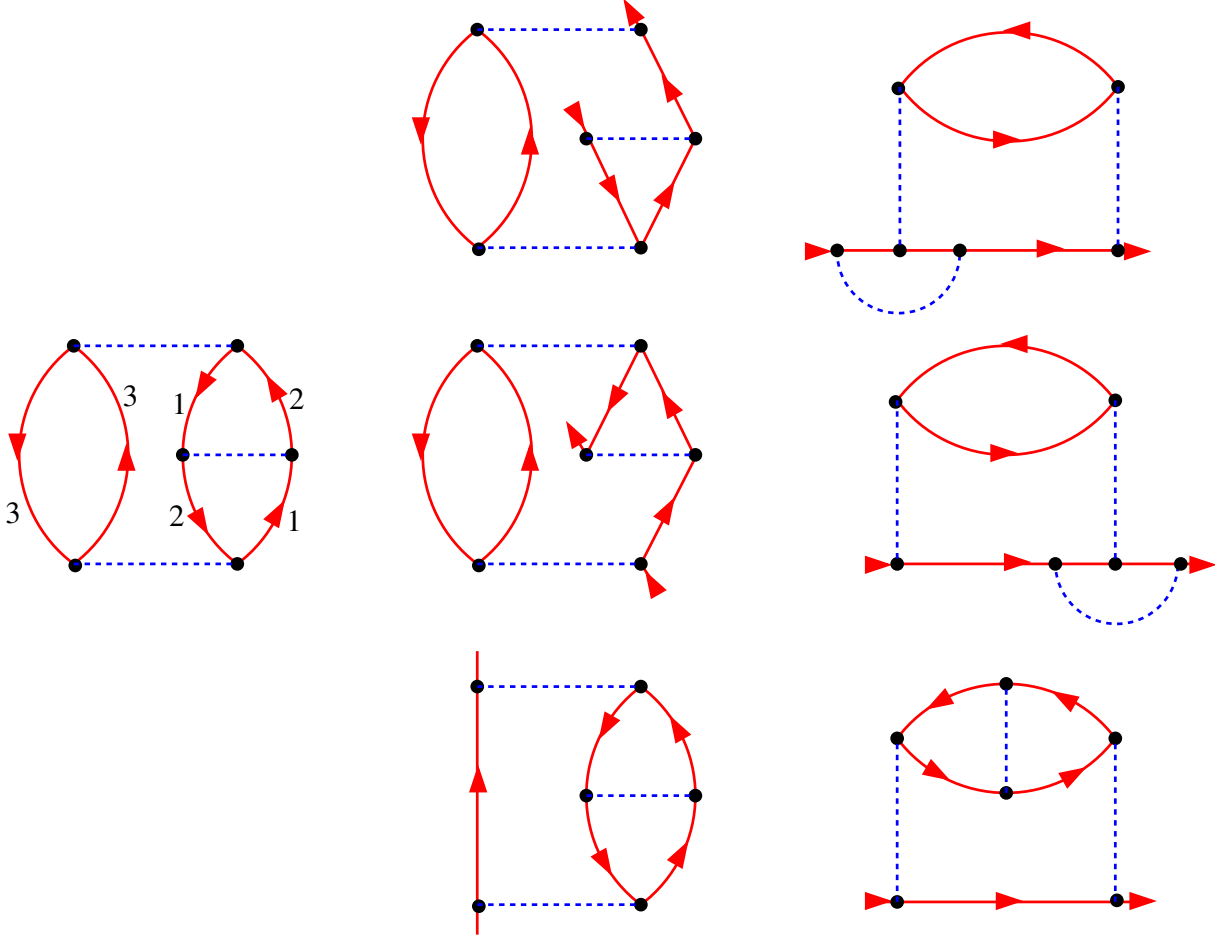


Fig. 7: The diagram on the left has $n = 3$ and $S = 2$ and accordingly 3 classes of symmetry-equivalent Green function lines. The lines are labeled by the number of the classes, compare Figure 5 and Table 1. Successively opening the lines of the diagram produces the three different self-energy diagrams in the center column and each of them is produced $S = 2$ times. The right column shows the diagrams redrawn to more look like self-energy diagrams.

We have seen that an n^{th} order diagram contributing to $\Phi[\mathbf{G}]$ with symmetry factor S produces $2n/S$ different skeleton-self-energy diagrams, and each of them S times and with a factor of (-1) , so that the remaining prefactor was $1/\beta$. Upon closing the Fermion line again, according to (33), each of these diagrams produces the original closed diagram (it is easy to see that for each self-energy diagram there is exactly one closed diagram from which it can be obtained). Since there are $2n/S$ self-energy diagrams originating from the original closed diagram the latter is produced $2n/S$ times and thus has the additional prefactor $-2n/S\beta$ (the factor of (-1) is due to the additional Fermion loop in the closed diagram). In the expansion of $\Phi[\mathbf{G}]$, however, the diagram would have had the prefactor $-1/S\beta$, or, put another way, closing the sum of all n^{th} order skeleton diagrams for Σ , according to (33), produces the n^{th} order contribution to $\Phi[\mathbf{G}]$ with an additional prefactor of $2n$ so that

$$\Phi^{(n)} = \frac{1}{2n\beta} \sum_{\nu, \mathbf{k}} \text{Tr} \mathbf{G}(\mathbf{k}, i\omega_\nu) \Sigma^{(s,n)}(\mathbf{k}, i\omega_\nu). \quad (34)$$

Lastly, we give one more comment on the symmetry factors. Readers who wish to study the original paper of Luttinger and Ward, which is highly recommended, will realize that no symmetry factors appear in this work. The reason is that LW carry out their derivation using what Negele-Orland (NO) call ‘labeled Feynman diagrams’. For example, LW’s equation (17) corresponds precisely to NO’s equation (2.96). The derivation in the present notes, however, uses what NO call ‘unlabeled Feynman diagrams’. The transition between these two types of diagrams and the emergence of the symmetry factors thereby is discussed in detail in section 2.3 of NO [7].

3.5 Calculation of $\partial\tilde{\Omega}/\partial\lambda$

We proceed to the final step of the proof and compute $\partial\tilde{\Omega}/\partial\lambda$. If we vary the interaction strength λ , there are two places in the expression for Ω' , eqn. (25), where this makes it self felt. Namely the self-energy Σ will change and moreover the interaction matrix elements V in the Luttinger-Ward functional (see, e.g., eqn. (31)) which have a prefactor of λ will also contribute to the variation. Let us first consider the variation of Σ and compute

$$\frac{\partial\Omega'}{\partial\Sigma_{\alpha,\beta}(\mathbf{k}, i\omega_\nu)}.$$

There are three terms in (25) and we consider them one after the other. The first two terms involve a sum over momentum and frequency and obviously only those terms with momentum $\mathbf{k}$ and frequency ω_ν will contribute. Accordingly, in the following equations we omit the arguments $(\mathbf{k}, i\omega_\nu)$ for brevity. Then we find by using the chain rule for differentiation

$$\begin{aligned} \frac{\partial}{\partial\Sigma_{\alpha,\beta}} \left(-\frac{1}{\beta} \ln \det (-\mathbf{G}^{-1}) \right) &= -\frac{1}{\beta} \sum_{\mu,\nu} \left(\frac{\partial}{\partial(-G_{\mu,\nu}^{-1})} \ln \det (-\mathbf{G}^{-1}) \right) \frac{\partial(-G_{\mu,\nu}^{-1})}{\partial\Sigma_{\alpha,\beta}} \\ &= -\frac{1}{\beta} \sum_{\mu,\nu} (-G_{\nu,\mu}) \delta_{\mu,\alpha} \delta_{\nu,\beta} \\ &= \frac{1}{\beta} G_{\beta,\alpha}. \end{aligned}$$

In going to the 2nd line we used the identity from Appendix B and the Dyson equation

$$-\mathbf{G}^{-1} = -\omega - \mu + \Sigma$$

from which it follows that

$$\frac{\partial(-G_{\mu,\nu}^{-1})}{\partial\Sigma_{\alpha,\beta}} = \delta_{\mu,\alpha} \delta_{\nu,\beta}.$$

We proceed to the second term,

$$\frac{\partial}{\partial\Sigma_{\alpha,\beta}} \left(-\frac{1}{\beta} \text{Tr } \Sigma \mathbf{G} \right) = \frac{\partial}{\partial\Sigma_{\alpha,\beta}} \left(-\frac{1}{\beta} \sum_{\mu,\nu} \Sigma_{\nu,\mu} G_{\mu,\nu} \right) = -\frac{1}{\beta} \left(G_{\beta,\alpha} + \sum_{\mu,\nu} \Sigma_{\nu,\mu} \frac{\partial G_{\mu,\nu}}{\partial\Sigma_{\alpha,\beta}} \right).$$

Lastly we consider the Luttinger-Ward functional. Using again the chain rule we find

$$\frac{\partial \Phi[G]}{\partial \Sigma_{\alpha,\beta}} = \sum_{\mu,\nu} \frac{\partial \Phi[G]}{\partial G_{\mu,\nu}} \frac{\partial G_{\mu,\nu}}{\partial \Sigma_{\alpha,\beta}} = \frac{1}{\beta} \sum_{\mu,\nu} \Sigma_{\nu,\mu} \frac{\partial G_{\mu,\nu}}{\partial \Sigma_{\alpha,\beta}}.$$

Adding up the three terms we thus obtain the important result

$$\frac{\partial \Omega'}{\partial \Sigma_{\alpha,\beta}(\mathbf{k}, i\omega_\nu)} = 0. \quad (35)$$

In other words: the expression Ω' , which will be seen to be equal to the grand potential Ω in a moment, is stationary with respect to variations of the self-energy! This is the stationarity condition for Σ which is the basis of the VCA.

First, however, we have to complete the proof and evaluate $\lambda \frac{\partial}{\partial \lambda} \Omega'(\lambda)$. Since there is no variation of Ω' due to a variation of Σ , the only remaining source of variation are the interaction lines in the Luttinger-Ward functional. Namely any n^{th} order diagram has the prefactor of λ^n so that

$$\lambda \frac{\partial}{\partial \lambda} \Phi^{(n)} = n \Phi^{(n)}$$

Using (34) we thus obtain

$$\begin{aligned} \lambda \frac{d\Omega'}{d\lambda} &= \sum_n n \Phi^{(n)} = \sum_n \frac{1}{2\beta} \sum_{\nu,\mathbf{k}} \text{Tr} \mathbf{G}_\lambda(\mathbf{k}, i\omega_\nu) \Sigma_\lambda^{(s,n)}(\mathbf{k}, i\omega_\nu) \\ &= \frac{1}{2\beta} \sum_{\nu,\mathbf{k}} \text{Tr} \mathbf{G}_\lambda(\mathbf{k}, i\omega_\nu) \left(\sum_n \Sigma_\lambda^{(s,n)}(\mathbf{k}, i\omega_\nu) \right) \\ &= \frac{1}{2\beta} \sum_{\nu,\mathbf{k}} \text{Tr} \mathbf{G}_\lambda(\mathbf{k}, i\omega_\nu) \Sigma_\lambda(\mathbf{k}, i\omega_\nu). \end{aligned}$$

Comparing with (29) we see that this is equal to $\lambda \frac{\partial}{\partial \lambda} \Omega(\lambda)$ which completes the proof.

Let us summarize the results which we have obtained:

1. The grand canonical potential Ω of an interacting Fermi system is given by (25).
2. The Luttinger-Ward functional is the generating functional of $\Sigma(\mathbf{k}, i\omega_\nu)$, see eqn. (32).
3. The Luttinger-Ward functional depends only on the interaction matrix elements $V_{\alpha\beta\delta\gamma}$ in the Hamiltonian and the Green function $\mathbf{G}$ which is the argument of the functional.
4. Ω is stationary under variations of $\Sigma(\mathbf{k}, i\omega_\nu)$ see (35).

Looking at the above proof one might worry about the fact that it assumes a continuous evolution of the system with increasing interaction strength λ , whereas we are also interested, e.g., in Mott-insulators where we have reason to believe that a phase transition occurs as a function of λ . However, Potthoff has recently given a nonperturbative proof of the theorem [12, 11] that means all of the above properties of the grand potential, the Luttinger-Ward functional and the self-energy remain valid in a strongly correlated electron system where a Feynman-diagram expansion of the Green function and the adiabatic continuity with the noninteracting system can no longer be assumed valid.

4 Conserving approximations

An important application of the Luttinger-Ward functional is the construction of conserving approximations. To discuss these, we will use a special form of the Hamiltonian for the remainder of this section. We assume that the set of quantum numbers reduces to $\alpha = \mathbf{r}$ that means we neglect spin for simplicity. To follow the notation common in the literature we denote the electron creation operator by $\Psi^\dagger(\mathbf{r})$ instead of $a_\alpha^\dagger$. The Hamiltonian for a homogeneous system is

$$H = \frac{\hbar^2}{2m} \int d\mathbf{r} \nabla \Psi^\dagger(\mathbf{r}) \cdot \nabla \Psi(\mathbf{r}) + \frac{1}{2} \int d\mathbf{r} \int d\mathbf{r}' \Psi^\dagger(\mathbf{r}) \Psi^\dagger(\mathbf{r}') V(\mathbf{r}-\mathbf{r}') \Psi(\mathbf{r}') \Psi(\mathbf{r}).$$

We assume that we can always perform a partial integration and drop the surface terms. In this way the first term in H can be brought to either of the two forms $-\frac{\hbar^2}{2m} \int d\mathbf{r} \Psi^\dagger(\mathbf{r}) (\nabla^2 \Psi(\mathbf{r}))$ or $-\frac{\hbar^2}{2m} \int d\mathbf{r} (\nabla^2 \Psi^\dagger(\mathbf{r})) \Psi(\mathbf{r})$. Since the system is homogeneous the electron density is n_0 , independent of position. We now assume that the system is acted upon by a perturbation of the form $H_p = \int d\mathbf{r} U(\mathbf{r}, t) n(\mathbf{r})$, i.e., a time dependent real potential U which couples to the electron density $n(\mathbf{r}) = \Psi^\dagger(\mathbf{r}) \Psi(\mathbf{r})$. Thereby we demand that $\int d\mathbf{r} U(\mathbf{r}, t) = 0$ for all t because a constant component would merely shift the zero of energy and not induce any response. As we have seen—or rather: quoted from FW—in the first section, the change of the expectation value of any operator $A(\mathbf{r})$ to first order in U is given by

$$\begin{aligned} \delta \langle A(\mathbf{r}) \rangle(t) &= \frac{1}{\hbar} \int d\mathbf{r}' \int_{-\infty}^{\infty} dt' G_{A,n}^R(\mathbf{r}t, \mathbf{r}'t') U(\mathbf{r}', t'), \\ G_{A,n}^R(\mathbf{r}t, \mathbf{r}'t') &= -i\Theta(t-t') \langle [A(\mathbf{r}, t), n(\mathbf{r}', t')] \rangle_{th}. \end{aligned}$$

Now we may choose $A(\mathbf{r})$ to be the operator $n(\mathbf{r})$ of electron density or the operator of electron current $\mathbf{j}(\mathbf{r}) = \frac{i\hbar}{2m} ((\nabla \Psi^\dagger(\mathbf{r})) \Psi(\mathbf{r}) - \Psi^\dagger(\mathbf{r}) \nabla \Psi(\mathbf{r}))$. Assuming that in the unperturbed state of the system there is no current and that the electron density is time independent, the induced changes must fulfill certain conservation laws:

$$\begin{aligned} \frac{\partial \delta n(\mathbf{r})}{\partial t} + \nabla \cdot \delta \mathbf{j}(\mathbf{r}) &= 0, \\ \frac{d}{dt} \int d\mathbf{r} m \delta \mathbf{j}(\mathbf{r}) &= \int d\mathbf{r} \left(-\nabla U(\mathbf{r}, t) \right) \delta n(\mathbf{r}, t) \\ \frac{d}{dt} \langle H \rangle &= \int d\mathbf{r} \left(-\nabla U(\mathbf{r}, t) \right) \cdot \delta \mathbf{j}(\mathbf{r}) \end{aligned}$$

The first line is the continuity equation, the second line states that the total momentum of the electron system changes according to Newton's law and the third equation states that the change of the total energy of the system is equal to the work done by the external force (in the second equation we have used that by partial integration the right hand side can be converted to $\int d\mathbf{r} U(\mathbf{r}) (\nabla n(\mathbf{r}))$ so that the constant component of n does not contribute). Since in general we have to make some approximation to compute the retarded Green function, however, it is not a priori clear that these equations are fulfilled. The imaginary time Green functions $G_{n,n}$ and

$G_{j,n}$ can be expressed as limits of the two-particle Green function G_2 and the self-energy also is related to G_2 , so that approximating Σ by a functional of the Green function gives us also an approximation for the response function. Building on the work of Luttinger and Ward, Baym has given a prescription to construct approximations for Σ which guarantee that the resulting response functions obey all conservation laws and we want to outline his ideas. To simplify the notation we introduce the ‘compound coordinate’ $x = (\mathbf{r}\tau)$ and it is understood that operators such as $n(x)$ which have x as an argument are imaginary-time Heisenberg operators. Moreover we denote

$$\int dx \cdots = \int_0^{\beta\hbar} d\tau \int d\mathbf{r} \cdots$$

We first generalize the definition of the imaginary-time single particle Green function

$$\begin{aligned} G(x, x')[U] &= -\frac{1}{Z[U]} \left\langle T \left[\Psi(x) \Psi^\dagger(x') \exp \left\{ \frac{1}{\hbar} \int dx_1 U(x_1) n(x_1) \right\} \right] \right\rangle_{th}, \\ Z[U] &= \left\langle T \left[\exp \left\{ \frac{1}{\hbar} \int dx_1 U(x_1) n(x_1) \right\} \right] \right\rangle_{th}. \end{aligned} \quad (36)$$

This modified Green function is a functional of the real and time dependent potential $U(x)$ and depends on τ and τ' separately. It is obvious that as $U \rightarrow 0$ it reduces to the ordinary Green function discussed so far. Moreover, using the cyclic property of the trace it can again be shown that this Green function obeys the same boundary condition $G(\beta\hbar) = -G(0)$ as the ordinary Green functions and thus has a Fourier expansion of the type (6). It should be stressed, that this Green function is defined for imaginary times and has no direct physical interpretation.

We now consider the functional derivative of G with respect to $U(x_1)$. This is defined as the change of $G[U]$ under an infinitesimal ‘ δ -spike’, $U(x) \rightarrow U(x) + \epsilon\delta(x-x_1)$. One may also think of the integrals in (36) as the limit of sums over grid points x_i , the functional derivative then is the limit of the derivative with respect to the value of U at the grid point closest to x_1 . Using the fact that in the argument of the time-ordering operator T the operators H or n can be commuted with each other and with both, Ψ and $\Psi^\dagger$, one finds

$$\hbar \frac{\delta G(x, x')[U]}{\delta U(x_1)} \Big|_{U=0} = -\langle T[\Psi(x) \Psi^\dagger(x') n(x_1)] \rangle_{th} + G(x, x') n_0.$$

Taking now the limit $x' \rightarrow x^+ = \mathbf{r}\tau^+$ we find

$$\begin{aligned} \lim_{x' \rightarrow x^+} \hbar \frac{\delta G(x, x')[U]}{\delta U(x_1)} \Big|_{U=0} &= -\langle T[n(x)n(x_1)] \rangle_{th} + n_0^2, \\ \lim_{x' \rightarrow x^+} \hbar \frac{i\hbar}{2m} (\nabla' - \nabla) \frac{\delta G(x, x')[U]}{\delta U(x_1)} \Big|_{U=0} &= -\langle T[\mathbf{j}(x)n(x_1)] \rangle_{th} + \langle \mathbf{j}(x) \rangle_{th} n_0. \end{aligned}$$

It should be noted that for $U = 0$ we have $\langle \mathbf{j}(x) \rangle_{th} = 0$ so that the second term in the second equation could have been omitted. The extra terms involving n_0 can be absorbed by re-defining $n(x) \rightarrow n(x) - n_0$, i.e., the operator of density fluctuations. We will assume that this has been

done from now on. The above shows that the time-ordered Green function of the operators $A = n(\mathbf{r})$ or $A = \mathbf{j}(\mathbf{r})$ and $B = n(\mathbf{r}_1)$ (which gives us the physical response function when continued analytically to above the real axis) can be obtained by functional differentiation of the modified Green function (36). In fact the only requirement is that A be quadratic in the field operators or its derivatives, moreover $n(\mathbf{r})$ can be replaced by any other operator. On the other hand, the Green function involves the self-energy Σ , which we assume to be approximated as a functional of the Green function itself: $\Sigma = \Sigma[G]$. The functional derivative—and hence the response functions—therefore depend on the form of $\Sigma[G]$. The question then is if it is possible to choose the functional $\Sigma = \Sigma[G]$ such, that the resulting response functions obey the conservation laws. Baym has shown that the answer to this question is affirmative and in fact the general prescription is to define an approximate Luttinger-Ward functional, which contains only a subset of all possible diagrams and to then construct Σ according to (32):

$$\Sigma_{\alpha,\beta}[G] = \frac{1}{K_B T} \frac{\delta \Phi[G]}{\delta G_{\beta,\alpha}} \Rightarrow \delta \Phi[G] = k_B T \sum_{\alpha,\beta} \Sigma_{\alpha,\beta}[G] \delta G_{\beta,\alpha}. \quad (37)$$

We show that a self-energy constructed in this way obeys the continuity equation: The equation of motion obeyed by the modified Green function is (we suppress the $[U]$ on all quantities)

$$\left(-\hbar \frac{\partial}{\partial \tau} + \frac{\hbar^2 \nabla^2}{2m} + \mu + U(\mathbf{r}, \tau) \right) G(\mathbf{r}\tau, \mathbf{r}'\tau') = \hbar \delta(\mathbf{r}-\mathbf{r}') \delta(\tau-\tau') \quad (38)$$

$$+ \int d\mathbf{r}_1 V(\mathbf{r}-\mathbf{r}_1) G_2(\mathbf{r}\tau, \mathbf{r}_1\tau, \mathbf{r}'\tau', \mathbf{r}_1\tau').$$

Here the modified G_2 in the presence of U is defined in a completely analogous way as G , see (36). We now assume that G_2 is approximated as a functional of G and replace the approximate G_2 by an approximate Σ . Equation (38) then becomes

$$\left(-\hbar \frac{\partial}{\partial \tau} + \frac{\hbar^2 \nabla^2}{2m} + \mu + U(x) \right) G(x, x') = \hbar \delta(x-x') + \hbar \int dx_1 \Sigma(x, x_1) G(x_1, x'). \quad (39)$$

Next we form $\int dx \int dx' G(x_2, x) \cdots G^{-1}(x', x_3)$ where $\cdots$ stands for either the right or left side of (39). We use partial integration to convert, e.g.

$$\int dx G(x_2, x) \left(-\hbar \partial_\tau + \frac{\hbar^2 \nabla^2}{2m} \right) G(x, x') = \int dx \left[\left(+\hbar \partial_\tau + \frac{\hbar^2 \nabla^2}{2m} \right) G(x_2, x) \right] G(x, x')$$

(here the property $G(\beta\hbar) = -G(0)$ of the modified Green function is essential!) and exchange $x_2 \rightarrow x$ and $x_3 \rightarrow x'$ in the resulting equation. In this way we obtain

$$\left(+\hbar \frac{\partial}{\partial \tau'} + \frac{\hbar^2 \nabla'^2}{2m} + \mu + U(x') \right) G(x, x') = \hbar \delta(x-x') + \hbar \int dx_1 G(x, x_1) \Sigma(x_1, x'). \quad (40)$$

Now we consider the change (with a real function $\Gamma(x)$)

$$G(x, x')[U] \rightarrow e^{i\Gamma(x)} G(x, x')[U] e^{-i\Gamma(x')}.$$

In other words, if we represent the Green function $G(x, x')[U]$ by a directed line from $x' \rightarrow x$ it is multiplied by $e^{-i\Gamma(x')}$ for the initial point and by $e^{i\Gamma(x)}$ for the endpoint. In a closed diagram, however, there is one incoming and one outgoing Green function line at every end of an interaction line $V(x-x')$ so that the two factors cancel and the value of the diagram remains unchanged whence $\delta\Phi = 0$. The crucial point is, that this holds true for each closed diagram individually, so that $\delta\Phi = 0$ remains true also for an approximate Luttinger-Ward functional in which only a subset of diagrams is kept. On the other hand, if we switch to an infinitesimal Γ we have $\delta\Phi = \Sigma\delta G$ and for infinitesimal Γ we have $\delta G(x, x')[U] = i(\Gamma(x) - \Gamma(x')) G(x, x')[U]$. It follows that (with $[U]$ again omitted)

$$\Sigma \delta G = i \int dx dx_1 (\Gamma(x_1) - \Gamma(x)) \Sigma(x, x_1) G(x_1, x) = 0.$$

We now interchange integration variables $x \leftrightarrow x_1$ in the term containing $\Gamma(x_1)$ and find

$$\int dx dx_1 \Gamma(x) (\Sigma(x, x_1) G(x_1, x) - \Sigma(x_1, x) G(x, x_1)) = 0$$

Since $\Gamma(x)$ is infinitesimal but arbitrary the x_1 integral must vanish:

$$\int dx_1 (\Sigma(x, x_1) G(x_1, x) - G(x, x_1) \Sigma(x_1, x)) = 0$$

We now subtract (39) from (40), let again $x' \rightarrow x^+$:

$$\lim_{x' \rightarrow x^+} \left\{ \left(\hbar \frac{\partial}{\partial \tau} + \hbar \frac{\partial}{\partial \tau'} \right) G(x, x')[U] + \frac{\hbar^2}{2m} (\nabla'^2 - \nabla^2) G(x, x')[U] \right\} = 0,$$

or, using $\nabla'^2 - \nabla^2 = (\nabla' + \nabla)(\nabla' - \nabla)$

$$\frac{\partial}{\partial \tau} G(x, x^+)[U] - i\nabla \cdot \left[\lim_{x' \rightarrow x^+} \frac{i\hbar}{2m} (\nabla' - \nabla) G(x, x')[U] \right] = 0.$$

Since $G(x, x^+) = n(x)$ is the electron density, and the expression in square brackets is the current density at $\mathbf{j}(x)$, the Green function obeys a kind of continuity equation. Now we take the functional derivative of both sides with respect to $U(x')$ and let $U \rightarrow 0$. This generates a relation between time-ordered Green functions:

$$\frac{\partial}{\partial \tau} G_{n,n}(\mathbf{r}\tau, \mathbf{r}'\tau') = i\nabla \cdot G_{\mathbf{j},n}(\mathbf{r}\tau, \mathbf{r}'\tau') \Rightarrow i\omega_\nu G_{n,n}(\mathbf{k}, i\omega_\nu) = \mathbf{k} \cdot G_{\mathbf{j},n}(\mathbf{k}, i\omega_\nu),$$

and after performing the analytic continuation $i\omega_\nu \rightarrow \omega + i\epsilon^+$ we obtain a relation between the physical response functions

$$i\omega G_{n,n}^{(R)}(\mathbf{k}, \omega) - i\mathbf{k} \cdot G_{\mathbf{j},n}^{(R)}(\mathbf{k}, \omega) = 0 \Rightarrow \frac{\partial}{\partial t} G_{n,n}^{(R)}(\mathbf{r}, t) + \nabla \cdot G_{\mathbf{j},n}^{(R)}(\mathbf{r}, t) = 0.$$

The last relation between the response functions, however, does guarantee the validity of the continuity equation for the fluctuations generated by an arbitrary perturbing potential.

Baym and Kadanaoff [3] have investigated under which conditions the momentum conservation

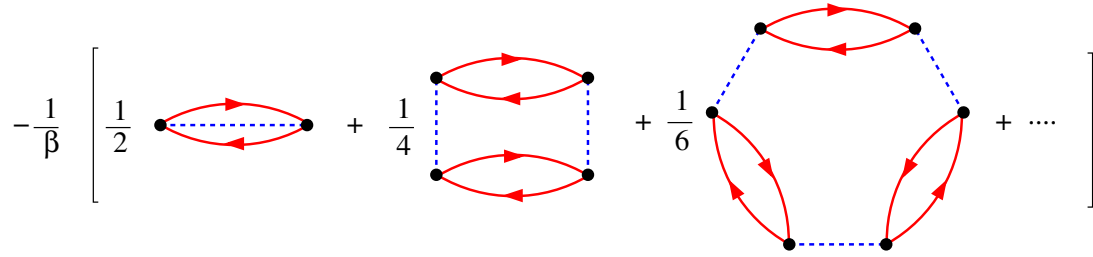


Fig. 8: *Approximate Luttinger-Ward functional for the GW approximation.*

law and the energy conservation law are obeyed as well, and Baym [4] has shown that this is always true for a self-energy derived from an approximate Luttinger-Ward functional according to (37). These derivations are more involved, however, so we do not present them here. A famous example of a conserving approximation is the GW-approximation [13]. This can be derived from an approximate Luttinger-Ward functional which contains only ‘bubble-diagrams,’ see Figure 8. As discussed by Negele-Orland, the symmetry factor of a bubble-diagram with n bubbles is $2n$, which explains the prefactors. Another famous conserving approximation is the fluctuation-exchange approximation (or FLEX) [14], which uses a Luttinger-Ward functional comprising bubbles and ladders and has been frequently applied to the Hubbard model.

5 Conclusion

To summarize, Luttinger and Ward found an expression for the grand canonical potential of interacting Fermi systems, whereby a key step was the introduction of the Luttinger-Ward functional. This turned out to have additional significance in that it allows the construction of conserving approximations. One issue that we did not touch upon in these notes is the relation to cluster methods which are widely used today. Potthoff has shown [12] that a wide variety of these cluster methods, such as Dynamical Mean-Field Theory or the Dynamical Cluster Approximation can be derived by making use of the stationarity of Ω under variation of Σ , equation (35). All of this shows that the Luttinger-Ward functional is a concept of central importance in many-body theory.

A A theorem on Fourier transforms

Let $f(t)$ be some function and $f(\omega)$ its Fourier transform. Then the Fourier transform of the function $g(t) = -i\Theta(\pm t)f(t)$ is

$$g(\omega) = \pm \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega' \frac{f(\omega')}{\omega - \omega' \pm i0^+}.$$

To see this we Fourier back-transform $g(\omega)$:

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega e^{-i\omega t} g(\omega) = \pm \frac{1}{4\pi^2} \int_{-\infty}^{\infty} d\omega e^{-i\omega t} \int_{-\infty}^{\infty} d\omega' \frac{f(\omega')}{\omega - \omega' \pm i0^+} \quad (41)$$

We first perform the integral over ω and use the trick of closing the integration path by an infinitely large semicircle and use the theorem of residues. We denote $\omega = \omega_1 + i\omega_2$ so that $e^{-i\omega t} = e^{\omega_2 t} e^{-i\omega_1 t}$. For $t > 0$ ($t < 0$) we therefore have to close in the lower (upper) half-plane to guarantee that the semicircle gives no contribution to the integral. Let us for simplicity consider the upper sign in (41). Then, as a function of ω the integrand has a pole at $\omega = \omega' - i0^+$ and if we close in the upper half-plane, the integration path does not enclose this pole so that the ω -integral vanishes. This happens for $t < 0$ so that the result will be proportional to $\Theta(t)$. If we close along the lower half-plane the integration path encloses the pole in clock-wise fashion so we get $-2\pi i$ times the residue:

$$-i\Theta(t) \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega' e^{-i\omega' t} f(\omega') = -i\Theta(t)f(t),$$

which proves the theorem.

B A theorem on determinants

Here we prove the identity

$$\frac{\partial \ln(\det A)}{\partial A_{ij}} = A_{ji}^{-1}.$$

We use Laplace's formula and expand $\det(A)$ in terms of minors

$$\det(A) = \sum_{l=1, n} (-1)^{i+l} A_{il} M_{il}.$$

Since none of the minors M_{il} contains the element A_{ij} , we find

$$\frac{\partial \ln(\det A)}{\partial A_{ij}} = \frac{(-1)^{i+j} M_{ij}}{\det(A)}$$

Next, the i^{th} column of A^{-1} is the solution of the equation system

$$Ac = e_i$$

where e_i is the i^{th} column of the unit matrix. This has all elements equal to zero, except for the i^{th} , which is one. We use Cramer's rule and find for the j^{th} element of the i^{th} column

$$A_{ji}^{-1} = \frac{\det(\bar{A}_j)}{\det(A)},$$

where $\bar{A}_j$ is the matrix where the j^{th} column has been replaced by e_i . Now we use again Laplace's formula for $\det(\bar{A}_j)$ and obtain

$$A_{ji}^{-1} = \frac{(-1)^{i+j} M_{ij}}{\det(A)}$$

which proves the theorem.

As an application we assume that the matrix elements of A are functions of some parameter α .

Then we find

$$\frac{\partial \ln(\det A)}{\partial \alpha} = \sum_{i,j} \frac{\partial \ln(\det A)}{\partial A_{ij}} \frac{\partial A_{ij}}{\partial \alpha} = \sum_{i,j} A_{ji}^{-1} \frac{\partial A_{ij}}{\partial \alpha} = \text{Tr} \left(A^{-1} \frac{\partial A}{\partial \alpha} \right).$$

References

- [1] J.M. Luttinger and J.C. Ward, Phys. Rev. **118**, 1417 (1960)
- [2] J.M. Luttinger, Phys. Rev. **119**, 1153 (1960)
- [3] G. Baym and L.P. Kadanoff, Phys. Rev. **124**, 287 (1961)
- [4] G. Baym, Phys. Rev. **127**, 1391 (1962)
- [5] A.A. Abrikosov, L.P. Gorkov and I.E. Dzyaloshinski: *Methods of Quantum Field Theory in Statistical Physics* (Prentice-Hall, New Jersey 1964)
- [6] A.L. Fetter and J.D. Walecka, *Quantum Theory of Many-Particle Systems* (McGraw-Hill, San Francisco, 1971)
- [7] J.W. Negele and H. Orland, *Quantum Many-Particle Systems* (Addison-Wesley, Redwood, 1988)
- [8] G. Baym and N.D. Mermin, J. Math. Phys. **2**, 232 (1961)
- [9] I. Dzyaloshinskii, Phys. Rev. B **68**, 085113 (2003)
- [10] J.M. Luttinger, Phys. Rev. **121**, 942 (1961)
- [11] M. Potthoff, Condens. Mat. Phys. **9**, 557 (2006)
- [12] M. Potthoff in A. Avella and F. Mancini (eds.): *Theoretical Methods for Strongly Correlated Systems* (Springer, 2011); see also preprint arXiv:11082183.
- [13] L. Hedin, Phys. Rev. **139**, A796 (1965)
- [14] N.E. Bickers, D.J. Scalapino, and S.R. White Phys. Rev. Lett. **62**, 961 (1989)

13 Introduction to Many-Body Green Functions In and Out Of Equilibrium

James K. Freericks

Department of Physics, Georgetown University
Reiss Science Building, 37th and O Streets, NW
Washington, DC 20057, USA

Contents

1	Introduction	2
2	Green functions in equilibrium and the Lehmann representation	3
3	Green functions out of equilibrium and the “contour”	5
4	The self-energy and the equation of motion	7
5	Incorporating an external field and the issue of gauge invariance	8
6	Nonequilibrium dynamical mean-field theory	11
7	Numerical strategies to calculate the Green function	13
8	Examples	18
9	Conclusions	24

1 Introduction

This chapter is all about Green functions. Many-body Green functions are a marvelous tool to employ in quantum-mechanical calculations. They require a rather large set of mathematical machinery to develop their theory and a sophisticated computational methodology to determine them for anything but the simplest systems. But what do we really use them for? Essentially, there are two main uses: the first is to compute the thermodynamic expectation value of $c^\dagger c$, which allows us to determine all interesting single-particle expectation values, like the kinetic energy, the momentum distribution, etc. The second use is to determine the many-body density of states, which tells us how the quantum states are distributed in momentum and energy. In large dimensions, Green functions can also be used to determine some two-particle correlation functions like the optical conductivity, because the vertex corrections vanish. It seems like this is a lot of work to end up with only a few concrete results. But that is the way it is. No one has figured out any way to do this more efficiently. If you determine how to, fame and fortune are likely to follow!

Before jumping into the full glory of the theory, we will give just a brief history. This is one that focuses on where, in my opinion, the critical ideas originated. It is not intended to be exhaustive or complete, and, of course, I may be wrong about where the different ideas came from. The equilibrium theory for Green functions was developed primarily in the 1950s. The Lehmann representation [1] was discovered in 1954. The Russian school developed much of the perturbative approach to Green functions, which is summarized in the monograph of Abrikosov, Gorkov, and Dzyaloshinski [2]. Joaquin Luttinger and John Ward developed a functional approach [3] and Matsubara determined how to solve for thermal Green functions [4]. But the reference I like the most for clarifying many-body Green functions both in and out of equilibrium is the monograph by Leo Kadanoff and Gordon Baym [5]. Their approach is one we will follow, at least in spirit, in this chapter. Of course, Keldysh's perspective [6] was also important. Serious numerical calculations of Green functions (in real time and frequency) only began with the development of dynamical mean-field theory in the 1990s [7]. The generalization to nonequilibrium began only in the mid 2000's [8–10], but it was heavily influenced by earlier work of Uwe Brandt [11–13].

Now that we have set the stage for you as to where we are going, get ready for the ride. The theory is beautiful, logical, abstract, and complex. Mastering it is a key in becoming a many-body theorist. We begin in section 2 with a discussion of equilibrium Green functions focused on the Lehmann representation. Section 3, generalizes and unifies the formalism to the contour-ordered Green function. Section 4 introduces the self-energy and the equation of motion. Section 5 illustrates how to include an external electric field. Section 6 introduces how one solves these problems within the nonequilibrium dynamical mean-field theory approach. Numerics are discussed in section 7, followed by examples in section 8. We conclude in section 9.

2 Green functions in equilibrium and the Lehmann representation

We begin with the definition of four Green functions in equilibrium: the retarded $G_{ij\sigma}^R(t)$, advanced $G_{ij\sigma}^A(t)$, lesser $G_{ij\sigma}^<(t)$ and greater $G_{ij\sigma}^>(t)$. These Green functions are time-dependent, complex thermal expectation values weighted by the thermal density matrix $\exp(-\beta\mathcal{H})/\mathcal{Z}$, with $\mathcal{Z} = \text{Tr} \exp(-\beta\mathcal{H})$ the partition function. They are defined in terms of the (anticommuting) fermionic creation (annihilation) operators for an electron at site i with spin σ : $c_{i\sigma}^\dagger$ ($c_{i\sigma}$). Time dependence is handled in the Heisenberg picture, where the operators depend on time and the states do not, given by $\mathcal{O}(t) = U^\dagger(t)\mathcal{O}U(t)$. For a time-independent Hamiltonian, the time-evolution operator $U(t)$ is simple and expressed as $U(t) = \exp(-i\mathcal{H}t)$, where we set $\hbar = 1$. The four Green functions then become

$$G_{ij\sigma}^R(t) = -i\theta(t)\frac{1}{\mathcal{Z}}\text{Tr}e^{-\beta\mathcal{H}}\{c_{i\sigma}(t), c_{j\sigma}^\dagger(0)\}_+, \quad (1)$$

$$G_{ij\sigma}^A(t) = i\theta(-t)\frac{1}{\mathcal{Z}}\text{Tr}e^{-\beta\mathcal{H}}\{c_{i\sigma}(t), c_{j\sigma}^\dagger(0)\}_+, \quad (2)$$

$$G_{ij\sigma}^<(t) = i\frac{1}{\mathcal{Z}}\text{Tr}e^{-\beta\mathcal{H}}c_{j\sigma}^\dagger(0)c_{i\sigma}(t), \quad (3)$$

$$G_{ij\sigma}^>(t) = -i\frac{1}{\mathcal{Z}}\text{Tr}e^{-\beta\mathcal{H}}c_{i\sigma}(t)c_{j\sigma}^\dagger(0). \quad (4)$$

Here, $\theta(t)$ is the Heaviside unit step function, which vanishes for $t < 0$, is equal to 1 for $t > 0$ while we choose it to be $1/2$ for $t = 0$, and the symbol $\{\cdot, \cdot\}_+$ denotes the anticommutator. Note that the fermionic operators satisfy the canonical anticommutation relations, namely $\{c_{i\sigma}, c_{j\sigma'}^\dagger\}_+ = \delta_{ij}\delta_{\sigma\sigma'}$. These Green functions can also be defined in momentum space, with the obvious changes of the subscripts from real space to momentum space—the only significant modification is that the Green functions will be *diagonal* in momentum for translationally invariant systems—which is all we will study in this chapter.

These definitions may look like they come out of the blue, but the key point to realize is that when $t = 0$, the lesser Green function provides precisely the expectation value we need to be able to calculate any single-particle expectation value. We will see in a moment that the retarded Green function also leads to the momentum-dependent spectral function and the density of states. The best way to understand and make sense of these Green functions is with the Lehmann representation [1]. This allows us to explicitly determine the Green functions as functions of frequency via a Fourier transformation:

$$G_{ij\sigma}(\omega) = \int_{-\infty}^{\infty} dt e^{i\omega t} G_{ij\sigma}(t). \quad (5)$$

The trick is to insert the resolution of unity ($\mathbb{I} = \sum_n |n\rangle\langle n|$ where the $|n\rangle$ are the eigenstates of $\mathcal{H}$ such that $\mathcal{H}|n\rangle = E_n|n\rangle$) in between the creation and annihilation operators in the definitions of the Green functions. We illustrate how this works explicitly for the retarded Green function, where we have

$$G_{ij\sigma}^R(t) = -i\theta(t)\frac{1}{\mathcal{Z}}\sum_{mn}\left[\langle m|e^{-\beta\mathcal{H}}c_{i\sigma}(t)|n\rangle\langle n|c_{j\sigma}^\dagger|m\rangle + \langle m|e^{-\beta\mathcal{H}}c_{j\sigma}|n\rangle\langle n|c_{i\sigma}(t)|m\rangle\right]. \quad (6)$$

We now use the fact that these states are eigenstates of $\mathcal{H}$ and the Heisenberg representation for the time dependent operators to find

$$G_{ij\sigma}^R(t) = -i\theta(t) \sum_{mn} \frac{e^{-\beta E_m}}{\mathcal{Z}} \left[e^{i(E_m - E_n)t} \langle m | c_{i\sigma} | n \rangle \langle n | c_{j\sigma}^\dagger | m \rangle + e^{-i(E_m - E_n)t} \langle m | c_{j\sigma}^\dagger | n \rangle \langle n | c_{i\sigma} | m \rangle \right]. \quad (7)$$

Next, we interchange $m \leftrightarrow n$ in the second term to give

$$G_{ij\sigma}^R(t) = -i\theta(t) \frac{1}{\mathcal{Z}} \sum_{mn} \left(e^{-\beta E_m} + e^{-\beta E_n} \right) e^{i(E_m - E_n)t} \langle m | c_{i\sigma} | n \rangle \langle n | c_{j\sigma}^\dagger | m \rangle. \quad (8)$$

It is now time to do the Fourier transform. We achieve this by shifting $\omega \rightarrow \omega + i0^+$ to make the integrand vanish in the large-time limit, which finally results in

$$G_{ij\sigma}^R(\omega) = \frac{1}{\mathcal{Z}} \sum_{mn} \frac{e^{-\beta E_m} + e^{-\beta E_n}}{\omega + E_m - E_n + i0^+} \langle m | c_{i\sigma} | n \rangle \langle n | c_{j\sigma}^\dagger | m \rangle. \quad (9)$$

This is already of the form that we can discover some interesting things. If we recall the Dirac relation $\frac{1}{\omega + i0^+} = \frac{\mathcal{P}}{\omega} - i\pi\delta(\omega)$ (with $\mathcal{P}$ denoting the principal value), we find the local density of states via

$$A_{i\sigma}(\omega) = -\frac{1}{\pi} \text{Im} G_{ii\sigma}^R(\omega) = \sum_{mn} \delta(\omega + E_m - E_n) \frac{e^{-\beta E_m} + e^{-\beta E_n}}{\mathcal{Z}} \left| \langle m | c_{i\sigma} | n \rangle \right|^2. \quad (10)$$

This density of states satisfies a sum rule. In particular,

$$\int_{-\infty}^{\infty} d\omega A_{i\sigma}(\omega) = \frac{1}{\mathcal{Z}} \sum_n e^{-\beta E_n} \left(c_{i\sigma} c_{i\sigma}^\dagger + c_{i\sigma}^\dagger c_{i\sigma} \right) = 1, \quad (11)$$

which follows from the anticommutation relation of the fermionic operators; note that you can also see this directly by taking the limit $t \rightarrow 0^+$ in Eq. (1). This result holds in momentum space too: the retarded Green function becomes

$$G_{k\sigma}^R(t) = -i\theta(t) \frac{1}{\mathcal{Z}} \text{Tr} e^{-\beta \mathcal{H}} \{ c_{k\sigma}(t), c_{k\sigma}^\dagger(t) \}_+ \quad (12)$$

$$G_{k\sigma}^R(\omega) = \frac{1}{\mathcal{Z}} \sum_{mn} \frac{e^{-\beta E_m} + e^{-\beta E_n}}{\omega + E_m - E_n + i0^+} \langle m | c_{k\sigma} | n \rangle \langle n | c_{k\sigma}^\dagger | m \rangle, \quad (13)$$

with the operators being the momentum space operators, determined by a Fourier transform from real space:

$$c_{k\sigma} = \frac{1}{|\Lambda|} \sum_{j \in \Lambda} c_{j\sigma} e^{-i\mathbf{k} \cdot \mathbf{R}_j} \quad \text{and} \quad c_{k\sigma}^\dagger = \frac{1}{|\Lambda|} \sum_{j \in \Lambda} c_{j\sigma}^\dagger e^{i\mathbf{k} \cdot \mathbf{R}_j} \quad (14)$$

Here, the lattice with periodic boundary conditions is denoted by $\Lambda = \{\mathbf{R}_i\}$ and $|\Lambda|$ is the number of lattice sites in the system. We usually do not use a bold font for the momentum label in the subscript of the creation or annihilation operator; one should be able to figure out

from the context whether we are working in real space or momentum space. The momentum-space Green functions have a similarly defined spectral function $A_{k\sigma}(k, \omega) = -\text{Im}G_{k\sigma}^R(\omega)/\pi$, which also satisfies a sum rule given by the integral over all frequency being equal to one. One can easily show that the advanced Green function in frequency space is equal to the Hermitian conjugate of the retarded Green function, because we need to shift $\omega \rightarrow \omega - i0^+$ to control the integrand at the lower limit. Hence, we have $G_{ij\sigma}^A(\omega) = G_{ji\sigma}^{R*}(\omega)$.

We also sketch how the calculation works for the lesser and greater Green functions. Here, we have to control the integrand at both endpoints of the integral. To do so, we split it at 0 and introduce the appropriate infinitesimal convergence factor in each piece. The rest of the calculation proceeds as above and one finds that the two pieces have their imaginary parts add, so we finally obtain

$$G_{ij\sigma}^<(\omega) = 2i\pi \frac{1}{\mathcal{Z}} \sum_{mn} \delta(\omega + E_m - E_n) e^{-\beta E_n} \langle m | c_{i\sigma} | n \rangle \langle n | c_{j\sigma}^\dagger | m \rangle. \quad (15)$$

The greater Green function similarly becomes

$$G_{ij\sigma}^>(\omega) = -2i\pi \frac{1}{\mathcal{Z}} \sum_{mn} \delta(\omega + E_m - E_n) e^{-\beta E_m} \langle m | c_{i\sigma} | n \rangle \langle n | c_{j\sigma}^\dagger | m \rangle. \quad (16)$$

Note that we can use the delta function to immediately infer an important identity, namely that $G_{ij\sigma}^<(\omega) = -\exp(-\beta\omega) G_{ij\sigma}^>(\omega)$. Combining this result with the fact that $\text{Im} G_{ij\sigma}^R(\omega) = \text{Im}[G_{ij\sigma}^<(\omega) - G_{ij\sigma}^>(\omega)]/2$, we finally learn that

$$G_{ij\sigma}^>(\omega) = -2if(\omega) \text{Im} G_{ij\sigma}^R(\omega) \quad \text{and} \quad G_{ij\sigma}^<(\omega) = 2i[1 - f(\omega)] \text{Im} G_{ij\sigma}^R(\omega) \quad (17)$$

with $f(\omega) = 1/(1 + \exp(-\beta\omega))$. This equilibrium relationship is sometimes called the fluctuation-dissipation theorem.

3 Green functions out of equilibrium and the “contour”

Now we move on to nonequilibrium where the Hamiltonian is time-dependent $\mathcal{H}(t)$. The evolution operator depends on two times and becomes $U(t, t') = \mathcal{T}_t \exp \left[-i \int_{t'}^t d\bar{t} \mathcal{H}(\bar{t}) \right]$. It satisfies the semigroup property $U(t, t'') U(t'', t') = U(t, t')$, with $U(t, t) = \mathbb{I}$ and $i\partial_t U(t, t') = \mathcal{H}(t) U(t, t')$. The Green functions in nonequilibrium are defined by the same equations as before, namely Eqs. (1–4), except now they depend on two times, with t' being the argument of the $c^\dagger$ operator. All of these Green functions require us to evaluate one of two matrix elements. We examine one of them here and assume $t > t'$ for concreteness

$$\frac{1}{\mathcal{Z}} \text{Tr} e^{-\beta\mathcal{H}} c_{i\sigma}(t) c_{j\sigma}^\dagger(t') = \frac{1}{\mathcal{Z}} \text{Tr} e^{-\beta\mathcal{H}} U^\dagger(t, -\infty) c_{i\sigma} U(t, -\infty) U^\dagger(t', -\infty) c_{j\sigma}^\dagger U(t', -\infty) \quad (18)$$

$$= \frac{1}{\mathcal{Z}} \text{Tr} e^{-\beta\mathcal{H}} U(-\infty, t) c_{i\sigma} U(t, t') c_{j\sigma}^\dagger U(t', -\infty) \quad (19)$$

where we used the facts that $U^\dagger(t_1, t_2) = U(t_2, t_1)$ and the semigroup property to show that $U(t, -\infty) U^\dagger(t', -\infty) = U(t, t')$. The time-evolution operators, including time evolution in

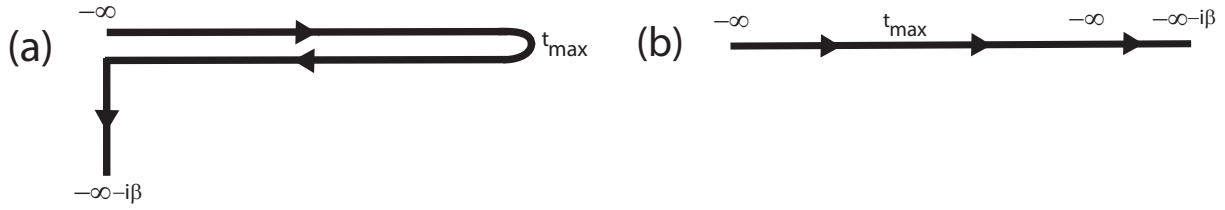


Fig. 1: (a) Kadanoff-Baym-Keldysh contour used in the construction of the contour-ordered Green function. The contour starts at $-\infty$, runs out to the maximum of t and t' , runs back to $-\infty$ and then runs parallel to the imaginary axis for a distance equal to β . (b) The contour can be “stretched” into a straight line as indicated here, which is convenient for properly implementing time-ordering along the contour.

the imaginary axis direction for the $\exp(-\beta\mathcal{H})$ term, can be seen to all live on the so-called Kadanoff-Baym-Keldysh contour, which is shown in Fig. 1. Starting with time at $-\infty$, we evolve forward to t' , then apply $c^\dagger$, evolve forward to time t , apply the c operator, and then evolve *backwards* in time to $-\infty$. The lesser and greater Green functions are easily determined in this fashion. The retarded or advanced Green functions can then be found by taking the appropriate differences of greater and lesser functions with the convention that $t > t'$ for the retarded case and $t < t'$ for the advanced case. Now that we have these Green functions, we can generalize the definition of the Green function to allow both times to lie anywhere on the contour. We take the time-ordered Green function, with the time-ordering taking place along the contour, and we call it the contour-ordered Green function

$$G_{ij\sigma}^c(t, t') = \begin{cases} -i \frac{1}{\mathcal{Z}} \text{Tr} e^{-\beta\mathcal{H}} c_{i\sigma}(t) c_{j\sigma}^\dagger(t') & \text{for } t >_c t' \\ i \frac{1}{\mathcal{Z}} \text{Tr} e^{-\beta\mathcal{H}} c_{j\sigma}^\dagger(t') c_{i\sigma}(t) & \text{for } t <_c t' \end{cases} \quad (20)$$

where the c subscript on the less-than or greater-than symbol is to denote whether one time is ahead of or behind the other on the contour, regardless of the numerical values of t and t' . This Green function is the workhorse of nonequilibrium many-body physics—we use it to calculate anything that can be calculated with the single-particle Green functions.

Note that in these lectures, we will work with the contour-ordered Green function itself, which is a continuous matrix operator in time. In many other works, the Green function is further decomposed into components where the times lie on the same or different branches of the real contour (2×2 matrix) [6] or of the two real and one imaginary branches of the contour (3×3 matrix) [14]. In the 3×3 case, one has real, imaginary and mixed Green functions, the last ones determining the so-called initial correlations. In the 2×2 case, it is common to transform the matrix to make it upper triangular, expressing the retarded, advanced and so-called Keldysh Green functions. In general, I find that these decompositions and transformations make the material more difficult than it actually is and it is best to defer using them until one has gained some mastery of the basics. Here, we will work solely with the contour-ordered functions.

4 The self-energy and the equation of motion

The next step in our analysis is to derive the equation of motion. We start from the definition of the contour-ordered Green function in Eq. (20) and write it using a unit step function on the contour, $\theta_c(t, t')$, which is one for t ahead of t' on the contour and zero for t behind t' . This theta function can be used to impose the two conditions. Its derivative with respect to time becomes a delta function on the contour $\partial_t \theta_c(t, t') = \delta_c(t, t')$. The delta function on the contour satisfies

$$\int_c dt' \delta(t, t') f(t') = f(t), \quad (21)$$

for any time t on the contour (including those on the imaginary spur). Now a simple derivative of the Green function with respect to time yields the equation of motion (we find it convenient to illustrate this in momentum space)

$$\begin{aligned} i \frac{\partial}{\partial t} G_{k\sigma}^c(t, t') &= \frac{1}{\mathcal{Z}} \text{Tr} e^{-\beta \mathcal{H}} \{c_{k\sigma}(t), c_{k\sigma}^\dagger(t')\}_+ \delta_c(t, t') \\ &\quad + i \theta_c(t, t') \frac{1}{\mathcal{Z}} \text{Tr} e^{-\beta \mathcal{H}} [\mathcal{H}(t), c_{k\sigma}(t)] c_{k\sigma}^\dagger(t') \\ &\quad - i \theta_c(t', t) \frac{1}{\mathcal{Z}} \text{Tr} e^{-\beta \mathcal{H}} c_{k\sigma}^\dagger(t') [\mathcal{H}(t), c_{k\sigma}(t)] \\ &\stackrel{!}{=} \delta_c(t, t') + (\varepsilon_k(t) - \mu) G_{k\sigma}^c(t, t') + \int_c dt'' \Sigma_c(t, t'') G_{k\sigma}^c(t'', t'). \end{aligned} \quad (22)$$

Here, we assumed that the Hamiltonian (in the Schrödinger picture) takes the form

$$\mathcal{H}_s(t) = \sum_{k\sigma} (\varepsilon_k(t) - \mu) c_{k\sigma}^\dagger c_{k\sigma} + V(t), \quad (24)$$

where $\varepsilon_k(t)$ is a time-dependent bandstructure and μ is the chemical potential. The symbol $V(t)$ is the potential energy operator of the Hamiltonian. The commutator of the potential with the fermionic destruction operator is complicated—we introduce the contour-ordered self-energy Σ_c to describe that cumbersome term—which can be viewed as the definition of the self-energy. The self-energy appears in a convolution with the Green function (over the entire contour) in Eq. (23). The equation of motion can be rearranged to look like

$$\int_c dt'' \left(\left(i \frac{\partial}{\partial t} - \varepsilon_k(t) + \mu \right) \delta_c(t, t'') + \Sigma_c(t, t'') \right) G_{k\sigma}^c(t'', t') = \delta_c(t, t'). \quad (25)$$

As we mentioned before, the Green functions are continuous matrix operators. Seen in that light, it should be clear that the object inside the outer braces is the matrix inverse of the Green function (since matrix multiplication, implied by the integration over the contour, has the matrix defined by the braces multiplied by G^c equaling the identity matrix).

At this stage it may look like we are pulling the wool over your eyes. In some sense we are. We have hidden the difficult question of how to evaluate the commutator of the potential with the destruction operator, by simply introducing a self-energy which makes the complications go away. This really would lead us nowhere if we did not have another means by which we can

determine the self-energy. Fortunately we do, and it will be introduced below. You will have to wait to see how that fits in. All of the challenge in many-body physics boils down to this question of “how can we calculate the self-energy?”

Now is also a good time to discuss the issue of how do we work with these complicated objects? If you look carefully, you will see we are working with matrices. Admittedly they are continuous matrix operator equations, but like any continuous object, we discretize them when we want to work with them on a computer. And indeed, that is what we will do here when we describe the DMFT algorithm below. We always extrapolate the discretization to zero to recover the continuous matrix operator results. It is important to recognize these facts now, as it will make your comfort level with the approach much higher as we develop the subject further.

5 Incorporating an external field and the issue of gauge invariance

I always found it odd in dealing with fields in quantum mechanics to learn that we do not always respect Maxwell’s equations. This should not come as a surprise to you, because Maxwell’s equations are relativistically invariant, and quantum mechanics only becomes relativistically invariant with a properly constructed relativistic quantum field theory. Nevertheless, we choose to neglect even more aspects of Maxwell’s equations than just relativistic invariance. This becomes necessary because the quantum problem cannot be solved unless we do this (and, fortunately, the neglected terms are small).

The main issue we face is that we want the field to be spatially uniform but varying in time. Maxwell says such fields do not exist, but we press on anyway. You see, a spatially uniform field means that I still have translation invariance and that is needed to make the analysis simple enough that it can be properly carried out. We justify this in part because optical wavelengths are much larger than lattice spacings and because effects of magnetic fields are small due to the extra $1/c$ factors multiplying them. Even with this simplification, properly understanding how to introduce a large amplitude electric field is no easy task.

The field will be described via the Peierls’ substitution [15]. This is accomplished by shifting $k \rightarrow k - eA(t)$, which becomes $k + eEt$, for a constant (dc) field. This seems like the right thing to do, because it parallels the minimal substitution performed in single-particle quantum mechanics when incorporating an electric field, but we have to remember that we are now working with a system projected to the lowest lying s -orbital band, so it may not be obvious that this remains the correct thing to do. Unfortunately, there is no rigorous proof that this result is correct, but one can show that if we work in a single band (neglecting interband transitions) and have inversion symmetry in the lattice, then the field generated by a scalar potential, and the field generated by the equivalent vector potential will yield identical noninteracting Green functions [16]. We then assume that adding interactions to the system does not change this correspondence, so we have an approach that is exact for a single-band model. One should bear in mind that there may be subtle issues with multiple bands and the question of how to precisely

work with gauge-invariant objects in multiband models has not been completely resolved yet. Fortunately, for model system calculations, working with a single band is often sufficient to illustrate the nonequilibrium phenomena. This is what we do here.

It is now time to get concrete in describing the Hamiltonian we will be working with, which is the spinless Falicov-Kimball model [17]. This model is nice to work with because it does possess strongly correlated physics (Mott transition and order-disorder transitions), but it can be solved exactly within DMFT in both equilibrium and nonequilibrium. Hence, it provides a nice test case to illustrate the approach we take here. The Hamiltonian, in an electric field is

$$\mathcal{H}_S(t) = \sum_k (\varepsilon(k - eA(t)) - \mu) c_k^\dagger c_k + E_f \sum_i f_i^\dagger f_i + U \sum_i c_i^\dagger c_i f_i^\dagger f_i, \quad (26)$$

where we employed the Peierls' substitution in the bandstructure. Here, the conduction electrons (c) are spinless and do not interact with themselves. They interact instead with the localized electrons (f), when both electrons are situated on the same lattice site— E_f is the local site energy of the f -electrons. Note that we use k for momentum and i for position in $\mathcal{H}$. All of the time dependence is in the kinetic-energy term, because the field does not couple to the localized electrons (due to the fact that they do not hop). Hence, we have an effective single-band model and the discussion above about the applicability of the Peierls' substitution applies.

It is instructive at this stage to determine the noninteracting Green function exactly. We do this by determining the Heisenberg creation and annihilation operators from which we form the Green function. In the case where $U = 0$, one can immediately find that the destruction operator (in the Heisenberg picture) becomes

$$c_{k\sigma}(t) = e^{-i \int_{-\infty}^t d\bar{t} (\varepsilon(k - eA(\bar{t})) - \mu)} c_{k\sigma}^\dagger c_{k\sigma} e^{-i \int_{-\infty}^t d\bar{t} (\varepsilon(k - eA(\bar{t})) - \mu)} c_{k\sigma}^\dagger c_{k\sigma} = e^{-i \int_{-\infty}^t d\bar{t} (\varepsilon(k - eA(\bar{t})) - \mu)} c_{k\sigma}. \quad (27)$$

There is no need for time ordering here, because the Hamiltonian commutes with itself at different times. *This simplification is what allows one to solve the problem easily.* By taking the Hermitian conjugate of Eq. (27), one can substitute in the definitions of the Green functions and immediately find that

$$G_{k\sigma}^<(t, t') = i \frac{1}{\mathcal{Z}} \text{Tr} e^{-\beta \mathcal{H}} e^{-i \int_{t'}^t d\bar{t} (\varepsilon(k - eA(\bar{t})) - \mu)} c_{k\sigma}^\dagger c_{k\sigma} = i e^{-i \int_{t'}^t d\bar{t} (\varepsilon(k - eA(\bar{t})) - \mu)} \langle n_{k\sigma} \rangle, \quad (28)$$

where $\langle n_{k\sigma} \rangle = \text{Tr} e^{-\beta \mathcal{H}} c_{k\sigma}^\dagger c_{k\sigma}$ is the *initial equilibrium* momentum distribution. We similarly find that the noninteracting retarded Green function in nonequilibrium is

$$G_{k\sigma}^R(t, t') = -i \theta(t - t') e^{-i \int_{t'}^t d\bar{t} (\varepsilon(k - eA(\bar{t})) - \mu)} \quad (29)$$

there is no equilibrium expectation value here because the anticommutator of the equal-time fermionic operators is 1. To find the corresponding local Green functions, we simply sum these expressions over all momenta.

We next discuss the issue of gauge invariance. This issue is important, because we are required to perform calculations in a specific gauge but any physically measurable quantity must be gauge-invariant. So how does this come about? To begin, one should notice that the Peierls' substitution guarantees that any quantity that is local, like the current, is gauge invariant, because the sum over all momenta includes every term, even if the momenta are shifted due to the Peierls substitution. Momentum-dependent quantities, however, appear to depend directly on the vector potential. To properly formulate a gauge-invariant theory requires us to go back and formulate the problem (using a complete basis set, not a single-band downfolding) properly to illustrate the gauge-invariant nature of the observables. No one has yet been able to do this for a system with a nontrivial bandstructure. For a single-band model, we instead directly enforce gauge invariance using the so-called gauge-invariant Green functions, introduced by Bertoncini and Jauho [18]. These Green functions are constructed to be manifestly gauge invariant—the change in the phase of the creation and destruction operators is chosen to precisely cancel the change in the phase from the gauge transformation. We determine them in momentum space by making a shift

$$k \rightarrow k(t, t') = k + \int_{-\frac{1}{2}}^{\frac{1}{2}} d\lambda A \left(\frac{t+t'}{2} + \lambda(t-t') \right). \quad (30)$$

Note that one needs to be careful in evaluating this integral in the limit as $t \rightarrow t'$. In this limit, we find $k \rightarrow k + A(t)$. The explicit formula for the gauge-invariant Green function, which uses a tilde to denote that it is gauge-invariant, is

$$\tilde{G}_{k\sigma}(t, t') = G_{k(t,t')\sigma}(t, t'). \quad (31)$$

The demonstration of the gauge-invariance is given in Ref. [18]. Note that the local Green functions are always manifestly gauge-invariant because the map $k \rightarrow k(t, t')$ is a “one-to-one and onto” map for any fixed set of times.

We end this section with a discussion of the current operator and how it is related to gauge invariance. The current operator is calculated in a straightforward way via the commutator of the Hamiltonian with the charge polarization operator. We do not provide the details here, but simply note that they can be found in most many-body physics textbooks. The end result is that

$$j = \sum_{\sigma} \sum_k v_k c_{k\sigma}^{\dagger} c_{k\sigma}. \quad (32)$$

Here, we have introduced the band velocity which is given by $v_k = \nabla \varepsilon_k$. The gauge-invariant form of the expectation value for the current is then

$$\langle j \rangle = -i \sum_{\sigma} \sum_k v_k \tilde{G}_{k\sigma}^{<}(t, t). \quad (33)$$

We can re-express in terms of the original Green functions, by shifting $k \rightarrow k - A(t)$. This yields the equivalent formula

$$\langle j \rangle = -i \sum_{\sigma} \sum_k v_{k-A(t)} G_{k\sigma}^{<}(t, t). \quad (34)$$

As you can easily see, these two results are identical by simply performing a uniform shift of the momentum in the summation. Generically, this is how different expectation values behave—in the gauge-invariant form, one evaluates them the same as in linear response, but uses the gauge-invariant Green functions, while in the non-gauge-invariant form, one evaluates them in a form similar to linear response, but appropriately shifts the momentum in the terms that multiply the Green functions (which is a beyond-linear-response “correction”).

6 Nonequilibrium dynamical mean-field theory

The premise of dynamical mean-field theory (DMFT) is that the self-energy of an impurity in an appropriately chosen time-dependent field is the same as the self-energy on the lattice. In general, we know this cannot be true, because the lattice self-energy depends on momentum, and hence is not local. But, if we choose the hopping matrix elements to scale like $t = t^*/2\sqrt{d}$ for a d -dimensional lattice, then one can rigorously show that in the limit $d \rightarrow \infty$, the self-energy does indeed become local [19]. So, in this limit, the DMFT assumption holds and it provides another limiting case where one can exactly solve the many-body problem (the other limit being that of $d = 1$, where one can use the Bethe ansatz).

The physical picture to keep in mind, for DMFT, is to think of looking at what happens just on a single site of the lattice. As a function of time, we will see electrons hop onto and off of the site. If we can adjust the time-dependent field for the impurity in such a way that it produces the same behavior in time on the impurity, then the impurity site will look like the lattice. This motivates the following iterative algorithm [20] to solve these models: (1) Begin with a guess for the self-energy ($\Sigma = 0$ often works); (2) sum the momentum-dependent Green function over all momenta to determine the local Green function

$$G_{ii\sigma}^c(t, t') = \sum_k \left((G_{k\sigma}^c(U=0))^{-1} - \Sigma \right)_{t,t'}^{-1}, \quad (35)$$

with $G_{k\sigma}^c(U=0)$ the noninteracting momentum-dependent contour-ordered Green function on the lattice and the subscript denoting the t, t' element of the matrix inverse); (3) Use Dyson's equation to determine the effective medium G_0 via

$$(G_{0\sigma}^c)^{-1}_{t,t'} = (G_{ii\sigma}^c)^{-1}_{t,t'} + \Sigma^c(t, t'); \quad (36)$$

(4) solve the impurity problem in the extracted effective medium (which determines the time evolution on the impurity); (5) extract the self-energy for the impurity using an appropriately modified Eq. (36); and (8) repeat steps (2–7) until the self-energy has converged to the desired level of accuracy. The difference from the conventional iterative DMFT algorithm is that the objects worked with are now continuous matrix operators in time rather than simple functions of frequency. Sometimes this creates a roadblock for students to follow how the algorithm works, but there is no reason why it should. Finally, we note that because each inverse of a Green function here implies a matrix inversion, we will want to organize the calculation in such a way as to minimize the inversions. This will be explained in more detail when we discuss numerical implementations below.

One of the simplifications in equilibrium DMFT is that the fact that the self-energy has no momentum dependence allows us to perform the summation over momentum via a single integral over the noninteracting density of states. Things are not quite as simple for the nonequilibrium case, as we now discuss. The bandstructure for a hypercubic lattice in $d \rightarrow \infty$ is given by

$$\varepsilon_k = - \lim_{d \rightarrow \infty} \frac{t^*}{\sqrt{d}} \sum_{i=1}^d \cos(k_i). \quad (37)$$

The central limit theorem tells us that ε_k is distributed according to a Gaussian distribution via

$$\rho(\varepsilon) = \frac{1}{\sqrt{\pi t^*}} e^{-\frac{\varepsilon^2}{t^{*2}}}. \quad (38)$$

In nonequilibrium, we have a second band energy to work with

$$\bar{\varepsilon}_k = - \lim_{d \rightarrow \infty} \frac{t^*}{\sqrt{d}} \sum_{i=1}^d \sin(k_i), \quad (39)$$

because the Peierls substitution shifts $\cos(k_i) \rightarrow \cos(k_i) \cos(A_i(t)) + \sin(k_i) \sin(A_i(t))$. The joint density of states becomes the product of two Gaussians, given by

$$\rho(\varepsilon, \bar{\varepsilon}) = \frac{1}{\pi t^{*2}} e^{-\frac{\varepsilon^2}{t^{*2}} - \frac{\bar{\varepsilon}^2}{t^{*2}}}. \quad (40)$$

This second band energy can be thought of as the projection of the band velocity onto the direction of the electric field. Hence, the computation of the local Green function in step (2) is complicated by requiring a two-dimensional rather than a one-dimensional integration. This does make the numerics require significant additional resources, especially because the integrands are matrix-valued.

The most challenging part of the algorithm is step (4)—solving the impurity problem in nonequilibrium. Unfortunately, there are few techniques available to do this. Monte Carlo methods suffer from the so-called phase problem, which might be able to be tamed using expansions about a perturbative solution and restricting updates to be nearly local in time via the inchworm algorithm [21]. The other choice is perturbation theory either in the interaction (which often does not work so well, except for electron-phonon coupling) or the hybridization of the impurity model (which works well at half filling). Here, we make a different choice, and choose a simplified model that can be solved exactly, the so-called spinless Falicov-Kimball model [17]

$$\mathcal{H} = \sum_{ij} t_{ij} c_i^\dagger c_j - \mu \sum_i c_i^\dagger c_i + E_f \sum_i f_i^\dagger f_i + U \sum_i c_i^\dagger c_i f_i^\dagger f_i. \quad (41)$$

This model involves the interaction of conduction electrons (c) and localized electrons (f). The conduction electrons can hop on the lattice (we usually take the hopping only between nearest neighbors on a hypercubic lattice in the $d \rightarrow \infty$ limit) and have a chemical potential μ . The localized electrons have a site energy E_f . Both electrons interact with an on-site interaction U . The Falicov-Kimball model describes a rich set of physics, but it should not be viewed as a paradigm for all strongly correlated electrons. This is because it has some aspects that are not

seen in more common models like the Hubbard model. For example, it is never a Fermi liquid when $U \neq 0$, the conduction electron density of states is independent of temperature in the normal state, and it is never a superconductor. But, it does display a lot of rich physics including a Mott metal-insulator transition, ordered charge-density-wave phases at low temperature and even phase separation when the densities of the two species are far enough apart. In addition, the f -electron density of states does exhibit strong temperature dependence, as expected in strongly correlated systems. It also displays orthogonality catastrophe behavior. The main interest in the nonequilibrium context has focused on the fact that it has a Mott transition. The most important aspect of the model is that it can be solved exactly within DMFT—both in equilibrium and in nonequilibrium.

We will not go through the details of the equation of motion for the Green functions in the Falicov-Kimball model to show how one can solve it. The procedure is most efficiently performed using a path-integral formulation because the time-dependent field on the impurity cannot be easily expressed in terms of a Hamiltonian (unless one introduces a bath for the impurity which does this). Instead, one can employ functional methods to exactly compute the functional derivative of the partition function with respect to the dynamical mean-field, which then is employed to extract the Green function. Then, because the f -electron number is conserved in the model, we obtain the Green function of the impurity by summing over appropriately weighted combinations of the solution with no f -electrons and with one f -electron. The end result is that the impurity Green function for the Falicov-Kimball model in an effective medium is the following:

$$G_0^c(t, t') = (1 - w_1)G_0^c(t, t') + w_1 [(G_0^c)^{-1} - U\mathbb{I}]_{t, t'}^{-1}. \quad (42)$$

The symbol $w_1 = \langle f^\dagger f \rangle$ is the filling of the localized electrons (per lattice site). Since the f -electron number operator commutes with $\mathcal{H}$, the localized electron number is a constant of motion and does not change with time, even in the presence of large electric fields. Note that we will work at half-filling where the electron densities for the conduction electrons and the localized electrons are each equal to $1/2$. This point corresponds to $\mu = U/2$ and $E_f = -U/2$. Details not presented here can be found in the original literature [8, 9].

7 Numerical strategies to calculate the Green function

Now we need to sort out just how we implement this algorithm. This discussion closely follows Ref. [9]. As mentioned above, the challenge is that we are working with continuous matrix operators, with integration over the contour corresponding to matrix multiplication of these operators. Such objects cannot be directly put onto a computer. Instead, we have to first discretize the problem, replacing the continuous operators by discrete matrices. There are a few subtle points that are involved corresponding to the integration measure over the contour when we do this, but we will precisely describe how we achieve this below. In order to recover the results for the continuous matrix operators, we need to extrapolate the discretized results to the limit where the discretization size goes to zero. This is done in the most mundane way possible—simply

repeat the calculation for a number of different Δt values and use Lagrange interpolation as an extrapolation formula to the $\Delta t \rightarrow 0$ limit. We finally check sum rules of the data to verify that the extrapolation procedure worked properly. Details follow.

We first need to be concrete about how the discretization is performed. This involves N_t points on the upper real branch (ranging from t_{min} to $t_{max} - \Delta t$), N_t points on the lower real branch (ranging from t_{max} to $t_{min} + \Delta t$), and 100 points along the imaginary axis (ranging from t_{min} to $t_{min} - i\beta + 0.1i$, with $\beta = 10$); hence $\Delta t = (t_{max} - t_{min})/N_t$. We often find that fixing the number of points on the imaginary time branch rather than scaling them in the same fashion as on the real axis does not cause any significant errors, but it allows the calculations to be performed with fewer total number of points. The discrete time values on the contour become

$$t_j = \begin{cases} -t_{min} + (j-1)\Delta t & \text{for } 1 \leq j \leq N_t \\ t_{max} - (j - N_t - 1)\Delta t & \text{for } N_t+1 \leq j \leq 2N_t \\ t_{min} - 0.1i(j - 2N_t - 1) & \text{for } 2N_t+1 \leq j \leq 2N_t+100 \end{cases} \quad (43)$$

where we used the fact that the discretization along the imaginary axis is always fixed at $\Delta\tau = 0.1$ in our calculations (and we pick the initial temperature to be $T = 0.1t^*$ or $\beta = 10$). We use a leftpoint rectangular integration rule for discretizing integrals over the contour, which is implemented as follows:

$$\int_c dt f(t) = \sum_{i=1}^{2N_t+100} W_i f(t_i), \quad (44)$$

where the weights satisfy

$$W_j = \begin{cases} \Delta t & \text{for } 1 \leq j \leq N_t \\ -\Delta t & \text{for } N_t+1 \leq j \leq 2N_t \\ -0.1i & \text{for } 2N_t+1 \leq j \leq 2N_t+100 \end{cases} \quad (45)$$

The leftpoint integration rule evaluates the function at the “earliest” point (in the sense of time ordering along the contour, see Fig. 1) in the time interval that has been discretized for the quadrature rule (which is the left hand side of the interval when we are on the upper real branch and right hand side when we are on the lower real time branch).

One of the important aspects of many-body Green functions is that they satisfy a boundary condition. This is what determines whether the Green function is bosonic or fermionic. For example, you should already be familiar with the fact that the thermal Green functions are antiperiodic on the imaginary time axis. The contour-ordered Green function also satisfies a boundary condition where we identify the points t_{min} with $t_{min} - i\beta$. One can show from the definition of the contour-ordered Green function and the invariance of the trace with respect to the ordering of operators that $G_{ii\sigma}^c(t_{min}, t') = -G_{ii\sigma}^c(t_{min} - i\beta, t')$ and $G_{ii\sigma}^c(t, t_{min}) = -G_{ii\sigma}^c(t, t_{min} - i\beta)$. The proof is identical to how it is done for the thermal Green functions. It involves cyclically moving a creation operator from the left to the right and commuting the $e^{-\beta\mathcal{H}}$ term through it which employs the fact that a cyclic permutation of elements in the product of a trace does not change the value of the trace.

The delta function changes sign along the negative real time branch, and is imaginary along the last branch of the contour in order to satisfy the property that $\int_c dt' \delta_c(t, t') f(t') = f(t)$. In addition, we find that the numerics work better if the definition of the delta function is done via “point splitting” (when we calculate the inverse of a Green function) so that the delta function does not lie on the diagonal, but rather on the first subdiagonal matrix (in the limit as $\Delta t \rightarrow 0$ it becomes a diagonal operator). Because we identify the times t_{min} and $t_{min} - i\beta$, the point splitting approach to the definition of the delta function allows us to incorporate the correct boundary condition into the definition of the discretized delta function. Hence, we define the discretized delta function in terms of the quadrature weights, in the following way

$$\delta_c(t_i, t_j) = \frac{1}{W_i} \delta_{i,j+1}, \quad \text{for integration over } j, \quad (46)$$

$$= \frac{1}{W_{i-1}} \delta_{i,j+1}, \quad \text{for integration over } i, \quad (47)$$

where t_i and t_j are two points on the discretized contour as described in Eq. (43), and W_i are the quadrature weights described in Eq. (45). We have a different formula for integration over the first variable versus integration over the second variable because we are using the leftpoint quadrature rule. Note that the formulas in Eqs. (46) and (47) hold only when $i \neq 1$. When $i = 1$, the only nonzero matrix element for the discretized delta function is the $1, j = 2N_t + 100$ matrix element, and it has a sign change due to the boundary condition that the Green function satisfies. The discretization of the derivative of the delta function on the contour is more complex. It is needed to determine the inverse of the G_0^c for the impurity. The derivative is calculated by a two-point discretization that involves the diagonal and the first subdiagonal. Since all we need is the discrete representation of the operator $(i\partial_t^c + \mu)\delta_c(t, t')$, we summarize the discretization of that operator as follows

$$(i\partial_t + \mu)\delta_c(t_j, t_k) = i \frac{1}{W_j} M_{jk} \frac{1}{W_k}, \quad (48)$$

with the matrix M_{jk} satisfying

$$M_{jk} = \begin{pmatrix} 1 & 0 & 0 & & \dots & & 1+i\Delta t\mu \\ -1-i\Delta t\mu & 1 & 0 & & \dots & & 0 \\ 0 & -1-i\Delta t\mu & 1 & 0 & & & \\ & & \ddots & & & & \\ & & 0 & -1+i\Delta t\mu & 1 & 0 & \\ & & & 0 & -1+i\Delta t\mu & 1 & \\ & & & & \ddots & & \\ & & & & & -1-\Delta\tau\mu & 1 \\ & & & & & \ddots & \\ & & & & & & -1-\Delta\tau\mu & 1 \end{pmatrix}, \quad (49)$$

where $\Delta\tau = 0.1$. The top third of the matrix corresponds to the upper real branch, the middle third to the lower real branch and the bottom third to the imaginary branch. Note that the operator $(i\partial_t^c + \mu)\delta_c$ is the inverse operator of the Green function of a spinless electron with a chemical potential μ . Hence the determinant of this operator must equal the partition function of a spinless electron in a chemical potential μ , namely $1 + \exp(\beta\mu)$. This holds because

$\det G_{\text{non}}^{-1} = \mathcal{Z}_{\text{non}}$ for noninteracting electrons. Taking the determinant of the matrix M_{jk} (by evaluating the minors along the top row) gives

$$\begin{aligned} \det M &= 1 + (-1)^{2N_t+N_\tau-1} (1 + i\Delta t\mu)(-1 - i\Delta t\mu)^{N_t-1} (-1 + i\Delta t\mu)^{N_t} (-1 - \Delta\tau\mu)^{N_\tau}, \\ &\approx 1 + (1 + \Delta\tau\mu)^{N_\tau} + O(\Delta t^2), \end{aligned} \quad (50)$$

which becomes $1 + \exp(\beta\mu)$ in the limit where $\Delta t, \Delta\tau \rightarrow 0$ (N_τ is the number of discretization points on the imaginary axis). This shows the importance of the upper right hand matrix element of the operator, which is required to produce the correct boundary condition for the Green function. It also provides a check on our algebra. In fact, we chose to point-split the delta function when we defined its discretized matrix operator precisely to ensure that this identity holds.

We also have to show how to discretize the continuous matrix operator multiplication and how to find the discretized approximation to the continuous matrix operator inverse. As we mentioned above, the continuous matrix operator is described by a number for each entry i and j in the discretized contour. We have to recall that matrix multiplication corresponds to an integral over the contour, so this operation is discretized with the integration weights W_k as follows

$$\int_c d\bar{t} A(t, \bar{t}) B(\bar{t}, t') = \sum_k A(t_i, t_k) W_k B(t_k, t_j). \quad (51)$$

Thus we must multiply the columns (or the rows) of the discrete matrix by the corresponding quadrature weight factors when we define the discretized matrix. This can be done either to the matrix on the left (columns) or to the matrix on the right (rows). To calculate the inverse, we recall the definition of the inverse for the continuous matrix operator

$$\int_c d\bar{t} A(t, \bar{t}) A^{-1}(\bar{t}, t') = \delta_c(t, t'), \quad (52)$$

which becomes

$$\sum_k A(t_i, t_k) W_k A^{-1}(t_k, t_j) = \frac{1}{W_i} \delta_{ij} \quad (53)$$

in its discretized form. Note that we do not need to point-split the delta function here. Hence, the inverse of the matrix is found by inverting the matrix $W_i A(t_i, t_j) W_j$, or, in other words, we must multiply the rows and the columns by the quadrature weights before using conventional linear algebra inversion routines to find the discretized version of the continuous matrix operator inverse. This concludes the technical details for how to discretize and work with continuous matrix operators.

The next technical aspect we discuss is how to handle the calculation of the local Green function. Since the local Green function is found from the following equation:

$$G_{ii}^c(t, t') = \frac{1}{\pi t^{*2}} \int d\varepsilon \int d\bar{\varepsilon} e^{-\frac{\varepsilon^2}{t^{*2}} - \frac{\bar{\varepsilon}^2}{t'^{*2}}} \left[(\mathbb{I} - G_{\varepsilon, \bar{\varepsilon}}^{c, \text{non}} \Sigma^c)^{-1} G_{\varepsilon, \bar{\varepsilon}}^{c, \text{non}} \right]_{t, t'} \quad (54)$$

with the noninteracting Green function discussed earlier and given by

$$G_{\varepsilon, \bar{\varepsilon}}^{c, \text{non}}(t, t') = i(f(\varepsilon - \mu) - \theta_c(t, t')) e^{-i \int_{t'}^t d\bar{t} (\cos(A(\bar{t}))\varepsilon - \sin(A(\bar{t}))\bar{\varepsilon})}. \quad (55)$$

Note that we must have $A(t) = 0$ before the field is turned on. Of course an optical pump pulse also requires $\int_{-\infty}^{\infty} A(t)dt = 0$, because a traveling electromagnetic wave has no dc field component. We did not compute the matrix inverse of the noninteracting Green function because re-expressing the formula in the above fashion makes the computation more efficient (because matrix multiplication requires less operations than a matrix inverse does).

This step is the most computationally demanding step because we are evaluating the double integral of a matrix-valued function and we need to compute one matrix inverse for each integrand. Fortunately, the computation for each $(\varepsilon, \bar{\varepsilon})$ pair is independent of any other pair, so we can do this in parallel with a master-slave algorithm. In the master-slave algorithm, one CPU controls the computation, sending a different $(\varepsilon, \bar{\varepsilon})$ pair to each slave to compute its contribution to the integral and then accumulating all results. The sending of the pairs to each slave is simple. One has to carefully manage the sending of the results back to the master, because the data is a large general complex matrix and they are all being received by one node. This is precisely the situation where a communication bottleneck can occur.

In order to use as few integration points as possible, we employ Gaussian integration, because the bare density of states for both the ε and the $\bar{\varepsilon}$ integral is a Gaussian. We found it more efficient to average the results with two consecutive numbers of points in the Gaussian integration (like 100 and 101) instead of just using 200 points (which would entail twice as much computation).

The rest of the DMFT loop requires serial operations and is performed on the master to reduce communications. One does have to pay attention to convergence. In metallic phases, and with strong fields, the code will converge quickly. But when the field is small or the interactions large, one might not be able to achieve complete convergence. Often the data is still good, nevertheless. One also should not make the convergence criterion too stringent. Usually 4 digits of accuracy is more than enough for these problems. In most cases, one will need to iterate as many as 50-100 times for hard cases. But many results can be obtained with ten iterations or less.

Finally, one has to repeat the calculations for different discretizations and extrapolate to the $\Delta t \rightarrow 0$ limit. In order to use as much data as possible, it is best to first use a shape-preserving Akima spline to put all data on the same time grid and then use Lagrange interpolation as an extrapolation method to the $\Delta t \rightarrow 0$ limit. It is critical that one employs a shape-preserving spline, otherwise it will ruin your results. In addition, we find that quadratic extrapolation usually works best, but sometimes had to look at higher-order extrapolations. It is best to extrapolate the data after the desired quantity has been determined as a function of t for all times on the grid. For example, one would first determine the current by using the lesser Green function and summing over all momenta for each time and then extrapolate the final current values as a function of time instead of extrapolating the Green functions, since doing the latter requires a two-dimensional Lagrange extrapolation and uses large amounts of memory.

Finally we mention that the accuracy can be tested by calculating spectral moment sum rules. These are known exactly for the retarded Green function (through third-order). At half-filling they simplify tremendously. We will explore how this works when we discuss some examples

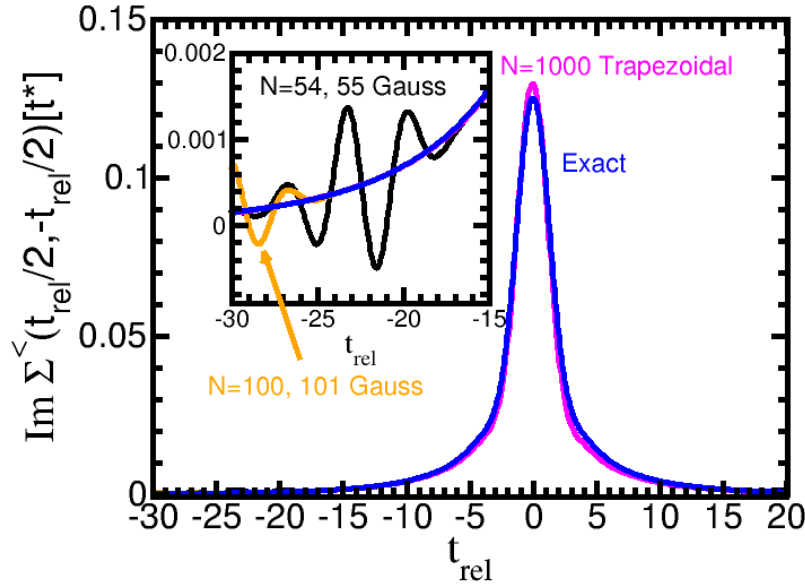


Fig. 2: Local lesser self-energy for the Falicov-Kimball model at half-filling, $\beta = 1$, and $U = 1$. The timestep was $\Delta t = 0.05$ and the time range ran from $-15 \leq t_{\text{rel}} \leq 15$. The blue curve is the exact result and the other colors are different methods for integration over the noninteracting density of states. The inset shows a blow up of the region at large negative times, where the different integration methods start to break down. This figure is adapted from Ref. [22].

in the next section. It is important to perform this test on *all* of your data. It is easy to do and ensures the data is high quality.

We end this section with a discussion of other methods. Tran showed how one can halve the dimension of the matrices by working with the retarded and lesser Green functions only [23], but such an algorithm is harder to code and you need to solve for twice as many Green functions, so it reduces the computational time by only a factor of two (it may allow for slightly larger matrices to be handled as well). There is an alternative way to discretize and solve this problem that integrates the Dyson equation in time in a causal fashion, producing the Green function along the boundary of the matrix as time increases [24]. It is very efficient, but requires you to be able to determine the self-energy in a causal fashion from your data at all *earlier* times. This is always possible when perturbation theory is used to determine the self-energy. But, while there should be a way to do this for the Falicov-Kimball model, an explicit method has not yet been discovered. For the Hubbard model, we have no exact methods available.

8 Examples

We now illustrate how all of this machinery works. The idea is not to provide an exhaustive review of the research work that has been completed, but rather to illustrate how these ideas are concretely put into action. As such, we will pick and choose the examples more for their didactic features than for their importance, or perhaps even their physical relevance.

To start, we focus on equilibrium. The DMFT for the Falicov-Kimball model was solved exactly in equilibrium by Brandt and Mielsch [11, 12] and is summarized in a review [25]. We take

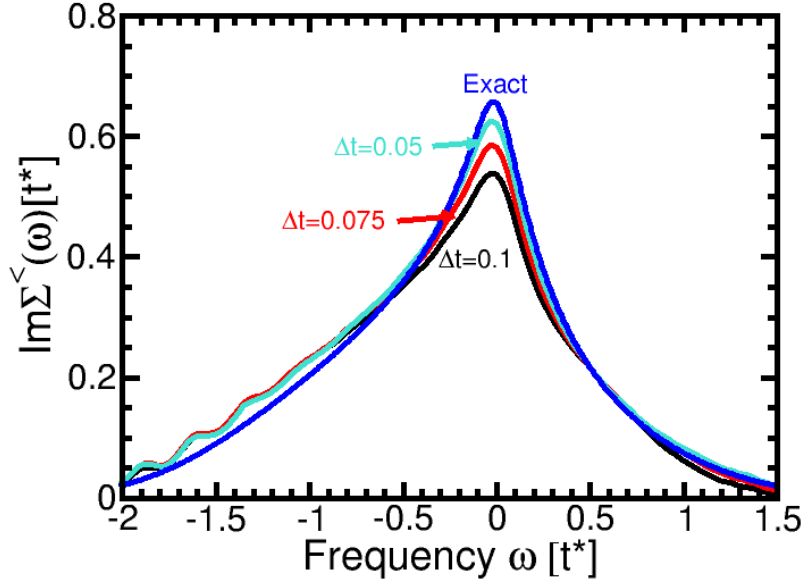


Fig. 3: Local lesser self-energy for the Falicov-Kimball model at half-filling, $\beta = 1$ and $U = 1$. Here, we Fourier transform the data from time to frequency. We use different time discretizations for the different curves, while the density of states integration used $N = 54$ and 55 points for Gaussian integration. The oscillations on the discrete calculations come from the oscillations in time shown in Fig. 2. This figure is adapted from Ref. [22].

the exact result as a function of frequency and Fourier transform to time. Then we compare that result with the result that comes out of the nonequilibrium formalism, to understand its accuracy. In all of this work, we solve the Falicov-Kimball model on a hypercubic lattice at half filling. In Fig. 2, we plot the (local) lesser self-energy $\Sigma^<(t)$. The calculation used a fixed $\Delta t = 0.05$ and has $U = 0.5$. The different curves are for various different integration schemes over the noninteracting density of states. In one case, we average the $N = 54$ and $N = 55$ Gaussian formulas. In another, we do the same, but with $N = 100$ and $N = 101$. We also compare to a much denser trapezoidal formula with 1000 points. Additionally, we plot the exact result. One can see the different integration schemes are quite close to each other for times near zero, but they begin to deviate at large times. The inset focuses on large negative time and one can quickly conclude that there is a maximum absolute value of time for which the results are accurate for any integration scheme. Beyond that, the system starts to generate increasing amplitude oscillations. The disagreement at short times from the exact result stems from the fact that these calculations are at a fixed Δt —no scaling to the $\Delta t \rightarrow 0$ limit were taken. This gives a sense of the accuracies we can hope to attain with this approach.

In Fig. 3, we plot the Fourier transform of the time data as a function of frequency. The number of Gaussian points used is $N = 54, 55$. The time steps are varied and one can see that they are approaching the exact result. If we used Lagrange extrapolation here, we would get quite close to the exact result, but we do not include that plot here because it would be too many lines close to each other. The oscillations in the tail of the data in Fig. 2 is responsible for the oscillations at negative frequency seen in the data. Those oscillations will remain even after scaling to the $\Delta \rightarrow 0$ limit. One can clearly see how the extrapolation method works for this case.

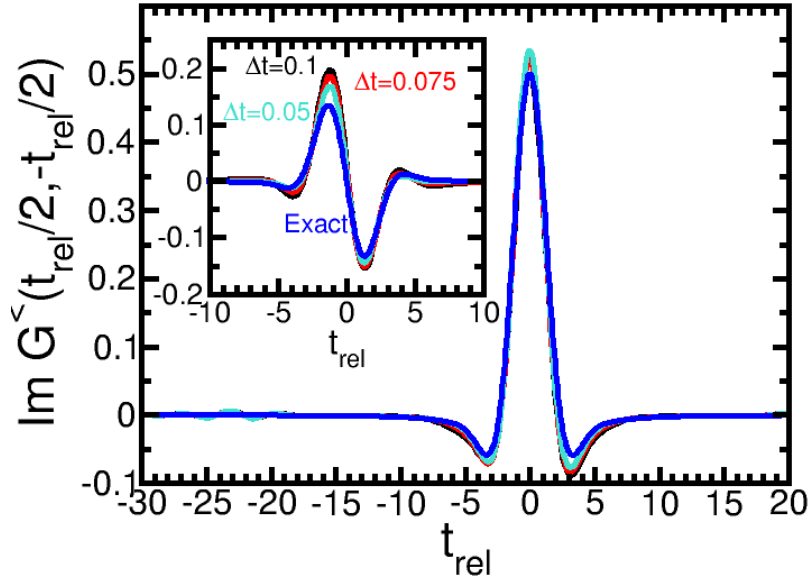


Fig. 4: Local lesser Green function for the Falicov-Kimball model at half-filling, $\beta = 1$ and $U = 1$. The parameters are the same as in the previous figure. The main plot is the imaginary part (which becomes symmetric for the exact result, and the inset is the real part, which becomes antisymmetric. This figure is adapted from Ref. [22].

In Fig. 4, we do a similar plot in the time domain for the lesser Green function. Note that the real part of the lesser Green function (inset) is an odd function when we are at half-filling due to particle-hole symmetry. Similarly, the imaginary part in the main panel is even for the exact result. One can clearly see how the extrapolation will work to approach the exact result if we did the full extrapolation of this data.

Now that we have gotten our feet wet with the numerics, we are ready to discuss some physics. At half-filling, there are enough conduction electrons to fill half of the sites of the lattice and similarly for the localized electrons. In this case, if the repulsion between the two species is large enough, they will avoid each other and the net effect is that the system becomes an insulator because charge motion is frozen due to the high energy cost for double occupancy. We can see this transition occur in Fig. 5, which is the local density of states at half-filling for different U values. As U is increased we evolve from the initial result, which is Gaussian for no interactions to results where a “hole” is dug into the density of states until it opens a gap at a critical value of U called the Mott transition. Beyond this point, the system has a gap to charge excitations.

Next, we illustrate how the density of states evolves in a transient fashion in time after a dc field is turned on. In Fig. 6, we show a series of frames at different time snapshots that plot the transient density of states (imaginary part of the local retarded Green function) as a function of frequency. One can see that the system starts off in a near Gaussian density of states and then develops features that are quite complex. The infinite-time limit is solved using Floquet nonequilibrium DMFT theory [27–29]; unfortunately, we are not able to describe the details for how that problem is solved. But there are a few important features to notice. Both before the field is applied and long after it has been applied, the system has a positive semidefinite

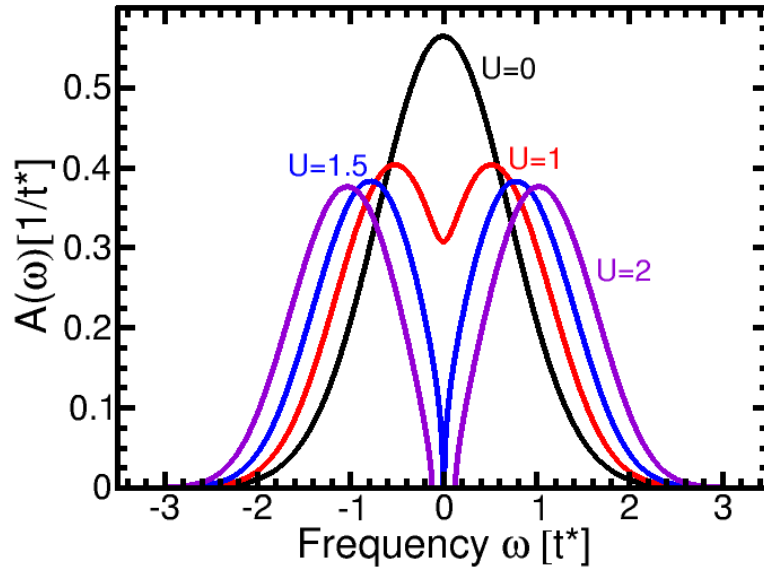


Fig. 5: Local density of states for the Falicov-Kimball model in equilibrium and at half-filling. The calculations are in equilibrium, where the density of states is temperature independent in the normal state. The curves are for different U . One can see as U increases we cross through a Mott transition at $U \approx \sqrt{2}$.

density of states. This is the same as in equilibrium and it allows the density of states to be interpreted probabilistically. But in the transient regime, it often becomes negative. Furthermore, because the density of states is determined via a Fourier transformation, one cannot just say it corresponds to a spectra at a specific average time. Instead it senses times near the average time governed by how rapidly the Green function decays in the time domain.

There is a lot of physics in these figures. If there were no interactions, the system would undergo Bloch oscillations, because the electrons are not scattering. The density of states then is the Fourier transform of a periodic function, which leads to delta function peaks forming a so-called Wannier-Stark ladder, with the separations given by the dc field E that is applied to the system. When interactions are added in, the Wannier-Stark ladder is broadened, but also split by the interaction U . This occurs because a delta function has a zero bandwidth and hence is highly susceptible to the Mott transition, even for relatively small interactions.

Next, we move on to examining the transient current (Fig. 7). We use the same case we have been examining throughout this brief summary— $U = 0.5$ and $\beta = 0.1$. Here the dc field is $E = 1$. One can see the current starts off as a weakly damped Bloch oscillations (underdamped). It dies off and remains quiescent for some time and then starts to recur at the longest times. The characteristic Bloch oscillation occurs for both metals and insulators, but it is damped much more rapidly in insulators because they interact more strongly. This is one of the common observables measured in a nonequilibrium experiment. But it is not measured directly, because oscilloscopes are not fast enough to see them.

Finally, we show how the sum rules hold and illustrate why it is important to scale results to the $\Delta t \rightarrow 0$ limit. We pick a case which is challenging to calculate, $E = 1$ and $U = 2$. This is in the Mott insulating phase, where the numerics are much more difficult to manage. We primarily use

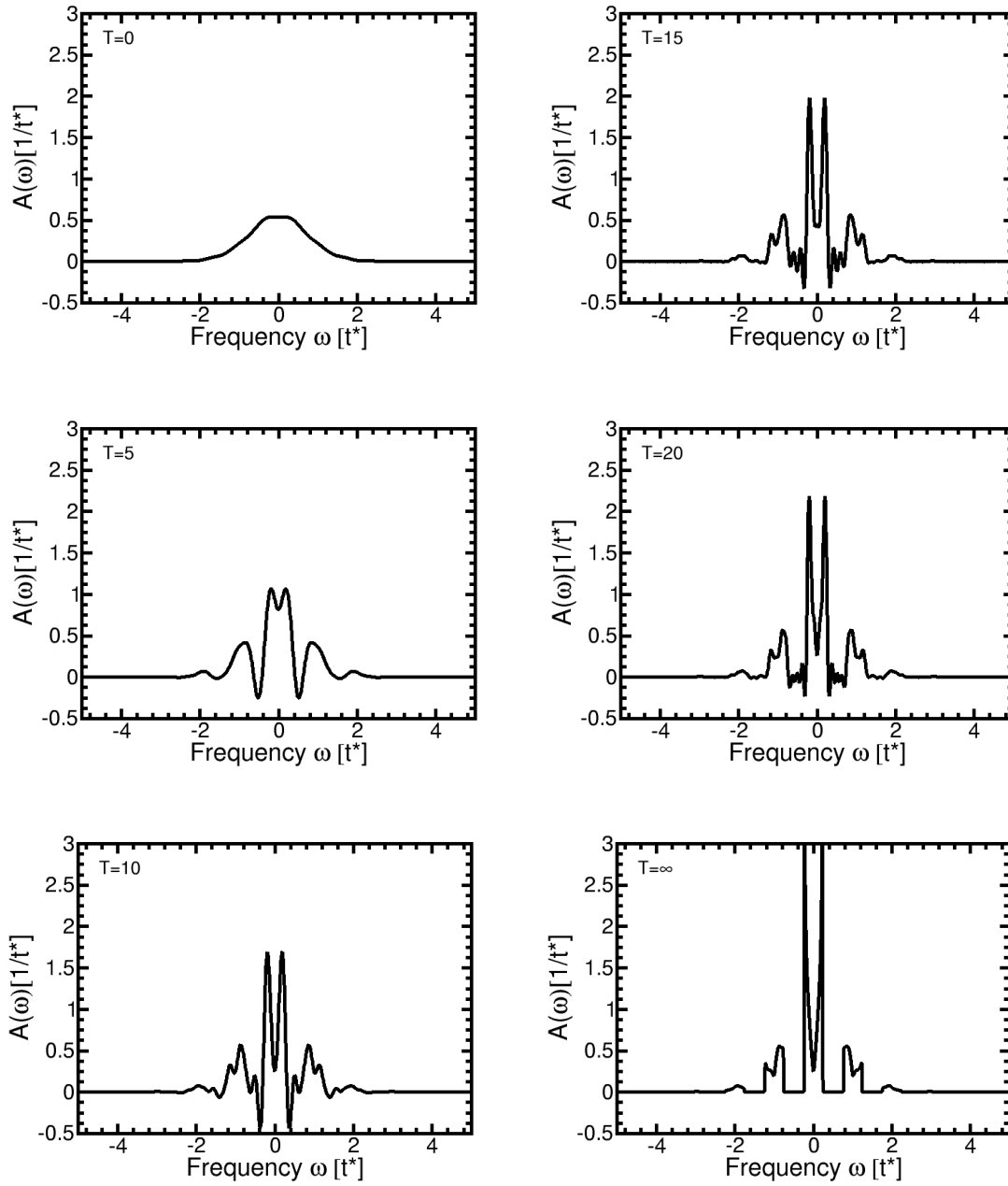


Fig. 6: Transient local density of states for the Falicov-Kimball model in a dc electric field with $U = 0.5$, $\beta = 0.1$ and $E = 1$ at half-filling. The different panels are labelled by the average time of the spectra. One can see that quite quickly after the dc field is turned on, the retarded Green function approaches the steady state. These results are adapted from [26].

the sum rules to indicate whether the calculations are accurate enough that they can be trusted. As shown in Fig. 8, we can see that the raw data can be quite bad, but the final scaled result ends up being accurate to 5% or less! Note how the results are worse on the equilibrium side (to the left) than on the nonequilibrium side (to the right) and they approach the exact results on the nonequilibrium side. Hence, these calculations are most accurate in moderate to large fields.

This ends our short representative tour of some numerical results that can be calculated with this nonequilibrium DMFT approach.

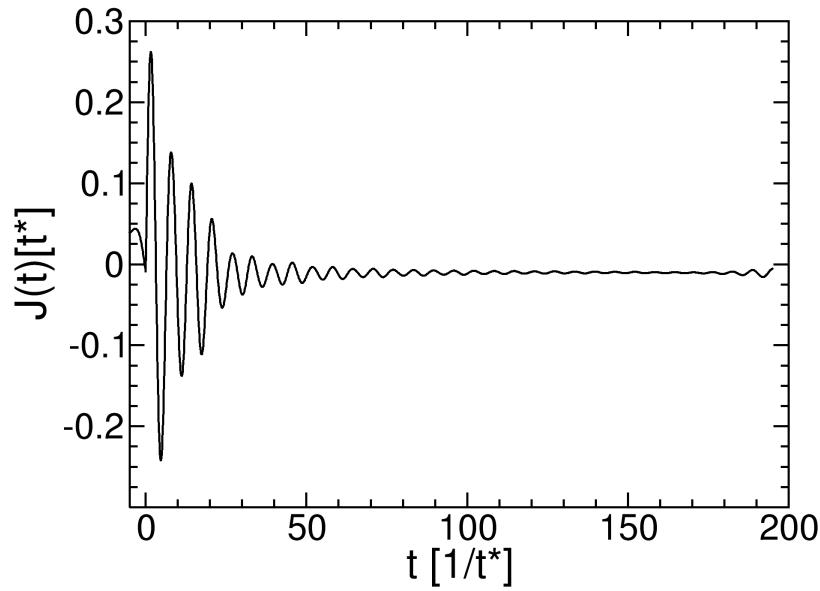


Fig. 7: Transient current for $E = 1$, $U = 0.5$ and $\beta = 0.1$. These results use $\Delta t = 0.1$ and are not scaled. Notice how they start off as a Bloch oscillation, but are damped by the scattering due to the interaction U . They die off and then start to have a recurrence at the longest times simulated. At the earliest times the current is nonzero simply because we have not scaled the data. Scaling is needed to achieve a vanishing current before the field is turned on. These results are adapted from Ref. [9].

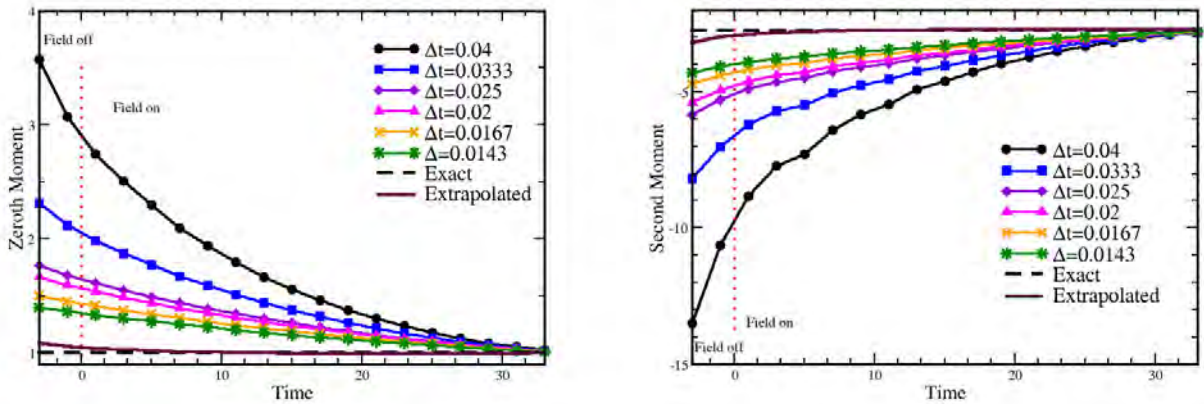


Fig. 8: Sum rules for the local retarded Green function for $E = 1$ and $U = 2$. Here, we illustrate how one can use sum rules to verify the scaling to $\Delta t \rightarrow 0$ has been done accurately. The zeroth moment sum rule equals 1 and the second moment sum rule equals $-(1/2 + U^2/4)$. We plot the raw data versus time, the exact result, and the extrapolated result. One can see that even if the raw data was off by a huge amount, the final extrapolated data works extremely well. These results are adapted from Ref. [30].

9 Conclusions

These lecture notes have been somewhat brief due to the page constraints of this contribution (and the time constraints of its author). Nevertheless, I have tried to present enough information here that you can follow the logic, reasoning, and even develop the formalism for yourself if you want to engage in these types of calculations. The field of nonequilibrium many-body physics is wide open. There are new and interesting experiments performed every day and the theory to describe them still needs further development. We do have a number of successes. The approach can be used for optical conductivity [10, 31], time-resolved angle-resolved photoemission [32], Raman scattering [33], x-ray photoemission spectroscopy and x-ray absorption spectroscopy [34], and resonant inelastic x-ray scattering [35]. The challenge is always about how far out in time can a simulation go. If you have a good idea for a nonequilibrium solver for a Hubbard impurity, I encourage you to give it a try. We really need it. I also hope you will enjoy working with many-body Green functions in the future. They are truly wonderful!

Acknowledgments

This work was supported by the Department of Energy, Office of Basic Energy Sciences, Division of Materials Science and Engineering under contract number DE-FG02-08ER46542. It was also supported by the McDevitt bequest at Georgetown University. I would also like to thank the organizers for the invitation to contribute to this school.

References

- [1] H. Lehmann, *Nuovo Cim.* **11**, 342 (1954)
- [2] A.A. Abrikosov, L.P. Gorkov, and I.E. Dzyaloshinski:
Methods of Quantum Field Theory in Statistical Physics (Prentice Hall, New York, 1963)
- [3] J.M. Luttinger and J.C. Ward, *Phys. Rev.* **118**, 1417 (1960)
- [4] T. Matsubara, *Prog. Theor. Phys.* **14**, 351 (1955)
- [5] L.P. Kadanoff and G. Baym: *Quantum statistical mechanics* (Benjamin, New York, 1962)
- [6] L.V. Keldysh, *Zh. Eksp. Teor. Fiz.* **47**, 1515 (1964) [*Sov. Phys. JETP* **20**, 1018 (1965)]
- [7] A. Georges, G. Kotliar, W. Krauth, and M.J. Rozenberg, *Rev. Mod. Phys.* **68**, 13 (1996)
- [8] J.K. Freericks, V.M. Turkowski, and V. Zlatić, *Phys. Rev. Lett.* **97**, 266408 (2006)
- [9] J K. Freericks, *Phys. Rev. B* **77**, 075109 (2008)
- [10] H. Aoki, N. Tsuji, M. Eckstein, M. Kollar, T. Oka, and P. Werner,
Rev. Mod. Phys. **86**, 779 (2014)
- [11] U. Brandt and C. Mielsch, *Z. Phys. B: Condens. Matter* **75**, 365 (1989)
- [12] U. Brandt and C. Mielsch, *Z. Phys. B: Condens. Matter* **79**, 295 (1990)
- [13] U. Brandt and M.P. Urbanek, *Z. Phys. B: Condens. Matter* **89**, 297 (1992)
- [14] M. Wagner, *Phys. Rev. B* **44**, 6104 (1991)
- [15] R.E. Peierls, *Z. Phys.* **80**, 763 (1933)
- [16] A. Joura: *Static and dynamic properties of strongly correlated lattice models under electric fields (Dynamical mean field theory approach)* (Ph.D. thesis, Georgetown Univ., 2014)
- [17] L.M. Falicov and J.C. Kimball, *Phys. Rev. Lett.* **22**, 997 (1969)
- [18] R. Bertoncini and A.P. Jauho, *Phys. Rev. B* **44**, 3655 (1991)
- [19] W. Metzner, *Phys. Rev. B* **43**, 8549 (1991)
- [20] M. Jarrell, *Phys. Rev. Lett.* **69**, 168 (1992)
- [21] G. Cohen, E. Gull, D.R. Reichman, and A.J. Millis, *Phys. Rev. Lett.* **115**, 266802 (2015)
- [22] J.K. Freericks, V.M. Turkowski, and V. Zlatić: *Real-time formalism for studying the non-linear response of ‘smart’ materials to an electric field*,
in D.E. Post (ed.): *Proceedings of the HPCMP Users Group Conference, Nashville, TN, June 28–30, 2005* (IEEE Computer Society, Los Alamitos, CA, 2005) pp. 25–34

- [23] M.-T. Tran, Phys. Rev. B **78**, 125103 (2008)
- [24] G. Stefanucci and R. van Leeuwen: *Nonequilibrium many-body physics of quantum systems: A modern introduction* (Cambridge University Press, 2013)
- [25] J.K. Freericks and V. Zlatić, Rev. Mod. Phys. **75**, 1333 (2003)
- [26] V. Turkowski and J.K. Freericks: *Nonequilibrium dynamical mean-field theory of strongly correlated electrons*, in J.M.P. Carmelo, J.M.B. Lopes dos Santos, V. Rocha Vieira, and P.D. Sacramento (eds.): *Strongly Correlated Systems: Coherence and Entanglement* (World Scientific, Singapore, 2007) pp. 187–210
- [27] A.V. Jura, J.K. Freericks, and T. Pruschke, Phys. Rev. Lett. **101**, 196401 (2008)
- [28] J.K. Freericks and A.V. Jura: *Nonequilibrium density of states and distribution functions for strongly correlated materials across the Mott transition*, in J. Bonca and S. Kruchinin (eds.): *Electron transport in nanosystems* (Springer, Berlin, 2008) pp. 219–236
- [29] N. Tsuji, T. Oka, and H. Aoki, Phys. Rev. B **78**, 235124 (2008)
- [30] J.K. Freericks, V.M. Turkowski, and V. Zlatić: *Nonlinear response of strongly correlated materials to large electric fields*, in D.E. Post (ed.): *Proceedings of the HPCMP Users Group Conference 2006, Denver, CO, June 26–29, 2006* (IEEE Computer Society, Los Alamitos, CA, 2006) pp. 218–226
- [31] A. Kumar and A.F. Kemper, *preprint* arXiv:1902.09549
- [32] J.K. Freericks, H.R. Krishnamurthy and T. Pruschke, Phys. Rev. Lett. **102**, 136401 (2009)
- [33] O.P. Matveev, A.M. Shvaika, T.P. Devereaux, and J.K. Freericks, Phys. Rev. Lett. **122**, 247402 (2019)
- [34] A.M. Shvaika and J.K. Freericks, Cond. Mat. Phys. **15**, 43701 (2012); *unpublished*
- [35] Y. Chen, Y. Wang, C. Jia, B. Moritz, A.M. Shvaika, J.K. Freericks, and T.P. Devereaux, Phys. Rev. B **99**, 104306 (2019)

14 Electronic Transport in Correlated Single-Molecule Junctions

Andrea Donarini
Institute of Theoretical Physics
University of Regensburg

Contents

1	Introduction	2
2	Spectroscopy and topography with atomic contact control	3
3	Liouville approach to correlated transport	4
3.1	General transport theory	4
3.2	Second-order perturbation	6
4	Examples of intrinsic correlated transport phenomena	8
4.1	Dynamical spin crossover	8
4.2	Apparent orbital reversal	15
5	Conclusions	22
A	Derivation of the generalized master equation	23

1 Introduction

The nanometer size, the perfect reproducibility and, at the same time, the huge variety of chemical compounds make molecular electronics an attractive bottom up approach to the quest of circuit miniaturization [1, 2]. A single-molecule junction (SMJ) is the archetypal device: ideally it only consists of a molecule contacted to metallic electrodes. Despite their deceptive simplicity, such devices explore a whole range of complex phenomena. Their electronic transport characteristics carry the fingerprints of the contacted molecule, primarily its correlated electronic structure, but also its mechanical excitations, spin dynamics, the response to external stimuli, as illumination, electrostatic gating, mechanical stress, or temperature gradient [3].

Here we will concentrate on the signatures of electronic correlation. Firstly, we can classify the effects of electronic correlation in extrinsic and intrinsic. Both of them stem from the interplay of electron-electron interaction on the molecule and the coupling to the electrodes. Extrinsic effects, though, (e.g. the Kondo [4–6], Yu-Shiba-Rusinov resonances [6–8], and many-body interference [9–12]) rely, for their appearance, on a specific property of the lead or of its coupling to the molecule. Intrinsic correlation phenomena, instead, arise directly from the entangled nature of the molecular eigenstates. The coupling to the leads is here only a tool to probe the electronic correlation and translate it into a specific transport signal. Examples of such effects are the excitonic dynamics [13], the equilibrium and non-equilibrium spin crossover [14–17], the magnetic anisotropy in single-molecule magnets [18–20], the charge dependent tunneling barrier modification [21] or the apparent orbital inversion discussed below [22]. Intrinsic correlation effects can only be captured within a many-body description of the molecular electronic structure. The latter is theoretically understood, in absence of major structural relaxations, as the complete set of eigenstates and eigenvalues of the Hamiltonian

$$\hat{H}_{\text{mol}} = \sum_i \varepsilon_i \hat{n}_i + \frac{1}{2} \sum_{ijkl} \sum_{\sigma\sigma'} V_{ijkl} \hat{d}_{i\sigma}^\dagger \hat{d}_{j\sigma'}^\dagger \hat{d}_{k\sigma} \hat{d}_{l\sigma}, \quad (1)$$

where the first term includes the (single particle) molecular orbital energies ε_i with the correspondent occupation operator $\hat{n}_i$ and the second describes the electron-electron interaction among the electrons. In its elementary form, the state of the molecule is given by the occupation of molecular orbitals according to the Pauli exclusion principle, thus neglecting electron-electron interaction. In the ground state all levels up to the highest occupied molecular level (HOMO) are filled, leaving empty the higher energy ones starting with the lowest unoccupied molecular level (LUMO). Transitions from an occupied to an empty level, or the addition (removal) of electrons, account for all electronic excitations. Such a description neglects correlations and assumes the electrons to be independent, thus giving many-body wave functions in the form of single Slater determinants. The simplest model able to describe the extrinsic correlated transport phenomena is the Anderson impurity model, in which the impurity Hamiltonian is of the form of Eq. (1) with a single orbital and the four associated eigenstates are the uncorrelated single Slater determinants

$$|0\rangle = |\emptyset\rangle, \quad |\uparrow\rangle = \hat{d}_\uparrow^\dagger |\emptyset\rangle, \quad |\downarrow\rangle = \hat{d}_\downarrow^\dagger |\emptyset\rangle, \quad |2\rangle = \hat{d}_\uparrow^\dagger \hat{d}_\downarrow^\dagger |\emptyset\rangle.$$

For more complex systems with a finite spatial extension different approximation schemes can be applied. In general, the long range, slow varying contribution of the Coulomb interaction can be incorporated within a mean-field approach, giving rise to occupation dependent shifts in the orbital energies and changes in the orbital shapes (orbital mixing). The short range, rapidly decaying contribution is instead responsible for exchange and correlation effects which can only be captured within a multi Slater-determinant description. A strong coupling to the metallic electrodes typically justifies (due to the hybridization and efficient metallic screening) a single-particle approach with independent electrons and mean field treatment. In the weak-coupling regime, electronic correlations become relevant as soon as the single-particle level splitting $\varepsilon_i - \varepsilon_j$ is comparable to the exchange V_{ijij} or pair hopping V_{iijj} energies. The discussion of intrinsic correlation phenomena arising in this regime is the main focus of these lecture.

2 Spectroscopy and topography with atomic contact control

A crucial role in a SMJ is played by the contacts. Different anchoring groups have been adopted to connect the molecule to the electrodes, ranging from thiol, to amine, or even fullerenes [23]. Also the direct contact of conjugated molecules with platinum electrodes have been investigated [24]. Correspondingly, a variety of mechanical stability and electronic coupling strengths have been reported. The common denominators and the technological challenge remains, though, the strong dependence of the transport characteristics on the precise contact configuration, down to the atomic scale.

As such, this strong contact dependence could undermine the reproducibility of the experimental results. Mainly, two approaches have been taken to overcome this problem, with completely opposite perspectives. On the one hand the break-junction technique averages over the contact fluctuations. Its goal is to identify the molecular fingerprints as persistent patterns in a statistical average of thousands of current traces [25–27]. On the contrary, ultra-high-vacuum low-temperature scanning tunneling microscopy minimizes the contact fluctuations. With this technique, also in combination with atomic force microscopy (AFM), one can achieve the most detailed characterization of both the tip [28] and the substrate electrodes. Even the position and orientation of the molecule with respect to the underlying crystal structure can be repeatedly monitored during the experiment.

Ideally, intrinsic electronic correlations are fully accessible only if both the spectrum and the eigenstates of a molecule are investigated within the same experiment. Scanning tunneling microscopy (STM) gives such a possibility [21, 22, 29–34] as one records the current by varying both the bias voltage and the tip position. Similarly to other electronic transport experiments [16, 18, 19, 35], the excitation spectrum is obtained by means of tunneling spectroscopy, i.e., by analyzing the peaks of the differential conductance. The current measured as a function of tip position gives, instead, access to the eigenstates. More precisely, the transitions between molecular many-body eigenstates are visualized, i.e., the quasiparticle wave function [36–39]

$$\phi(\mathbf{r}; N, E, E') = \langle N, E | \hat{\psi}(\mathbf{r}) | N+1, E' \rangle, \quad (2)$$

where N is the particle number on the molecule before the transition connecting states with energies E and E' , while $\hat{\psi}(\mathbf{r})$ is the electronic field operator. The quasiparticle wave function plays also an important role in the Liouville approach to transport theory which we will present in Sec. 3. If we express the field operator $\hat{\psi}(\mathbf{r})$ in terms of the complete set of the molecular orbitals (i.e., the eigenfunctions of the single-particle Hamiltonian), $\hat{\psi}(\mathbf{r}) = \sum_i \psi_i(\mathbf{r}) \hat{d}_i$, we see that the quasi-particle wave function coincides, for uncorrelated systems, with the molecular orbital, while deviations from this picture are expected as far as correlation steps in.

Particularly interesting to this extend, have been the realization of an STM with a thin insulating layer (a few monolayers) interposed between the molecule and the underlying metallic substrate. This technique, pioneered in the group of prof. G. Meyer [29], allows to select a single many-body transition and to visualize the corresponding quasiparticle wave function [22, 33]. The partial decoupling of the molecule from the metallic substrate enhances the intrinsic properties of the molecule. Recently, the concept has been further developed and measurements on insulators have been realized. The read-out of the electronic transitions is performed in the context of non contact atomic force microscopy [40].

3 Liouville approach to correlated transport

The problem of calculating the transport characteristics for a strongly interacting system coupled to leads still lacks a unifying solution able to cope with non-equilibrium boundary conditions in all transport regimes, despite great efforts to fill this gap [41–45]. Non-equilibrium Green's functions [46] remain the method of choice for systems with strong tunneling coupling to the leads, the interaction on the system been treated perturbatively. In the opposite limit, Liouville approaches [47] are the natural choice when the goal is the exact treatment of the strong correlated system, at the prize of a perturbative treatment of the tunneling coupling.

3.1 General transport theory

We briefly review here the second approach, which takes its name from the Liouville-von Neumann equation for the density matrix

$$\dot{\hat{\rho}} = -\frac{i}{\hbar} [\hat{H}, \hat{\rho}] \equiv \mathcal{L}\hat{\rho}, \quad (3)$$

where $\hat{\rho}$ is the density operator, $\mathcal{L}$ the Liouville superoperator and $\hat{H}$ the total Hamiltonian for the nanojunction. The latter can be split into the three components

$$\hat{H} = \hat{H}_{\text{mol}} + \hat{H}_{\text{leads}} + \hat{H}_{\text{tun}}, \quad (4)$$

where the $\hat{H}_{\text{mol}}$ describes the (strongly interacting) molecule, $\hat{H}_{\text{leads}}$ the (non interacting) leads and $\hat{H}_{\text{tun}}$ takes care of the tunneling coupling which transfers electrons between the leads and the molecule. Within this formal separation one envisages the purpose of the Liouville approach to derive, from the exact but intractable Eq. (3), an effective equation for the reduced

density operator $\hat{\rho}_{\text{red}} \equiv \text{Tr}_{\text{leads}} \hat{\rho}$, which describes only the evolution of the molecule, although in presence of the leads.

The molecule and the leads are taken as initially independent $\hat{\rho}(0) = \hat{\rho}_{\text{mol}}(0) \otimes \hat{\rho}_{\text{leads}}$ with the left and right lead in local thermal equilibrium

$$\hat{\rho}_{\text{leads}} = \frac{1}{Z_{G,L}} e^{-\beta(\hat{H}_L - \mu_L \hat{N}_L)} \otimes \frac{1}{Z_{G,R}} e^{-\beta(\hat{H}_R - \mu_R \hat{N}_R)}, \quad \mu_L - \mu_R = eV_{\text{bias}}, \quad (5)$$

where $\hat{H}_{\text{leads}} = \hat{H}_L + \hat{H}_R$, $\hat{N}_\eta$ is the particle number operator of lead $\eta = L, R$ with electrochemical potential μ_η , $Z_{G,\eta}$ is the grand canonical partition function, and V_{bias} is the potential drop across the junction. Entanglement between the molecule and the leads is induced by the tunneling dynamics. We concentrate, though, only on its influence onto the separable component of the density operator $\mathcal{P}\hat{\rho}(t) = \hat{\rho}_{\text{red}}(t) \otimes \hat{\rho}_{\text{leads}}$. The latter fulfills the following integro-differential equation of motion

$$\mathcal{P}\dot{\hat{\rho}}(t) = \mathcal{L}_{\text{mol}}\mathcal{P}\hat{\rho}(t) + \int_0^t \mathcal{K}(t-s)\mathcal{P}\hat{\rho}(s) ds \quad (6)$$

with the propagator kernel given by

$$\mathcal{K}(t-s) = \mathcal{P}\mathcal{L}_{\text{tun}} e^{(\mathcal{L}_{\text{mol}} + \mathcal{L}_{\text{leads}} + \mathcal{Q}\mathcal{L}_{\text{tun}}\mathcal{Q})(t-s)} \mathcal{L}_{\text{tun}} \mathcal{P}, \quad (7)$$

where we have introduced the projector operator $\mathcal{Q} = 1 - \mathcal{P}$ and the Liouvillian superoperators $\mathcal{L}_i = -\frac{i}{\hbar}[\hat{H}_i, \bullet]$ for each component of the Hamiltonian. See Appendix A for a derivation of Eq. (6) and (7). The latter are still exact and indicate, through the time dependent kernel, the emergence of memory in the molecular dynamics due to the coupling to the leads. This contribution, dissipative in the limit of large leads, adds to the time local coherent evolution of the isolated molecule represented by $\mathcal{L}_{\text{mol}}$. Differently from the Pauli master equation obtained by further approximations, Eq. (6), also called quantum master equation or Generalized Master Equation (GME) keeps coherences among molecular states essential, for example to capture interference effects even in the weak coupling limit [9, 11, 12, 48, 49].

The goal of a transport theory is the calculation of the electrical current through the system. The latter is formally derived starting from its definition of variation of the average particle number in the lead, $I_\eta(t) = e \frac{d\langle \hat{N}_\eta \rangle}{dt}$ with $\eta = L, R$, and reads

$$I_\eta(t) = \text{Tr} \left[\int_0^t \mathcal{K}_{I_\eta}(t-s) \mathcal{P}\hat{\rho}(s) ds \right], \quad (8)$$

where the trace is taken over the molecule *and* the leads Fock spaces. Moreover, the current kernel is written

$$\mathcal{K}_{I_\eta}(t-s) = \mathcal{P}\hat{I}_\eta e^{(\mathcal{L}_{\text{mol}} + \mathcal{L}_{\text{leads}} + \mathcal{Q}\mathcal{L}_{\text{tun}}\mathcal{Q})(t-s)} \mathcal{L}_{\text{tun}} \mathcal{P}, \quad (9)$$

where $\hat{I}_\eta$ is the time derivative of the number operator of lead η in the Heisenberg picture calculated at the initial time. From the expressions above it is clear how the formalism is ready to incorporate transient time dependent phenomena which arise due to the non-equilibrium condition induced by the leads. With some more effort even time dependent Hamiltonians can be included, but this goes beyond the scope of this lecture. We will analyze instead the

stationary limit (DC current) associated to a time independent bias voltage across the single-molecule junction. Due to the convolutive form of the kernels in Eq. (7) and (9) is it useful to go into the Laplace space. The stationary reduced density matrix is obtained, thanks to the final value theorem, as the solution of the equation

$$\left[\mathcal{L}_{\text{mol}} + \tilde{\mathcal{K}}(0) \right] \mathcal{P} \hat{\rho}(t \rightarrow \infty) = 0, \quad (10)$$

where, by definition of the Laplace transform, $\tilde{\mathcal{K}}(0) = \lim_{\lambda \rightarrow 0^+} \int_0^\infty e^{-\lambda t} \mathcal{K}(t) dt$. The stationary current thus reads

$$I_\eta(t \rightarrow \infty) = \text{Tr} \left[\tilde{\mathcal{K}}_{I_\eta}(0) \mathcal{P} \hat{\rho}(t \rightarrow \infty) \right]. \quad (11)$$

An explicit expression of the propagator and the current kernels depends on the specific form of the Hamiltonian. At this point it is useful to note that a systematic expansion of Eq. (7) and (9) in powers of the tunneling Liouvillian $\mathcal{L}_{\text{tun}}$ yields a consistent perturbation theory to be handled order by order within a diagrammatic approach [50]. Non perturbative effects like (energy dependent) level broadening due to quantum fluctuations, negative tunneling magnetoresistance [51, 52], and even precursors of the Kondo physics [42, 53, 54] have been obtained by infinite resummations of certain classes of diagrams.

3.2 Second-order perturbation

We consider here the second-order perturbation in the tunneling Hamiltonian which, as can be seen from Eq. (6), is the lowest non vanishing order in the propagator kernel. To fix the ideas, let us consider the following leads and tunneling Hamiltonians

$$\hat{H}_{\text{leads}} = \sum_{\eta \mathbf{k} \sigma} \varepsilon_{\eta \mathbf{k} \sigma} \hat{c}_{\eta \mathbf{k} \sigma}^\dagger \hat{c}_{\eta \mathbf{k} \sigma}, \quad (12)$$

with the momentum $\mathbf{k}$ and the spin σ of the quasi free electron in the lead η and

$$\hat{H}_{\text{tun}} = \sum_{\eta \mathbf{k} \sigma n p} p t_{\eta \mathbf{k} \sigma n}^p \hat{d}_{n \sigma}^p \hat{c}_{\eta \mathbf{k} \sigma}^{\bar{p}}, \quad (13)$$

where $p = \pm$ with the convention $\hat{c}^+ = \hat{c}^\dagger$ and $\hat{c}^- = \hat{c}$ and analogously for the operators $\hat{d}_{n \sigma}$ associated to the molecular orbital n with spin σ . The tunneling amplitudes $t_{\eta \mathbf{k} \sigma n}^p$, being proportional to the overlap of the lead and the molecule wave functions, keep track of the geometry of the contact. For them we define $t^+ = (t^-)^*$. Once the tunneling and the leads Hamiltonian are defined, we can calculate explicitly the action of the projection operators $\mathcal{P}$ in Eq. (7) and obtain the expression of the second-order kernel

$$\tilde{\mathcal{K}}^{(2)}(0) = -\frac{i}{2\pi} \sum_{\alpha_1 \alpha_2} \sum_{p \eta \sigma} \sum_{nm} \int \alpha_1 \alpha_2 \Gamma_{nm}^p(\varepsilon, \eta \sigma) \mathcal{D}_{n \sigma \alpha_2}^{\bar{p}} \frac{f_\eta^{(p \alpha_1)}(\varepsilon)}{p \varepsilon - i \hbar \mathcal{L}_{\text{mol}} + i 0^+} \mathcal{D}_{m \sigma \alpha_1}^p d\varepsilon. \quad (14)$$

The superoperators $\mathcal{D}_{n \sigma \alpha}^{\bar{p}}$ describe the action of a creation (annihilation) operator of the molecule when acting from the left ($\alpha = +1$) or from the right ($\alpha = -1$), i.e.,

$$\mathcal{D}_{n \sigma +}^p \hat{\rho} \equiv \hat{d}_{n \sigma}^p \hat{\rho}, \quad \mathcal{D}_{n \sigma -}^p \hat{\rho} \equiv \hat{\rho} \hat{d}_{n \sigma}^p.$$

The second-order kernel accounts for tunneling events in which only one lead at a time is involved. The direction of the tunneling event (to or from the lead maintained in local thermal equilibrium) is reflected in the Fermi distribution function, $f_{\eta}^{+}(\varepsilon) = 1/(1+e^{\beta(\varepsilon-\mu_{\eta})})$ and $f_{\eta}^{-}(\varepsilon) = 1-f_{\eta}^{+}(\varepsilon)$. Finally, the geometry of the contact is included in the single-particle rate-matrix $\Gamma^p(\varepsilon, \eta\sigma)$ which reads

$$\Gamma_{nm}^p(\varepsilon, \eta\sigma) = \frac{2\pi}{\hbar} \sum_{\mathbf{k}} t_{\eta\mathbf{k}\sigma n}^{\bar{p}} t_{\eta\mathbf{k}\sigma m}^p \delta(\varepsilon - \varepsilon_{\eta\mathbf{k}\sigma}). \quad (15)$$

By tracing Eq. (6) over the leads degrees of freedom and projecting on the (correlated) many-body eigenstates of $\hat{H}_{\text{mol}}$ we obtain a set of equation for the populations *and* coherences of the reduced density matrix, i.e., its diagonal and off-diagonal elements, respectively. The kernel $\tilde{\mathcal{K}}^{(2)}(0)$ and eventually also the current through the system is thus expressed in terms of many-body rates. For example,

$$R_{NE \rightarrow N+1 E'}^{\eta\sigma} = \sum_{nm} \langle NE | d_{n\sigma} | N+1 E' \rangle \Gamma_{nm}^{+}(E' - E, \eta\sigma) \langle N+1 E' | d_{m\sigma}^{\dagger} | NE \rangle f^{+}(E' - E - \mu_{\eta}) \quad (16)$$

is the rate of change of the population of the state with N particles and energy E due to transitions towards the state with $N+1$ particles and energy E' . For an STM set up with an s -symmetry tip, according to Chen's derivative rule [55], the tunneling amplitude is proportional to the molecular orbital at the tip position. The many-body rate in Eq. (16) is proportional to the modulus square of the quasi-particle wave function given in Eq. (2). Since the tip transitions are, typically, the bottle-neck of the electronic dynamics, the spatial dependence of the current will thus resemble the quasi-particle wave function. The latter reduces to the molecular orbital in absence of intrinsic correlations. Admixtures of p or d symmetry components of the tip wave-function enhance the corrugation in the imaging of the molecular states by introducing components of the current proportional to higher derivatives of the quasiparticle wave function [56]. Coherences in the density matrix become relevant only in presence of quasi-degenerate many-body levels ($\Delta E < \hbar\Gamma$) and for tunneling matrices associated to the leads which cannot be simultaneously diagonalized, i.e., in the absence of independent transport channels. It is under these conditions that many-body interference dominates the transport, leading to the formation of coherent trapping and dark states [12]. Under such conditions, the kernel $\tilde{\mathcal{K}}^{(2)}$ is also responsible of an effective internal dynamics, which is not described in terms of tunneling, but rather as a Lamb shift correction which adds to the molecule Hamiltonian [10, 48, 57, 58].

Summarizing, the most general equation of motion for the reduced density matrix, written up to second-order perturbation in the tunneling to the leads, reads

$$\dot{\hat{\rho}}_{\text{red}} = -\frac{i}{\hbar} [\hat{H}_{\text{mol}} + \hat{H}_{\text{LS}}^{(2)}, \hat{\rho}_{\text{red}}] + \mathcal{L}_{\text{tun}}^{(2)} \hat{\rho}_{\text{red}}. \quad (17)$$

Eq. (17) represents the starting point for the study of the intrinsic correlated transport phenomena presented in the next section.

The calculation of the spectrum and the many-body eigenstates of the molecular Hamiltonian is addressed within a mixed approach which aims, as much as possible, to combine the accuracy of

the ab-initio methods with the simplicity of the model. As a first step, a set of molecular orbitals is extracted from a tight binding model (or a DFT calculation). A limited set of N frontier orbitals is subsequently chosen for setting up a fully interacting Hamiltonian Eq. (1) which is diagonalized numerically in the complete 4^N dimensional Fock space. In this correlated many-body basis we evaluate the GME for the reduced density matrix, Eq.(17), and discuss the transport phenomena which characterize the different single-molecule junctions.

4 Examples of intrinsic correlated transport phenomena

4.1 Dynamical spin crossover

As a first example of a correlated phenomenon we consider the spin crossover induced on a metal-organic molecule under non-equilibrium conditions. Spin-crossover metal-organic compounds play a prominent role [14–16, 19, 59] in the emergent field of molecular spintronics. Spin crossover is the transition between metastable spin states under the influence of external stimuli [60]. The many-body exchange interaction of the d -electrons on the metal center, in combination with the crystal field generated by the surrounding ligand, determines their spin state. In three-terminal devices, the gate electrode used to tune the charge also governs the associated spin state [16, 19, 59]. Metal-organic molecules have also come in the focus of STM experiments [14, 15, 32]. More generally, the role itself of many-body effects in STM single-molecule junctions is receiving increasing attention, both theoretically [33, 37–39, 57] and experimentally [14, 15, 21, 33].

In this section we demonstrate the appearance of a non-equilibrium high-spin state in CuPc on an insulating substrate caused by many-body correlations. We illustrate that, for a given substrate work-function, it is possible to control the effective ground state of the molecule by varying the tip position or the bias voltage across the junction. The only requirements for this genuine many-body effect are an asymmetry between tip and substrate tunneling rates, which is naturally inherent to STM setups, and an energetic proximity of an excited neutral state of the molecule to its anionic ground state. As discussed below, the experimental set-up is similar to that of Ref. [21], but with a slightly larger work-function for the substrate. Control over the work-function can be achieved by choosing different materials or crystallographic orientation for the substrate, with effects analogous to a discrete gating of the molecule. Several approaches to gate an STM junction have been also recently investigated [61–63].

Many-body Hamiltonian and spectrum of CuPc: To properly describe the many-body electronic structure of CuPc is by itself a nontrivial task, since the relatively large size of the molecule makes it impossible to diagonalize exactly a many-body Hamiltonian written in a local, atomic basis as done for smaller molecules [9, 64, 65]. STM transport experiments on single-molecules, however, are restricted to an energy window involving only the low-lying states of the molecule in its neutral, cationic and anionic configuration, with the equilibrium configuration at zero bias set by the work-function ϕ_0 of the substrate [21]. This allows one to use a restricted basis of frontier orbitals to construct the many-body Hamiltonian [20, 66]. For example, for a copper

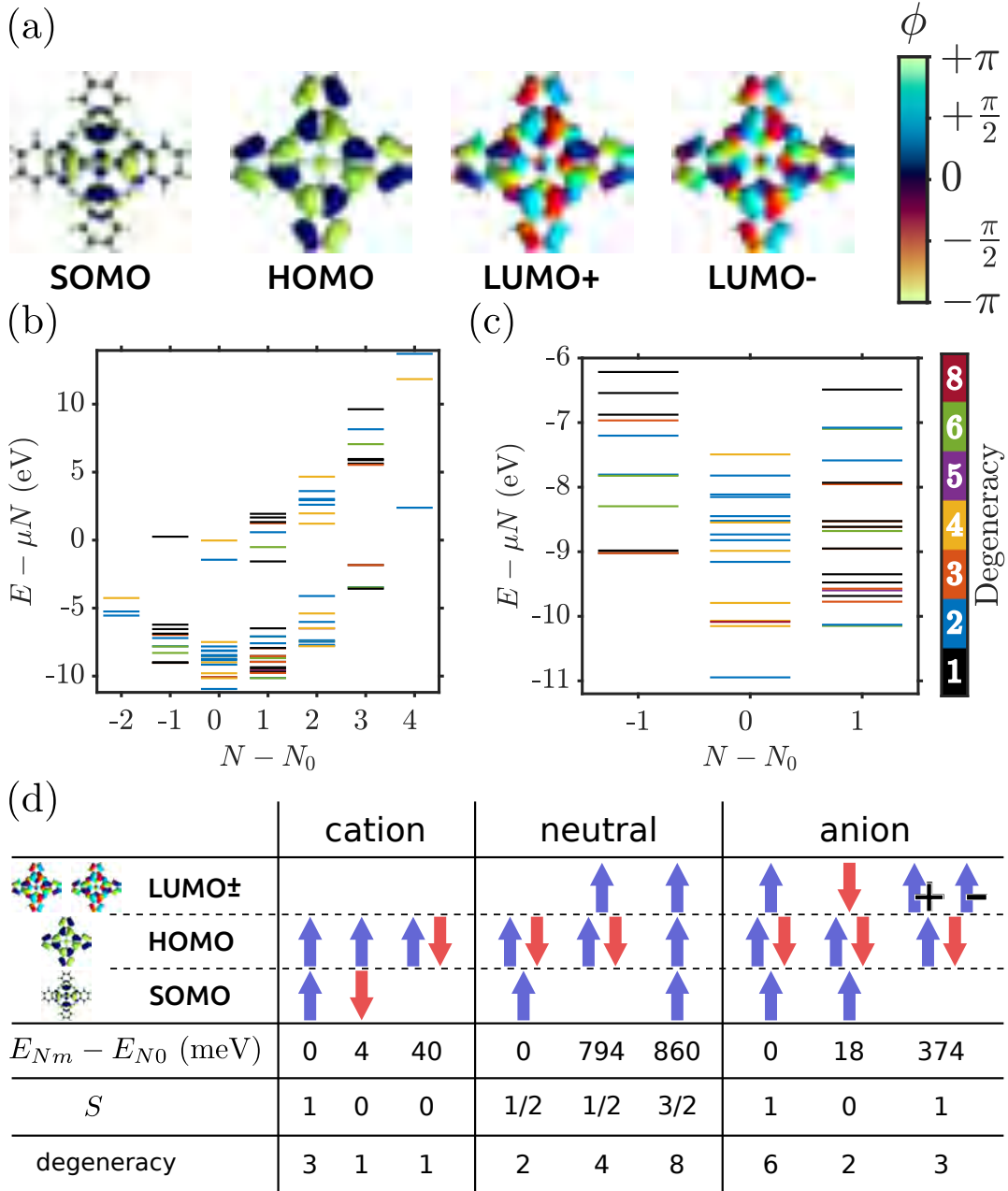


Fig. 1: (a) Frontier orbitals used for the many-body calculation, in their complex representation. The color code shows the phase of the wavefunctions. (b), (c) Full and low-energy cutout, respectively, of the many-body spectrum of CuPc at chemical potential $\mu = -4.65$ eV. (d) Scheme of the lowest-lying many-body states. From [17].

substrate as in [21] is $\phi_0 = 4.65$ eV, and CuPc in equilibrium is in its neutral ground state. Thus, in the following we only retain four frontier orbitals of CuPc, the SOMO (S), the HOMO (H) and the two degenerate LUMO ($L^\pm$) orbitals, see Fig. 1(a). In equilibrium, the molecule contains $N_0 = 3$ frontier electrons. In this basis, *all* matrix elements of the Coulomb interaction are retained. Hence, besides Hubbard-like density-density interaction terms, our model also includes exchange and pair hopping terms, which ultimately are important for the structure and spin configuration of the molecular excited states.

The Hamiltonian of CuPc in the basis of the four single-particle frontier orbitals reads

$$\hat{H}_{\text{mol}} = \sum_i \tilde{\varepsilon}_i \hat{n}_i + \frac{1}{2} \sum_{ijkl} \sum_{\sigma\sigma'} V_{ijkl} \hat{d}_{i\sigma}^\dagger \hat{d}_{j\sigma'}^\dagger \hat{d}_{k\sigma'} \hat{d}_{l\sigma}, \quad (18)$$

where $i = S, H, L\pm$ and σ is the spin degree of freedom. The energies $\tilde{\varepsilon}_i = \varepsilon_i + \Delta_i$ contain the single-particle molecular energies ε_i obtained from diagonalizing the single-particle Hamiltonian $\hat{H}_0$ of CuPc, $\varepsilon_S = -12.0$ eV, $\varepsilon_H = -11.7$ eV, and $\varepsilon_{L\pm} = -10.7$ eV. The parameters Δ_i account for crystal-field corrections and the ionic background of the molecule, since the atomic onsite energies in $\hat{H}_0$ come from Hartree-Fock calculations for isolated atoms [67]. The Δ_i are free parameters of the theory. Isolated CuPc has D_{4h} symmetry; the four molecular orbitals $|i\sigma\rangle$ that make up the basis of Eq. (18) transform like its b_{1g} (S), a_{1u} (H) and e_u ($L^\pm$) representations. As a consequence, they acquire distinct phases ϕ_i when rotated by 90 degrees around the main symmetry axis of the molecule, as illustrated in Fig. 1(a). This yields an easy rule to determine the nonvanishing Coulomb matrix elements V_{ijkl} in Eq. (18): $V_{ijkl} \neq 0$ if $\phi_i + \phi_j - \phi_k - \phi_l = 0 \bmod 2\pi$, i.e., non-vanishing contributions are only possible if the phases of the corresponding molecular orbitals add up to multiples of 2π . These considerations remain true in the presence of a homogenous substrate, which reduces the symmetry to C_{4v} . For a detailed discussion concerning the parametrization of Eq. (18) we refer to the supplemental material of [17]. Exact numerical diagonalization of $\hat{H}_{\text{mol}}$ finally yields the many-body eigenenergies E_{Nm} and eigenstates $|Nm\rangle$ of the molecule, labeled by particle number N and state index m .

Since the molecule is in contact with the substrate and is able to exchange electrons, it is necessary to consider a grand canonical ensemble $\hat{H}_{\text{mol}} - \mu\hat{N}$, where μ is the chemical potential of the substrate, which is given by its negative work-function, $\mu = -\phi_0$. Moreover, the presence of the leads renormalizes the Hamiltonian $\hat{H}_0$ due to image-charge effects [65, 68]. We model these effects with an effective Hamiltonian $\hat{H}_{\text{mol-env}} = -\delta_{\text{ic}}(\hat{N} - N_0)^2$, with $\hat{N}$ the particle-number operator on the system and δ_{ic} obtained from electrostatic considerations. To fit our spectrum to the experiment of Swart et al. [21], which was taken on a copper substrate Cu(100) ($\phi_0 = 4.65$ eV) on a trilayer of NaCl, we used a constant shift $\Delta_i = \Delta = 1.83$ eV, a dielectric constant $\varepsilon_{\text{mol}} = 2.2$ in the evaluation of the matrix elements V_{ijkl} , and an image-charge renormalization $\delta_{\text{ic}} = 0.32$ eV.

Figures 1(b), (c) show the cationic, neutral, and anionic subblocks of the many particle spectrum and their degeneracies. A schematic depiction of these states is shown in Fig. 1(d). As the actual states are linear combinations of several Slater determinants, only dominant contributions are shown. The neutral groundstate has a doublet structure (with total spin $S = \frac{1}{2}$) coming from the doubly filled HOMO and the unpaired spin in the SOMO. The cationic and anionic groundstates have triplet structures ($S = 1$). The former has a singly filled HOMO, the latter a singly filled LUMO orbital which form spin triplets (and singlets, $S = 0$, for the first excited states) with the singly filled SOMO. Importantly, the orbital degeneracy of the LUMO makes up for an additional twofold multiplicity of the anionic ground and first excited states. The first excited state of the neutral molecule is found to be also a doublet ($S = \frac{1}{2}$) with additional twofold

orbital degeneracy. Finally, the second excited state shows a spin quadruplet structure ($S = \frac{3}{2}$) together with twofold orbital degeneracy.

Transport dynamics and spin crossover: The full system is characterized by the Hamiltonian $\hat{H} = \hat{H}_{\text{mol}} + \hat{H}_{\text{mol-env}} + \hat{H}_{\text{S}} + \hat{H}_{\text{T}} + \hat{H}_{\text{tun}}$, where $\hat{H}_{\text{S}}$ and $\hat{H}_{\text{T}}$ are describing non-interacting electronic reservoirs for substrate (S) and tip (T). The tunneling Hamiltonian is $\hat{H}_{\text{tun}} = \sum_{\eta\mathbf{k}i\sigma} t_{\eta\mathbf{k}i} \hat{c}_{\eta\mathbf{k}\sigma}^\dagger \hat{d}_{i\sigma} + \text{h.c.}$, where $\hat{c}_{\eta\mathbf{k}\sigma}^\dagger$ creates an electron in lead η with spin σ and momentum $\mathbf{k}$. The tunneling matrix elements $t_{\eta\mathbf{k}i}$ are obtained analogously to Ref. [57]. The dynamics is calculated via a generalized master equation for the reduced density operator $\rho_{\text{red}} = \text{Tr}_{\text{S,T}}(\rho)$, see Refs. [9, 57]. In particular, we are interested in the state ρ_{red}^∞ solving the stationary equation $\mathcal{L}[\rho_{\text{red}}] = 0$, where $\mathcal{L}$ is the Liouvillian superoperator.

In analogy to Ref. [69] we included a phenomenological relaxation term $\mathcal{L}_{\text{rel}}$ in the Liouvillian¹

$$\mathcal{L}_{\text{rel}}[\hat{\rho}] = -\frac{1}{\tau} \left(\hat{\rho} - \sum_{Nm} \rho_{mm}^{\text{th},N} |Nm\rangle \langle Nm| \sum_n \rho_{nn}^N \right). \quad (19)$$

It is proportional to the deviation of the reduced density matrix from the thermal one, ρ^{th} , which is given by the Boltzmann distribution $\rho_{mm}^{\text{th},N} \sim \exp\left(-\frac{E_{Nm}}{k_B T}\right)$ with $\sum_m \rho_{mm}^{\text{th},N} = 1$. Since $\mathcal{L}_{\text{rel}}$ describes relaxation processes which conserve the particle number on the molecule, it does not contribute directly to the current. The relaxation factor $1/\tau$ is taken of the same order of magnitude as the tip tunneling rate. The stationary current through the system is evaluated from

$$\langle \hat{I}_{\text{S}} + \hat{I}_{\text{T}} \rangle = \frac{d}{dt} \langle \hat{N} \rangle = \text{Tr}_{\text{mol}} \left(\hat{N} \mathcal{L}[\hat{\rho}_{\text{red}}^\infty] \right) \equiv 0. \quad (20)$$

The Liouvillian $\mathcal{L} = \mathcal{L}_{\text{rel}} + \sum_{\eta} \mathcal{L}_{\eta}$ decomposes into the relaxation term and sub-Liouvillians for each lead. Sorting the occurring terms in Eq. (20) into substrate and tip contributions yields the current operator of the respective lead η as $\hat{I}_{\eta} = \hat{N} \mathcal{L}_{\eta}$.

Results of our transport calculations are presented in Fig. 2. In panels (a,d,g) we show constant height current maps, constant current STM images in (b,e,h) and in (c,f,i) maps of the expectation value of the total spin of the molecule depending on tip position, $S_{\text{rT}} = \sqrt{\langle \hat{S}^2 \rangle_{\text{rT}} + \frac{1}{4}} - \frac{1}{2}$ where $\langle \hat{S}^2 \rangle_{\text{rT}} = \text{Tr}_{\text{mol}} \left(\hat{S}^2 \hat{\rho}_{\text{red}}^\infty(\mathbf{r}_{\text{T}}) \right)$. The constant height and spin maps are each taken at a tip-molecule distance of 5 Å. The upper three panels (a,b,c) are for a work-function of $\phi_0 = 4.65$ eV and a bias voltage of $V_b = -2.72$ V. At this position the cationic resonance is occurring. Since the difference between neutral and cationic groundstate is the occupation of the HOMO (see Fig. 1(d)), tunneling occurs via this orbital, and the current maps (a,b) resemble its structure. With the same work function, $\phi_0 = 4.65$ eV, the anionic resonance is taking place at the positive bias $V_b = 0.81$ V, see Fig. 2(d,e). For equivalent reasons as in the former case, tunneling is happening via the LUMO and the spatial dependence of the current resembles the topography of this orbital. Panels (g,h,i) instead are recorded at $\phi_0 = 5.2$ eV, again at the anionic resonance which is now shifted to $V_b = 1.74$ V due to the larger work-function. Panel (g) is puzzling. Despite being an anionic resonance, it closely resembles the HOMO, cf. panels (a),(b). A closer

¹Differently from Ref. [69] we have included in Eq. (19) also the coherences. $\mathcal{L}_{\text{rel}}$ accounts thus also for dephasing. For simplicity, we assume the same phenomenological rate for dephasing and dissipation.

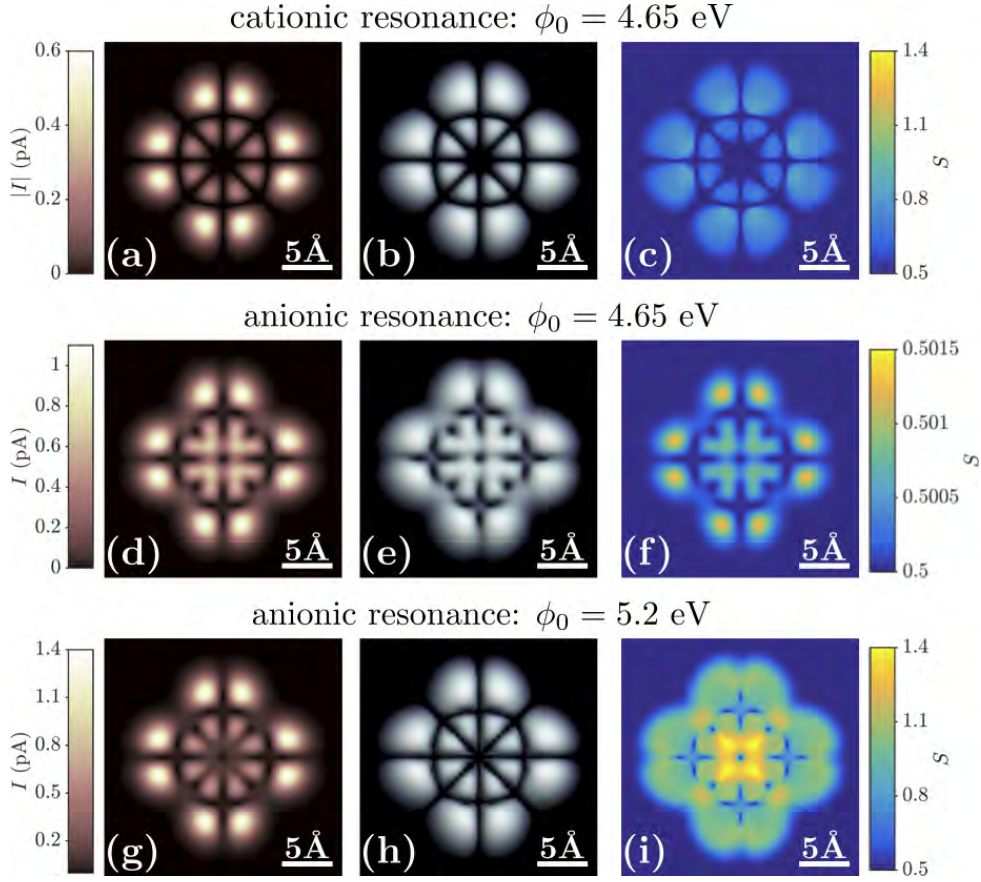


Fig. 2: (Constant height current maps (a,d,g), constant current maps (b,e,h) and maps of the system's total spin S (c,f,i). Constant height and spin maps are taken at a tip-molecule distance of 5 Å, constant current maps at currents $I = 0.5, 0.75$, and 1.0 pA for panels (c), (f), and (i), respectively. From [17].

inspection reveals also a likeness to the LUMO (see panel (d)) but with additional diagonal nodal planes, matching the nodal plane structure of the HOMO. When observing the constant current map in panel (h), and comparing it with panels (b) and (e), this statement becomes more evident. This anomalous topography can not be explained by single orbital tunneling.

Panels (c), (f) and (i) reveal the tip-position dependent expectation value of the total spin. At the standard anionic transition, panel (f), the spin remains essentially constant. At the standard cationic transition, panel (c), the rather homogeneous enhancement of the molecular spin is due to small populations of a large number of excited states, made accessible by the large resonance bias ($V_{\text{res}} = -2.7$ V). The anomalous anionic transition, panel (i), shows the largest variation of the molecular spin, concentrated at the positions of the anomalous current suppression, compare panels (g) and (d). To explain the unconventional properties shown in Fig. 2, we examine bias traces taken at different tip positions and values of the work-function. Figure 3(a) shows a shift of the anionic resonant peak in the $\frac{dI}{dV}$ for the anomalous case. The value V_{res} at which the peak is expected is given by

$$V_{\text{res}}(\phi_0) = \frac{1}{\alpha_{\text{T}}|e|} (E_{N_0+1,0} - E_{N_0,0} - \delta_{\text{ic}} + \phi_0), \quad (21)$$

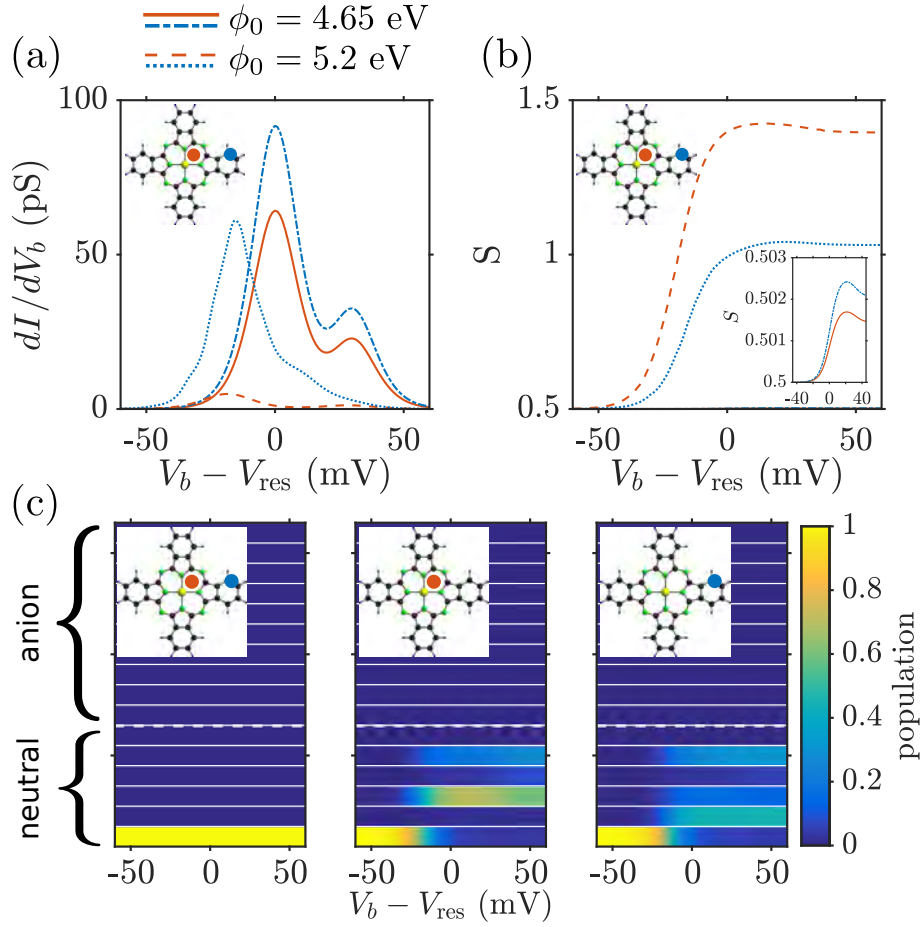


Fig. 3: (a) Differential conductance and (b) total spin curves taken at different tip positions and work-functions around the bias $V_{\text{res}}(\phi_0)$ of the anionic resonance. The inset in (b) shows the change of the spin for the standard case in magnification. (c) Populations of the density matrix around $V_{\text{res}}(\phi_0)$. Left panel: standard case, $\phi_0 = 4.65$ eV. Middle (right) panel: anomalous case, $\phi_0 = 5.2$ eV, with tip near the center (outer on the ligand). From [17].

where α_T is the fraction of bias drop between tip and molecule, and $E_{N,0}$ is the energy of the N -particle ground state. The shift of the resonance to lower biases seen in Fig. 3(a) suggests the appearance of a population inversion from the neutral ground state to an excited state. Transitions from the latter to the anionic ground state open in fact at much lower biases. Also the evolution of the spin of the molecule shown in Fig. 3(b) reinforces this proposition. In the anomalous case, the change of the system from a low to a high spin state, as well as the saturation of the spin, can be clearly seen. This contrasts with the normal anionic transition, where only a marginal change is observable. In Fig. 3(c) we show the evolution of the eigenvalues of the stationary density matrix ρ_{red}^∞ , i.e., the populations of the physical basis [9], around the anionic resonance $V_{\text{res}}(\phi_0)$, depending on work-function and tip position. In the standard case (left panel of Fig. 3(c)), the ground state of the system is always the neutral ground state. For the anomalous case (middle and right panels of Fig. 3) however, the picture changes dramatically, as there is a remarkable depopulation of the neutral ground state in favor of different excited states, depending on the position of the tip.

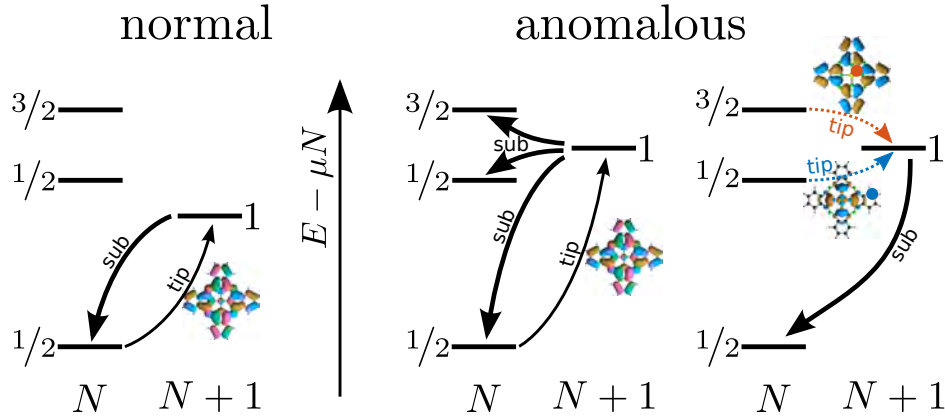


Fig. 4: Simplified sketch of the tunneling processes at the anionic resonance for the standard ($\phi_0 = 4.65$ eV) and the anomalous ($\phi_0 = 5.2$ eV) case. In the latter, population inversion takes place. The colors of the arrows denote tip positions where the corresponding transition acts as a bottleneck: Orange (blue) stands for the center (the outer ligand) of CuPc. From [17].

We focus now on the mechanism yielding the population inversion with associated spin crossover. In the standard case, at sufficiently high bias, the transition from the neutral to the anionic groundstate is opening, and tunneling of an electron into the LUMO brings the molecule into the anionic ground state. By consecutive tunneling to the substrate, the system goes back into its neutral ground state, see Fig. 4 for a simple sketch. Since the tunneling rates to the substrate are much larger than their tip counterparts, the system stays essentially in the neutral ground state with spin $S = \frac{1}{2}$. Also in the anomalous case an initial tunneling event brings the molecule into the anionic ground state. However, from there, due to finite temperature and proximity of the many-body eigenenergies, the system has a finite probability to go into a neutral excited state by releasing an electron to the substrate. The position of the tip and the structure of these excited states themselves then determine the stationary state: The molecule can only return to its neutral ground state by successive transitions to the anionic ground state via the tip, and from there to the neutral ground state via the substrate. However, the former process acts as a bottleneck and depends on the tip position. Leaving the first excited state ($S = \frac{1}{2}$) requires tunneling into the SOMO, while leaving the second excited state ($S = \frac{3}{2}$) would require tunneling into the HOMO. Additionally, near the center of the molecule the HOMO is vanishing, whereas on the outer ligand part the SOMO has little to no amplitude. Therefore, tunneling into these orbitals at the respective positions is strongly suppressed and the system ultimately ends up in the corresponding neutral excited states.

In synthesis, for an experimentally accessible substrate work-function of $\phi_0 = 5.2$ eV, we predict the appearance, in proximity to the anionic resonance, of a population inversion between the neutral ground and excited states of CuPc. Depending on the tip position, the molecule is triggered into a low-spin ($S=1/2$) to high-spin ($S=3/2$) transition which is mediated by this population inversion. This inversion is experimentally observable via dramatic changes in the topographical properties of constant height and constant current STM images, compared to a

standard LUMO-mediated anionic transition. Direct observation of the spin crossover might be accessible using spin-polarized scanning probe microscopy techniques. [70] The effect is also robust against moderate charge conserving relaxation processes. The quantitative accuracy of the spectroscopic and topographical results presented here is limited by the adopted semiempirical model. The spin crossover with the associated anomalous topography of the anionic resonance depends, however, on qualitative properties of the many-body spectrum and of the molecular orbitals. Thus, despite our focus on CuPc, they should be observable also in other molecules with comparable frontier orbital structure.

4.2 Apparent orbital reversal

In general, Coulomb charging energies strongly depend on the localization of electrons and hence on the spatial extent of the orbitals they occupy. Therefore the orbital sequence of a given molecule can reverse upon electron attachment or removal, if some of the frontier orbitals are strongly localized while others are not, like in phthalocyanines [30, 71–74]. Coulomb interaction may also lead to much more complex manifestations such as quantum entanglement of delocalized molecular orbitals.

Here we show, that the energy spacing of the frontier orbitals in a single-molecular wire of individual dicyanovinyl-substituted quinquethiophene (DCV5T) can be engineered to achieve near-degeneracy of the two lowest lying unoccupied molecular orbitals, leading to a strongly-entangled ground state of DCV5T^{2-} . These orbitals are the lowest two of a set of particle-in-a-box-like states and differ only by one additional nodal plane across the center of the wire. Hence, according to the fundamental oscillation theorem of Sturm-Liouville theory their sequence has to be set with increasing number of nodal planes, which is one of the basic principles of quantum mechanics [75, 76]. This is evidenced and visualized from scanning tunneling microscopy (STM) and spectroscopy (STS) of DCV5T on ultrathin insulating films. Upon lowering the substrate's work function, the molecule becomes charged, leading to a reversal of the sequence of the two orbitals. The fundamental oscillation theorem seems strikingly violated since the state with one *more* nodal plane appears *lower* in energy. This contradiction can be solved, though, by considering intramolecular correlation leading to a strong entanglement in the ground state of DCV5T^{2-} .

We refer to experiments carried out in the group of J. Repp, with a home-built combined STM/atomic force microscopy (AFM) using a qPlus sensor [77] operated in ultra-high vacuum. Bias voltages are applied to the sample. All AFM data, dI/dV spectra and maps, were acquired in constant-height mode. Calculations of the orbitals and effective single-particle electronic structure were performed within the density functional theory (DFT) as implemented in the SIESTA code [78] and are based on the generalized gradient approximation (GGA-PBE). The many-body eigenstates are determined from a diagonalization of the many-body model Hamiltonian H_{mol} , similar to the one introduced in Eq. (1). Based on these, STM-image and spectra simulations were performed within a Liouville approach for the density matrix ρ , as presented in section 3.

Control of the level spacing: The molecular structure of DCV5T, shown in Fig. 5a, consists of a quinquethiophene (5T) backbone and a dicyanovinyl (DCV) moiety at each end. The delocalized electronic system of polythiophene and oligo-thiophene enables conductance of this material [79–81]. The lowest unoccupied orbital of each of the thiophene rings couples electronically to its neighbors and forms a set of particle-in-a-box-like states [82,83]. The LUMO to LUMO+1 level spacing of the quinquethiophene (5T) backbone is approx. 0.7 eV [82], which is in good agreement with the energy difference calculated for free 5T based on DFT, as shown in Fig. 5a, left. This DFT-based calculation also confirms the nature of the LUMO and LUMO+1 orbitals, both deriving from the single thiophene’s LUMOs and essentially differing only by one additional nodal plane across the center of the molecule. To enable the emergence of correlation and thus level reordering, we have to bring these two states closer to each other. This is achieved by substituting dicyanovinyl moieties with larger electron affinity at each end of the molecular wire. As the orbital density of the higher lying particle-in-a-box-like state, namely LUMO+1, has more weight at the ends of the molecule, it is more affected by this substitution than the lowest state, the LUMO. This is evidenced by corresponding calculations of DCV5T, for which the LUMO to LUMO+1 energy difference is reduced by more than a factor of two, see Fig. 5a, left. The increased size of DCV5T may also contribute to the reduced level spacing. For the rest of this section, we concentrate on the LUMO and LUMO+1 orbitals only. To avoid confusion, we refrain from labeling the orbitals according to their sequence but instead according to their symmetry with respect to the mirror plane perpendicular to the molecular axis, as symmetric (S) and antisymmetric (AS). Hence, the former LUMO and the LUMO+1 are the S and AS states, respectively.

Apparent orbital reversal: To study the energetic alignment of the orbitals as well as their distribution in real space, ultrathin NaCl insulating films were employed to electronically decouple the molecules from the conductive substrate [29]. It has been shown that in these systems the work function can be changed by using different surface orientations of the underlying metal support [21, 29, 84]. Importantly, this does not affect the (100)-terminated surface orientation of the NaCl film, such that the local chemical environment of the molecule remains the same, except for the change of the work function.

However, in the present case, this alone has a dramatic effect on the electronic structure of the molecular wires as is evidenced in Fig. 5b. There, the STM images are shown for voltages corresponding to the respective lowest lying molecular resonances at positive sample voltage for DCV5T adsorbed on NaCl/Cu(111) (top panel) and NaCl/Cu(311) (bottom panel). They both show a hot-dog like appearance of the orbital density, in which a central lobe is hidden inside outer lobes, as was discussed, for example, in [13, 82]. However, whereas the orbital density of DCV5T/NaCl/Cu(311) shows a clear depression at the center of the molecule, indicating a nodal plane, DCV5T/NaCl/Cu(111) does not. Apparently, the energetically lowest lying state is not the same for the two cases, but S for DCV5T/NaCl/Cu(111) and AS in the case of DCV5T/NaCl/Cu(311). In contrast, STM images, acquired at voltages well below the first resonance, reflect the geometry of the molecule in both cases as wire-like protrusion (see insets of Fig. 5b).

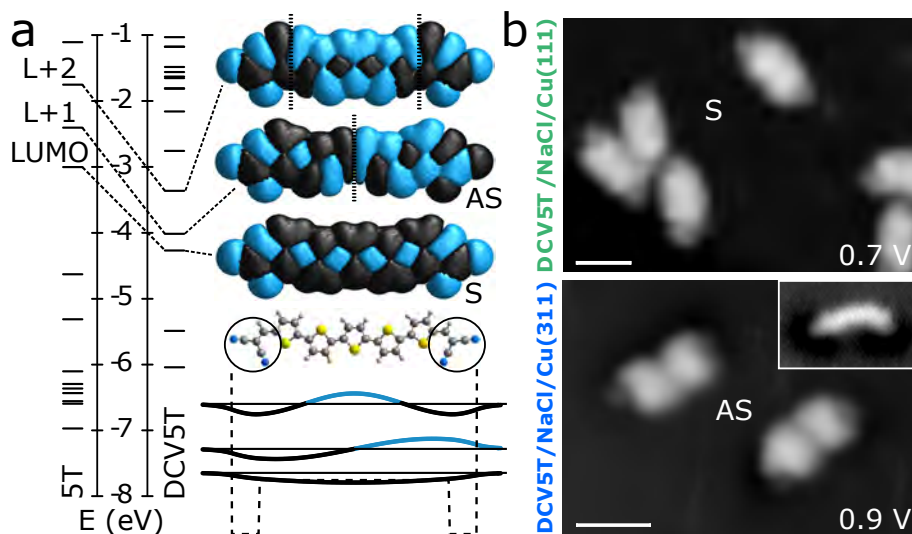


Fig. 5: (a) Molecular structure and density-functional theory based calculations of the electronic structure of 5T and DCV5T. The panel depicts the molecular structure, the calculated orbitals and energies for the LUMO, LUMO+1 and LUMO+2 as indicated. The orbitals are depicted as contours of constant probability density. The LUMO and LUMO+1 orbitals derive from the thiophene subunit's LUMO. They are the lowest two of a set of particle-in-a-box-like states and differ only by one additional nodal plane. Whereas the LUMO to LUMO+1 energy difference is approx. 0.7 eV for 5T, this difference is drastically reduced for DCV5T. The basic principle of level engineering is illustrated for a one-dimensional quantum box. (b) STM images of the first DCV5T electronic resonance on NaCl/Cu(111) (top) and NaCl/Cu(311) (bottom). The inset shows an STM image at a voltage below the first molecular resonance. From [22].

We hence assume that the molecules are neutral on NaCl/Cu(111) and that the S state corresponds to the LUMO. According to the literature, changing the copper surface orientation from Cu(111) to Cu(311) results in a lowering of the work function by approximately 1 eV [29, 85, 86]. Hence, one may expect that the former LUMO, initially located 0.7 eV above the Fermi level E_F in the case of NaCl/Cu(111) will shift to below the Fermi level [21, 30] for NaCl/Cu(311) such that the molecule becomes permanently charged.

To obtain a systematic understanding of the level alignment of the S and AS states of the molecule on both substrates, differential conductance (dI/dV) spectra and dI/dV -maps on DCV5T molecules have been acquired. Typical spectra measured at the center and the side of the molecule are shown in Figs. 6a and b on NaCl/Cu(111) and NaCl/Cu(311), respectively. DCV5T exhibits two dI/dV resonances at positive bias but none at negative voltages down to -2.5 V. According to the dI/dV maps and consistent with the different intensities in the spectra acquired on and off center of the molecule, the S state at $\simeq 0.7$ V is lower in energy than the AS state occurring at $\simeq 1.1$ V. The energy difference of $\simeq 0.4$ eV is in rough agreement to our calculations (see Fig. 5a). As discussed above, in the case of NaCl/Cu(311), DCV5T exhibits the AS state as the lowest resonance at positive bias voltages, this time at $\simeq 0.9$ V. This is additionally evidenced by the constant-current STM image and the corresponding dI/dV map in Fig. 6b. The S state is now located at higher voltages, namely at $\simeq 1.3$ V, as seen in the

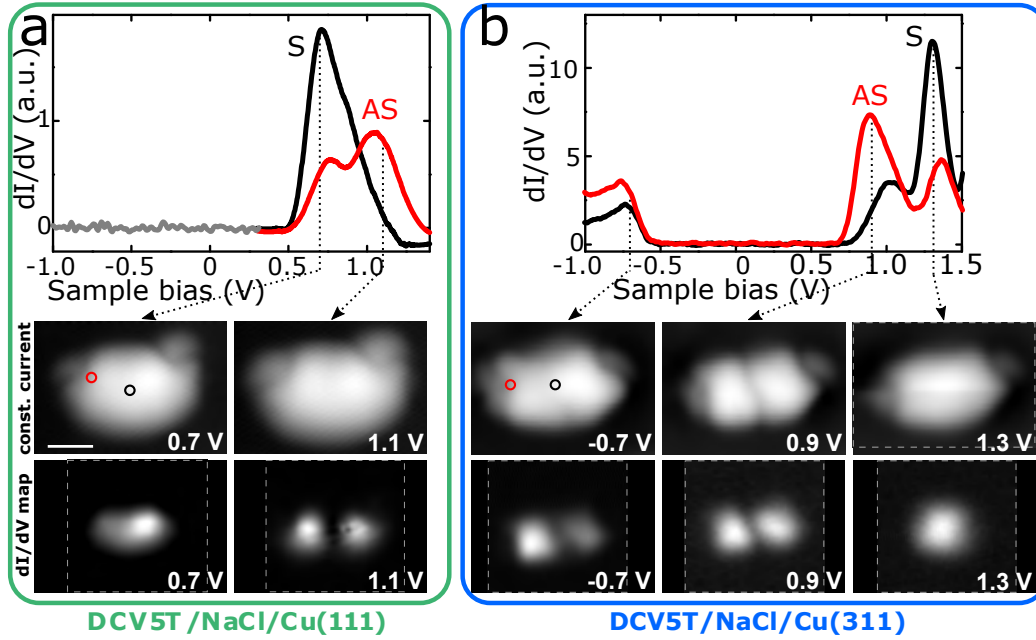


Fig. 6: dI/dV spectra (top panels), constant-current STM images (center panels) and dI/dV maps (bottom panels) on the individual molecule DCV5T on NaCl/Cu(111) (a) and NaCl/Cu(311) (b) respectively. The resonances are labelled with S and AS, referring to the symmetric and antisymmetric states, respectively. dI/dV spectra were recorded on (black) and off (red) the center of the molecule as indicated by dots in the STM images. The spatial distribution of orbitals gives rise to the different intensities at different tip positions as depicted in the inset. To not miss any small dI/dV signals in the low-bias range, a corresponding spectrum (grey) was measured with the tip being $\simeq 2 \text{ \AA}$ closer to the surface compared to the other two (red and black). All spectra were slightly low-pass filtered. The negative differential conductance beyond the AS peak (black curve) can be attributed to the increase in the tunneling barrier height with increasing bias voltage [29]. The images are resized to have the same size and scale, whereby the area of measured data is indicated with white dashed rectangles. Constant current images $I = 2.4 \text{ pA}$; bias voltage as indicated. Scale bar 1 nm. From [22].

spectrum and the dI/dV map. Obviously, the two states are reversed in their sequence. In this case, at negative bias voltages, a peak in dI/dV indicates an occupied state in equilibrium, in stark contrast to DCV5T/NaCl/Cu(111) but in agreement with the assumption of the molecule being negatively charged. The constant-current image acquired at -0.7 V , corresponding to the first peak at negative bias, seems to be a superposition of both the S and AS states.

The experimentally observed reversal of the orbital sequence is in striking disagreement with the fundamental oscillation theorem. To understand this apparent orbital reversal we go beyond the single-particle picture and invoke the role of electronic correlations. In the double-barrier tunneling junction geometry employed here, the resonances in dI/dV are associated with a temporary change of electron number on the molecule. In this terms the two peaks of DCV5T/NaCl/Cu(111) at positive bias are $\text{DCV5T} \leftrightarrow \text{DCV5T}^-$ transitions (see Fig. 7), and, in the same spirit, the ones of DCV5T/NaCl/Cu(311) at positive and at negative bias should be interpreted as $\text{DCV5T}^{2-} \leftrightarrow \text{DCV5T}^{3-}$ and $\text{DCV5T}^{2-} \leftrightarrow \text{DCV5T}^-$ transitions, respectively.

Many-body Hamiltonian and entangled ground state: Both the topographical and the spectroscopic data presented so far suggest that the electronic transport through DCV5T involves, in the present bias and work-function ranges, only the symmetric (S) and the antisymmetric (AS) orbitals. We concentrate on them and freeze the occupation of the other lower (higher) energy orbitals to 2 (0). In terms of these S and AS *frontier* orbitals we write the minimal interacting Hamiltonian for the isolated molecule

$$H_{\text{mol}} = \varepsilon_S \hat{n}_S + \varepsilon_{AS} \hat{n}_{AS} + \frac{U}{2} \hat{N}(\hat{N} - 1) + J \left(\sum_{\sigma\sigma'} \hat{d}_{AS\sigma}^\dagger \hat{d}_{S\sigma'}^\dagger \hat{d}_{AS\sigma'} \hat{d}_{S\sigma} + \hat{d}_{AS\uparrow}^\dagger \hat{d}_{AS\downarrow}^\dagger \hat{d}_{S\downarrow} \hat{d}_{S\uparrow} + \hat{d}_{S\uparrow}^\dagger \hat{d}_{S\downarrow}^\dagger \hat{d}_{AS\downarrow} \hat{d}_{AS\uparrow} \right), \quad (22)$$

where $\hat{d}_{S(AS)\sigma}^\dagger$ creates an electron with spin σ in the symmetric (antisymmetric) orbital, $\hat{n}_i$ counts the number of electrons in the orbital with $i = S, AS$ and $\hat{N}$ represents the total number of electrons occupying the two frontier orbitals. The interaction parameters $U = 1.4$ eV and $J = 0.75$ eV are obtained from the DFT orbitals by direct calculation of the associated Coulomb integrals and assuming a dielectric constant $\varepsilon_r = 2$ which accounts for the screening introduced by the underlying frozen orbitals [66, 17]. As expected from their similar (de-)localization, the Coulomb integrals of the S and AS states are almost identical.² Besides a constant interaction charging energy U , the model defined in Eq. (22) contains exchange interaction and pair-hopping terms, both proportional to J , which are responsible for the electronic correlation. The electrostatic interaction with the substrate is known to stabilize charges on atoms and molecules [87, 65, 84] due to image charge and polaron formation. We account for this stabilization with the additional Hamiltonian $H_{\text{mol-env}} = -\delta \hat{N}^2$. The orbital energies $\varepsilon_S = -3.1$ eV and $\varepsilon_{AS} = -2.8$ eV as well as the image-charge renormalization $\delta = 0.43$ eV are obtained from the experimental resonances of the neutral molecule and previous experimental results on other molecules as detailed in the supplemental material of [22].

Many-body interaction manifests itself most strikingly for the ground state DCV5T^{2-} , which will therefore be discussed at first. Consider the two many-body states, in which the two extra electrons both occupy either the S or the AS state: They differ in energy by the energy 2Δ , where $\Delta = \varepsilon_{AS} - \varepsilon_S$ is the single-particle level spacing between the S and the AS state. These two many-body states interact via pair-hopping of strength J , leading to a level repulsion. As long as $\Delta \gg J$, this effect is negligible. In DCV5T, though, the single-particle level spacing Δ is small compared to the pair-hopping J , leading to an entangled ground state of DCV5T^{2-} as

$$|2, 0\rangle = \cos \theta \hat{d}_{S\uparrow}^\dagger \hat{d}_{S\downarrow}^\dagger |0, 0\rangle + \sin \theta \hat{d}_{AS\uparrow}^\dagger \hat{d}_{AS\downarrow}^\dagger |0, 0\rangle, \quad (23)$$

with $|N, m\rangle$ the m^{th} excited N particle state of DCV5T and $\theta = \arctan(J/\Delta)/2$. Note that here, as $J/\Delta \approx 2.6$, this state shows more than 30% contribution from both constituent states, is strongly entangled, and therefore it can not be approximated by a single Slater determinant. The first excited state of DCV5T^{2-} is a triplet with one electron in the S and one in the AS orbital at about 54 meV above the ground state, as shown in Fig. 7.

²For the Coulomb integrals we obtain $U_{S-S} = 1.37$ eV, $U_{AS-AS} = 1.43$ eV, $U_{S-AS} = 1.37$ eV.

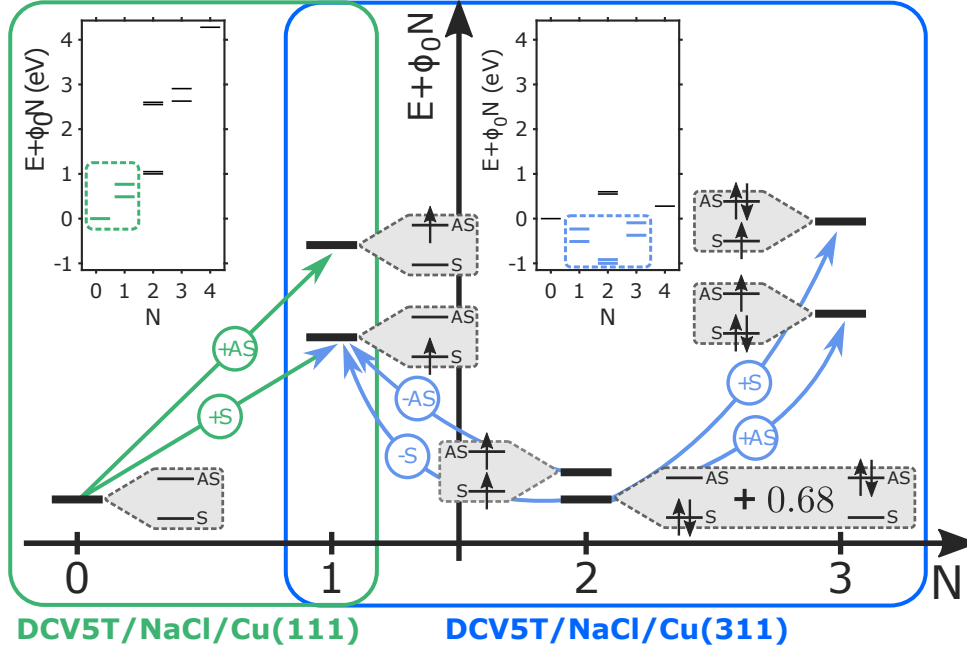


Fig. 7: Scheme of the many-body transitions associated to the measured resonances. In the green framed panel the transition between the neutral and the singly charged DCV5T^- are illustrated ($\text{DCV5T/NaCl/Cu(111)}$). In the blue framed panel the transitions involving DCV5T^- , DCV5T^{2-} and DCV5T^{3-} are analyzed ($\text{DCV5T/NaCl/Cu(311)}$). The electronic structure associated to the different many-body states is explicitly given in the gray labels. In the insets, the many-body spectra of the molecule on the two corresponding substrates are plotted. From [22].

The level repulsion in DCV5T^{2-} mentioned above leads to a significant reduction of the ground state energy by roughly 0.5 eV. This effect enhances the stability of the doubly charged molecule to the disadvantage of DCV5T^- , which has just a single extra electron and therefore does not feature many-body effects. Quantitatively, this is captured by the addition energies $E_N^{\text{add}} = E_{N+1,0} - 2E_{N,0} + E_{N-1,0}$ being a measure of the stability of the N particle ground state. Within the framework of the many-body theory, as sketched in Fig. 7, the apparent orbital reversal between Fig. 6a and Fig. 6b is naturally explained. To this end, as mentioned above, tunneling events in the STM experiments have to be considered as *transitions* between the many-body states of different charges N (see arrows in Fig. 7). The spatial fingerprints of the transitions and hence their appearance in STM images is given by the orbital occupation *difference* between the two many-body states and is indicated by the labels S and AS in Fig. 7. When on NaCl(2ML)/Cu(111) , the DCV5T molecule is in its neutral ground state, see green panel in Fig. 7. A sufficiently large positive sample bias triggers transitions to the singly charged DCV5T^- : The S and AS transitions subsequently become energetically available in the expected order of the corresponding single-particle states. A fast tunneling of the extra electron to the substrate restores the initial condition enabling a steady-state current.

When on NaCl(2ML)/Cu(311) the molecule is doubly charged and in the entangled ground state described by Eq. (23), see Fig. 7. At sufficiently high positive sample bias the transitions to DCV5T^{3-} are opening, enabling electron tunneling from the tip to the molecule. The topog-

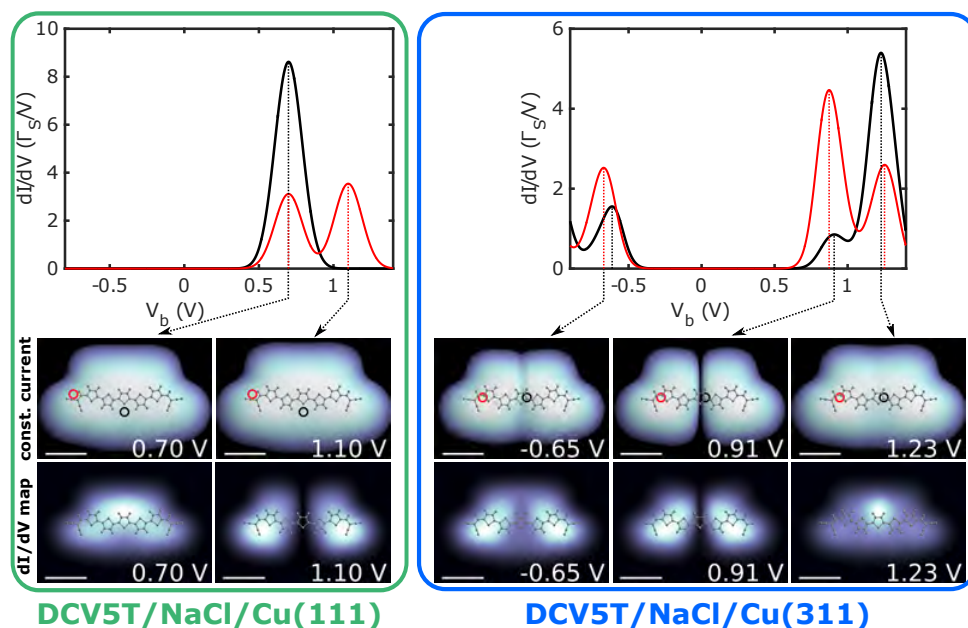


Fig. 8: Theoretical simulations of dI/dV spectra (top), constant-current STM images (center) and dI/dV maps (bottom) on the individual molecule DCV5T on NaCl/Cu(111) (a) and NaCl/Cu(311) (b) respectively. dI/dV spectra were recorded on (black) and off (red) center of the molecule as indicated by dots in the constant-current STM images. From [22].

raphy of these transitions is again obtained by comparing the 2- and the 3- (excess) electron states of DCV5T (cf. Fig. 7). The transition to the 3-particle *ground* state occurs by the population of the AS state and involves the *first* component of the entangled 2-electron ground state only. The second component cannot contribute to this transition, which is bound to involve only a *single* electron tunneling event. Correspondingly, at a larger bias the *first excited* 3-particle state becomes accessible, via a transition involving the *second* component of the 2-particle ground state only. This transition has a characteristic S state topography. Hence, although the electronic structure of the 3-electron states does follow the Aufbau principle, the entanglement of the 2-particle ground state leads to the apparent reversal of the orbital sequence.

In addition to the many-body spectrum we calculated the full dynamics of subsequent tunneling processes for all relevant situations, resulting in the calculated dI/dV characteristics, constant current maps and constant height dI/dV maps for a DCV5T single-molecule junction presented in Fig. 8. A qualitative agreement with the experimental results of Fig. 6 can be observed both for the relative strength of the spectral peaks and the dI/dV maps. The above discussed apparent orbital reversal is fully consistent with the calculations.

The experimental data of DCV5T on the Cu(311) substrate at negative bias also show a non-standard feature. The dI/dV map at resonance resembles a superposition of the S and AS orbital, see Fig. 6b. The effect is also reproduced in the theoretical simulations presented in Fig. 8. This can be rationalized in terms of a non-equilibrium dynamics associated to a population inversion predicted for Cu-Phthalocyanine [17]. The transition between the 2- and the 1-particle ground states has the topography of an S orbital. Its close vicinity to the ground state

implies that also the 2-particle first excited state should participate in transport. Moreover, the transition which connects the latter to the 1-particle ground state vanishes if the tip is in the vicinity of the molecular center, thus producing a nodal plane in the dI/dV topography and a large population of the 2-particle excited state. Though, this non-equilibrium induced population inversion is in competition with relaxation processes which conserve the molecular charge. This experiment indirectly proves that, although not negligible, this relaxation is by far not the fastest process for STM on thin insulating films: in the opposite case the nodal plane at the negative bias resonance would not have been observed.

In conclusion, we showed that a reduction of the single-particle level spacing of two frontier orbitals enables the manifestation of strong electron-correlation effects in single molecules. Here, the single-particle level spacing engineered by dicyanovinyl-substitution is leading to an apparent reversal of orbital sequence and a strongly-entangled ground state of DCV5T^{2-} . The many body description of the electronic transport is capable of reconciling the experimental observations of the orbital reversal with the fundamental oscillation theorem of quantum mechanics and shows how to achieve quantum entanglement of frontier orbitals in molecules.

5 Conclusions

The transport characteristics of correlated single-molecule junctions show a fascinating spectrum of different phenomena stemming from the interplay between the electron-electron correlation on the molecule and the tunneling coupling to the electrodes. A crucial role is played by the contacts. The control over their conformation down to atomic scale precision is often required to achieve quantitative reproducibility of the single current voltage characteristics. Scanning tunneling microscopy can achieve these standards and, with its unique capability of combining spectroscopy and topography, has proven a valuable tool to investigate correlation phenomena in single-molecule junctions. We classified the latter in extrinsic and intrinsic and concentrated on the second class, which originate directly from the entangled nature of the molecular many-body states. After introducing the Liouville approach to the transport problem, we have applied it to two examples of intrinsic correlation phenomena: the non-equilibrium spin crossover and the apparent orbital reversal. The transport characteristics were calculated for relatively simple models, based on a very small number of frontier orbitals, for which the full interaction Hamiltonian has been taken into account. In both cases, the topographical and spectral fingerprints of the correlation have been identified and criteria have been given to define the class of molecules in which similar phenomena can occur. We think that the study of correlation effects in single-molecule junctions is an open research fields with many challenges. Theoretically a unified approach to the transport problem is still lacking, while experimentally the advantages of different experimental techniques should be combined to reach a better control of the molecular many-body states and to study the signatures of their complex and fascinating dynamics.

A Derivation of the generalized master equation

The starting point is the Liouville-von Neumann equation Eq. (3), which we project on the separable and non-separable components via the operators $\mathcal{P}$ and $\mathcal{Q} = 1 - \mathcal{P}$, respectively

$$\begin{cases} \mathcal{P}\dot{\hat{\rho}} = \mathcal{P}\mathcal{L}\mathcal{P}\hat{\rho} + \mathcal{P}\mathcal{L}\mathcal{Q}\hat{\rho} \\ \mathcal{Q}\dot{\hat{\rho}} = \mathcal{Q}\mathcal{L}\mathcal{P}\hat{\rho} + \mathcal{Q}\mathcal{L}\mathcal{Q}\hat{\rho} \end{cases} \quad (24)$$

We formally solve the second equation with the help of the propagator: $\mathcal{G}_{\mathcal{Q}}(t, s) = e^{\mathcal{Q}\mathcal{L}(t-s)}$

$$\mathcal{G}_{\mathcal{Q}}(0, t) \mathcal{Q}\dot{\hat{\rho}} - \mathcal{G}_{\mathcal{Q}}(0, t) \mathcal{Q}\mathcal{L}\mathcal{Q}\hat{\rho} = \frac{d}{dt} [\mathcal{G}_{\mathcal{Q}}(0, t) \mathcal{Q}\hat{\rho}] = \mathcal{G}_{\mathcal{Q}}(0, t) \mathcal{Q}\mathcal{L}\mathcal{P}\hat{\rho},$$

which, by integration implies

$$\mathcal{G}_{\mathcal{Q}}(0, t) \mathcal{Q}\dot{\hat{\rho}} - \mathcal{Q}\hat{\rho}(0) = \int_0^t \mathcal{G}_{\mathcal{Q}}(0, s) \mathcal{Q}\mathcal{L}\mathcal{P}\hat{\rho}(s) ds$$

and, by multiplication on the left by $\mathcal{G}_{\mathcal{Q}}(t, 0)$

$$\mathcal{Q}\hat{\rho}(t) = \mathcal{G}_{\mathcal{Q}}(t, 0) \mathcal{Q}\hat{\rho}(0) + \int_0^t \mathcal{G}_{\mathcal{Q}}(t, s) \mathcal{Q}\mathcal{L}\mathcal{P}\hat{\rho}(s) ds. \quad (25)$$

We now insert (25) into the first of Eq. (24) taking a separable initial condition $\hat{\rho}(0) = \hat{\rho}_{\text{mol}}(0) \otimes \hat{\rho}_{\text{leads}}$, which implies $\mathcal{Q}\hat{\rho}(0) = 0$. Further, the following relations follow from the definition of $\mathcal{P}$ and the fact that the tunneling Hamiltonian does not conserve the lead particle number

$$\mathcal{P}\mathcal{L}_{\text{tun}}\mathcal{P} = 0, \quad [\mathcal{P}, \mathcal{L}_{\text{mol}}] = [\mathcal{P}, \mathcal{L}_{\text{leads}}] = \mathcal{P}\mathcal{L}_{\text{leads}} = 0. \quad (26)$$

With all these observations, we can write the closed integro-differential equation for $\mathcal{P}\hat{\rho}$

$$\mathcal{P}\dot{\hat{\rho}}(t) = \mathcal{L}_{\text{mol}}\mathcal{P}\hat{\rho}(t) + \int_0^t \mathcal{P}\mathcal{L}\mathcal{G}_{\mathcal{Q}}(t, s) \mathcal{Q}\mathcal{L}\mathcal{P}\hat{\rho}(s) ds, \quad (27)$$

from which we can already identify (remember that $\mathcal{P}^2 = \mathcal{P}$) a propagator kernel of the form

$$\mathcal{K}(t-s) = \mathcal{P}\mathcal{L}e^{\mathcal{Q}\mathcal{L}(t-s)}\mathcal{Q}\mathcal{L}\mathcal{P}.$$

We further manipulate the expression above with the idea of better emphasizing the perturbative expansion in $\mathcal{L}_{\text{tun}}$. By expanding the exponential, we obtain

$$\mathcal{P}\mathcal{L}e^{\mathcal{Q}\mathcal{L}t}\mathcal{Q}\mathcal{L}\mathcal{P} = \sum_{n=0}^{\infty} \frac{1}{n!} \mathcal{P}\mathcal{L}\mathcal{Q}(\mathcal{Q}\mathcal{L}\mathcal{Q})^n \mathcal{Q}\mathcal{L}\mathcal{P}t^n$$

On the other hand, with the help of the relations among Liouvillian and projectors (26) one easily obtains

$$\mathcal{P}\mathcal{L}\mathcal{Q} = \mathcal{P}\mathcal{L}_{\text{tun}}, \quad \mathcal{Q}\mathcal{L}\mathcal{P} = \mathcal{L}_{\text{tun}}\mathcal{P}, \quad \mathcal{Q}\mathcal{L}\mathcal{Q} = \mathcal{Q}(\mathcal{L}_{\text{mol}} + \mathcal{L}_{\text{leads}} + \mathcal{Q}\mathcal{L}_{\text{tun}}\mathcal{Q}).$$

Since it also clear, from the definition of $\mathcal{Q}$ that $[\mathcal{Q}, \mathcal{Q}\mathcal{L}_{\text{tun}}\mathcal{Q}] = 0$ we can write

$$\mathcal{K}(t) = \sum_{n=0}^{\infty} \frac{1}{n!} \mathcal{P}\mathcal{L}_{\text{tun}}\mathcal{Q}(\mathcal{L}_{\text{mol}} + \mathcal{L}_{\text{leads}} + \mathcal{Q}\mathcal{L}_{\text{tun}}\mathcal{Q})^n t^n \mathcal{L}_{\text{tun}}\mathcal{P} \quad (28)$$

where, according to Eq. (26), the first $\mathcal{Q}$ projector from the left can be omitted and, by resumming the series, one obtains to the desired result of Eq. (7).

References

- [1] M. Ratner, *Nat. Nano* **8**, 378 (2013)
- [2] M. Thoss and F. Evers, *J. Chem. Phys.* **148**, 030901 (2018)
- [3] S.V. Aradhya and L. Venkataraman, *Nat. Nanotechnol.* **8**, 399 EP (2013)
- [4] W. Liang, M.P. Shores, M. Bockrath, J.R. Long, and H. Park, *Nature* **417**, 725 (2002)
- [5] A. Zhao, Q. Li, L. Chen, H. Xiang, W. Wang, S. Pan, B. Wang, X. Xiao, J. Yang, J.G. Hou, and Q. Zhu, *Science* **309**, 1542 (2005)
- [6] K.J. Franke, G. Schulze, and J.I. Pascual, *Science* **332**, 940 (2011)
- [7] N.Y. Yao, L.I. Glazman, E.A. Demler, M.D. Lukin, and J.D. Sau, *Phys. Rev. Lett.* **113**, 087202 (2014)
- [8] J.O. Island, R. Gaudenzi, J. de Bruijckere, E. Burzurí, C. Franco, M. Mas-Torrent, C. Rovira, J. Veciana, T.M. Klapwijk, R. Aguado, and H.S.J. van der Zant, *Phys. Rev. Lett.* **118**, 117001 (2017)
- [9] D. Darau, G. Begemann, A. Donarini, and M. Grifoni, *Phys. Rev. B* **79**, 235404 (2009)
- [10] A. Donarini, G. Begemann, and M. Grifoni, *Nano Lett.* **9**, 2897 (2009)
- [11] A. Donarini, G. Begemann, and M. Grifoni, *Phys. Rev. B* **82**, 125451 (2010)
- [12] A. Donarini, M. Niklas, M. Schafberger, N. Paradiso, C. Strunk, and M. Grifoni, *Nat. Comm.* **10**, 381 (2019)
- [13] L. Bogner, Z. Yang, M. Corso, R. Fitzner, P. Bäuerle, K.J. Franke, J.I. Pascual, and P. Tegeder, *Phys. Chem. Chem. Phys.* **17**, 27118 (2015)
- [14] T.G. Gopakumar, F. Matino, H. Naggert, A. Bannwarth, F. Tuczek, and R. Berndt, *Angew. Chem. Int. Ed.* **51**, 6262 (2012)
- [15] T. Miyamachi, M. Gruber, V. Davesne, M. Bowen, S. Boukari, L. Joly, F. Scheurer, G. Rogez, T.K. Yamada, P. Ohresser, E. Beaurepaire, and W. Wulfhekel, *Nat. Commun.* **3**, 938 (2012)
- [16] V. Meded, A. Bagrets, K. Fink, R. Chandrasekar, M. Ruben, F. Evers, A. Bernand-Mantel, J.S. Seldenthuis, A. Beukman, and H.S.J. van der Zant, *Phys. Rev. B* **83**, 245415 (2011)
- [17] B. Siegert, A. Donarini, and M. Grifoni, *Phys. Rev. B* **93**, 121406(R) (2016)
- [18] M.H. Jo, J.E. Grose, K. Baheti, M.M. Deshmukh, J.J. Sokol, E.M. Rumberger, D.N. Hendrickson, J.R. Long, H. Park, and D.C. Ralph, *Nano Lett.* **6**, 2014 (2006)
- [19] A.S. Zyazin, J.W.G. van den Berg, E.A. Osorio, H.S.J. van der Zant, N.P. Konstantinidis, M. Leijnse, M.R. Wegewijs, F. May, W. Hofstetter, C. Danieli, and A. Cornia, *Nano Lett.* **10**, 3307 (2010)

- [20] A. Chiesa, S. Carretta, P. Santini, G. Amoretti, and E. Pavarini, Phys. Rev. Lett. **110**, 157204 (2013)
- [21] I. Swart, T. Sonleitner, and J. Repp, Nano Lett. **11**, 1580 (2011)
- [22] P. Yu, N. Kocić, J. Repp, B. Siegert, and A. Donarini, Phys. Rev. Lett. **119**, 056801 (2017)
- [23] E. Leary, A. La Rosa, M.T. González, G. Rubio-Bollinger, N. Agraït, and N. Martín, Chem. Soc. Rev. **44**, 920 (2015)
- [24] M. Kiguchi, O. Tal, S. Wohlthat, F. Pauly, M. Krieger, D. Djukic, J.C. Cuevas, and J.M. van Ruitenbeek, Phys. Rev. Lett. **101**, 046801 (2008)
- [25] B. Xu and N.J. Tao, Science **301**, 1221 (2003)
- [26] S. Wagner, F. Kißlinger, S. Ballmann, F. Schramm, R. Chandrasekar, T. Bodenstein, O. Fuhr, D. Secker, K. Fink, M. Ruben, and H.B. Weber, Nat. Nanotechnol. **8**, 575 (2013)
- [27] A. Borges, J. Xia, S.H. Liu, L. Venkataraman, and G.C. Solomon, Nano Lett. **17**, 4436 (2017)
- [28] T. Hofmann, F. Pielmeier, and F.J. Giessibl, Phys. Rev. Lett. **112**, 066101 (2014)
- [29] J. Repp, G. Meyer, S.M. Stojković, A. Gourdon, and C. Joachim, Phys. Rev. Lett. **94**, 026803 (2005)
- [30] C. Uhlmann, I. Swart, and J. Repp, Nano Lett. **13**, 777 (2013)
- [31] B.W. Heinrich, L. Braun, J.I. Pascual, and K.J. Franke, Nano Lett. **15**, 4024 (2015)
- [32] M. Gruber, V. Davesne, M. Bowen, S. Boukari, E. Beaurepaire, W. Wulfhekel, and T. Miyamachi, Phys. Rev. B **89**, 195415 (2014)
- [33] F. Schulz, M. Ijäs, R. Drost, S.K. Hämäläinen, A. Harju, A.P. Seitsonen, and P. Liljeroth, Nat. Phys. **11**, 229 (2015)
- [34] B. de la Torre, M. Švec, G. Foti, O.C.V. Krejčí, P. Hapala, A. Garcia-Lekue, T. Frederiksen, R. Zbořil, A. Arnau, H. Vázquez, and P. Jelínek, Phys. Rev. Lett. **119**, 166001 (2017)
- [35] S. Ballmann, R. Härtle, P.B. Coto, M. Elbing, M. Mayor, M.R. Bryce, M. Thoss, and H.B. Weber, Phys. Rev. Lett. **109**, 056801 (2012)
- [36] M. Rontani and E. Molinari, Phys. Rev. B **71**, 233106 (2005)
- [37] D. Toroz, M. Rontani, and S. Corni, J. Chem. Phys. **134**, 024104 (2011)
- [38] A. Donarini, B. Siegert, S. Sobczyk, and M. Grifoni, Phys. Rev. B **86**, 155451 (2012)
- [39] D. Toroz, M. Rontani, and S. Corni, Phys. Rev. Lett. **110**, 018305 (2013)
- [40] L.L. Patera, F. Queck, P. Scheuerer, and J. Repp, Nature **566**, 245 (2019)
- [41] N.S. Wingreen and Y. Meir, Phys. Rev. B **49**, 11040 (1994)
- [42] J. König, H. Schoeller, and G. Schön, Phys. Rev. Lett. **76**, 1715 (1996)

- [43] O. Karlström, C. Emary, P. Zedler, J.N. Pedersen, C. Bergenfeldt, P. Samuelsson, T. Brandes, and A. Wacker, *J. Phys. A* **46**, 065301 (2013)
- [44] R. Härtle, G. Cohen, D.R. Reichman, and A.J. Millis, *Phys. Rev. B* **88**, 235426 (2013)
- [45] F. Chen, M.A. Ochoa, and M. Galperin, *J. Chem. Phys.* **146**, 092301 (2017)
- [46] J. Rammer and H. Smith, *Rev. Mod. Phys.* **58**, 323 (1986)
- [47] S. Nakajima, *Prog. Theor. Phys.* **20**, 948 (1958)
- [48] M.G. Schultz, *Phys. Rev. B* **82**, 155408 (2010)
- [49] M. Misiorny, M. Hell, and M.R. Wegewijs, *Nat. Phys.* **9**, 801 EP (2013)
- [50] S. Koller, M. Grifoni, M. Leijnse, and M.R. Wegewijs, *Phys. Rev. B* **82**, 235307 (2010)
- [51] S. Koller, M. Grifoni, and J. Paaske, *Phys. Rev. B* **85**, 045313 (2012)
- [52] A. Dirnaichner, M. Grifoni, A. Prüfling, D. Steininger, A.K. Hüttel, and C. Strunk, *Phys. Rev. B* **91**, 195402 (2015)
- [53] J. Kern and M. Grifoni, *Eur. Phys. J. B* **86**, 384 (2013)
- [54] D. Mantelli: *Analytical and numerical study of quantum impurity systems in the intermediate and strong coupling regimes* (Ph.D. thesis, University of Regensburg, 2016)
- [55] C.J. Chen, *Phys. Rev. B* **42**, 8841 (1990)
- [56] L. Gross, N. Moll, F. Mohn, A. Curioni, G. Meyer, F. Hanke, and M. Persson, *Phys. Rev. Lett.* **107**, 086101 (2011)
- [57] S. Sobczyk, A. Donarini, and M. Grifoni, *Phys. Rev. B* **85**, 205408 (2012)
- [58] M. Braun, J. König, and J. Martinek, *Phys. Rev. B* **70**, 195345 (2004)
- [59] E.A. Osorio, K. Moth-Poulsen, H.S.J. van der Zant, J. Paaske, P. Hedegard, K. Flensberg, J. Bendix, and T. Bjornholm, *Nano Lett.* **10**, 105 (2010)
- [60] P. Gütllich, Y. Garcia, and H.A. Goodwin, *Chem. Soc. Rev.* **29**, 419 (2000)
- [61] I. Fernández-Torrente, D. Kreikemeyer-Lorenzo, A. Stróżecka, K.J. Franke, and J.I. Pascual, *Phys. Rev. Lett.* **108**, 036801 (2012)
- [62] S. Bouvron: *Gate-controlled scanning tunneling spectroscopy of CoPc molecules on graphene* (Ph.D. thesis, University of Konstanz, 2014)
- [63] J. Martínez-Blanco, C. Nacci, S.C. Erwin, K. Kanisawa, E. Locane, M. Thomas, F. von Oppen, P.W. Brouwer, and S. Fölsch, *Nat. Phys.* **11**, 640 EP (2015)
- [64] M.H. Hettler, W. Wenzel, M.R. Wegewijs, and H. Schoeller, *Phys. Rev. Lett.* **90**, 076805 (2003)
- [65] K. Kaasbjerg and K. Flensberg, *Phys. Rev. B* **84**, 115457 (2011)

- [66] D.A. Ryndyk, A. Donarini, M. Grifoni, and K. Richter, *Phys. Rev. B* **88**, 085404 (2013)
- [67] J.B. Mann: *Atomic Structure Calculations I. Hartree-Fock Energy Results* (Tech. Rep. LA-3690, Los Alamos Scientific Laboratory of the Univ. of California, 1967)
- [68] M.L. Perrin, C.J.O. Verzijl, C.A. Martin, A.J. Shaikh, R. Eelkema, J.H. van Esch, J.M. van Ruitenbeek, J.M. Thijssen, H.S.J. van der Zant, and D. Dulic, *Nat. Nanotechnol.* **8**, 282 (2013)
- [69] J. Koch and F. von Oppen, *Phys. Rev. Lett.* **94**, 206804 (2005)
- [70] F. Pielmeier and F.J. Giessibl, *Phys. Rev. Lett.* **110**, 266101 (2013)
- [71] M.S. Liao and S. Scheiner, *J. Chem. Phys.* **114**, 9780 (2001)
- [72] K.A. Nguyen and R. Pachter, *J. Chem. Phys.* **118**, 5802 (2003)
- [73] S.W. Wu, N. Ogawa, and W. Ho, *Science* **312**, 1362 (2006)
- [74] S.W. Wu, N. Ogawa, G.V. Nazin, and W. Ho, *J. Phys. Chem. C* **112**, 5241 (2008)
- [75] E. Hueckel, *Zeitschrift fur Physik* **70**, 204 (1931)
- [76] W. Barford: *Electronic and Optical Properties of Conjugated Polymers* (Oxford University Press, 2005)
- [77] F.J. Giessibl, *Appl. Phys. Lett.* **76**, 1470 (2000)
- [78] J.M. Soler, E. Artacho, J.D. Gale, A. García, J. Junquera, P. Ordejón, and D. Sánchez-Portal, *J. Phys. Condens. Matter* **14**, 2745 (2002)
- [79] W.J. Feast, J. Tsibouklis, K.L. Pouwer, L. Groenendaal, and E.W. Meijer, *Polymer* **37**, 5017 (1996)
- [80] R. Yamada, H. Kumazawa, T. Noutoshi, S. Tanaka, and H. Tada, *Nano Lett.* **8**, 1237 (2008)
- [81] R. Fitzner, E. Reinold, A. Mishra, E. Mena-Osteritz, H. Ziehlke, C. Körner, K. Leo, M. Riede, M. Weil, O. Tsaryova, A. Weiss, C. Uhrich, M. Pfeiffer, and P. Bäuerle, *Adv. Funct. Mater.* **21**, 897 (2011)
- [82] J. Repp, P. Liljeroth, and G. Meyer, *Nat. Phys.* **6**, 975 (2010)
- [83] D.A. Kisilitsyn, B.N. Taber, C.F. Gervasi, L. Zhang, S.C.B. Mannsfeld, J.S. Prell, A. Briseno, and G.V. Nazin, *Phys. Chem. Chem. Phys.* **18**, 4842 (2016)
- [84] F.E. Olsson, S. Paavilainen, M. Persson, J. Repp, and G. Meyer, *Phys. Rev. Lett.* **98**, 176803 (2007)
- [85] P.O. Gartland, S. Berge, and B. Slagsvold, *Phys. Rev. Lett.* **28**, 738 (1972)
- [86] F.E. Olsson, M. Persson, J. Repp, and G. Meyer, *Phys. Rev. B* **71**, 075419 (2005)
- [87] K. Kaasbjerg and K. Flensberg, *Nano Lett.* **8**, 3809 (2008)

15 Diagrammatic Monte Carlo

Nikolay Prokof'ev

University of Massachusetts

666 North Pleasant Street, Amherst, MA 01003, USA

Contents

1	Introduction	2
2	Diagrammatic Monte Carlo	2
2.1	Updates: general principles	3
2.2	Normalization	6
3	Fröhlich polaron	7
3.1	Type-I updates	8
3.2	Type-II updates. Data structure	11
3.3	Illustrative results: polaron Green function	14
4	Fermionic sign blessing I	16
4.1	Convergence of diagrammatic series for fermions	16
4.2	Resummation techniques	17
4.3	Shifted action	19
5	Fermionic sign blessing II	21
5.1	Sum or sample?	21
5.2	Determinant method for connected diagrams	22
5.3	Computational complexity problem for interacting fermions and its solution . .	23
5.4	Illustrative results: Hubbard model and unitary Fermi-gas	24
6	Conclusions	26

1 Introduction

This contribution reviews the principles and key ideas behind the Diagrammatic Monte Carlo method (DiagMC), as well as some technical details important for its efficient practical implementation. In short, DiagMC is a set of generic rules for unbiased sampling of the configuration space that involves a varying number of continuous variables. Applications of the method include series of connected Feynman diagrams and non-linear integral equations, lattice- and continuous-space path-integrals, continuous-time impurity-solvers, and any problem where the answer can be formally represented by the sum of multi-dimensional integrals. Once the configuration space to be simulated is defined, the DiagMC method will ensure that stochastic sampling is performed without systematic bias, leaving statistical error bars as the only source of uncertainty on the final answer.

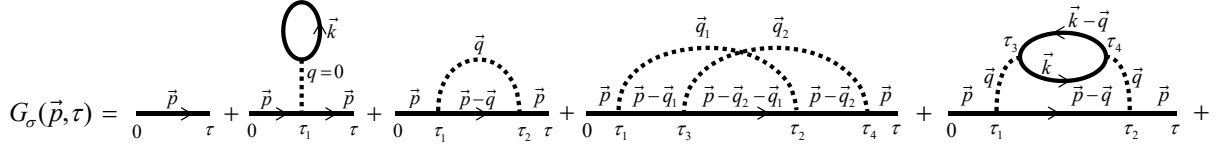
Properties of large systems cannot be obtained by direct enumeration of the exponentially growing configuration/Hilbert space, or ν -space, for brevity. A variety of numerical schemes rely on mathematical formulations instead, which, if solved, would reproduce the same statistical predictions as the original model. Path integrals, high-temperature expansions, and Feynman diagrams belong to this category of methods. I will focus on the Monte Carlo (MC) sampling technique [1], which is, arguably, among the most powerful universal tools designed to deal with large and complex ν -spaces, and explain in detail how it works in the space of connected Feynman diagrams. While each implementation is model and representation specific, most rules and considerations are generic.

2 Diagrammatic Monte Carlo

In the most abstract form one is interested in knowing some quantity $Q(\mathbf{y})$ as a function of variable $\mathbf{y}$ (in general, the multi-dimensional variable $\mathbf{y}$ may include both continuous and discrete components) when the answer is expressed as a series of multi-dimensional integrals/sums

$$Q(\mathbf{y}) = \sum_{n=0}^{\infty} \sum_{\mathbb{T}} \int \cdots \int d\mathbf{x}_1 \cdots d\mathbf{x}_n D(n, \mathbb{T}; \mathbf{x}_1, \dots, \mathbf{x}_n; \mathbf{y}), \quad (1)$$

with D being some known function of its arguments. The “diagram order” n controls the number of “internal” integration/summation variables, $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$, and the “topology” index $\mathbb{T}$ labels different terms of the same order in the series. The most familiar physics example, as the name of the technique suggests, would be Feynman diagrams for the many-body system illustrated in Fig. 1. Strict diagrammatic rules relate the graphical representation to the corresponding mathematical expression for the function D : up to a phase factor, it is given by the product of functions associated with the graph lines (often called propagators), $D = \prod_{\text{lines}} F_{\text{line}}$. For example, in momentum-imaginary time representation for the system of electrons interacting via the pairwise potential $V(\mathbf{r})$, the dotted lines are associated with the Fourier transform, $V(\mathbf{q})$, of the interaction potential and the solid lines with the single-particle propagators $G_0(\mathbf{p}_i, \tau)$.



$$G_o(\vec{p}, \tau) = \frac{\vec{p}}{0 \rightarrow \tau} + \frac{\vec{p}}{0 \rightarrow \tau_1} \frac{\vec{p}}{\tau_1 \rightarrow \tau} + \frac{\vec{p}}{0 \rightarrow \tau_1} \frac{\vec{p}-\vec{q}}{\tau_1 \rightarrow \tau_2} \frac{\vec{p}}{\tau_2 \rightarrow \tau} + \frac{\vec{p}}{0 \rightarrow \tau_1} \frac{\vec{p}-\vec{q}_1}{\tau_1 \rightarrow \tau_3} \frac{\vec{p}-\vec{q}_2-\vec{q}_1}{\tau_3 \rightarrow \tau_2} \frac{\vec{p}-\vec{q}_2}{\tau_2 \rightarrow \tau} \frac{\vec{p}}{\tau_2 \rightarrow \tau} + \frac{\vec{p}}{0 \rightarrow \tau_1} \frac{\vec{p}-\vec{q}}{\tau_1 \rightarrow \tau_3} \frac{\vec{p}-\vec{q}}{\tau_3 \rightarrow \tau_4} \frac{\vec{p}-\vec{q}}{\tau_4 \rightarrow \tau_2} \frac{\vec{p}-\vec{q}}{\tau_2 \rightarrow \tau} \frac{\vec{p}}{\tau_2 \rightarrow \tau} + \dots$$

Fig. 1: Graphical representation of the diagrammatic expansion for the Green function of an interacting many-body system.

There are well-established diagrammatic series for other quantities of interest such as self-energies, polarization operators, pair-propagators, current-current and other correlation functions, etc. Numerous alternative representations of quantum and classical models, such as path integrals and impurity solvers, are mathematically identical to Eq. (1). Thus, regardless of the origin of Eq. (1), it can be viewed as a mathematical expression for the answer in terms of a series of multi-dimensional integrals. The real challenge is to evaluate it with high accuracy.

Let us denote the collection of all external and internal parameters that lead to a complete characterization of the diagram as $\nu = (n, \mathbb{T}; \mathbf{x}_1, \dots, \mathbf{x}_n; \mathbf{y})$, and call it the “configuration space;” a particular set of parameters has to be viewed as a point in $\{\nu\}$. Accordingly, the modulus of D_ν will be called the configuration “weight.” Since, in general, the D -function is not sign-positive, we will need to introduce also the configuration “phase,” $\varphi_\nu = \arg D_\nu$ (the diagram phase is not necessarily equal to 0 or π).

2.1 Updates: general principles

The MC process of generating diagrams with probabilities proportional to their weight is based on the conventional Markov-chain updating scheme [2–4] implemented directly in the space of continuous variables. All updates are broadly classified as type-I and type-II. The number of continuous variables is not changed in type-I updates that perform sampling of diagrams of the same order n . Typical examples are shown in Fig. 2. They are based on the simplest possible local modifications of the topology and line parameters allowed by the rules and conservation laws. Their implementation is straightforward; e.g., for the update illustrated in Fig. 2(a) select at random any pair of consecutive interaction vertices and exchange their places. An acceptance ratio for the corresponding update, $R_{\nu \rightarrow \nu'}$, is given by the ratio of the diagram weights,

$$R_{\nu \rightarrow \nu'} = |D_{\nu'} / D_\nu|, \quad (2)$$

which is easily calculated, since D_ν is the product of F_{line} -functions and only three of them change their values in this update. Changing internal or external variables, see Figs. 2(b) and 2(c), is also standard. For example, one may select at random some interaction vertex and propose a new value for its time variable, $\tau_i \rightarrow \tau'_i$, from the (arbitrary) normalized probability density $P(\tau'_i)$. The acceptance ratio for this update is given by the ratio of probabilities for suggesting the $\nu \rightarrow \nu'$ and $\nu' \rightarrow \nu$ moves times the ratio of the diagram weights

$$R_{\nu \rightarrow \nu'} = \left| \frac{D_{\nu'}}{D_\nu} \right| \frac{P(\tau_i) d\tau}{P(\tau'_i) d\tau} = \left| \frac{D_{\nu'}}{D_\nu} \right| \frac{P(\tau_i)}{P(\tau'_i)}. \quad (3)$$

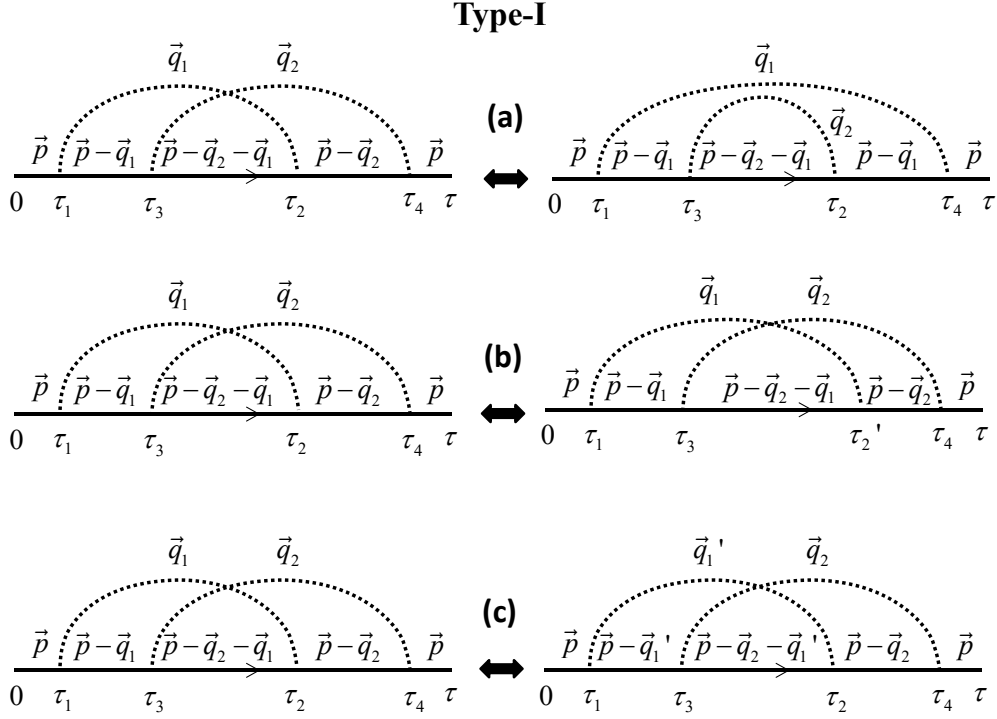


Fig. 2: Typical type-I updates in the configuration space of Feynman diagrams for polarons: (a) changing the diagram topology by permuting the end-points of two dashed lines; (b) changing the value of the internal variable τ_2 to τ_2' ; (c) changing the momentum transfer along the dashed line from $\mathbf{q}_1$ to $\mathbf{q}_1'$.

The simplest implementation of this update would be to have non-zero $P(\tau_i)$ only on the time interval determined by the times of the previous and following interaction vertexes (times τ_3 and τ_4 in Fig. 2(b)). The probability distribution $P(\tau)$ should be optimized for the best acceptance ratio without compromising one's ability to use it for fast generation of random variables (more details will be provided when we discuss the practical implementation of the technique for a Fröhlich polaron).

Clearly, there are numerous other possibilities for type-I updates which are standard for MC simulations of multidimensional integrals. For pedagogical reasons I will keep mentioning differential measures when I first state the acceptance ratio in order to see explicitly how they cancel out in the final answer.

Type-II updates change the diagram order $n \leftrightarrow n+m$ (they form complementary pairs of updates) and thus require that new variables be proposed from some (arbitrary) normalized probability density distribution $W(\nu; \mathbf{x}_{n+1}, \dots, \mathbf{x}_{n+m})$ when going from n to $n+m$, or erased from the diagram when going from $n+m$ to n . For example, to implement the transformation illustrated in Fig. 3 we need to propose time positions τ_3 and τ_4 and the momentum transfer $\mathbf{q}_2$ for the new dashed line. In the reverse update, these variables need to be erased. The detailed balance equation for a pair of updates which increase/decrease the diagram order by m reads

$$r_{n \rightarrow n+m} u_{n \rightarrow n+m} |D_\nu| W(\nu; \mathbf{x}_{n+1}, \dots, \mathbf{x}_{n+m}) (d\mathbf{x})^{n+m} = r_{n+m \rightarrow n} u_{n+m \rightarrow n} |D_{\nu'}| (d\mathbf{x})^{n+m}, \quad (4)$$

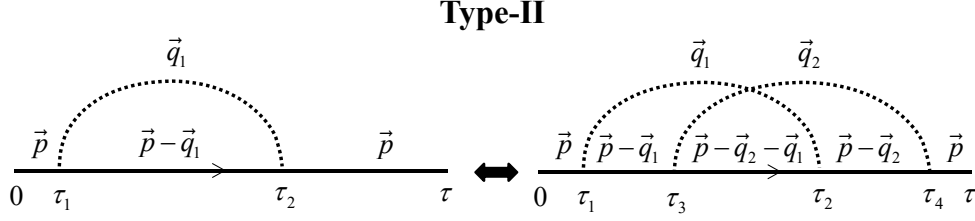


Fig. 3: Type-II updates that increase/decrease the diagram order by one.

where $u_{n \rightarrow n+m}$ and $u_{n+m \rightarrow n}$ are algorithm-specific probabilities of deciding which new diagram elements will be added or removed (for specific details, see the Fröhlich polaron Section), respectively, while $r_{n \rightarrow n+m}$ and $r_{n+m \rightarrow n}$ are the probabilities of accepting the update. An acceptance ratio, $R_{\nu \rightarrow \nu'} = r_{n \rightarrow n+m} / r_{n+m \rightarrow n}$, to go from configuration ν of order n to configuration ν' of order $n + m$ is then

$$R_{\nu \rightarrow \nu'} = \frac{u_{n \rightarrow n+m}}{u_{n+m \rightarrow n}} \left| \frac{D_{\nu'}}{D_{\nu} W(\nu; \mathbf{x}_{n+1}, \dots, \mathbf{x}_{n+m})} \right|. \quad (5)$$

As expected, all differential measures cancel in the acceptance ratio meaning that sampling of the configuration space with arbitrary and ever changing number of continuous variables can be done without encountering systematic errors. Note that the ratio of the diagram weights, $|D_{\nu'} / D_{\nu}|$, is some model-specific function of ν and the new variables $\mathbf{x}_{n+1}, \dots, \mathbf{x}_{n+m}$. The optimal choice of W is then a compromise between the efficiency (and programming convenience) of using it for seeding new variables and the largest (on average) acceptance ratio.

It is relatively straightforward to design a set of type-I and type-II updates that satisfies the ergodicity requirement: given two arbitrary configurations ν and ν' contributing to the answer, it should take the algorithm a finite number of updates with non-zero acceptance ratios to transform one configuration into another. At this point I would like to stress that DiagMC is fundamentally different from enumerating/listing all diagrams with orders $n \leq n_{\max}$ and then computing the corresponding multidimensional integrals for each diagram separately using classical MC methods. In DiagMC the diagram order, its topology, and all internal and external variables are treated on equal footing and are sampled stochastically from the probability distribution D_{ν} . From the DiagMC perspective, each diagram represents a point, not an integral, in the configuration space ν , see Fig. 4, and each sampled point, no matter the diagram order n , contributes equally to the statistics of the final result. For example, every diagram shown in Fig. 1 (with all internal variables specified) contributes $e^{i\varphi_{\nu}}$ to the statistics of $G(\mathbf{p}, \tau)$. One may wonder where did all the integrals go and why do the configuration space points with different differential measures contribute equally? Formally, this is what the detailed balance equation (5) is telling us. The other way to answer the question is as follows. One may pretend that *all* diagrams *are* of the same order(!) by interpreting unity factors in terms of the normalization integrals for W -functions in Eq. (5)

$$\int \dots \int d\mathbf{x}_1 \dots d\mathbf{x}_n D_{\nu} \times 1 = \int \dots \int d\mathbf{x}_1 \dots d\mathbf{x}_{n+m} D_{\nu} \times W(\nu; \mathbf{x}_{n+1}, \dots, \mathbf{x}_{n+m}). \quad (6)$$

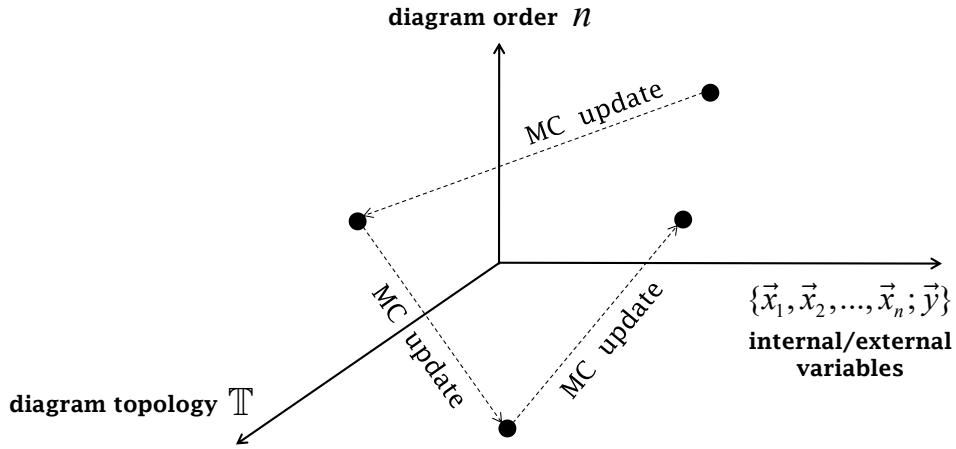


Fig. 4: An illustration of the DiagMC process: all configuration parameters are treated in the simulation protocol on an equal footing and are subject to local Markov-chain updates.

This point of view literally reduces type-II updates to type-I updates with the only caveat that one is free to consider any normalized W function for performing this “match of dimensions.” One last note. In DiagMC the autocorrelation time is almost never a problem. When the algorithm takes the configuration to the lowest-order diagram (at this point nearly all variables are erased), it can be de-correlated in $O(1)$ updates. Given that most many body simulations are done for expansion orders $\lesssim 10$, the autocorrelation times is measured in fractions of a millisecond for local updating schemes.

2.2 Normalization

Statistics collected for $Q(\mathbf{y})$, let us denote it as $Q_{MC}(\mathbf{y})$, grows linearly with the simulation time, and needs to be properly normalized to produce a physically meaningful result. Note that Eq. (1) is not based on a ratio of two quantities and, thus, normalization is done differently from the textbook example of the Ising model. Suppose that the lowest-order term in Eq. (1) for some value $\mathbf{y}_0$ is known because it does not involve any integrals and reduces to the analytic expression for $|Q^{(0)}(\mathbf{y}_0)| = |D(0; \mathbf{y}_0)|$. Stochastic sampling eventually brings the configuration to the lowest order, and this is when we update the normalization counter, $Z_N = Z_N + \delta_{n,0} \delta(\mathbf{y} - \mathbf{y}_0)$. We also realize that full statistics is the sum of contributions collected from different orders, $Q_{MC}(\mathbf{y}) = Q_{MC}(n = 0, \mathbf{y}) + Q_{MC}(n > 0, \mathbf{y})$, and if we were to determine the lowest-order contribution to the modulus of $Q(\mathbf{y}_0)$ we would get Z_N . This immediately tells us that the properly normalized answer for the final result is

$$Q(\mathbf{y}) = |Q^{(0)}(\mathbf{y}_0)| \frac{Q_{MC}(\mathbf{y})}{Z_N}. \quad (7)$$

The entire protocol can be called “normalization to known result.”

If there is some doubt that statistics for a given point $\mathbf{y}_0$ is representative, one can generalize the above idea by considering an integral, $I_N = \int |Q^{(0)}(\mathbf{y})| d\mathbf{y}$, which is assumed to be known either analytically or numerically (to any degree of accuracy). Each time the sampled diagram

order is zero, we add unity to the normalization counter, $Z_N = Z_N + \delta_{n,0}$, which is subsequently used to obtain the final result as

$$Q(\mathbf{y}) = I_N \frac{Q_{MC}(\mathbf{y})}{Z_N}. \quad (8)$$

In most cases, the lowest-order contributions are indeed trivial and there is no problem with implementing the normalization protocol. If none of the contributions to $Q(\mathbf{y})$ is known, we can still employ this protocol by adding a “fake diagram” with positive-definite weight, $D_F(\mathbf{y}) \delta_{n,0}$, and known normalization integral, $I_N = \int D_F(\mathbf{y}) d\mathbf{y}$, to the configuration space, and updating the Z_N counter each time this fake diagram is sampled. Equation (8) works for this setup without any modifications with the understanding that the fake diagram is used for normalization purposes only and is not contributing to $Q_{MC}(\mathbf{y})$. It is also worth noting that normalization can be always done to some positive-definite quantity; i.e., you will not face the sign-problem in the denominator.

3 Fröhlich polaron

Let us now focus on the Fröhlich polaron model and see in detail how the DiagMC technique can be used for obtaining the polaron Green function at zero temperature, see also [4]. The model Hamiltonian $H = H_e + H_{ph} + H_{e-ph}$ contains three terms where

$$H_e = \sum_{\mathbf{p}} (\epsilon(\mathbf{p}) - \mu) a_{\mathbf{p}}^\dagger a_{\mathbf{p}}, \quad H_{ph} = \sum_{\mathbf{q}} \omega(\mathbf{q}) b_{\mathbf{q}}^\dagger b_{\mathbf{q}}, \quad H_{e-ph} = \sum_{\mathbf{p}, \mathbf{q}} V(\mathbf{q}) (b_{\mathbf{q}}^\dagger - b_{-\mathbf{q}}) a_{\mathbf{p}-\mathbf{q}}^\dagger a_{\mathbf{p}}, \quad (9)$$

$$\epsilon(\mathbf{p}) = \frac{p^2}{2M}, \quad \omega(\mathbf{q}) = \Omega, \quad V(\mathbf{q}) = i \sqrt{2^{3/2} \pi \Omega^{3/2} \alpha / M^{1/2} q^2}, \quad (10)$$

with standard notations for creation and annihilation operators. The chemical potential, $\mu < 0$, is introduced here solely for the purpose of controlling the statistics at large times because otherwise it would diverge. The coupling between electrons and optical phonons with energy Ω can no longer be assumed weak when the dimensionless constant $\alpha > 1$. The Fröhlich polaron is the canonical model used to describe non-degenerate charge carriers in ionic semiconductors; Figure 5 explains why lattice effects can be neglected and one may proceed with the continuum description based on the parabolic dispersion relation, $\epsilon(p) = p^2/2M$, for the electron and a dispersionless optical phonon, $\omega(p) = \Omega = \text{const.}$

Connected diagrams for the Green function in imaginary time representation, $G(\mathbf{p}, \tau) = \langle c_{\mathbf{p}}(\tau) c_{\mathbf{p}}^\dagger \rangle$, are shown in Fig. 6. To convert the graphics into mathematical expressions (1) one has to use the following “conversion” rules:

- A straight line with momentum $\mathbf{p}_c$ between the time points τ_b and τ_a is associated with the bare polaron propagator, $G_0 = e^{-(\epsilon(p_c) - \mu)(\tau_b - \tau_a)}$.
- An ark connecting two dots is associated with the product of the phonon propagator, $D = e^{-\omega(p_c)(\tau_b - \tau_a)}$, the modulus of the coupling vertex squared, $|V(p_c)|^2$, and the momentum space integration factor $(2\pi)^{-3}$.

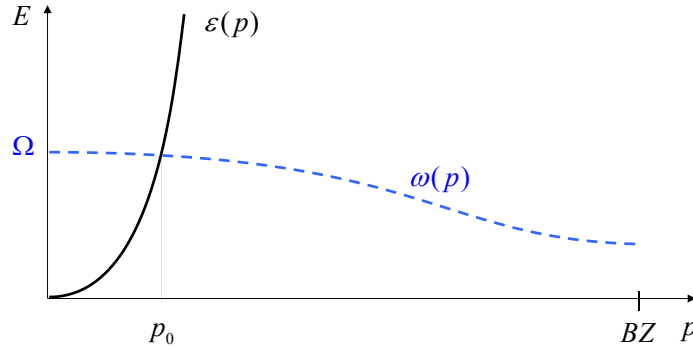


Fig. 5: Dispersion relations for electrons and phonons are such that their intersection and important physics effects take place at momenta $p \sim p_0 = \sqrt{2M\Omega}$, much smaller than the Brillouin zone boundary.

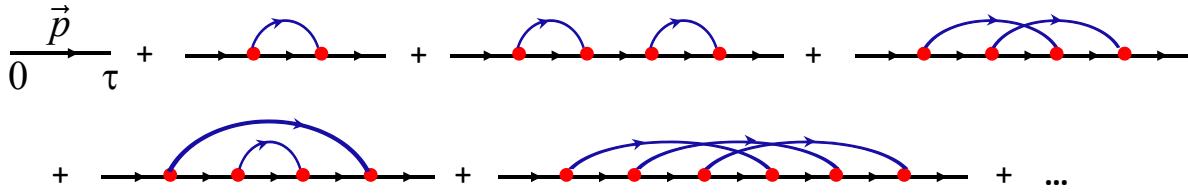


Fig. 6: Green function diagrams obtained by expanding in the number of phonon lines.

Thus, for this problem all diagrams are sign-positive and the series are convergent for any value of τ . Sign-positive series cannot be resummed; they are either meaningless or convergent. In our case, there are $(2n-1)!!$ diagrams of order n . On the other hand, the integration measure of $2n$ time-ordered points is $\tau^{2n}/(2n)!$, and this is sufficient to guarantee series convergence for Fröhlich polarons. The rest is a straightforward implementation of the generic DiagMC protocol. My set of updates is just one possible realization among many available.

3.1 Type-I updates

- *Global τ update.* The simplest update changing the global time variable τ is shown in Fig. 7(a). The probability density for the new value is a simple exponential

$$P(\tau') = E e^{-E(\tau' - \tau_{\text{last}})}, \quad E = p^2/2M - \mu, \quad (11)$$

and the acceptance ratio is unity because $D_{\nu'}/D_{\nu} = P(\tau')/P(\tau)$, see Eq. (3). [If r is a random number uniformly distributed on the interval $(0, 1)$, then $\tau' = \tau_{\text{last}} - E^{-1} \ln(r)$ by the transformation method.] Strictly speaking, this is the only type-I update required for ergodicity! Below I present several other type-I updates that can be added in order to (i) improve efficiency and reduce autocorrelations, (ii) have an over-complete set of updates for meaningful tests of the detailed balance, (iii) have fun.

- *Internal τ update.* Changing the time variable of the interaction vertex is equally easy, see Fig. 7(b). The probability density for the new value of τ'_b is a simple exponential

$$P(\tau'_b) = \frac{E e^{-E(\tau'_b - \tau_a)}}{1 - e^{-E(\tau_c - \tau_a)}}, \quad E = (p_a^2 - p_b^2)/2M \pm \Omega, \quad (12)$$

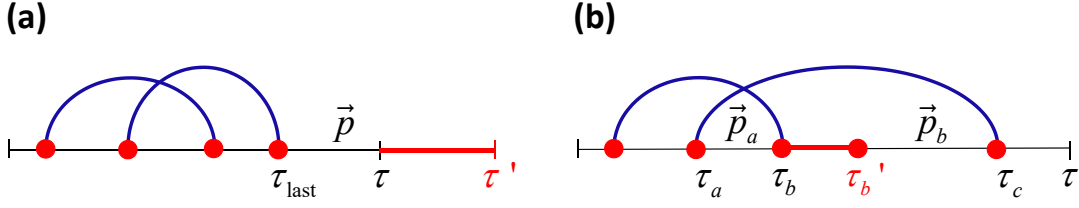


Fig. 7: Type-I updates changing the external (a) and internal (b) time variables:

and the acceptance ratio is unity for exactly the same reason as for the *global* τ update.

[$\tau'_b = \tau_a - E^{-1} \ln[1 - r(1 - e^{-E(\tau_c - \tau_a)})]$ by the transformation method.]

• *Rescaling all τ variables update.* By introducing dimensionless time variables $s_i = \tau_i/\tau$, and paying attention that all propagators are exponential functions of time, we realize that the diagram dependence on the global time τ is given by the Poisson distribution. If we were to use

$$P(\tau') = \frac{E(E\tau')^{2n}}{(2n)!} e^{-E\tau'}, \quad E = \sum_{i=1}^{2n+1} \Delta s_i E_i - \mu, \quad (13)$$

where the sum is over all time intervals in the graph, and E_i is the energy associated with each interval (counting both the polaron energy and the energy of all phonon lines covering it), to propose new values of τ' , we would always accept the update. Unfortunately, for this distribution the transformation method cannot be used. However, for large $E\tau$ and n , the Poisson distribution is approaching the Gaussian. The idea then is to use

$$P(\tau') = \frac{E}{\sqrt{4\pi n}} e^{-(E\tau' - 2n)^2/4n}, \quad (14)$$

instead. [From now on I will stop mentioning the transformation method and how it is used (sometimes with tricks) for well-know distributions.] Obviously, the update is rejected, when the proposed τ' is negative—one should not be afraid of proposing unphysical values if this leads to radical simplifications of the numerical procedure with only a minor loss of efficiency. The acceptance ratio is given by

$$R = \exp \left(2n \ln \frac{\tau'}{\tau} - E(\tau' - \tau) + \frac{(E\tau' - 2n)^2 - (E\tau - 2n)^2}{4n} \right), \quad (15)$$

and is close to unity (on average) for large diagram orders.

• *Internal $|q|$ update.* To change the modulus of the phonon momentum (the direction is preserved) we select at random any of the phonon lines (probability of selecting a particular one is $1/n$) and propose the new value for q from the Gaussian probability density

$$P(q') = \frac{1}{\sqrt{2\pi s^2}} e^{-(q' - q_0)^2/2s^2}, \quad (16)$$

where $q_0 = \langle \mathbf{p} \rangle \cdot \mathbf{q}/q$, $s^2 = M/(\tau_b - \tau_a)$, and $\langle \mathbf{p} \rangle = (\tau_b - \tau_a)^{-1} \int_{\tau_a}^{\tau_b} d\tau [\mathbf{p}(\tau) + \mathbf{q}]$ is the “average” electron momentum on the time interval (τ_a, τ_b) in the absence of the updated phonon propagator, see Fig. 8(a). This update is always accepted (provided q' is non-negative) because $P(q')$ is

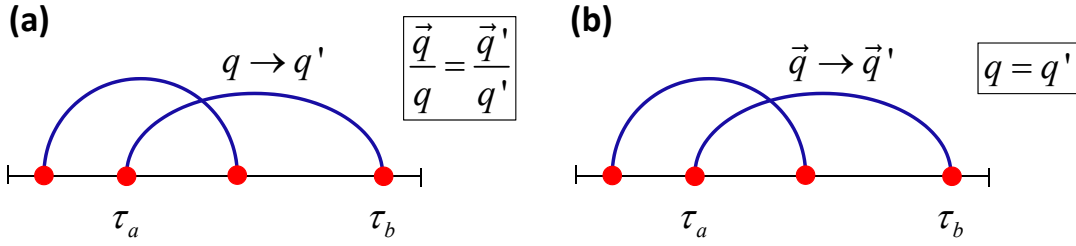


Fig. 8: Type-I updates changing the momentum variable.

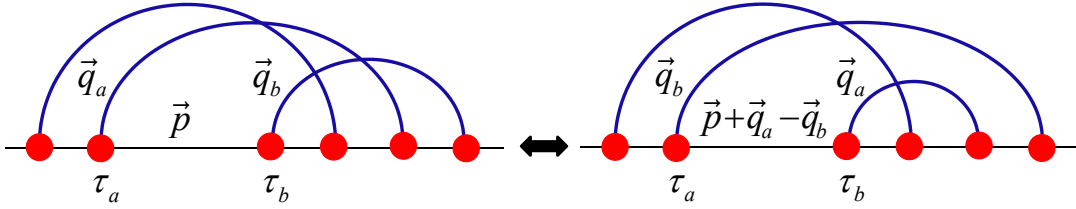


Fig. 9: Type-I update changing the local diagram topology.

reproducing precisely the diagram weight for the Fröhlich polaron. Indeed,

$$\frac{1}{2M} \int_{\tau_a}^{\tau_b} d\tau [(\mathbf{p}(\tau) + \mathbf{q} - \mathbf{q}')^2 - p^2] = \text{const.} + \frac{\tau_b - \tau_a}{2M} (q' - q_0)^2. \quad (17)$$

- *Internal $\mathbf{q}/q$ update.* To change the phonon momentum direction while keeping its modulus fixed, we select at random any of the phonon lines and propose the new value for $\mathbf{q}/q$ from the uniform distribution for the azimuthal angle φ and exponential distribution for the cosine of the polar angle θ

$$P(\varphi, \theta) = \frac{A \sin(\theta)}{4\pi \sinh(A)} e^{A \cos(\theta)}, \quad (18)$$

where $A = (\tau_b - \tau_a)|\langle \mathbf{p} \rangle|q/M$. Both angles are defined relative to the axis set by the vector $\langle \mathbf{p} \rangle$ defined in the previous update, see also Fig. 8(b). This update is always accepted because (18) reproduces the functional dependence of (17) on updated angles.

- *Topology change.* Here the idea is to select at random any nearest-neighbor (n.n.) pair of vertices and swap their places, see Fig. 9. The momenta of phonon propagators remain fixed except when the selected pair is connected by the phonon line, in which case it changes sign. This proposal will change the momentum of the polaron line to $\mathbf{p}' = \mathbf{p} + \mathbf{q}_a - \mathbf{q}_b$. The acceptance ratio is directly related to the ratio of diagram weights

$$R = \exp\{-(\tau_b - \tau_a)[\epsilon(p') - \epsilon(p) \pm \omega(q_a) \pm \omega(q_b)]\}, \quad (19)$$

where the proper $\pm$ option has to be chosen depending on whether the corresponding phonon propagator is getting longer or shorter in time.

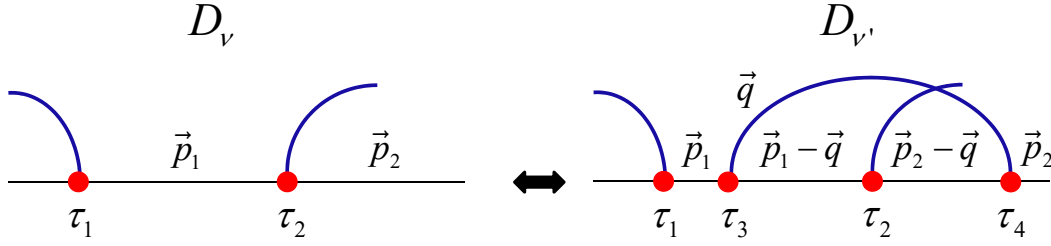


Fig. 10: Type-II updates changing the diagram order.

3.2 Type-II updates. Data structure

The design of type-II updates, especially in view of Eq. (6), is as flexible. I will only describe one of them following a particular strategy of selecting new variables. Formally, this type-II update and the first update described in the previous subsection, constitute an ergodic set of MC procedures capable of simulating the Green function dependence on time.

• *Increase/decrease the diagram order by one.* Type-II updates typically come in complementary pairs that satisfy the detailed balance condition within each pair. Let $p_{n \rightarrow n+1}$ and $p_{n+1 \rightarrow n}$ be the probabilities to make a decision to apply one of these two updates. In the *increase* update propose the following steps:

- Select one of the polaron propagators at random (corresponding probability: $1/(2n+1)$); let the parameters of the selected interval be $\mathbf{p}_1, \tau_1, \tau_2$, see Fig. 10.
- propose the first new time variable τ_3 from the uniform probability density $1/\Delta\tau$, where $\Delta\tau = \tau_2 - \tau_1$.
- propose the momentum for the new phonon propagator from the probability distribution

$$\frac{\sin(\theta)}{4\pi} \frac{1}{p_0(1 + q/p_0)^2}, \quad (20)$$

where $p_0 = \sqrt{2M\Omega}$, see Fig. 5. It is uniform for the solid angle of $\mathbf{q}$ and has an easy to handle power law for the modulus of $\mathbf{q}$.

- propose the new time variable $\tau_4 > \tau_3$ from the distribution

$$\Omega(1 + q/p_0)^2 e^{-\Omega(1+q/p_0)^2(\tau_4-\tau_3)}. \quad (21)$$

As usual, the update is rejected if $\tau_4 > \tau$. This distribution is not a perfect match to the ratio of the diagram weights in the proposed setup

$$\left| \frac{D_{v'}}{D_v} \right| = |V(q)|^2 e^{-\Delta E(\tau_4-\tau_3)} \frac{q^2 \sin(\theta)}{(2\pi)^3} \propto e^{-\Delta E(\tau_4-\tau_3)} \sin(\theta), \quad (22)$$

where the energy change on the updated interval is given by $\Delta E = \Omega + [q^2 - 2\mathbf{q} \cdot \langle \mathbf{p} \rangle]/2M$, and $\langle \mathbf{p} \rangle = (\tau_4 - \tau_3)^{-1} \int_{\tau_3}^{\tau_4} d\tau \mathbf{p}(\tau)$. However, it is “good enough” in terms of the average acceptance ratio.

The *decrease* update is conceptually very simple: Select any of the existing phonon propagators (probability is $1/n$, where n is the *current* diagram order), and propose to remove it. We are all set to formulate the detailed balance equation

$$r_{n \rightarrow n+1} |D_\nu| \frac{p_{n \rightarrow n+1}}{2n+1} \frac{\Omega \sin(\theta)}{4\pi \Delta\tau p_0} e^{-\Omega(1+q/p_0)^2(\tau_4 - \tau_3)} = r_{n+1 \rightarrow n} |D_{\nu'}| \frac{p_{n+1 \rightarrow n}}{n+1} \quad (23)$$

and its solution for the *increase* update

$$R_{n \rightarrow n+1} = \frac{p_{n+1 \rightarrow n}}{p_{n \rightarrow n+1}} \frac{2n+1}{n+1} \frac{2\alpha\Omega\Delta\tau}{\pi} e^{[qp_0 + \mathbf{q} \cdot \langle \mathbf{p} \rangle](\tau_4 - \tau_3)/M}. \quad (24)$$

Notice that the second ratio contains $(n+1)$, the number of phonon propagators in the proposed configuration, in the denominator. The solution of the same detailed balance equation for the *decrease* update acceptance ratio reads:

$$R_{n \rightarrow n-1} = \frac{p_{n-1 \rightarrow n}}{p_{n \rightarrow n-1}} \frac{2n-1}{n} \frac{\pi}{2\alpha\Omega\Delta\tau} e^{-[qp_0 + \mathbf{q} \cdot \langle \mathbf{p} \rangle](\tau_4 - \tau_3)/M}. \quad (25)$$

Again, one has to be careful in formulating it in terms of the current, order n , and proposed, order $n-1$, configuration parameters. In particular, $\Delta\tau$ is the duration of the polaron interval where the removed polaron propagator starts *after* the corresponding phonon propagator is removed (it may be the case that both τ_3 and τ_4 are smaller than τ_2 in Fig. 10). Also, the average polaron momentum $\langle \mathbf{p} \rangle$ needs to be computed for the proposed configuration.

- As far as normalization is conserved, the easiest way would be to normalize to the known integral of the bare Green function

$$I_N = \int_0^\infty d\tau e^{-(p^2/2M - \mu)\tau} = \frac{1}{p^2/2M - \mu}. \quad (26)$$

- At this point it is worth saying a couple of words about the data structure because recovering the necessary information for performing updates is often crucial for the efficiency of the algorithm. Most updates are designed to modify the diagram structure and its parameters locally (in terms of graph connections); i.e., only a few parameters and propagators are involved in each update. [This requirement does not apply to DiagMC algorithms based on exact summation of all diagram topologies that are discussed in the next Section.] Correspondingly, the data structure should be implemented in such a way that updates can be completed after performing $O(1)$ operations in the limit of $(n, \tau) \rightarrow \infty$. Here is one possible structure:

1. Every polaron propagator (or “interval”) in the diagram has a unique label $\ell > 0$.
2. This label is used to retrieve information about the propagator momentum, as well as its initial, and final times, using $p(\ell, 1:3)$, $\tau_i(\ell)$, and $\tau_f(\ell)$ arrays, respectively. One can introduce additional arrays, if necessary, and update the corresponding information; e.g., $N_{\text{ph}}(\ell)$ returns the number of phonon propagators covering the ℓ interval (number of phonons in the virtual state at time $\in (\tau_i(\ell), \tau_f(\ell))$).

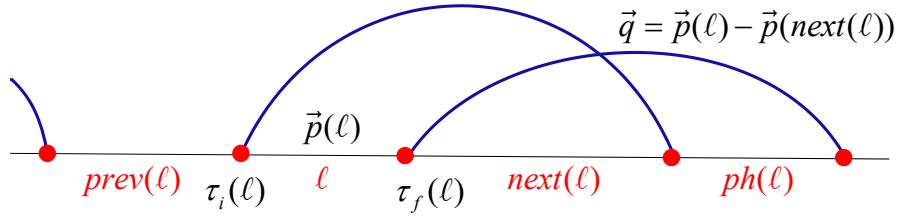


Fig. 11: Data structure for efficient implementation of local updates.

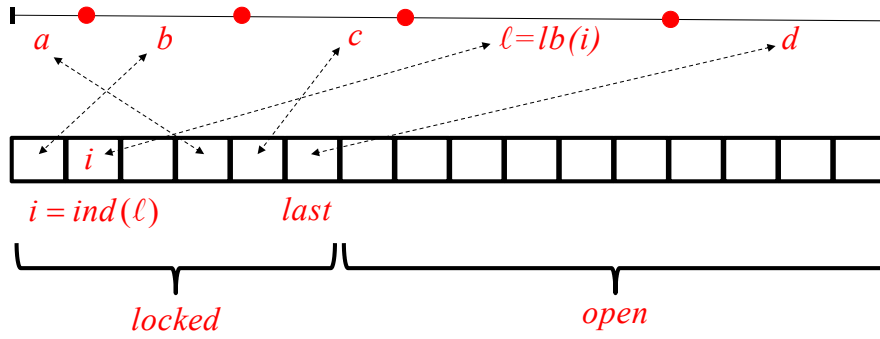


Fig. 12: Efficient management of a finite set of labels.

3. All labels are linked pairwise in two ways. Arrays $prev(\ell)$ and $next(\ell)$, see Fig. 11, allow one to get labels of intervals immediately preceding and following ℓ . Obviously, $next(prev(\ell)) = \ell$, and $prev(next(\ell)) = \ell$. An array $ph(\ell)$ establishes a link between the left ends of the intervals connected by the phonon propagator. Again, $ph(ph(\ell)) = \ell$. It is easy to see that any local (around ℓ) information about the graph properties can be quickly recovered without knowing the global structure. An ordered array $1, 2, 3, \dots, 2n+1$ of labels cannot be used because if some label ℓ is removed from the middle of the list (see also text below), all labels with values $> \ell$ must be updated, in violation of the “locality” principle.
4. In a long simulation run new/existing intervals will be created/erased trillions of times. It is thus not practical to never use the same label twice. There is, however, a simple programming trick that allows one to use the same set of labels forever, and manage it with $O(1)$ operations. At any moment it is known which $n+1$ labels are “locked” because they are already utilized for linking the graph intervals and which ones are “open” for labeling new elements. This is achieved with the help of the $ind(\ell)$ and $lb(i)$ arrays. By definition, labels $lb(1), lb(2), \dots, lb(last = 2n+1)$ are “locked” and labels $lb(> 2n+1)$ are “open,” while $ind(lb(i)) = i$ establishes a connection between the graph labels and an ordered array of indexes, see an illustration in Fig. 12. When a new interval is added to the diagram it is given a label $lb(last+1)$ and the variable $last$ is increased by one. When some interval and its label ℓ are removed, one first determines its index $i = ind(\ell)$. If $i = last$, the value of $last$ is decreased by one; otherwise, it is necessary to swap an association between indexes i and $last$ and labels ℓ and $lb(last)$, and only then decrease the value of $last$ by one.

3.3 Illustrative results: polaron Green function

One can employ DiagMC for in-depth studies of polaron properties, including direct access to every coefficient (by modulus) in the Lehmann expansion of the exact wave function with momentum $\mathbf{p}$

$$|\Psi_{\mathbf{p}}\rangle = c_{\mathbf{p}} a_{\mathbf{p}}^{\dagger} |0\rangle + \sum_{\mathbf{q}} c_{\mathbf{p},\mathbf{q}} b_{\mathbf{q}}^{\dagger} a_{\mathbf{p}-\mathbf{q}}^{\dagger} |0\rangle + \sum_{\mathbf{q}_1 \mathbf{q}_2} c_{\mathbf{p},\mathbf{q}_1,\mathbf{q}_2} b_{\mathbf{q}_1}^{\dagger} b_{\mathbf{q}_2}^{\dagger} a_{\mathbf{p}-\mathbf{q}_1-\mathbf{q}_2}^{\dagger} |0\rangle + \dots \quad (27)$$

With this one can determine the probability of finding an electron along with a given number of phonons in the polaron cloud [4]. Here I will only review how one computes the Green function and extracts the polaron Z -factor, $Z_{\mathbf{p}} = |c_{\mathbf{p}}|^2$, and energy from its asymptotic expression

$$G_{\mathbf{p}}(\tau \rightarrow \infty) \rightarrow Z_{\mathbf{p}} e^{-(E_{\mathbf{p}} - \mu)\tau}. \quad (28)$$

The simplest way of collecting statistics for $G_{\mathbf{p}}(\tau)$ would be to split the entire τ axis into small bins $\Delta_i = \tau_i - \tau_{i-1}$ and update bin counters for an integral of the function over the bin size (for brevity I will suppress the momentum index)

$$I^{(i)} = \int_{\tau_{i-1}}^{\tau_i} d\tau G(\tau), \quad (29)$$

$$I_{MC}^{(i)} = I_{MC}^{(i)} + 1, \quad \text{if } \tau \in (\tau_i, \tau_{i-1}). \quad (30)$$

For small bins, an estimate for the Green function at point $\bar{\tau}_i = (\tau_i + \tau_{i-1})/2$ can be made as

$$G(\bar{\tau}_i) \approx \frac{I_N}{Z_N} \frac{I_{MC}^{(i)}}{\Delta_i}. \quad (31)$$

While integrals (29) are free of systematic errors, the most straightforward final step (31) does involve a finite bin size error.

There are several ways for eliminating this systematic error. The bin hierarchy method [5, 6] takes the limit of very small bins and overcomes the problem of large statistical noise for small bins by restoring the entire function $G(\tau)$ using splines with self-adaptive nodes. This protocol is rather technical to reproduce here but it is very efficient and ensures that the systematic errors are always smaller than the statistical ones.

One may also keep the bin sizes fixed, but collect several integrals over bins to improve the accuracy of restoring $G(\tau)$. Any smooth function on a given interval can be expanded as

$$G(\tau) = \sum_{j=1}^{\infty} \alpha_j^{(i)} e_j^{(i)}(\tau), \quad \text{if } \tau \in (\tau_i, \tau_{i-1}), \quad (32)$$

where $\{e_j^{(i)}(\tau)\}$ is an ortho-normal basis (ONB) on the interval (τ_i, τ_{i-1}) . It may be Legendre polynomials, but any other ONB with an inner product defined as

$$\langle f|g \rangle = \int_{\tau_{i-1}}^{\tau_i} d\tau w^{(i)}(\tau) f(\tau) g(\tau), \quad (33)$$

can be used instead. In practice, the series is truncated at some finite order N_i , that is determined to provide an accurate description of $G(\tau)$ with unmeasurable (within statistical errors) systematic bias. To account for divergencies in $G(\tau)$ one may need to use singular basis functions and $w > 0$ weights. For example, if $G(\tau \rightarrow 0) \propto 1/\tau^{1/2}$, the first basis function in the set may be $e_1(\tau) = A_a/\tau^{1/2}$. In this case, the $w(\tau)$ is required to ensure that the normalization integral for e_1 is finite, and $w(\tau) = \tau^{1/2}$ will do the job. It is also possible to consider infinite size bins; if the leading asymptotic decay is of power-law type $G(\tau \rightarrow \infty) \propto 1/\tau^a$, the ONB on the time interval (τ_h, ∞) may be constructed from the set of functions $e_j \propto 1/\tau^{a+b_j}$ with $b_1 = 0$ and $b_{j>1} > 0$ to account for the dominant term and several subleading corrections.

According to the theory of Hilbert spaces, the coefficients of expansion are determined by the integrals

$$\alpha_j^{(i)} = \int_{\tau_{i-1}}^{\tau_i} d\tau w^{(i)}(\tau) e_j^{(i)}(\tau) G(\tau), \quad (34)$$

with unbiased MC estimators

$$\alpha_{j,MC}^{(i)} = \alpha_{j,MC}^{(i)} + w^{(i)}(\tau) e_j^{(i)}(\tau), \quad \text{if } \tau \in (\tau_i, \tau_{i-1}). \quad (35)$$

After appropriate normalization of statistics, one obtains the Green function from

$$G(\tau) = \frac{I_N}{Z_N} \sum_{j=1}^{N_i} \alpha_{j,MC}^{(i)} e_j^{(i)}(\tau). \quad (36)$$

Obviously, the conventional procedure described by Eqs. (29)-(31) is nothing but the special case when there is only one constant basis function.

Finally, one can use the reweighing method for an unbiased estimate of the function at a specified set of points $\bar{\tau}_i$. For each point one decides on the interval (τ_i, τ_{i-1}) that will be used to collect statistics for $G(\bar{\tau}_i)$; there are no formal restrictions on the sizes and locations of these intervals or their overlaps for different points. An optimal choice would be to have $\bar{\tau}_i$ roughly in the middle of the interval, and the interval width Δ_i to be small enough to avoid multi-scale variations of G within the interval. For any simulated point than falls within the interval, a factor $D_\nu(\bar{\tau}_i)/D_\nu(\tau)$ accounts for the difference between the $G(\bar{\tau}_i)$ and $G(\tau)$ functions. Thus

$$G(\bar{\tau}_i) \Delta_i = \int_{\tau_{i-1}}^{\tau_i} d\tau G(\tau) \frac{D_\nu(\bar{\tau}_i)}{D_\nu(\tau)}, \quad (37)$$

implying that an unbiased estimator for $G(\bar{\tau})$ is given by

$$G_{MC}(\bar{\tau}_i) = G_{MC}(\bar{\tau}_i) + \frac{D_\nu(\bar{\tau}_i)}{D_\nu(\tau) \Delta_i}, \quad \text{if } \tau \in (\tau_i, \tau_{i-1}). \quad (38)$$

The normalization of statistics using I_N/Z_N does not change.

Once the data for $G(\tau)$ are collected they are analyzed according to Eq. (28), see left panel in Fig. 13. The dependence of the quasiparticle residue on the coupling constant is shown in the right panel of Fig. 13. It is not unusual to have statistical errors for this problem at the level of 10^{-6} in relative units. Many more results and direct MC estimators for polaron properties can be found in Ref. [4].

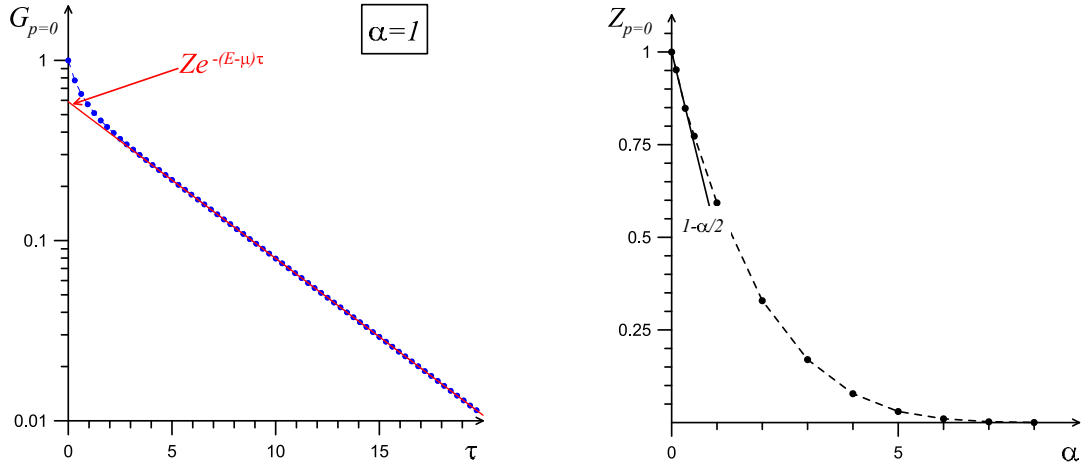


Fig. 13: Left panel: Fröhlich polaron Green function at zero momentum for $\alpha = 1$ (we use units such that $M = 1$ and $\Omega = 1$) and its asymptotic single exponential behavior. Right panel: Quasiparticle residue at $p = 0$ as a function of the coupling constant.

4 Fermionic sign blessing I

DiagMC for a generic interacting fermionic system follows the same rules. Clearly, the diagrams themselves are different, see Fig. 1, the differences in the interaction Hamiltonian. Also, the diagrams are no longer sign-positive. Even in the absence of gauge fields, the diagram sign may change for the following reasons: (i) for repulsive interactions and expansion in the interaction potential the diagram sign contains $(-U)^n$, (ii) fermionic propagators are subject to anti-periodic boundary conditions, $G(\tau < 0) = -G(\tau + \beta)$, (iii) the diagram sign contains $(-1)^L$ where L is the number of fermionic loops in a given topology. This raises two important questions: “Are the diagrammatic series convergent and under what conditions?” and “What is the role of the fermionic sign?” In what follows, I will explain that the two questions are closely related. The bottom line is that the DiagMC technique works and can be made very efficient thanks to the fermionic sign; i.e., it is a “blessing,” not a problem.

4.1 Convergence of diagrammatic series for fermions

Dyson’s argument for perturbative expansions in the coupling constant states that the convergence radius of the series is zero for continuous-space systems, no matter whether for bosons or fermions. Indeed, if it were finite, system properties would be analytic functions of the coupling constant near the origin. However, when the sign of the interaction is flipped from repulsion to attraction, continuous space systems collapse to a point (infinite density state) because even for fermions the kinetic energy increase $\propto n^{5/3}$ cannot overcome the potential energy gain $-|g|n^2$. To understand how this physical intuition is reflected in the mathematical structure of the series, notice that the number of different topologies for diagrams of order $n \gg 1$ is proportional to $n!$. Since $n!$ always beats the power law dependence on the coupling constant g^n , there is also a formal reason for suspecting that the convergence radius might be zero.

It appears then that Dyson's argument makes the entire diagrammatic approach for many body systems nearly useless in the strongly correlated regime. This is where the Pauli principle and the fermionic sign come to rescue. To begin with, the collapse argument does not apply to lattice systems, such as the famous single-band tight-binding Hubbard model. At best, one can place two, not more, fermions with opposite spins on the site and the system cannot collapse to infinite density. Thus, we expect that finite temperature properties of the Hubbard model are analytic functions of the coupling constant U in the limit of $U \rightarrow 0$ and this is confirmed by exact mathematical considerations. Likewise, Dyson's argument can be refuted for continuous space fermions with high momentum cutoff that acts similarly to the Brillouin zone boundary. Extrapolating converged calculations to the infinite cutoff limit may be a well-defined procedure. Finally, the diagrammatic technique admits an infinite number of alternative formulations when certain geometric series of the original diagrams are accounted for right from the beginning are incorporated self-consistently into the new "expansion point;" typical examples include mean-field and dynamic mean-field theories, ladder summations, screening, and solutions for the low-order skeleton set. The new expansion is no longer in terms of the coupling constant, and the original Dyson argument does not apply directly. Self-consistent mean-field and skeleton set solutions, on top of which the new expansion is made, can easily incorporate non-analytic dependence on the bare coupling constant, as, e.g., in the famous BCS solution. Going back to the mathematical structure of the diagrammatic expansion, we realize that the only possibility for the series to converge, despite factorial scaling of the number of allowed topologies, is to have massive cancellations between the diagrams within the same order. This is "sign blessing I:" the DiagMC method relies on the fermionic sign because this is the necessary condition for having series with nice properties. In what follows I will discuss simple illustrative examples demonstrating how resummation techniques allow one to extract accurate answers from divergent sign-alternating series. [If series converge, it is time to publish the solution.]

4.2 Resummation techniques

One way to deal with divergent sign-alternating series, $Q = \sum_{n=1}^{\infty} d_n$, outside their finite convergent radius is as follows. Introduce a smooth function, $f(n, \epsilon)$, that satisfies two conditions: for $\epsilon \rightarrow 0$ and finite n it approaches unity, $f(n, 0) = 1$; for $n \rightarrow \infty$ and finite ϵ it goes to zero faster than an exponential function, $f(n \rightarrow \infty, \epsilon) a^n \rightarrow 0$ for any $a > 1$. The resummed series

$$Q_{\epsilon} = \sum_{n=1}^{\infty} d_n f(n, \epsilon) , \quad (39)$$

is guaranteed to converge because $f(n, \epsilon)$ suppresses the geometrical divergence of the original series, while in the limit of $\epsilon \rightarrow 0$ the original and resummed series coincide. By extrapolating Q_{ϵ} to $\epsilon = 0$ one effectively performs an analytic continuation of the sign-alternating series outside of its convergence radius, see Fig. 14(a). The $f(n, \epsilon)$ function is up to you to design, because apart from the conditions specified it is rather arbitrary; different choices for f provide a good estimate for the error introduced by extrapolation from finite values of ϵ to zero.

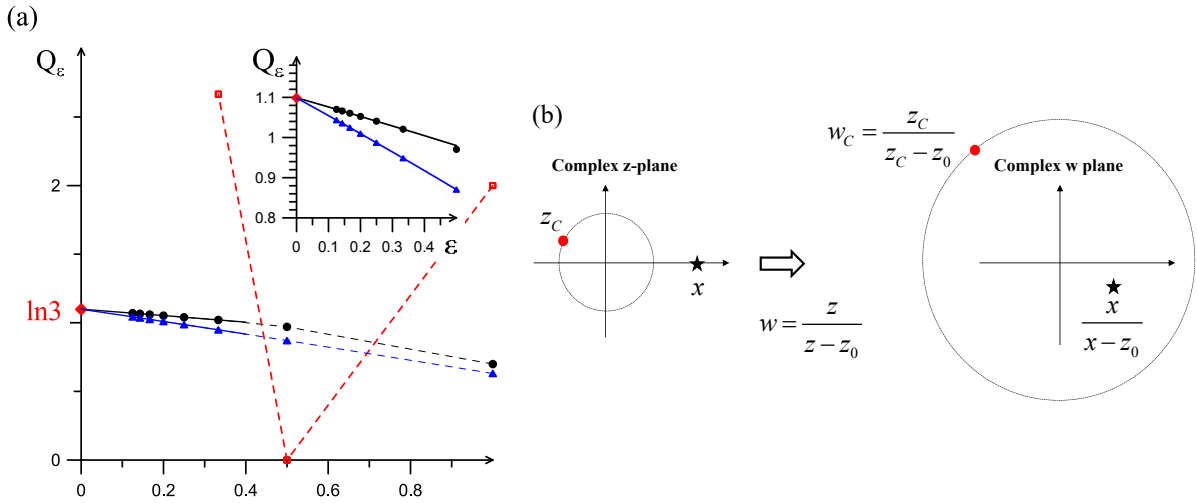


Fig. 14: (a) Resummation of divergent series for $\ln(1+x)$ with $x = 2$ using $f(n, \epsilon) = e^{-\epsilon n^2}$ (circles) and $f(n, \epsilon) = e^{-\epsilon n^{3/2}}$ (triangles). Extrapolation to $\epsilon = 0$ was performed using parabolic fits. Partial sums $\sum_{n=1}^{[\epsilon]} (-x)^{n+1}/n$ are shown by open squares. The value $\ln 3$ is marked by the diamond on the vertical axis. (b) Moving a simple-pole and increasing the convergence radius using conformal mapping.

The above protocol is blind to specific properties of the series and may require knowledge of many terms in the series for reliable extrapolation, especially for less “aggressive” f -functions. More efficient methods exist when the reason for reaching the convergence radius is known better. Suppose that the series behaves as $d_n = \gamma_n x^n$ with $\gamma_n = (-1)^n$ and $M = 10$ terms are known. The goal is get an answer for $x = 3$, well outside of the radius of convergence. From available information one can roughly estimate the convergence radius, and produce constant phase lines $y(x)$ for the complex function $Q(z = x + iy) = \sum_n^M \gamma_n z^n$ to establish that a simple pole is located close to the real axis, say at $z_0 \approx -1.05$. Next, one performs a conformal mapping $w = z/(z - z_0)$, or $z = -wz_0/(1 - w)$ and constructs the Taylor series for $Q_w(w) = \sum_n^M \sigma_n w^n$. The final answer is given by $Q_w(x/(x - z_0))$; with extraordinary accuracy it reproduces $1/(1 + x) = 0.25$. Under conformal mapping the singularities are moved away from the origin of the expansion and the point of interest ends up well within the radius of convergence, see the illustration in Fig. 14(b).

Similarly, it is possible to handle poles of higher order or several poles, but high accuracy rests on the number of known terms in the series. A slightly different version of the method is known as extrapolation by Padé approximants. One assumes that the function behind the series is given by the ratio of two polynomials, $Q(z) = P_k(z)/P_m(z)$, with $k + m \leq M$. For each (k, m) pair the polynomials are determined by matching the coefficients of the Taylor series for the ratio to γ_n . The final answer is determined by examining how $P_k(x)/P_m(x)$ depends on (k, m) when we increase the order of polynomials.

Conformal mappings can be also used to improve the convergence properties of series by moving branch cuts away from the origin. The ratio of polynomials can be replaced by the ratio of hypergeometric functions to achieve efficient extrapolation in cases when the convergence radius is limited by the branch cuts [7]. The mathematical and physical literature on the topic

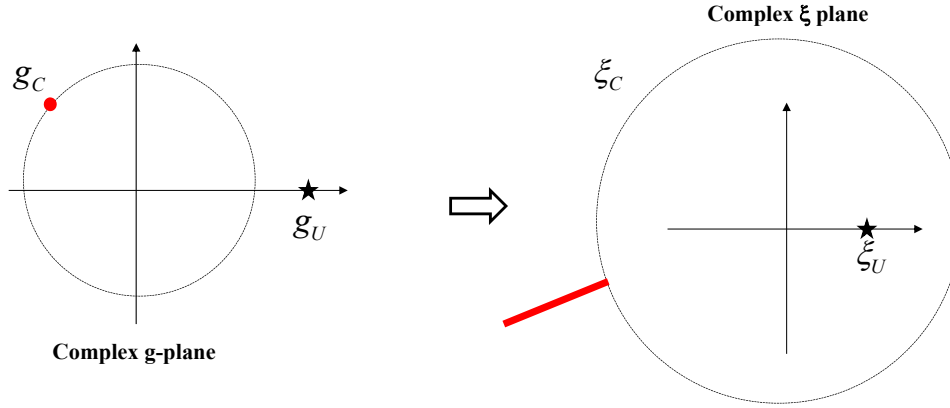


Fig. 15: *Illustration of the shifted action trick. It amounts to changing the origin of expansion and introducing a different expansion parameter.*

is vast, and many methods carry names of famous mathematicians. The bottom line is that divergent series outside their radius of convergence are almost as valuable for extracting the required information as convergent series, provided enough terms in the series are known and singularities are reasonably well understood.

To conclude this subsection, I would like to mention that series with convergence radius zero (e.g., when branch cuts originate from the center of expansion) are also subject to efficient resummation methods that guarantee that the final answer is unique in the limit. The analysis starts with establishing the asymptotic behavior of the Taylor series coefficients γ_n for large n by employing the method developed by Lipatov [8]. Since all singularities are encoded in $\gamma_{n \rightarrow \infty}$ (any finite number of terms is just a polynomial), the corresponding information should be used for designing the optimal resummation method and guaranteeing that it performs the unique analytic continuation for the physical parameters of the problem. A beautiful example of such an analysis can be found in Ref. [9].

4.3 Shifted action

Yet another way of manipulating series convergence is shifting the expansion point, as illustrated in Fig. 15. Of course, it can not be done by solving exactly an interacting problem for some value of the coupling constant, say g_1 , and then expanding in $g - g_1$ (but this protocol may be tried numerically if g_1 is inside the convergence radius). However, it is relatively “cheap” to expand on top of the mean-field solution or solutions based on a limited subset of diagrams, such as ladders, random phase approximation, and self-consistent skeleton graphs, as well as any set of dressed propagators. The new expansion still provides an exact solutions to the problem. The tool goes under the name of “shifted action.”

I will explain how shifted action works by considering the case of the Green function. Further generalizations are mentioned in Ref. [10]; in brief, the shifted action protocol can be applied at the level of any multi-point correlation function (with the help of Stratonovich-Hubbard transformations). Let the original interacting problem be described by the action (in terms of the

Grassmann field ψ)

$$S[\psi] = \langle \psi | G_0^{-1} | \psi \rangle + g S_{\text{int}}[\psi], \quad (40)$$

where G_0 is the bare fermion propagator, and g is the coupling constant. [For brevity, I will use vector-space notations to suppress space, time, spin, etc., indices and integrals/sums over them, and loosely call the corresponding kernels “functions.”] Instead of expanding e^{-S} in powers of g , one can introduce an auxiliary action

$$S_\xi^{(\mathcal{N})}[\psi] = \langle \psi | \tilde{G}_\mathcal{N}^{-1} + \xi A_1 + \dots + \xi^\mathcal{N} A_\mathcal{N} | \psi \rangle + \xi g S_{\text{int}}[\psi] \quad (41)$$

with auxiliary expansion parameter ξ . Despite the fact that the first term is harmonic, we will treat $\tilde{G}_\mathcal{N}$ as a new “shifted” bare propagator and expand $e^{-S_\xi^{(\mathcal{N})}}$ in powers of ξ using all ξ -dependent terms. For the two actions to represent the same physical system at $\xi = 1$ we demand that

$$\tilde{G}_\mathcal{N}^{-1} + \sum_{n=1}^{\mathcal{N}} A_n = G_0^{-1}. \quad (42)$$

Given that the final answer for the Green function can be expressed in terms of the proper self-energy Σ as

$$G^{-1} = G_0^{-1} - \Sigma = \tilde{G}_\mathcal{N}^{-1} - \left[\Sigma - \sum_{n=1}^{\mathcal{N}} A_n \right], \quad (43)$$

the $\{A_n\}$ functions act as counter-terms with respect to the n -th order proper self-energy diagrams generated by the interaction term $g S_{\text{int}}[\psi]$. There are no restrictions on the number of counter terms or their functional dependence; this freedom can be used to optimize the convergence of the series for $\xi = 1$. Even a simple self-energy shift such as $A_1 = \mu_1$, equivalent to a change in the chemical potential, can help to solve the problem by moving the $\xi = 1$ point inside the radius of convergence [11].

Of special practical interest is the case when the counter-term A_n is exactly the n -th order contribution to the self-energy coming from S_{int} . The resulting expansion—standard for effective-field theories—is then identical to the semi-skeleton series based on Dyson summation of infinite sets of irreducible diagrams associated with the first $\mathcal{N}$ orders of the perturbative expansion of the original action (40). For example, if this protocol is followed for $\mathcal{N} = 1$, then the expansion will be done on top of the self-consistent Hartree-Fock solution. If shifted action is applied to the screening channel as well, by selecting the $\mathcal{N} = 1$ skeleton set to define counter-terms one is setting the expansion on top of the self-consistent GW -approximation. Next, one can account for the leading vertex corrections, etc. In view of the exact cancellation of all contributions up to order $\mathcal{N}$, the expansion starts at order $n = \mathcal{N} + 1$ and only then the counter-terms A_n enter the diagrammatic expansion explicitly. To find $\tilde{G}_\mathcal{N}$ and all counter terms for a given G_0 one has to perform the so-called “bold” DiagMC simulation, or BDMC. This leads to the numeric protocol consisting of two independent parts:

- Part I is the BDMC simulation of the truncated order- $\mathcal{N}$ skeleton sum with the goal of solving for $\tilde{G}_\mathcal{N}$ and $\{A_n[\tilde{G}_\mathcal{N}]\}$ satisfying Eq. (42);
- Part II is the DiagMC simulation of higher-order terms using $\tilde{G}_\mathcal{N}$ as the bare propagator.

5 Fermionic sign blessing II

Apart from the massive cancellation of contributions from diagrams of the same order, the fermionic sign is also key for having efficient algorithms to account for all possible topologies. Indeed, consider the case of density-density interparticle interactions. If diagrams are formulated in the real-space, imaginary-time representation (to eliminate restrictions imposed by the energy-momentum conservation laws) then the sum over all possible graph topologies, both connected and disconnected, has the form of a determinant for each spin component

$$\det |G_\sigma(\mathbf{r}_i, \tau_i; \mathbf{r}_j, \tau_j)|. \quad (44)$$

Thus, $n!$ terms can be summed in $O(n^3)$ number of operations. As I will discuss below, it takes much longer to compute contributions from *connected* diagrams, but it is still possible to do it much faster than in $n!$ operations, see [12]. Ultimately, this observation allows one to say that DiagMC for convergent series or series subject to resummation, generically solves the computational complexity problem for interacting fermions.

5.1 Sum or sample?

So far we discussed the original MC approach to sampling the configuration space of connected Feynman diagrams. As illustrated in Fig. 4, one option is to sample various topologies within a given order. However, knowing that the sum of all topologies features massive cancellations of contributions, this is not necessarily the best strategy. To make the point, consider the problem of determining an answer for a large sum of sign-alternating terms

$$A = Z^{-1} \sum_{i=1}^M c_i, \quad Z = \sum_{i=1}^M |c_i|. \quad (45)$$

In M operations one can know the answer with machine precision by performing the sum. If the sum is sampled from the probability distribution $p_i = |c_i|/Z$, the error bar on the result after M operations will be (assuming that the algorithm is perfect in eliminating the autocorrelation time problem)

$$\sigma_A = \sqrt{(1-A^2)/M}. \quad (46)$$

Since we assume in this discussion that $|A| \ll 1$, the answer to the dilemma of what is the best method of dealing with (45) is clear: for $|A| \ll M^{-1/2}$ one is better off by doing the sum; otherwise, the answer will be known with good accuracy faster by sampling.

Enumerating terms and performing the sum over all topologies for high-order graphs is certainly possible for $n \leq 10$. At the same time, for convergent series, it is expected that the average sign is factorially small. It is thus plausible that in many cases problem (1) has to be reformulated as

$$Q(\mathbf{y}) = \sum_{n=0}^{\infty} \int \cdots \int d\mathbf{x}_1 \cdots d\mathbf{x}_n \bar{D}(n; \mathbf{x}_1, \dots, \mathbf{x}_n; \mathbf{y}), \quad (47)$$

$$\bar{D}(n; \mathbf{x}_1, \dots, \mathbf{x}_n; \mathbf{y}) = \sum_{\mathbb{T}} D(n, \mathbb{T}; \mathbf{x}_1, \dots, \mathbf{x}_n; \mathbf{y}). \quad (48)$$

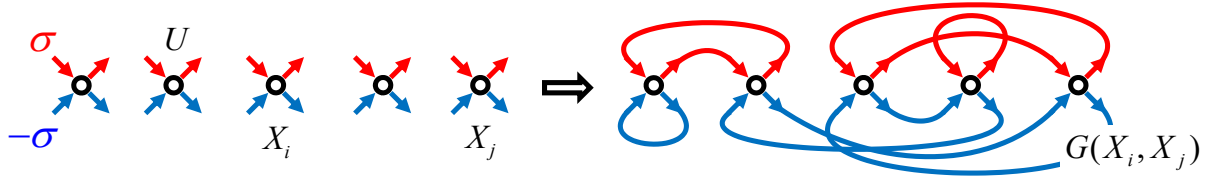


Fig. 16: By connecting all outgoing arrows to incoming ones with the same spin index, one obtains a Feynman diagram for the partition function. Free energy density diagrams must form a connected graph.

5.2 Determinant method for connected diagrams

An efficient method for computing $\bar{D}$ for expansions in the coupling constant was developed by Rossi in Ref. [12]. It rests on the simple observation that the sum of all connected topologies can be obtained from the sum of all topologies by subtracting disconnected ones. To be specific, consider the fermionic Hubbard model and Feynman diagrams for the free energy density. Given space-time positions of interaction vertexes $X_1, \dots, X_n$, where $X_i = (\mathbf{r}_i, \tau_i)$, all topologies are generated by establishing pairwise associations between the incoming and outgoing arrows with the same spin index, as in Fig. 16. Apart from the global factor $(-U)^n$, the diagram contribution is given by the product of all Green functions and the sign rule based on the number of fermionic loops. According to this rule, each time one swaps the destination points for two propagators the number of loops changes by ± 1 and this leads to an additional factor of -1 . Thus, the sum over all possible topologies forms a determinant (44).

Let us introduce a short-hand notation for the collection of all vertex coordinates, $V = \{X_i\}$, any proper subset of coordinates, $S \subsetneq V$, the sum over all topologies (determinant) for a given set of coordinates, $\det(V)$, and the sum over all connected topologies, $C(V)$. Then, by subtracting from $\det(V)$ all disconnected cases, we obtain $C(V)$

$$C(V) = \det(V) - \sum_{S \subsetneq V} C(S) \det(V \setminus S). \quad (49)$$

This is a set of recursive equations for connected contributions after similar equations are written for subsets of V . Its coefficients are based on determinants and the cost of computing all of them scales as $n^3 2^n$, where 2^n comes from the combinatorial number of possible proper subsets, $\sum_{m=1}^{n-1} n!/m!(n-m)!$. The number of arithmetic operations required to solve these recursive equations is $\propto 3^n$ —in the large n limit this is the main computational cost.

In this scheme, the effort is exponential in the diagram order and this is certainly an enormous improvement compared to the $(n!)^2$ scaling of the total number of connected graphs. After summation over $\{X_i\}$ one should not forget to divide the n -th order contribution by $n!$ to account for the indistinguishability of the vertices. One can use this scheme (or its generalizations) to compute connected diagrams for any correlation function [12], proper self-energy [13, 14], and even semi-skeleton series; in the latter case, however, the computational cost will increase to roughly 6^n .

One final remark. When the cost of computing the diagram weight is minimal, as for Fröhlich polarons, local updates changing a couple of variables with large acceptance ratio are preferred because of their efficiency and simplicity. However, when getting the diagram weight is computationally very costly, there is nothing wrong in spending at least as much CPU time on designing efficient global updates changing all diagram variables. The gain in reduced autocorrelation time may more than overcompensate the loss in the acceptance ratio. This is where machine learning techniques hold a great promise for further improving the efficiency of DiagMC simulations. By learning typical model-dependent statistical properties of connected Feynman diagrams with n multi-dimensional coordinates, such as “gyration radius,” “dipole,” “quadrupole,” and “multi-pole” correlations, as well as asymptotic laws for moving one or more vertices well outside of the gyration radius, global updates can achieve large enough acceptance ratios. I am not aware of systematic work done in this direction for connected Feynman diagrams.

5.3 Computational complexity problem for interacting fermions and its solution

To begin with, interacting fermions do not suffer from the sign problem, as any experimentalist measuring their properties is certainly aware of. It is a problem for some, but not all, theoretical methods used to simulate their properties. In general, sign-alternation of contributions simulated by MC methods is neither sufficient nor necessary to state that the problem is intractable. For precise quantitative discussion, one needs to define the “computational complexity problem” (CCP). The most relevant practical question to answer is “How easily can one increase the accuracy of the computed thermodynamic-limit answer?” This leads to the following definition of the CCP that can be applied to any numerical scheme. Let Q be the intensive quantity of interest in the thermodynamic limit.

A numerical scheme has a CCP if the computational time t required to obtain Q with error ϵ diverges faster than any polynomial function of $\epsilon^{-1} \rightarrow \infty$. The CCP is considered to be solved if

$$t(\epsilon) = O(\epsilon^{-a}). \quad (50)$$

I discuss only unbiased methods, for which the difference between the computed and exact values can be made arbitrary small. For methods containing an unknown systematic bias, the accuracy cannot be increased indefinitely (but it may be very small).

I will skip the discussion of how methods that suffer from the sign-problem also generically have CCP in dimensions $d \geq 2$, see Ref. [15] for details. On the contrary, for convergent (or subject to resummation) diagrammatic series, the CCP is solved by DiagMC. Indeed, for convergent series the accuracy of truncated sums $Q_N = \sum_{n=0}^N d_n$ is improved exponentially fast with the largest diagram order accounted for, $|Q - Q_N| \propto Q\alpha^N$, with $0 < \alpha < 1$. Thus, given some small value of ϵ , the required accuracy is reached by simulating diagrams up to order $N_\epsilon \sim \ln(\epsilon)/\ln(\alpha)$. For the determinant scheme described above the simulation time required to compute all diagrams up to order N_ϵ is an exponential function of N_ϵ , leading to an estimate

$$t(\epsilon) = t(N_\epsilon) \propto b^{N_\epsilon} = \epsilon^{\ln(b)/\ln(\alpha)} \longrightarrow \text{CCP solved} . \quad (51)$$

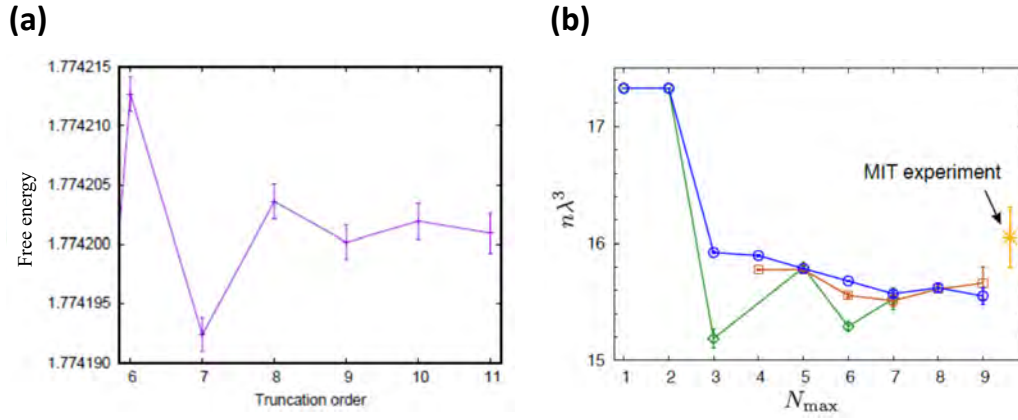


Fig. 17: (a) Free energy density for the fermionic Hubbard model as a function of truncation order, at $T/t = 0.125$, $U/t = 2$, and $n = 0.87500(2)$ (reproduced from Ref. [12]). (b) Density of the unitary fermi gas (λ is the thermal wavelength) vs. maximal diagram order at $T/\mu = 0.5$ (or $T/T_F = 0.2$). The bold diagrammatic series is resummed by three variants of the conformal-Borel transformation (see Ref. [9]).

Apart from the cost of the solving recursive equations for connected contributions, the value of b may also include the cost of performing an integral over the space-time variables.

5.4 Illustrative results: Hubbard model and unitary Fermi-gas

It is rare to see sign-free path integral simulations for bosons to be performed with accuracy better than 5 significant digits even for finite size systems. The remarkable plot in Fig. 17(a) proves that for convergent series one can reach an accuracy of 6 significant digits for a generic interacting fermionic system. In this example, the determinant method for connected diagrams was used to simulate the fermionic Hubbard model away from half-filling at density $n = 0.87500(2)$ and relatively low temperature $T/t = 0.125$, where t is the n.n. hopping amplitude. Better convergence was achieved through the Hartree diagram shift of the chemical potential (the convergence radius in the on-site coupling U was estimated to be about $5.1 t$). The selected parameter set corresponds to the Fermi liquid regime; for larger values of U and on approach to half-filling the situation is much worse and further work is required to improve the performance of DiagMC. Nevertheless, a number of interesting results concerning the nature of magnetic correlations and the pseudogap regime were already obtained.

Fig. 17(b) shows another remarkable outcome of the BDMC simulations done for the unitary fermi gas. This system features a number of universal properties and is relevant for understanding properties of ultra-cold atomic gases and dilute neutron matter. Microscopically, imagine that fermions interact via a short-range attractive potential of radius R and strength U_0 that is fine tuned to the threshold of having a shallow bound state. For two-body collisions at zero energy this situation corresponds to a large s -wave scattering length, $a_s \gg R$. Next, consider a many-body system at finite density, $n = k_F^3/3\pi^2$, where k_F is the Fermi momentum, in the so-call “zero-range” limit, $k_F R \rightarrow 0$ when the interparticle distance vastly exceeds the potential radius. In this limit system properties become universal in the sense that all microscopic potentials with the same $k_F a_s$ parameter should be considered equivalent to each other. If $k_F a_s$

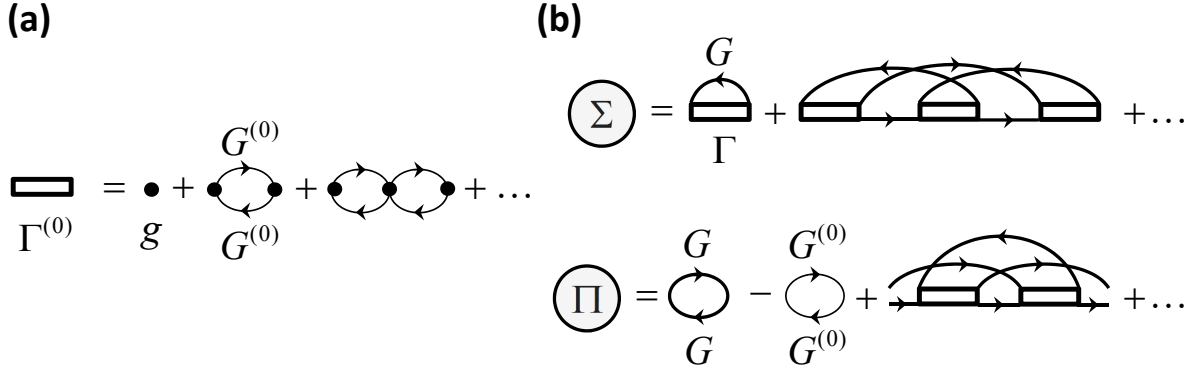


Fig. 18: (a) Bare pair propagator $\Gamma^{(0)}$ based on the summation of ladder diagrams in terms of the zero-range potential $g\delta(\mathbf{r})$ (dots) and non-interacting Green functions (thin lines). (b) An order n skeleton diagram consists of n fully dressed pair propagators and Green functions (bold lines) connecting them.

is finite, one talks about resonant fermions; this is the canonical model for discussing the BCS-BEC crossover within the superfluid fermionic state.

The unitary fermi gas corresponds to $k_F a_s = \infty$. At low temperature $T < T_F$ it is a strongly interacting system (every spin-up fermion is “contemplating” to form a bound state with every other spin-down fermion) without small parameters to justify a perturbative or mean-field treatment because the only meaningful length (energy) scale in the problem is k_F (T_F). Its solution by the DiagMC method involves nearly all the tricks mentioned in this contribution:

- (i) To eliminate ultra-violet divergences and to take the zero-range limit analytically, one has to perform summation of the ladder diagrams prior to the DiagMC simulation and formulate the expansion in the number of pair propagators, see Fig. 18(a).
- (ii) To reduce the number of sampled topologies, the simulation is performed for proper self energies $\Sigma[G, \Gamma]$ and $\Pi[G, \Gamma]$ within the self-consistent skeleton formulation, see Fig. 18(b). The self-consistent loop is closed by Dyson equations:

$$1/G = 1/G^{(0)} - \Sigma[G, \Gamma], \quad 1/\Gamma = 1/\Gamma^{(0)} - \Pi[G, \Gamma].$$

- (iii) Asymptotic properties of Γ and G in the limit of short time and distance should be taken care of using exact analytic relations, see Ref. [16] for details.
- (iv) Since the resulting series have zero convergence radius, one has to study the nature of non-analytic behavior at the origin of the expansion by employing Lipatov’s technique and construct the appropriate conformal Borel resummation method, see Ref. [9].

The amount of analytic and numeric work may seem daunting, but it can hardly be avoided: the goal is to produce results with guaranteed accuracy bounds. Currently, theoretical error bars are smaller than the most precise experimental data [17] for the equation of state at low temperature right above the transition point to the superfluid state, see Fig. 17(b). At high temperature, $T > T_F$, the BDMC results are far more accurate than experimental data and provide the most stringent test for high-order virial expansion coefficients.

6 Conclusions

This contribution reviewed key principles of DiagMC, as well as some of the recent developments for interacting fermionic (or fermionized) systems that radically improve the efficiency of simulations for high orders of expansion. The number of successful applications is already very large and covers both lattice and continuous-space systems, short- and long-range interaction potentials, effective field theories, interacting topological insulators, frustrated quantum magnets, Bose and Fermi polarons, continuous-time impurity solvers, lattice- and continuous-space path integrals, point-contact dynamics, etc.

There are no known fundamental restrictions on the applicability of the method. However, its efficiency strongly depends on the convergence properties of the series and a deep theoretical/mathematical understanding of the singularities that control this convergence. Gaining such analytic understanding is, arguably, the most important and urgent direction for future work. Numerically, the field will expand in terms of applications to cover models with gauge fields, spin-orbit coupling, and more complex forms of interaction potentials. Codes have to be developed and tested for a variety of effective field theories, shifted action protocols, multi-point correlation functions, and, ultimately, for material science applications.

References

- [1] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953)
- [2] N.V. Prokof'ev, B.V. Svistunov, and I.S. Tupitsyn, Sov. Phys. JETP **87**, 310 (1998); *ibid.* Phys. Lett. A **238**, 253 (1998)
- [3] N.V. Prokof'ev and B.V. Svistunov, Phys. Rev. Lett. **81**, 2514 (1998)
- [4] A.S. Mishchenko, N.V. Prokof'ev, A. Sakamoto, and B.V. Svistunov, Phys. Rev. B **62**, 6317 (2000)
- [5] O. Goulko, A. Gaenko, E. Gull, N. Prokof'ev, and B. Svistunov, Comp. Phys. Comm. 0010-4655 (2018)
- [6] O. Goulko, N. Prokof'ev, and B. Svistunov, Phys. Rev. E **97**, 053305 (2018)
- [7] H. Mera, T.G. Pedersen, and B.K. Nikolić, Phys. Rev. B **94**, 165429 (2016); *ibid.* arXiv:1802.06034
- [8] L.N. Lipatov, Zh. Eksp. Teor. Fiz. **72**, 411 (1977) [Sov. Phys. JETP **45**, 216 (1977)]
- [9] R. Rossi, T. Ohgoe, K. Van Houcke, and F. Werner, Phys. Rev. Lett. **121**, 130405 (2018)
- [10] R. Rossi, F. Werner, N. Prokof'ev, and B. Svistunov, Phys. Rev. B **93**, 161102(R) (2016)
- [11] W. Wu, M. Ferrero, A. Georges, and E. Kozik, Phys. Rev. B **96**, 041105 (2017)
- [12] R. Rossi, Phys. Rev. Lett. **119**, 045701 (2017)
- [13] F. Šimkovic IV., E. Kozik, arXiv:1712.10001
- [14] R. Rossi, arXiv:1802.04743
- [15] R. Rossi, N. Prokof'ev, B. Svistunov, K. Van Houcke, and F. Werner, EPL **118**, 10004 (2017)
- [16] K. van Houcke, F. Werner, T. Ohgoe, N. Prokof'ev, and B. Svistunov, Phys. Rev. B **99**, 035140 (2019)
- [17] M.J.H. Ku, A. Sommer, L.W. Cheuk, and M.W. Zwierlein, Science **335**, 563 (2012)

16 Stochastic Series Expansion Methods

Anders W. Sandvik

Department of Physics, Boston University

Boston, Massachusetts, USA

Institute of Physics, Chinese Academy of Sciences

Beijing, China

Contents

1	Overview of stochastic series expansion methods	2
1.1	Path integrals and series expansions	2
1.2	Continuous time in the power-series representation	5
1.3	Stochastic series expansion for ground states	8
1.4	Quantum annealing with generalized stochastic series expansion	10
2	Sampling algorithms and expectation values	11
2.1	Configuration representations and diagonal updates	11
2.2	Loop and cluster updates	16
3	Estimators for expectation values	20
3.1	Diagonal correlations and susceptibilities	20
3.2	Off-diagonal observables	22
4	Recent applications	23
4.1	Quantum phases and criticality in spin systems	24
4.2	Dynamic response of quantum magnets	26
4.3	Disordered systems	26

The Stochastic Series Expansion (SSE) technique is a quantum Monte Carlo (QMC) method that is especially efficient for many quantum spin systems and boson models. It was the first generic method free from the discretization errors affecting previous path integral based approaches. These lecture notes will serve only as a brief overview of the SSE method, and readers who would like to implement it for some specific model are recommended to consult some of the cited references for more details.

In the introductory section, the representation of quantum statistical mechanics by the power series expansion of $e^{-\beta H}$ will be compared with the likely more familiar path integrals in discrete and continuous imaginary time. Extensions of the SSE approach to ground state projection and quantum annealing in imaginary time will also be briefly discussed. The later sections will introduce efficient sampling schemes (loop and cluster updates) that have been developed for many classes of models. A summary of generic forms of estimators for important observables will be given. Some recent applications will also be reviewed.

1 Overview of stochastic series expansion methods

1.1 Path integrals and series expansions

The most obvious way to construct a Monte Carlo (MC) simulation scheme for a lattice Hamiltonian H is to start from the path integral formulation of quantum statistical mechanics, where $e^{-\beta H}$ is written as a product of imaginary-time evolution operators $e^{-\Delta_\tau H}$ with a small “time slice,” $\Delta_\tau = \beta/L$, for some large number L of slices [1, 2]. The partition function $Z = \text{Tr}\{e^{-\beta H}\}$ can then be written as

$$Z_{\text{PI}} = \sum_{\{\alpha\}} \langle \alpha_0 | e^{-\Delta_\tau H} | \alpha_{L-1} \rangle \cdots \langle \alpha_2 | e^{-\Delta_\tau H} | \alpha_1 \rangle \langle \alpha_1 | e^{-\Delta_\tau H} | \alpha_0 \rangle, \quad (1)$$

where the sum is over L complete sets of states in a suitably chosen basis. Because the time step Δ_τ of the slices is small, the matrix elements can be computed to some approximation with a total error of order Δ_τ^p , where p depends on the approximation used, e.g., $p = 2$ with the lowest-order split-operator (Suzuki-Trotter) decomposition of the exponential operators, $e^{-\Delta_\tau(H_A+H_B)} \approx e^{-\Delta_\tau H_A} e^{-\Delta_\tau H_B}$ (for generic non-commuting terms H_A and H_B). Several variants of such path-integral based QMC methods, often called *world line* (WL) methods, were developed in the 1970s and 1980s [3, 4].

There was also an earlier proposal by Handscomb, dating back to the early 1960s [5], to instead start from the power series expansion of $e^{-\beta H}$. Handscomb made the observation that the Heisenberg exchange interaction between two $S = 1/2$ spins, $H_{ij} = \mathbf{S}_i \cdot \mathbf{S}_j$, is a permutation operator, and that traces of strings of these operators for different spin pairs can be easily evaluated analytically. In the case of a ferromagnetic coupling (i.e., with a negative sign in front of $\mathbf{S}_i \cdot \mathbf{S}_j$), the minus sign in the exponential is canceled. The powers of the Hamiltonian can be further decomposed into strings of two-body (bond) operators, $H_{b_n} \cdots H_{b_2} H_{b_1}$, with all combinations of the operators $H_b = H_{i(b),j(b)}$, $b \in \{1, \dots, N_b\}$ on a finite lattice in any number of

space dimensions D (where $N_b = DL$ for nearest-neighbor interactions on a D -dimensional simple cubic lattice). Denoting a string of n operator indices $b_1, b_2, \dots, b_n$ by S_n , the partition function is then

$$Z_H = \sum_{n=0}^{\infty} \frac{\beta^n}{n!} \sum_{S_n} \text{Tr}\{H_{b_n} \cdots H_{b_2} H_{b_1}\}, \quad (2)$$

where the traces of the strings of permutation operators are positive definite. With a simple scheme to evaluate the traces in terms of permutation cycles, the operator strings can be sampled using a Metropolis MC method and various thermodynamic expectation values can be evaluated [5, 6]. Among these expectation values, the total internal energy is particularly important as it takes the simple form $E = \langle H \rangle = -\langle n \rangle / \beta$, which shows that the mean number of operators n in the expansion during the sampling process (which includes updates increasing and decreasing the string length n) scales as βN , N being the total number of spins. Thus, the expansion is convergent (as is also guaranteed since H on a finite lattice has a bounded energy spectrum) and there is in practice a “fuzzy bound” on the expansion order n , proportional to βN , that will never be exceeded during the sampling process. Both the computer storage requirement and the processor time for completing a full MC updating sweep then also scale as βN .

It is illustrative to compare Handscomb’s expansion with the simple power series for the exponential of a positive number x , $e^x = \sum_m x^m / m! = \sum_m W(m)$, which is always convergent and where the mean of the distribution $P(m) = W(m)e^{-x}$ (i.e., the Poisson distribution) is $\langle m \rangle = x$. In light of this fact, the distribution of expansion orders in the case of $\text{Tr}\{e^{-\beta H}\}$ can be easily understood in the limit $\beta \rightarrow \infty$, where $e^{-\beta E_0}$ (E_0 being the ground state energy) is the only surviving contribution to the trace and $\beta|E_0|$ therefore corresponds to x above; $\langle n \rangle = \beta|E_0|$. At any temperature, the fluctuations of n are related to the heat capacity; $C = \langle n^2 \rangle - \langle n \rangle^2 = \langle n \rangle$ [5]. Therefore, the variance of the distribution at $T = 0$ is also exactly the same as for the Poisson distribution; $\langle n^2 \rangle - \langle n \rangle^2 = \langle n \rangle$.

Handscomb’s method is certainly elegant, but in its original formulation it was very limited in applicability, as there is only a small number of models for which the traces can be computed analytically [7]. The expansion is also normally not positive definite. The problem of mixed positive and negative terms (i.e., an incarnation of the QMC sign problem) appears already for the antiferromagnetic Heisenberg interaction, and it took some time until it was realized that this sign problem was easily solvable for bipartite lattices by simply adding a suitable constant to the interaction [8]. The traces can then still be computed and the strings can be sampled in a way similar to the original formulation. However, in practice the modified method did not perform as well as the path-integral based methods that had been developed by then for the Heisenberg antiferromagnet and many other models. Though some further applications were reported [9], Handscomb’s method was essentially abandoned, as it became clear that it was not possible in general to compute traces of operator strings efficiently, and, in the cases where the traces can be evaluated, the existing sampling schemes were also often inefficient.

The dismissal of the idea of starting from the series expansion was premature. It is not clear why it was not realized earlier that it is not necessary to compute the traces analytically—they also can be sampled in a chosen basis along with the operator strings [10, 11]. The sam-

pling weight in the extended space comprising states and operator strings has matrix elements $\langle \alpha | H_{b_n} \cdots H_{b_2} H_{b_1} | \alpha \rangle$ in place of the full traces in Eq. (2);

$$Z_{\text{SSE}} = \sum_{n=0}^{\infty} \frac{\beta^n}{n!} \sum_{S_n} \sum_{\alpha} \langle \alpha | H_{b_n} \cdots H_{b_2} H_{b_1} | \alpha \rangle. \quad (3)$$

Here the Hamiltonian has been defined as $H = -\sum_b H_b$, so that no negative signs appear explicitly. Of course the string of operators can still produce negative signs, and the class of models for which this can be avoided is, as always, limited, but includes many important systems worthy of study (some recent examples will be discussed in Sec. 4). It can be noted here that sign problems originating from the diagonal part of the interaction can always be avoided by adding suitable constants to some of the H_b operators. Signs appearing with off-diagonal operations (either explicitly from negative prefactors of some H_b or due to fermion anti-commutation) are in general difficult to avoid [12], except in special cases, e.g., the aforementioned bipartite antiferromagnets, where the number of minus signs in the product is always even.

Methods based on sampling the traces in Eq. (3) were first developed for general- S Heisenberg models [10] and 1D Hubbard-type models [11] almost 30 years after the advent of Handscomb's original method, and the extended series scheme eventually became known as the SSE. Over the years, many further improvements of these algorithms have been made—some inspired by developments within other methods and some proposed first within the SSE scheme and later adopted into other techniques. The SSE method was the first broadly applicable exact QMC scheme (i.e., not affected by any systematical errors such as those from time discretization) and it was also the forerunner and inspiration to later algorithms based on other series expansions, in particular, the perturbation series leading to the continuous-time *worm algorithm* [13].

The SSE method can in principle be used for any model written in a discrete basis, though in practice sign problems restrict applications to the same classes of models as the WL methods. Models that have been successfully studied include many spin Hamiltonians, boson models, and 1D fermion systems (see Sec. 4). Both the WL and SSE approaches normally have insurmountable sign problems for fermions in higher dimensions (unlike, in some cases, auxiliary-field fermion determinant methods [14, 15]). Models with controllable sign problems accessible to SSE include certain frustrated impurities in antiferromagnets, where the signs arise only locally at the impurity and the bulk operators do not involve any signs [16].

The sampled basis in SSE simulations is typically the z components S_i^z for spin systems or the site occupation numbers $n_i = a_i^\dagger a_i$ for particle models. More complicated states can also be used, e.g., the basis of singlets and triplets on spin dimers has been used to solve certain sign problems [17, 18]. The primary advantage of SSE over discrete-time WL methods for sign-free models is the absence of time-discretization error—the discrete dimension corresponding to the location in the SSE operator string constitutes a faithful representation of continuous imaginary time [11, 19], as we will discuss below. Compared to more recently developed continuous-time WL methods [20, 13], the discreteness of the SSE representation often allows for more efficient sampling schemes. For many models (especially spin systems), the SSE outperforms all other known QMC methods. Typically the SSE method is also easier to implement.

1.2 Continuous time in the power-series representation

Before discussing how to sample the SSE configuration space (in Sec. 2) and compute expectation values of interest (in Sec. 3), it is useful to consider some formal relationships between path integrals and the series representation of statistical mechanics. At the same time we will introduce some concepts that will be useful later in algorithm definitions and implementations. We can insert complete sets of states between all the operators in Eq. (3) to write the SSE partition function in the form

$$Z_{\text{SSE}} = \sum_{n=0}^{\infty} \frac{\beta^n}{n!} \sum_{S_n} \sum_{\{\alpha\}} \langle \alpha_0 | H_{b_n} | \alpha_{n-1} \rangle \cdots \langle \alpha_2 | H_{b_2} | \alpha_1 \rangle \langle \alpha_1 | H_{b_1} | \alpha_0 \rangle, \quad (4)$$

where one can see a similarity with the path integral in Eq. (1). The inserted states (except for $|\alpha_0\rangle$) are redundant in the SSE method, however, because the operators H_b should be defined such that, for every possible combination of operator H_b and basis state $|\alpha\rangle$, the state $H_b|\alpha\rangle$ is proportional to a single basis state; $H_b|\alpha\rangle = h_{b\alpha}|\alpha'\rangle$, $|\alpha'\rangle \in \{|\alpha\rangle\}$ (and $h_{b\alpha}$ is the trivial matrix element $\langle \alpha' | H_b | \alpha \rangle$). Then the operator string itself uniquely defines how the state $|\alpha_0\rangle$ in Eq. (4) is propagated through a series of basis states (similar to a path in the WL formulation) and eventually arrives back to the same state $|\alpha_0\rangle$ for an operator string contributing to the partition function (with the periodicity reflecting the original trace operation). Clearly the vast majority of the strings violate the periodicity condition and for those we have $\langle \alpha_0 | H_{b_n} \cdots H_{b_1} | \alpha_0 \rangle = 0$. The strings should therefore be sampled so that the periodic “time” boundary condition is always maintained in any attempted update.

To ensure the “no-branching” condition $H_b|\alpha\rangle \propto |\alpha'\rangle$, the diagonal and off-diagonal parts of all the terms in H have to be separated. The index b then does not just refer to sites or bonds (or larger units of sites for interactions involving more than two sites) but enumerates separately the diagonal and off-diagonal terms. This is sometimes accomplished (formally and in actual program implementations) by introducing an additional index, so that $H_{1,b}$ refers to, say, the diagonal part and $H_{2,b}$ is the off-diagonal part of the interaction on bond b . In some cases the off-diagonal operators have to be formally split up further, e.g., in the case of $S > 1/2$ Heisenberg spin systems the off-diagonal part of the interaction, $S_i^x S_j^x + S_i^y S_j^y = \frac{1}{2}(S_i^+ S_j^- + S_i^- S_j^+)$ has to be regarded as two different operators, $\frac{1}{2}S_i^+ S_j^-$ and $\frac{1}{2}S_i^- S_j^+$. In contrast, for $S = 1/2$ the sum of the two terms can be considered as a single operator, since only one of them can give a non-zero result when acting on a basis state.

With the no-branching condition ensured, for a given operator string S_n , a propagated state $|\alpha(p)\rangle$ is defined in the SSE method as the state obtained when the first p operators have acted on the ket state $|\alpha_0\rangle = |\alpha(0)\rangle$ in Eq. (4),

$$|\alpha(p)\rangle = r \prod_{i=1}^p H_{b_i} |\alpha_0\rangle, \quad (5)$$

where r formally is a normalization constant (which does not have to be computed in practice). In the sum over complete sets of states in Eq. (4), there is now a single contributing component

for each p , namely, $|\alpha_p\rangle = |\alpha(p)\rangle$, and the operator string uniquely defines the path of states evolving from $|\alpha_0\rangle$.

The propagated states also have a simple relationship to imaginary-time evolution in path integrals, where starting from some state $|\phi(0)\rangle$ at imaginary time $\tau = 0$ we have the evolved state $|\phi(\tau)\rangle = re^{-\tau H}|\phi(0)\rangle$, where again we formally have a normalization constant r because of the non-unitary evolution. For an SSE string of length n , we roughly have the correspondence $(p/n)\beta \approx \tau$ (which becomes increasingly accurate with increasing n).

The precise relationship between the discrete index p and the continuous time variable τ can be seen in the SSE expression for a time dependent correlation function of some operators A and B . With the a time dependent operator given by $B(\tau) = e^{\tau H} B e^{-\tau H}$, by series expanding the thermal expectation value $\langle B(\tau) A(0) \rangle$ we obtain

$$\langle B(\tau) A(0) \rangle = \frac{1}{Z} \sum_{\alpha} \sum_{m,p} \frac{(\tau - \beta)^m (-\tau)^p}{m! p!} \langle \alpha | H^m B H^p A | \alpha \rangle, \quad (6)$$

which can also be expanded into a summation over operator strings as we did in the partition function above. We defer until later the question of what types of operators can be considered in correlation functions in practice (but note that diagonal operators can always be treated easily). Here we just consider the general formal relationship between the index p in Eq. (6) and the time difference time τ . We can see that the expansion of $e^{-\tau H}$ in powers H^p is independent of the expansion of $e^{-(\beta-\tau)H}$ in powers H^m , and to estimate the distribution $P(p)$ roughly we can just replace H^p by $\langle H \rangle^p$. Then we again have a Poisson distribution with mean $\langle p \rangle = \tau |\langle H \rangle|$, which also equals $(\tau/\beta) \langle n \rangle$, where n is the total expansion power of a given term; $n = m + p$. Stated differently, propagation by p operators in SSE corresponds roughly to a displacement $\tau = (p/n)\beta$ in imaginary time, as already mentioned at the end of the preceding paragraph. Eq. (6) shows more precisely that the relationship between any given τ involves summation over a sharply peaked distribution of states propagated by powers H^p of the Hamiltonian.

The similarity between the series approach and the path integral becomes even more obvious if the exponentials in Eq. (1) are expanded to linear order, $e^{-\Delta_\tau H} \approx 1 - \Delta_\tau H$;

$$Z_{\text{PI}} \approx \sum_{\{\alpha\}} \langle \alpha_0 | 1 - \Delta_\tau H | \alpha_{L-1} \rangle \cdots \langle \alpha_2 | 1 - \Delta_\tau H | \alpha_1 \rangle \langle \alpha_1 | 1 - \Delta_\tau H | \alpha_0 \rangle, \quad (7)$$

where the total error arising from all slices is of order Δ_τ , worse than the Δ_τ^2 error when the Trotter decomposition is used. We will in the end take the limit $\Delta\tau \rightarrow 0$ and the treatment becomes exact. Next we again write H as a sum over its individual terms H_b with a minus sign taken out in front, $H = -\sum_b H_b$. Furthermore, we can introduce a new name for the unit operators 1 in the matrix elements, defining $H_0 = 1$. Then, similar to the SSE formalism, we can write the path integral as a sum over index sequences $S_L = b_L, \dots, b_2, b_1$ but now with the new index $b = 0$ also allowed and with the number L of indices fixed;

$$Z_{\text{PI}} = \sum_{\{\alpha\}} \sum_{S_L} \Delta_\tau^n \langle \alpha_0 | H_{b_L} | \alpha_{L-1} \rangle \cdots \langle \alpha_2 | H_{b_2} | \alpha_1 \rangle \langle \alpha_1 | H_{b_1} | \alpha_0 \rangle, \quad (8)$$

where n is the number of elements in S_L that are not equal to 0 (i.e., the number of times actual terms of H appear in a given operator product).

In the case of the SSE partition function (4), there is an explicit sum over expansion orders n that does not appear in the path integral expression (8), but since we know that the series expansion is convergent we can introduce a cut-off, $n_{\max} = M$, and for expansion orders lower than M we can use “fill-in” unit operators, defining $H_0 = 1$ as above. If we further consider all possible ways of distributing n Hamiltonian terms within a product of M operators out of which $M - n$ are unit operators, the SSE partition function becomes

$$Z_{\text{SSE}} = \sum_{\{\alpha\}} \sum_{S_M} \frac{\beta^n (M - n)!}{M!} \langle \alpha_0 | H_{b_M} | \alpha_{M-1} \rangle \cdots \langle \alpha_2 | H_{b_2} | \alpha_1 \rangle \langle \alpha_1 | H_{b_1} | \alpha_0 \rangle, \quad (9)$$

where we have divided the weight by in Eq. (4) by the combinatorial factor $M!/[n!(M - n)!]$ to compensate for overcounting of identical contributions with different locations of the unit operators. Note again that n is the number of non-0 elements in the operator-index string, and a summation over n in Eq. (9) is implicit in the summation over all fixed-length sequences.

For a given common string length $M = L$, we now see that the path integral and SSE partition functions in Eqs. (8) and (9) involve exactly the same configuration space, but the weighting of the configurations is slightly different. However, the weights become identical in the limit where $\Delta\tau \rightarrow 0$ is taken in the path integral, since $\Delta^n = \beta^n / L^n$ and for $M \rightarrow \infty$ we have $\beta^n (M - n)! / M! \rightarrow \beta^n / M^n$ in the SSE weight. Thus, we conclude that the two approaches are really identical if the limit $M = L \rightarrow \infty$ is taken. An important difference is that the SSE approach is in practice exact (i.e., the truncation error is exponentially small and not detectable in practice) already for some M of order βN , while the path integral, in the approximation used above, is affected by an error of order β / L . An exceedingly large number of slices L would have to be used for the error to become completely negligible. In a sense, the SSE method automatically finds the optimal number of slices for given N and β .

Of course, the path integral approach as described above should not be used in an actual WL algorithm, because simple approximants of $e^{-\Delta\tau H}$ with smaller errors are available, i.e., the Trotter decomposition. The reason for using the linear approximation here was simply to obtain the most direct relationship between the SSE and path-integral forms of Z . In the case of the SSE, while it is not necessary to introduce the fill-in operators $H_0 = 1$ [11], in practice it is actually convenient and efficient to use this device to achieve a fixed length of the index sequence. The cut-off M can be easily adjusted automatically during the equilibration part of a simulation, in such a way that the number n of Hamiltonian operators always stays below M by a safe margin, as will be further discussed when updating schemes are described in Sec. 2.

The more recently developed continuous-time WL methods can be formally obtained by taking the limit $\Delta\tau \rightarrow 0$ in Eq. (1). This is equivalent to the form (8) of the partition function, where the events correspond to the operators H_{b_i} . However, in this case it is better to keep the diagonal operators in the exponential form instead of expanding them to linear order, and then the paths of events dictated by the off-diagonal operators formally correspond to the perturbation expansion in the interaction representation [13, 19]. In a computer program, only the “events”—where and how the paths change—need to be stored [20, 13].

To see the equivalence with the perturbation series more clearly, we can divide the H into its diagonal part H_0 in the chosen basis and an off-diagonal part V . As an example, for a generic Heisenberg spin system we could choose $H_0 = \sum_{ij} J_{ij} S_i^z S_j^z$ and $V = \sum_{ij} J_{ij} (S_i^+ S_j^- + S_i^- S_j^+)$. Then, by considering V as a perturbation to H_0 (though eventually there will be no restriction on the strengths of the two terms) and carrying out the perturbation series to all orders we obtain the partition function as an integral in continuous imaginary time

$$Z_{\text{CT}} = \sum_{n=0}^{\infty} (-1)^n \int_0^{\beta} d\tau_1 \int_0^{\tau_1} d\tau_2 \cdots \int_0^{\tau_{n-1}} d\tau_n \text{Tr}\{e^{-\beta H_0} V(\tau_n) \cdots V(\tau_2) V(\tau_1)\}, \quad (10)$$

where the time-evolved operator in the interaction representation is $V(\tau) = e^{\tau H_0} V e^{-\tau H_0}$. Like in the SSE method, we can now sample the trace in the chosen basis and write the product $V(\tau_n) \cdots V(\tau_2) V(\tau_1)$ as a string of time evolved operators V_b (now only including off-diagonal terms). When inserting complete sets of states between all the operators and summing over diagonal terms for the trace, the exponentials just become numbers, and the weight of a given configuration (a state $|\alpha_0\rangle$ acted on with an operator string) for a given set of times $\tau_1, \dots, \tau_n$ can be easily computed. The integrals over time can also be sampled in an MC procedure in which all the degrees of freedom are updated together efficiently [20, 13]. The relationships between the SSE and continuous time formulation have been discussed in more detail in Ref. [19].

Like the SSE method, the perturbation expansion converges with a mean string length of order βN , regardless of the relative strengths of H_0 and V . When H_0 dominates the energy the prefactor will be smaller in the perturbation expansion, which then is more economical. When H_0 is not very dominant (which is common with quantum spin models, where the diagonal and off-diagonal energies are often similar in magnitude) the SSE approach has an advantage owing to the fact that it is formulated in a completely discrete representation, while in the continuous-time formulation floating-point numbers (the imaginary time points) have to be processed. Recently, the SSE scheme was also further developed for systems where H_0 dominates [21] (which often correspond to a nearly classical statistical system), essentially by integrating out all the diagonal terms from the SSE operator strings. This approach may be as efficient as the continuous-time WL approaches (or even more efficient in some cases) when there is a large imbalance in the strengths of the diagonal and off-diagonal terms. The original SSE approach should still be better (because of the simplicity of the algorithm) when the terms are similar in strength.

1.3 Stochastic series expansion for ground states

In order to study the ground state of a system, one can take the limit $T \rightarrow 0$ within one of the methods discussed above. In practice, this means $T \ll \Delta$, where Δ is the smallest finite-size excitation gap of the system. For systems with gaps closing as $1/L$ (where from now on L will refer to the linear size of the lattice), the ground state can be routinely reached on systems with thousands, in many cases even tens of thousands, of spins or bosons.

Ground states can also be studied using “projector methods,” where, instead of tracing over a complete basis, the imaginary-time evolution operator $e^{-\beta H}$ propagates some initial state $|\Psi(0)\rangle$

(often called “trial state,” though this term can be misleading, since the final result should be independent of the choice of initial state) that overlaps with the ground state; $|\Psi(\beta)\rangle = e^{-\beta H}|\Psi(0)\rangle$. By expanding in energy eigenstates, one can readily confirm that $|\Psi(\beta)\rangle$ approaches the ground state for sufficiently large β (up to a normalization factor).

In this case, the MC sampling is of terms contributing to the normalization $\langle\Psi(\beta)|\Psi(\beta)\rangle$. The numerator of an expectation value,

$$\langle A \rangle = \frac{\langle\Psi(\beta)|A|\Psi(\beta)\rangle}{\langle\Psi(\beta)|\Psi(\beta)\rangle}, \quad (11)$$

is similarly expressed to obtain estimators to be averaged in the MC process. Here one can proceed as in the path integral approach, by introducing a discrete slicing $\Delta\tau = \beta/L$ of $e^{-\beta H}$ or taking the limit $\Delta\tau = 0$ as in the continuous time formulation. This type of method is known generically as the *path integral ground state* (PIGS) approach; for a review see Ref. [22]. One can also proceed with the ground state expectation value (11) as in the SSE method by series expansion and sampling strings of operators. In either case, the main difference from the $T > 0$ methods is the boundary condition in the time direction—at $T > 0$ we have periodic boundaries, reflecting the trace operation, while in the ground state approach the boundary condition is dictated by the starting state $|\Psi(0)\rangle$. This state is normally chosen such that the time boundary condition is convenient for the updating process used in the MC sampling [23], and it can also sometimes be optimized in some way, so that it already is a good approximation to the ground state. Note that the projection procedure is variational, so that $\langle H \rangle$ always approaches the true ground state energy monotonically from above.

Instead of projecting out the ground state with $e^{-\beta H}$, a high power $(-H)^m$ of the Hamiltonian can also be used (where the negative sign is just included for convenience as the ground state energy is normally negative). For sufficiently large m , $(-H)^m|\Psi(0)\rangle$ approaches the eigenstate whose eigenvalue is the largest in magnitude, i.e., either the ground state of H or of $-H$. Convergence to the ground state of H can be ensured by adding a suitable constant to H .

We can now ask, what is the more efficient approach, using $e^{-\beta H}$ or $(-H)^m$? Proceeding as we did above when discussing the distribution of expansion orders in the SSE, we see that the power m required to reach the ground state is related to the β value required with the exponential operators as $m \approx \beta|E_0|$, where E_0 is the total ground state energy (which has been discussed in detail in Ref. [24]). As in the SSE method, $(-H)^m$ is expanded out into strings of the terms of the Hamiltonian, and these are sampled along with the starting state. There is no major difference between the two approaches, as the summation over n required with the series expansion is accomplished essentially for free when using the fixed string-length SSE approach, Eq. (9). The sampling of the operator strings does not differ significantly between the two approaches.

The ground state projector method is particularly useful when starting states $|\Psi(0)\rangle$ can be used that are tailored to the sampling method used for the operator sequences. For Heisenberg quantum spin models, good variational starting states can be written in the valence-bond basis [25, 26], i.e., the overcomplete basis of singlet pairs, and these are ideally matched with

loop-algorithm sampling schemes [23]. The valence-bond states have total spin $S_{\text{tot}} = 0$ and momentum 0, which is the ground state sector for the models where the approach is suitable. Fixing these quantum numbers from the outset can help to reduce the projection time β (or power m)

We will not discuss further any specifics of ground-state projector methods, but just note again that the differences with respect to $T > 0$ methods are very minor, and typically it is very easy to change a program from one case to the other.

1.4 Quantum annealing with generalized stochastic series expansion

In projector QMC calculations, results obtained from a projected state $|\Psi(\beta)\rangle$ at first sight have no obvious use if the projection “time” β is not sufficiently large for achieving convergence to the ground state. However, when considering $e^{-\beta H}$ as a time evolution operator in imaginary time, i.e., with time $t = -i\beta$ in $U(t) = e^{-itH}$, the projection corresponds to a quantum quench from the initial state $|\Psi(0)\rangle$ where at $t = 0$ the Hamiltonian is suddenly changed from the one which has $|\Psi(0)\rangle$ as its ground state to whatever H is considered in the simulation. Though imaginary and real time evolutions are quite different, one can sometimes extract real-time dynamic information from such imaginary-time quenches [27].

Real-time quantum dynamical calculations are in general very difficult, with any existing method, and it is interesting to ask what information may be gleaned from imaginary time evolution, where, for sign free models, QMC calculations can be carried out for various out-of-equilibrium situations. The aforementioned quantum quench in imaginary time is an extreme case of a more general setup where the Hamiltonian has some time dependence; $H = H(\tau)$ in imaginary time. The time evolution operator, starting at $\tau = 0$, is then

$$U(\tau) = T \exp \left(- \int_0^\tau d\tau' H(\tau') \right), \quad (12)$$

where T indicates time-ordering, and one may consider expectation values

$$\langle A(\tau) \rangle = \frac{\langle \Psi(0) | U(\tau) A U(\tau) | \Psi(0) \rangle}{\langle \Psi(0) | U(\tau) U(\tau) | \Psi(0) \rangle}. \quad (13)$$

Here one can again proceed with a time-slicing approach or apply a series expansion. Using the latter, the time evolution operator can be written as

$$U(\tau) = \sum_{n=0}^{\infty} (-1)^n \int_{\tau_{n-1}}^{\beta} d\tau_n \cdots \int_{\tau_1}^{\beta} d\tau_2 \int_0^{\beta} d\tau_1 H(\tau_n) \cdots H(\tau_2) H(\tau_1), \quad (14)$$

and one can proceed as in the several other cases discussed above and expand further into strings of terms of the Hamiltonian. Now, however, the Hamiltonian depends on imaginary time and the string is always time ordered. One can sample the strings and the initial state $|\Psi(0)\rangle$ with schemes very similar to the SSE method, and the integrals can be taken into account by sampling sequences of ordered time points [28].

There is also an alternative approach based on a product of Hamiltonians $H(\tau_m) \cdots H(\tau_2)H(\tau_1)$ for a fixed number of operators m , with a fixed spacing between the time points and no integrals over time [29]. This results in a dynamics slightly different from the Schrödinger dynamics in imaginary time, but for scaling purposes, e.g., when investigating dynamical quantum-criticality (i.e., the dependence on observables on the rate at which the Hamiltonian is changed in time close to a critical point), the two approaches both give correct results. It should be noted that these approaches really probe Hamiltonian quantum dynamics, and not the stochastic dynamics of the QMC sampling methods (as is the case with the other methods often referred to as “simulated quantum annealing” [30]) [31, 32].

2 Sampling algorithms and expectation values

We consider two classes of important $S = 1/2$ quantum spin models to illustrate SSE sampling schemes. First, the Heisenberg antiferromagnet defined by

$$H = J \sum_{\langle ij \rangle} \mathbf{S}_i \cdot \mathbf{S}_j = J \sum_{\langle ij \rangle} [S_i^z S_j^z + \frac{1}{2}(S_i^+ S_j^- + S_i^- S_j^+)], \quad (15)$$

where $\langle ij \rangle$ refers to the nearest-neighbor site pairs on an arbitrary bipartite (to avoid sign problems) lattice. Second, we consider the transverse-field Ising model (which will hereafter be referred to as just the Ising model), which is often defined with Pauli matrices,

$$H = - \sum_{ij} J_{ij} \sigma_i^z \sigma_j^z - h \sum_i (\sigma_i^+ + \sigma_i^-). \quad (16)$$

Here the Hamiltonian is written with a generic Ising coupling J_{ij} with no restriction on the range of the interactions, and we will consider both short-range and long-range cases. The SSE method can also be used with long-range interactions in the Heisenberg case, though with the limitation that there can be no frustration in order to maintain positive-definite sampling weights. With anisotropic Heisenberg interactions, which we will consider later, the diagonal part can also be frustrated. It is sometimes better to study the Ising model in a rotated basis, or, equivalently, to use the same σ^z basis as above but with the Hamiltonian written as

$$H = - \sum_{ij} J_{ij} \sigma_i^x \sigma_j^x - h \sum_i \sigma_i^z, \quad (17)$$

where $\sigma_i^x = \sigma_i^+ + \sigma_i^-$. Whichever version is better depends on details of the system (e.g., whether the interactions are frustrated, if disorder is present, etc.) and what physical observables are to be calculated (noting that diagonal observables are often easier to access, as we will see in Sec. 3). We will here only discuss implementations of algorithms with the version in Eq. (16).

2.1 Configuration representations and diagonal updates

As discussed in Sec. 1.1, a configuration in the SSE algorithm comprises a state $|\alpha_0\rangle$ and an index sequence $S_M = b_1, b_2, \dots, b_M$, the latter referring to a product of operators (and we

will use the terms “operators” and “indices” interchangeably). Here we take the expression of the partition function in Eq. (9), where the index sequence is of a fixed length M (which can be determined in a self-consistent way by the program during the equilibration part of a simulation). Removing the redundant complete sets of states we have

$$Z_{\text{SSE}} = \sum_{S_M} \frac{\beta^n (M - n)!}{M!} \sum_{\{\alpha\}} \langle \alpha | H_{b_M} \cdots H_{b_2} H_{b_1} | \alpha \rangle, \quad (18)$$

where in the products of M operators there are n terms of the Hamiltonian along with $M - n$ randomly distributed unit operators represented by the index $b = 0$. The number n changes in updates that are called *diagonal updates*, because they involve replacing a unit operator H_0 by a diagonal term of H (which we simply refer to as an operator insertion), whence $n \rightarrow n + 1$, or vice versa (an operator removal), in which case $n \rightarrow n - 1$. The generic way to carry out a sequence of diagonal updates is to go through the elements in S_M one-by-one and to attempt a replacement of the index whenever it does not correspond to an off-diagonal operator.

Off-diagonal operators cannot be updated individually while maintaining the periodicity constraint $|\alpha(M)\rangle = |\alpha(0)\rangle$ for the propagated states defined according to Eq. (5). How to carry out updates with the off-diagonal operator will be discussed in detail in the next section, but here we note that the general strategy is to replace some number of diagonal operators by off-diagonal ones, or vice versa, without changing the lattice location of the operator. With some models, the original Hamiltonian contains operators that make this possible, e.g., in the case of the Heisenberg interaction there are diagonal and off-diagonal operators on each bond, while in other cases certain constant diagonal operators have to be added just for the purpose of enabling the necessary operator replacements. The operator replacements also imply changes in the states, and normally the SSE algorithm is ergodic even if no other explicit state updates are included (though at high temperatures it can be useful to also carry out additional updates only on the stored state $|\alpha\rangle$, keeping the operator string unchanged).

To carry out an update at position p in the sequence requires the propagated state $|\alpha(p-1)\rangle = |\alpha(p)\rangle$ on which the diagonal operator H_{b_p} acts. The operator-index string is examined for each $p = 1, \dots, M$, and diagonal updates are attempted where possible. Whenever an off-diagonal operator is encountered at a position p , the stored propagated state is advanced; $|\alpha(p)\rangle = H_{b_p} |\alpha(p-1)\rangle$. If an encountered index $b_p = 0$, one out of a number N_d of diagonal operators can be chosen at random and inserted with a Metropolis acceptance probability, which depends on the matrix element $\langle \alpha(p) | H_{b'_p} | \alpha(p) \rangle$, where b'_p is the new generated diagonal index. If the current index $b_p \neq 0$, a replacement with $b'_p = 0$ is attempted, and the acceptance probability then depends on the current matrix element $\langle \alpha(p) | H_{b_p} | \alpha(p) \rangle$. These updates change the expansion power n by $+1$ and -1 , respectively, and this change also enters in the acceptance probability due to the factor $(M - n)!$ in the configuration weigh in Eq. (18). Importantly, to maintain detailed balance, the acceptance probabilities must also compensate for the inherent imbalance stemming from the fact that there are N_d ways of tentatively replacing an index $b_p = 0$ by a non-zero index (some of which may correspond to vanishing matrix elements, but that is irrelevant at this stage) when the number of Hamiltonian terms n is increased by one,

while for removing an operator there is only one way of replacing $b_p \neq 0$ by 0.

With all the relevant factors and imbalance taken into account, the following are the correct generic acceptance probabilities for a single-operator diagonal update with the SSE partition function (18):

$$P(0 \rightarrow b_p) = \frac{\beta N_d \langle \alpha(p) | H_{b_p} | \alpha(p) \rangle}{M - n}, \quad (19a)$$

$$P(b_p \rightarrow 0) = \frac{M - n + 1}{\beta N_d \langle \alpha(p) | H_{b_p} | \alpha(p) \rangle}, \quad (19b)$$

where n is the number of Hamiltonian operators in S_M before the update is carried out. The matrix elements are trivial, and often some fraction of them vanish.

As an alternative to Metropolis-type updates as described above, one can carry out the diagonal updates according to a heat-bath scheme, where the relative probabilities of all diagonal operators, including the fill-in unit operators, are considered and a choice is made among them according to their relative probabilities (instead of choosing them with equal probability in the scheme above, irrespective of the values of the matrix elements). In many cases it would take too long to compute these relative probabilities for each new state $|\alpha(p)\rangle$, but for sufficiently simple models it is possible to avoid this step, at the cost of some rejected attempts, as will be discussed below in the context of systems with long-range interactions.

Let us now be more specific and discuss how to represent the SSE configurations for the Heisenberg and Ising models. Illustrations are provided in Fig. 1. We denote the spins in the stored state by σ_i , $i = 1, 2, \dots, N$, and in a computer implementation they can be stored as integers, $\sigma_i = \pm 1$. In both cases, to define the terms H_b appearing in the operator products it is convenient to regard the subscript b as formally representing two indices (which can still for efficiency be packed back into a single integer in a program), but in slightly different ways for the two models considered. In both cases we use the notation $H_{0,0} = 1$ for the fill-in unit operators. For the antiferromagnetic Heisenberg model (15) we set $J = 1$ and define the following diagonal and off-diagonal operators:

$$H_{1,b} = \frac{1}{4} - S_i^z S_j^z, \quad H_{2,b} = \frac{1}{2}(S_i^+ S_j^- + S_i^- S_j^+). \quad (20)$$

Then $H = -\sum_b [H_{1,b} - H_{2,b}] + N_b/4$, where N_b is the total number of interacting spin pairs, e.g., $N_b = DN$ for N spins on a D -dimensional simple cubic lattice with periodic boundary conditions. Here a constant $1/4$ has been included in the diagonal operators, and they can therefore act with a non-zero outcome only on two antiparallel spins. There is then a useful (as we will see) similarity with the off-diagonal terms, which also can only act on antiparallel spins. The non-zero matrix elements of the Hamiltonian terms $H_{1,b}$ and $H_{2,b}$ are all $1/2$. The weight of an allowed SSE configuration in Eq. (18) is therefore $W(S_M) = (\beta/2)^n (M - n)!$, where the unimportant overall factor $1/M!$ has been omitted and there are never any minus signs (for bipartite interactions) because the number of off-diagonal operators in the string has to be even. Note that there is no explicit dependence of the weight on the state $|\alpha\rangle$ in Eq. (18), but the state imposes constraints on the operator string as only operations on antiparallel spins are allowed.

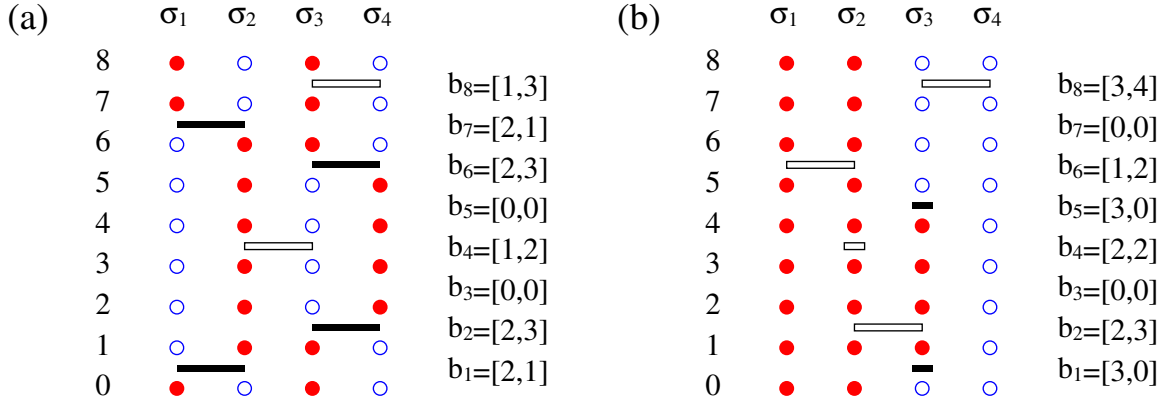


Fig. 1: Graphical representations of SSE configurations for (a) the Heisenberg model and (b) the Ising model, in both cases for a system of four spins and with the SSE cutoff $M = 8$. Up and down spins correspond to solid and open circles. All the propagated states $|\alpha(0)\rangle, \dots, |\alpha(M)\rangle$, with $|\alpha(M)\rangle = |\alpha(0)\rangle$, are shown along with the operators H_{b_p} . The number of Hamiltonian terms for both systems is $n = 6$, and the two cases of empty slots between propagated states correspond to fill-in unit operators $H_{0,0}$ at these locations. In (a) the solid and open bars represent, respectively, off-diagonal and diagonal parts of the Heisenberg exchange operators. In (b) the ferromagnetic Ising interactions are likewise represented by open bars, and the off-diagonal single-spin flip operators are represented by short solid bars. The short open bars correspond to the constant site-indexed operators.

An example of a very small Heisenberg SSE configuration is shown in Fig. 1(a). Note again that the mean number of operators is $\propto \beta N$, and in large-scale simulations the number can be up to many millions.

In the diagonal update, if an encountered index pair at the current location p is $b_p = [0, 0]$, a bond index b is generated at random among all the choices. If the spins at the sites $i(b), j(b)$ connected by bond b are antiparallel in the currently stored state $|\alpha(p)\rangle$, i.e., $\sigma_i \neq \sigma_j$, then the operator $H_{1,b}$ is allowed and the index pair is set to $b_p = [1, b]$ with probability given by (19a), where the matrix element equals $1/2$. If the two spins are parallel nothing is changed and the process moves to the next position, $p \rightarrow p + 1$. Each time an off-diagonal operator $[2, b]$ is encountered, in which case no diagonal update can be carried out, the stored state is propagated; $\sigma_i \rightarrow -\sigma_i$ and $\sigma_j \rightarrow -\sigma_j$.

For the Ising model (16), where the Ising interactions J_{ij} are of arbitrary range, we define the following operators [33]:

$$H_{i,j} = |J_{ij}| - J_{ij}\sigma_i^z\sigma_j^z \quad (i \neq j), \quad H_{i,i} = h, \quad H_{i,0} = h(\sigma_i^+ + \sigma_i^-). \quad (21)$$

Here the constant site-indexed operators $H_{i,i}$ serve as an example of how trivial diagonal terms can be added to the Hamiltonian for the purpose of carrying out off-diagonal updates—as we will see in the next section, updates will be based on replacements $H_{ii} \leftrightarrow H_{i,0}$. In the diagonal updates, the trivial constants will be inserted and removed along with the Ising operators $H_{i,j}$ ($i, j \neq 0, i \neq j$). In the Ising operators, the presence of the constant $|J_{ij}|$ implies that only an operation on a parallel pair of spins (for a ferromagnetic coupling J_{ij}) or an antiparallel pair (for

antiferromagnetic coupling) is allowed. This choice of the added constant is again motivated by its convenience for constructing the off-diagonal (cluster) update, as we will see further below. The example of an SSE Ising configuration in Fig. 1(b) only includes nearest-neighbor ferromagnetic interactions.

For a D -dimensional simple cubic lattice and only nearest-neighbor interactions J included, there are now $N_d = (D + 1)N$ diagonal operators, of which DN are Ising terms and N are the constant operators $H_{i,i}$. When these are all generated with equal probability in attempts to insert operators, with the Metropolis acceptance probability given by Eq. (19a), there is an inherent inefficiency if h and J are very different. For example, if $J = 1$ and $h \ll 1$, most of the attempted h -operator insertions will be rejected.

The rejection problem becomes worse with long-range interactions, e.g., if J_{ij} decays with the separation r_{ij} as a power law, $|J_{ij}| \propto r_{ij}^{-\alpha}$. Then there are $N(N-1)/2 + N = N(N+1)/2$ diagonal operators to generate at random, and those with very small J_{ij} will be rejected almost always. This problem can be overcome easily by generating the diagonal Hamiltonian terms along with the fill-in operators H_{00} using a heat-bath method [33]. Instead of treating operator insertions and removals as different types of updates, these are now combined and carried out at all positions p at which the current operator H_{b_p} is not off-diagonal. The method needs a precomputed table of integrated relative probabilities P_{ij} of all the different diagonal operators, where it is tentatively assumed that all operators are allowed. The probabilities are calculated from Eq. (18) and the definitions in (21), and, for efficiency, mapped into a single-index ordered table P_k , $k = 1, \dots, N(N+1)/2$. In each diagonal update, a random number $r \in [0, 1)$ is generated and the corresponding operator is identified in the table, using the bisection method to search in the ordered table for the corresponding probability window $P_{k-1} \leq r < P_k$, thus finding the correct operator indices $i(k), j(k)$ in $\propto \ln(N)$ operations. A new Ising operator has to be rejected if the spins σ_i, σ_j are incompatible with the sign of the interaction J_{ij} .

A very useful aspect of this approach is that it renders an algorithm with processing time scaling as $N \ln(N)$ for long-range interactions, instead of the naively expected N^2 scaling; details are described in Ref. [33]. This efficient method for the diagonal updates can also be used for the Heisenberg model with long-range interactions, and for many other cases as well. Even with short-range interactions, the heat-bath approach may be slightly more efficient than the Metropolis update, though the difference in efficiency is likely minimal in most cases.

The probability of finding a 0-element (fill-in operator) and attempting an operator insertion in the string clearly depends on how the cutoff M of the series expansion is chosen. The cutoff naturally should be high enough for the contributions from terms with $n > M$ to be completely negligible, so that the truncation of the series expansion is no approximation in practice. In an SSE simulation this can be ensured by always requiring M to be significantly larger than the largest n that has been reached so far during the equilibration part of the simulation (with any adjustments done after each completed sweep of diagonal update during equilibration). In practice, M exceeding the maximum n by 30-50% is a suitable choice; clearly sufficient for causing no systematical error and also enough to allow a large number of operator insertion attempts. Normally M (and $\langle n \rangle$) converges very quickly at the initial stages of a simulation.

2.2 Loop and cluster updates

In classical MC simulations of spin models, cluster updates [34, 35] have played a major role in reaching system sizes sufficiently large for reliable finite-size scaling studies. These methods also have generalizations for some quantum spin models [36–38, 20, 39, 40, 33], including the Heisenberg and Ising systems discussed here. Within the SSE approach, the loop and cluster updates are carried out in the operator string, which at this stage is regarded as a network of connected *vertices* comprising the operators and their associated “incoming” and “outgoing” states (i.e., the information needed to compute the weight of the operator string). The general strategy is to update a set of vertices (which are connected to each other to form a loop or a cluster) but maintain their lattice locations (which dictate their connectivity). The connectivity is changed only as a consequence of the diagonal updates. The off-diagonal updates also can change the stored state $|\alpha_0\rangle$, since the loops or clusters can span across the periodic time boundary represented by the stored state.

2.2.1 Linked vertex list

The loop or cluster updates are carried out in the linked vertex list, which after a full sweep of updates is mapped back into the simple operator-index string and state $|\alpha\rangle$ (or, in some cases, it is better to do this mapping-back continually during each loop or cluster flipping procedure). A vertex comprises the spin states before and after an operator has acted, and these states are associated with the *legs* of the vertex, e.g., for a two-body interactions there are four legs for each vertex; two before and two after the operator has acted on the two spins. Thus, the vertex is associated with a matrix element of the form $\langle\sigma_i(p), \sigma_j(p)|H_{b_p}|\sigma_i(p-1), \sigma_j(p-1)\rangle$, where i and j are the sites involved in the interaction term H_{b_p} . The way the legs of different vertices are linked to each other corresponds directly to the segments of unchanged spins between operators in the “time” direction in Fig. 1. These lines of spins are redundant when representing the changes that can be made in a configuration by changing the type of some of the operators (e.g., diagonal to off-diagonal or vice versa) and making associated changes in the spin state (represented by the changes at the vertex legs). In a computer program, these lines of constant spins are represented by bidirectional links connecting the relevant vertices, enabling direct jumps between any of the connected vertex legs without regard for the intermediate irrelevant spins. The linked list can be constructed according to a simple and efficient scheme discussed in detail in Ref. [41].

2.2.2 Loop and cluster construction

Some relevant vertices are illustrated in Fig. 2, along with lines indicating how the legs of vertices can be affected by a loop or cluster flip. The general idea is to change the spin state at one vertex leg and then move either to another leg of the same vertex or follow a link to another vertex, until the process closes. The allowed processes depend on the Hamiltonian. In the simplest cases, the *deterministic loops* for Heisenberg models and Ising cluster updates, there is a unique way to move at each step. The entire system can then be divided into a unique set

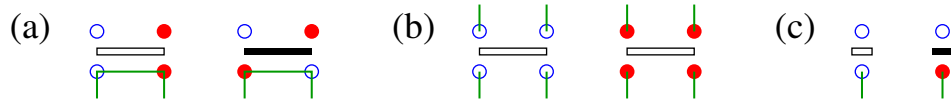


Fig. 2: Examples of the elementary vertex processes by which loop and cluster updates are carried out for Heisenberg and Ising models. The green line segments represent parts of a loop or cluster. The spins and operators on a given vertex correspond to the state before the loop [in (a)] or cluster [in (c) and (d)] has been flipped, and the adjacent vertex shows the state after the loop or cluster has been flipped.

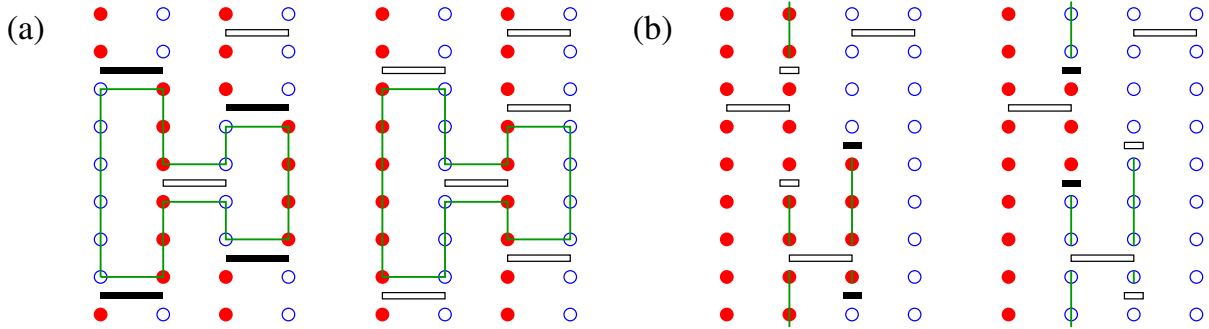


Fig. 3: Examples of entire loop (a) and cluster (b) updates based on the elementary processes shown in Fig. 2. The pairs of configurations correspond to the states before and after the loop or cluster (indicated by green lines) has been flipped. Note that in (b) the cluster spans across the periodic time boundary. In cases where the loop and cluster algorithms are efficient, the vertex weight (matrix element) is the same before and after the flip.

of loops or clusters, each of which can be flipped (i.e., changing the spin states at the affected vertex legs) without affecting the configuration weight. Each loop or cluster can then be flipped independently with probability $1/2$, as in the classical Swendsen-Wang algorithm.

The condition that the loop or cluster flip must not change the configuration weight is the main restriction of this type of update, and is very similar to the limitation of classical cluster updates, e.g., the Swendsen-Wang cluster update for the Ising model [34] does not work in the presence of an external magnetic field. In the next section we will discuss more complicated *directed loop updates*, which partially overcome this limitation.

We will not discuss the details of the SSE loop [39,41] and cluster updates [33] here, but refer to the literature. We just note that the loop update, also called *operator-loop* update within the SSE framework to emphasize that everything is formulated with a network of connected vertices (operators), corresponds to moving along a one-dimensional, non-branching path in the network of vertices. At some point, this process necessarily will lead back to the vertex leg at which the process started, at which point the loop is completed and a new one can be started. An example of a loop is shown in Fig. 3(a). Normally these loops are constructed one-by-one, with the random decision of whether or not to flip the loop made before each new loop is started, and in each case the traversed vertex legs are flagged as visited, so that each new loop can be started from a not previously visited leg until all legs have been visited.

The cluster update for the Ising model differs from the Heisenberg loop update in the important sense of being branching. Since each two-body vertex must have all four legs in the same spin state, because of the choice of the added constant in the Ising operators H_{ij} in Eq. (21), once one of those spins has been flipped the other three must be flipped as well to “heal” the vertex. The cluster therefore can branch out each time a two-body vertex is encountered. However, no branches extend out from vertex legs that have been previously visited, and eventually the cluster therefore does not grow further, and all defective vertices generated in the process have been healed. Branches that point to a single-site operator also terminate, because by flipping the type of single-site vertex from $H_{i,0}$ to $H_{i,i}$, or vice versa, the propagating defect is healed and no further changes need to be made at that branch. Fig. 3(b) shows one example, which in this small configuration involves a single two-body operator and four one-body operators. In general, the branching-out can propagate to a very large number of vertices.

In a program implementation, the successive branching-out is handled by adding new branches to a stack, from which branches to follow are picked one-by-one until the stack is empty (the cluster is completed). By starting each new cluster from a vertex leg not previously visited, the entire system will eventually be subdivided into clusters (each of which was flipped by probability 1/2), as in the Swendsen-Wang algorithm.

2.2.3 Directed-loop updates

The loop update described above can be generalized to Heisenberg systems with anisotropic interactions and uniform magnetic fields, i.e., pair interactions of the form

$$H_{ij} = J_{ij}[\Delta_{ij}S_i^zS_j^z + \frac{1}{2}(S_i^+S_j^- + S_i^-S_j^+) + h_iS_i^z + h_jS_j^z], \quad (22)$$

to which a suitable negative constant may have to be added in order to avoid a sign problem originating from the diagonal terms. In the directed-loop algorithm [42], the one-dimensional path through the vertex list is no longer unique, in contrast to the deterministic loop update, but a choice on how to continue the path has to be made at each vertex. A set of equations, called the *directed-loop equations* relate the possible vertex weights to the probabilities of the different paths through a vertex, and by solving these equations a process maintaining detailed balance is ensured.

The paths through the system taken in the directed-loop algorithm are similar to those in the continuous-time worm algorithm [13], but the probabilities are different. The directed loops often lead to a much more efficient evolution of the configuration. The directed loops can also be directly formulated within the continuous-time framework [42], and implementations of continuous- and discrete-time WL methods now often rely on the directed loop ideas [43].

2.2.4 Sweeping cluster update

A very interesting recent development of efficient updates within the SSE approach is the *sweeping cluster update* developed for highly constrained models such as the quantum dimer

model [44]. It is somewhat similar to an earlier *multibranch cluster update* developed in order to enhance the performance of a loop algorithm for a bosonic model with a constrained ring-exchange interaction [45], but is more directly tailored to strict geometric restrictions. Simulation results for the square-lattice quantum dimer model indicate that this is a promising approach, overcoming at least some of the rather severe limitations of previous methods for this important class of models.

2.2.5 Extended-cell approach

In the original SSE formulation discussed in Sec. 1.1, the terms H_b are defined as they appear in the Hamiltonian, e.g., they are the single-site operators such as $S_i^+ + S_i^-$ or two-body terms such as $S_i^+ S_j^- + S_i^- S_j^+$. A greater degree of freedom can be given to loop and cluster updates by enlarging the cell on which the operators and vertices are defined [46]. For instance, one can define an off-diagonal operator on three sites as $(S_i^+ S_j^- + S_i^- S_j^+) I_k$, where I_k is a unit operator with a site index. This allows the off-diagonal part to move during an update, e.g., a vertex with the mentioned operator can be updated (within some scheme involving more than one operator) to $(S_i^+ S_k^- + S_i^- S_k^+) I_j$. This trick has proved to be very helpful for speeding up SSE simulations of systems where the diagonal interactions are highly frustrated [47–50].

2.2.6 Loop and cluster updates at $T = 0$

As we saw in Sec. 1.3, the $T > 0$ SSE algorithm can be very easily modified for ground-state projection, with the only essential difference being the change from periodic time boundary conditions to boundary conditions dictated by the trial state. How this change affects the loop and cluster algorithms depends on the type of trial state used.

In the case of spin-isotropic Heisenberg models, a particularly good choice of trial state is one written in the valence-bond basis, i.e., the overcomplete basis of singlet pairs, with the two members of each singlet occupying different sublattices on a bipartite lattice [23]. The valence bonds then serve as segments of loops touching the trial state, and there is no weight change, as before, when flipping a loop. The valence-bond trial state can also in many cases be variationally optimized, for improved convergence [25, 26]. Typically the valence-bond trial state is a superposition of valence-bond coverings, and the simulation involves a simple update for reconfiguring the bonds.

A convenient type of trial state to use in combination with the Ising cluster update is the ferromagnetic product state in the x spin direction, i.e., $|\Psi(0)\rangle = \prod (\uparrow_i + \downarrow_i)$, where $\uparrow$ and $\downarrow$ correspond to $\sigma_i^z = \pm 1$. This state corresponds to a completely open time boundary condition, where cluster branches simply terminate at the trial state, and boundary clusters can be flipped as any other cluster without any change to the configuration weight. At the opposite extreme, it may sometimes be useful to use a fully polarized ferromagnetic state as the trial state [29], in which case any cluster touching the boundary cannot be flipped at all.

3 Estimators for expectation values

Expectation values suitable for estimation with MC simulation are normally written in the form

$$\langle A \rangle = \frac{\sum_C A_C P_C}{\sum P_C}, \quad (23)$$

where $\{C\}$ is some configuration space, P_C is the probability (or relative probability) of configuration C , and A_C is the estimator of the quantity A . In the QMC algorithms discussed in the preceding sections the sum of weights or probabilities in the denominator is the partition function at $T > 0$ and the wave-function norm in $T = 0$ projector methods. When importance-sampling the configurations, so that their probability of being generated equals P_C , the simulation result for $\langle A \rangle$ is the mean value $\langle A_C \rangle$, with a statistical error computed using some data binning technique. Here we discuss the form of the estimator A_C for some important classes of physical observables.

We should first note that, in classical MC simulations, A_C is normally just a trivial function of the sampled degrees of freedom, and that is also the case with SSE and WL methods when the operator A is diagonal in the basis used. However, for off-diagonal operators the situation is more complicated, as then the configurations C contributing to $\langle A \rangle$ are not necessarily the same as those that contribute to the sampled normalization in Eq. (23). We will here first discuss diagonal operators and then briefly touch on the topic of off-diagonal correlation functions.

3.1 Diagonal correlations and susceptibilities

Equal-time expectation values, e.g., two-point or multi-point correlation functions, of diagonal operators in the basis used are normally trivial, and for improved statistics they can be averaged over the SSE propagation index p ;

$$\langle A \rangle = \left\langle \frac{1}{n} \sum_{p=0}^{n-1} A(p) \right\rangle, \quad (24)$$

where $A(p) = \langle \alpha(p) | A | \alpha(p) \rangle$ is the value of A computed in propagated state p in a given SSE configuration. Here when defining the propagated state $|\alpha(p)\rangle$ as in Eq. (5) we have implicitly disregarded the fill-in unit operators H_0 in the formulation of the method with a fixed string length M . It is also correct to include the unit operators and average over M propagated states. In many cases it is too time consuming to average over all states, and since the propagated states are highly correlated there is also no loss of statistics in practice to average over a small fraction of the states, e.g., only including every N th or so state in Eq. (24). Similarly, spatial averaging may be carried out fully or partially to improve the statistics.

We are often interested in correlation functions in Fourier space of a periodic system, e.g., with the transformed diagonal spin operator

$$S_q^z = \frac{1}{\sqrt{N}} \sum_{\mathbf{r}} e^{-\mathbf{q} \cdot \mathbf{r}} S_{\mathbf{r}}^z. \quad (25)$$

The real part of the correlation function (often called the structure factor) is $S(\mathbf{q}) = \langle S_{-\mathbf{q}}^z S_{\mathbf{q}}^z \rangle$ and the imaginary part vanishes by symmetry. If all or many values of the wave-vector $\mathbf{q}$ are needed, the most efficient way is often to evaluate (25) using the *Fast Fourier Transformation* method (see, for example, Ref. [51]). However, if only a small number of values are needed it can be better to use the basic summation formula just for those cases.

Time dependent (in imaginary time) correlations can also be calculated, and by numerical analytic continuation they can provide real-frequency dynamic spectral functions of experimental interest, e.g., the dynamic spin structure factor measured in inelastic neutron scattering (see Ref. [52] for a recent discussion of analytic continuation and spectral functions). The way imaginary time is treated within the SSE method was already considered in Eq. (6); a fixed value τ of imaginary time corresponds to a summation over separations p between propagated states. Because the weights of the different separations follow a very narrow distribution, only a small fraction of them has to be summed over in practice, and one can use a precomputed table of safe lower and upper bounds on the separations for the different τ values considered.

Alternatively, time dependent quantities can be more easily calculated if the SSE is formulated in combination with time slicing, where first the exponential operator is written as $(e^{-\Delta H})^m$, with $\Delta = \beta/m$, and the factors are series expanded individually with equal cutoffs M_Δ . This is no approximation, since the slicing does not involve any separation of non-commuting operators. The operator string consists of m segments, $i = 1, \dots, m$, containing n_i Hamiltonian terms. Therefore, in the diagonal updates, β and n in Eqs. (19a) and (19b) are replaced by Δ and n_i , with i corresponding to the segment in which an update is carried out. The off-diagonal updates are not affected at all by the slicing, since they do not change the numbers n_i . Time-dependent correlations of diagonal operators can now be evaluated easily by just considering the propagated states at the boundaries between slices, which correspond to sharp time displacements in multiples of Δ .

An interesting aspect of the SSE method is that rather simple estimators for static susceptibilities can be constructed. Such susceptibilities are in general given by Kubo integrals of the form

$$\chi_{AB} = \int_0^\beta d\tau \langle A(\tau) B(0) \rangle, \quad (26)$$

where $h_A A$ can be regarded as a perturbation added to the Hamiltonian (the number h_A being a corresponding field strength) and $\langle B \rangle$ is the linear response of the operator B to this perturbation; $\langle B \rangle = h_A \chi_{AB}$. If A and B commute with H , then $\chi_{AB} = \beta \langle AB \rangle$, an important example of which is the uniform magnetic susceptibility of a system in which the magnetization M is conserved; then $A = B = M$ and $\chi_u = \beta \langle M^2 \rangle$ as in a classical system. In the more common case where A and B do not commute, e.g., if they are individual spins at different sites; $A = S_i^z$, $B = S_j^z$, the integral over the time dependent correlation function has to be evaluated. Within SSE, the integral can be computed for each individual SSE configuration, with the result [10,11]

$$\chi_{AB} = \left\langle \frac{\beta}{n(n+1)} \left[\left(\sum_{p=0}^{n-1} A(p) \right) \left(\sum_{p=0}^{n-1} B(p) \right) + \sum_{p=0}^{n-1} A(p) B(p) \right] \right\rangle, \quad (27)$$

where $A(p)$ and $B(p)$ are the eigenvalues of the respective diagonal operators computed in propagated state p . Often these susceptibilities are also needed in Fourier space, but it can still be better to do the computations in real space (depending again on how many points in momentum space are needed, etc.) and only take the Fourier transform as the last step on the final averaged real-space susceptibilities.

3.2 Off-diagonal observables

We have already encountered two estimators for off-diagonal observables in the SSE method; the total energy and heat capacity

$$E = \langle H \rangle = -\frac{\langle n \rangle}{\beta}, \quad C = \frac{dE(T)}{dT} = \langle n^2 \rangle + \langle n \rangle^2 - \langle n \rangle, \quad (28)$$

which depend only on the number n of Hamiltonian terms in the sampled operator sequences. These expressions are exactly the same as those in Handscomb's method. Note that the internal energy E contains any constants that have been added to H for the sampling scheme.

For any operator H_b in the Hamiltonian, diagonal or off-diagonal, its expectation value is simply given by

$$\langle H_b \rangle = -\frac{\langle n_b \rangle}{\beta}, \quad (29)$$

where n_b is the number of instances of the index b in the sampled sequences S_M . Thus, the energy estimator in Eq. (28) corresponds to summing over all b . The simplicity of off-diagonal observables that can be related to the terms in the Hamiltonian also carries over to correlation functions [19];

$$\langle H_a H_b \rangle = \frac{(n-1)\langle n_{ab} \rangle}{\beta^2}, \quad (30)$$

where n_{ab} is the number of times the indices a and b appear next to each other in S_n , where S_n is the sequence obtained from S_M when all zero indices are disregarded. This result can also be generalized to time dependent correlations.

Kubo integrals of the form (26) also have their direct SSE estimators. For $A = H_A$ and $B = H_B$ (terms in the Hamiltonian corresponding to the operators H_b in the formalism above), after integrating over time for each SSE configuration we have [11]

$$\chi_{AB} = \beta^{-1}(\langle n_A n_B \rangle - \delta_{AB} \langle n_A \rangle), \quad (31)$$

which is very easy to evaluate. Two important off-diagonal quantities of this type can be mentioned (see the cited references for details):

The spin stiffness is the second-order energy (or free energy at $T > 0$) response to a boundary twist imposed on the spins, or, alternatively, to a continuous twist field analogous to a vector potential. The spin stiffness is also analogous to the superfluid stiffness of a boson system. The stiffness (spin or superfluid) has an interesting estimator related to fluctuations of winding numbers [53], essentially the currents circulating around a periodic system. The SSE estimator for

the stiffness is similar to Eq. (31), involving the counting of off-diagonal operators transporting spin (or charge) to “left” or “right” [54].

Recently, a simple algorithm for computing the fidelity susceptibility was proposed [55] (likely more efficient than a different method proposed earlier [56]). This observable, which quantifies at what rate a state changes when some parameter is varied, plays an important role in quantum information theory and is also useful in studies of quantum phase transitions. The estimator for the fidelity susceptibility is similar to a correlator of the form (31), with n_A and n_B corresponding to operators related to the infinitesimally varied coupling constant. These operators are counted in different halves of the time-periodic SSE configurations.

Two other SSE-accessible information-theory inspired quantities can also be mentioned; the *entanglement entropy* [57] and the *mutual information* [58], both in their Renyi versions. They are evaluated using a so-called swap operator, or a modified space-time lattice corresponding to a ratio of two different partition functions.

To evaluate general off-diagonal correlation functions, which cannot be expressed simply with the terms of the Hamiltonian, one has to go outside the space of the terms contributing to the partition function (or wave function norm in $T = 0$ approach). An efficient scheme was first devised within the worm algorithm [13], and a simple generalization to the SSE framework is also known [59]. We will not discuss details here, but just note that in the context of loop, directed-loop, or worm algorithms, the loop or worm building process involves two point defects that can be associated with the raising or lowering operators, S_i^+ and S_j^- (or creation and destruction operators in particle models). Space-time correlation functions involving these operators, e.g., $\langle S_i^+(\tau) S_j^-(0) \rangle$ are therefore directly related to the probability distribution of the separation between the defects.

4 Recent applications

MC simulations have played, and continue to play, an important role in studies of phase transitions in classical statistical physics. In a similar way, QMC simulations of quantum lattice models are now helping to push the boundaries of knowledge in the field of quantum many-body physics, uncovering various quantum states and elucidating the nature of various quantum phase transitions (i.e., transitions between different types of ground states and associated scaling behaviors at $T > 0$) using sign free “designer Hamiltonians” [60]. Some selected applications of SSE methods to different classes of quantum-critical models will be very briefly reviewed in Sec. 4.1, as a guide to recent works and mainly reflecting the author’s own interests. Works aimed at extracting dynamic spectral functions from imaginary-time correlations, using numerical analytic continuation methods, will be discussed in Sec. 4.2. Finally, SSE works on disorder (randomness) effects are reviewed in Sec. 4.3. The emphasis is on SSE applications, and references to works using other methods are therefore very incomplete.

4.1 Quantum phases and criticality in spin systems

One of the first successes of WL-type QMC simulations in quantum magnetism was the convincing demonstration of long-range order at $T = 0$ in the 2D $S = 1/2$ Heisenberg antiferromagnet [61] (the currently most precise results were obtained with the SSE [54] and valence-bond projector [23] methods). Following this important result (of which there is still no rigorous analytical proof), the focus shifted to ways in which the long-range order can be destroyed by perturbing the Heisenberg model in various ways. Many studies were devoted to statically dimerized 2D Heisenberg models, e.g., the SSE studies in Refs. [62–65], where there is a pattern of nearest-neighbor couplings of two strengths, J_1 (inter-dimer) and J_2 (intra-dimer), such that each spin belongs uniquely to a dimer. As a function of the ratio $g = J_2/J_1$, there is then a loss of antiferromagnetic order at $T = 0$ when g reaches a critical value g_c . The physical mechanism of this transition is that the density of singlets on the dimer increases with g , and eventually the ground state becomes close to a product state of dimer singlets. The transition is continuous and belongs to the 3D O(3) universality class, with the third dimension corresponding to imaginary time. In some cases, confirming this universality was challenging [64], because of, as it turns out [65], effects of competing scaling corrections. The $T = 0$ quantum-critical point is associated with a so-called *quantum-critical fan* extending out in the (T, g) plane from g_c and which is associated with various scaling laws of physical quantities [66]. SSE and other QMC studies have, among other things, established the range of validity of these asymptotic scaling behaviors, and also tested the applicability of various approximate analytical calculations [24], e.g., the $1/N$ expansion, where N is the number of spin components.

The O(3) transition driven by dimerization can be realized experimentally in the 3D spin-dimer system TiCuCl_3 under pressure [67] and this has motivated SSE simulations of this phase transition also in 3D generalizations of the 2D Heisenberg systems discussed above. In a 3D Heisenberg system antiferromagnetic long-range order can survive also at $T > 0$ (which is excluded by the Mermin-Wagner theorem in 2D). An empirical universal scaling form of the critical temperature was found in Ref. [68] and further studied in Ref. [69]. Multiplicative logarithmic corrections at the $T = 0$ and $T > 0$ phase transitions have also been studied in detail [70]

In the statically dimerized 2D and 3D systems, the paramagnetic phase is a unique quantum state with no spontaneous symmetry breaking—the singlets simply form, with some fluctuations, at the dimers imposed by the Hamiltonian itself. A more interesting case is where also the paramagnetic state breaks additional symmetries spontaneously. It was discovered by SSE simulations that a certain planar [XY, or U(1) symmetric] $S = 1/2$ spin model could go through a transition from XY magnetized to spontaneously dimerized in what appeared to be a continuous quantum phase transition [71]. Shortly thereafter, a theory was proposed for a new type of quantum phase transition, beyond the Landau-Ginzburg-Wilson (LGW) paradigm, between 2D quantum antiferromagnetic and spontaneously dimerized states (also called valence-bond-solids, VBSs). In this theory of *deconfined quantum critical points* [72], the two different order parameters arise out of the same objects—spinons and gauge fields—instead of being described by separate fields corresponding to the two order parameters. This theory stimulated many fur-

ther works on various 2D quantum antiferromagnets with VBS transitions, and these studies have uncovered a rich variety of phenomena beyond the original DQCP proposal.

Traditionally, VBS states were discussed in the context of frustrated Heisenberg models, such as the model with first J_1 and second J_2 neighbor interactions on the square lattice. These models have sign problems and are not accessible to QMC simulations. Another class of models, was proposed to study the DQCP phenomenon with QMC without sign problems—the J - Q model, in which the Heisenberg exchange is supplemented by a multi-spin interaction built out of $S = 1/2$ singlet operators $P_{ij} = 1/4 - \mathbf{S}_i \cdot \mathbf{S}_j$ [73]. This interaction by itself, $-JP_{ij}$, is equivalent to the antiferromagnetic Heisenberg exchange. Products of two or more of the singlet projectors make up the competing Q interaction, e.g., terms of the form $-QP_{ij}P_{kl}$ with the involved sites i, j, k, l suitably arranged on plaquettes (without breaking the symmetry of the square lattice).

When the ratio Q/J is large, the correlated singlets favored by the multi-spin interactions cause many of these models to undergo quantum phase transitions into four-fold degenerate columnar VBS states. This phase transition can be investigated in detail only with QMC simulations. SSE studies have established what seems like a continuous transition [73–78], similar to the proposed DQCP but with anomalous scaling corrections that so far only have a phenomenological explanation [79]. It has also been proposed that the transition is actually weakly first-order, with the true DQCP only existing outside the space that can be reached with lattice models (e.g., in a fractal dimension or on the complex plane) [80]. Though the ultimate nature of the phase transition is still unsettled, it is already clear that the length scale associated with the transition is very large, and the DQCP phenomenology applies.

By perturbing the J - Q model sufficiently so that the Q terms form a checker-board pattern, the VBS state can be modified from a four-fold columnar to a two-fold degenerate plaquette-singlet state. The transition from the antiferromagnet into the plaquette-singlet state is clearly first-order according to $T > 0$ SSE and $T = 0$ projector QMC studies [81]. However, the transition point is associated with an unexpected higher symmetry, combining the $O(3)$ magnetic order parameter and the scalar Z_2 plaquette order parameter into an $O(4)$ vector. A similar phenomenon with emergent $SO(5)$ symmetry has been studied with SSE simulations in a spin-1 J - Q model [82]. The plaquette-singlet state is of relevance in the frustrated 2D magnet $\text{SrCu}_2(\text{BO}_3)_2$ [83], and SSE simulations aimed at explaining the phase transitions driven by high pressure in this system have been reported very recently [84].

Building on the idea of the J - Q model, Kaul and collaborators have constructed several other classes of sign-free “designer Hamiltonians” [85–89]. The original J - Q models and these extended variants provide unique opportunities to further explore many interesting quantum phases and quantum phase transitions.

Highly frustrated Ising interactions, supplemented with various off-diagonal terms, can also be studied with SSE simulations (though the sampling is more challenging [49] and system sizes as large as those for Heisenberg and J - Q interactions can not be reached). The focus of these studies is typically to explore different incarnations of Z_2 quantum spin liquids and their quantum phase transitions [90–92].

4.2 Dynamic response of quantum magnets

To connect numerical simulations of lattice models to experiments, dynamic response functions are the most useful, e.g., in quantum magnets the dynamic spin structure factor $S(q, \omega)$ can be measured directly in inelastic neutron scattering experiments, and the low-energy structure factor is accessed, e.g., in NMR experiments. In QMC calculations, dynamic spectral functions can only be accessed in the form of imaginary-time dependent correlation functions, and these have to be analytically continued to the real time (or frequency) domain using some numerical scheme [93]. Analytic continuation in its own right is an interesting and challenging technical problem subject to ongoing research activities; see Ref. [94,52] for a recent example of methods developed by the author and collaborators. While all numerical analytical continuation methods have natural limitations in the frequency resolution that can be achieved, due to the statistical noise in the QMC data (even when the noise level is exceedingly small), important spectral features can be resolved, and sometimes it is possible to compare with experiments in a very detailed manner [52,95].

While the static properties of the 2D Heisenberg model have been well understood for some time, there has been a long-standing unresolved mystery in the dynamic response: At and close to the equivalent wave-vectors $q = (\pi, 0)$ and $(0, \pi)$, the excitation energy is reduced and the spectral line shape of $S(q, \omega)$ is anomalously broadened. The anomalies cannot be easily explained within spin-wave theory. In recent work based on SSE and analytic continuation, it was found that the phenomenon is a precursor to a DQCP that can be reached by adding other interactions [52]. Spectral functions at the DQCP of the J - Q model have also been studied and are in excellent agreement with a field-theory treatment based on the so-called π -flux state [96]. The simpler dimerized Heisenberg models with $O(3)$ transitions also have interesting dynamical response functions. In particular, the amplitude model (also often called the Higgs mode) of near-critical systems (on the ordered side of the phase transition) have been studied in 2D [97] and 3D [98,99] and compared with experimental results for TiCuCl_3 [67].

In highly frustrated quantum Ising systems, spectral functions can give important insights into the nature of exotic excitations. Recent SSE-studied examples include the identification of separate photon and spinon modes in a quantum spin-ice system [100] and a two-spinon continuum in a kagome-lattice model [92].

4.3 Disordered systems

Randomness (quenched disorder) can fundamentally affect quantum states and quantum phase transitions. Many SSE studies have been devoted to the effects of random couplings in the ordered Heisenberg antiferromagnet [101] and at the $O(3)$ transition in the dimerized Heisenberg systems mentioned above in Sec. 4.1 [102–104]. A still open issue is why the Harris criterion for the relevance or irrelevance of disorder appears to be violated in some cases [105]. Systems with dilution have also been frequently studied [106,107], and interesting excitations with a very low energy scale, that was not anticipated, have been found at the percolation point on the 2D square lattice [108].

Very recently, effects of various types of disorder at the DQCP in the J - Q model have been investigated [109], and it was found that the VBS state is replaced by a critical phase similar to the *random singlet state* that is well known in random $S = 1/2$ Heisenberg chains (and which have also been studied with SSE simulations [110,111]). SSE simulations have also been applied to the Bose-Hubbard model with site randomness [112], where interesting differences were found between weak and strong disorder.

Disordered frustrated Ising models, which often have spin-glass phases, are of interest in the field of optimization with quantum annealing. Excitation gaps have been extracted from imaginary-time correlations computed with the SSE method [113], and the generalized SSE method for imaginary-time quantum annealing has also been used to study the dynamics of the quantum phase transition between a paramagnet and a quantum spin glass [32]. Recent further developments of the SSE algorithm for highly anisotropic (near-classical) frustrated Ising models were specifically aimed at quantum annealing applications [21].

Acknowledgments

Some of the work reported here was supported by the NSF under Grant No. DMR-1710170, by a Simons Investigator Award, and by computer time allocations on the Shared Computing Cluster administered by Boston University's Research Computing Services.

References

- [1] R.P. Feynman, Phys. Rev. **91**, 1291 (1953)
- [2] M. Suzuki, Prog. Theor. Phys. **56**, 1454 (1976)
- [3] M. Suzuki, S. Miyashita, and A. Kuroda, Prog. Theor. Phys. **58**, 1377 (1977)
- [4] J.E. Hirsch, R.L. Sugar, D.J. Scalapino, and R. Blankenbecler, Phys. Rev. B **26**, 5033 (1982)
- [5] D.C. Handscomb, Proc. Cambridge Philos. Soc. **58**, 594 (1962); **60** 115 (1964)
- [6] J.W. Lyklema, Phys. Rev. Lett. **49**, 88 (1982)
- [7] S.C. Chakravarty and D.B. Stein, Phys. Rev. Lett. **49**, 582 (1982)
- [8] D.H. Lee, J.D. Joannopoulos, and J.W. Negele, Phys. Rev. B **30**, 1599 (1984)
- [9] E. Manousakis and R. Salvador, Phys. Rev. B **39**, **575** (1989)
- [10] A.W. Sandvik and J. Kurkijärvi, Phys. Rev. B **43**, 5950 (1991)
- [11] A.W. Sandvik, J. Phys. A **25**, 3667 (1992)
- [12] P. Henelius and A.W. Sandvik, Phys. Rev. B **62**, 1102 (2000)
- [13] N.V. Prokof'ev, B.V. Svistunov, and I.S. Tupitsyn, Zh. Eks. Teor. Fiz. **114**, 570 (1998) [JETP **87**, 310 (1998)]; arXiv:cond-mat/9703200.
- [14] R. Blankenbecler, D.J. Scalapino, and R.L. Sugar, Phys. Rev. B **24**, 2278 (1981)
- [15] J.E. Hirsch, Phys. Rev. B **31**, 4403 (1985)
- [16] C.-W. Liu, S. Liu, Y.-J. Kao, A.L. Chernyshev, and A.W. Sandvik, Phys. Rev. Lett. **102**, 167201 (2009)
- [17] F. Alet, K. Damle, and S. Pujari, Phys. Rev. Lett. **117**, 197203 (2016)
- [18] A. Honecker, S. Wessel, R. Kerkdyk, T. Pruschke, F. Mila, and B. Normand, Phys. Rev. B **93**, 054408 (2016)
- [19] A.W. Sandvik, R.R.P. Singh, and D.K. Campbell, Phys. Rev. B **56**, 14510 (1997)
- [20] B. Beard and U.J. Wiese, Phys. Rev. Lett. **77**, 5130 (1996)
- [21] T. Albash, G. Wagenbreth, and I. Hen, Phys. Rev. E **96**, 063309 (2017)
- [22] Y. Yan and D. Blume, J. Phys. B **50**, 223001 (2017)

- [23] A.W. Sandvik and H.G. Evertz, Phys. Rev. B **82**, 024407 (2010)
- [24] A. Sen, H. Suwa, and A.W. Sandvik, Phys. Rev. B **92**, 195145 (2015)
- [25] J. Lou and A.W. Sandvik, Phys. Rev. B **76**, 104432 (2007)
- [26] Y.-C. Lin, Y. Tang, J. Lou, and A.W. Sandvik, Phys. Rev. B **86**, 144405 (2012)
- [27] P. Weinberg and A.W. Sandvik, Phys. Rev. B **96**, 054442 (2017)
- [28] C. De Grandi, A. Polkovnikov, and A.W. Sandvik, Phys. Rev. B **84**, 224303 (2011)
- [29] C.-W. Liu, A. Polkovnikov, and A.W. Sandvik, Phys. Rev. B **87**, 174302 (2013)
- [30] S. Boixo, T.F. Rønnow, S.V. Isakov, Z. Wang, D. Wecker, D.A. Lidar, J.M. Martinis, and M. Troyer, Nat. Phys. **10**, 218 (2014)
- [31] C. De Grandi, A. Polkovnikov, and A.W. Sandvik, J. Phys.: Condens. Matter **25**, 404216 (2013)
- [32] C.-W. Liu, A. Polkovnikov, and A.W. Sandvik, Phys. Rev. Lett. **114**, 147203 (2015)
- [33] A.W. Sandvik, Phys. Rev. E **68**, 056701 (2003)
- [34] R.H. Swendsen and J.S. Wang, Phys. Rev. Lett. **58**, 86 (1987)
- [35] U. Wolff, Phys. Rev. Lett. **62**, 361 (1989)
- [36] H.G. Evertz, G. Lana, and M. Marcu, Phys. Rev. Lett. **70**, 875 (1993)
- [37] H.-P. Ying, U.-J. Wiese, and D.-R. Jia, Phys. Lett. A **183**, 441 (1993)
- [38] N. Kawashima and J.E. Gubernatis, Phys. Rev. Lett. **73**, 1295 (1994)
- [39] A.W. Sandvik, Phys. Rev. B **59**, 14157 (1999)
- [40] H.G. Evertz, Adv. Phys. **52**, 1 (2003)
- [41] A.W. Sandvik, AIP Conf. Proc. **1297**, 135 (2010)
- [42] O.F. Syljuåsen and A.W. Sandvik, Phys. Rev. E **66**, 046701 (2002)
- [43] J.E. Gubernatis, N. Kawashima, and P. Werner: *Quantum Monte Carlo Methods: Algorithms for Lattice Models* (Cambridge University Press, 2016)
- [44] Z. Yan, Y. Wu, C. Liu, O.F. Syljuåsen, J. Lou, and Y. Chen, Phys. Rev. B **99**, 165135 (2019)
- [45] R.G. Melko and A.W. Sandvik, Phys. Rev. E **72**, 026702 (2005)

- [46] K. Louis and C. Gros, Phys. Rev. B **70**, 100410(R) (2004)
- [47] D. Heidarian and K. Damle, Phys. Rev. Lett. **95**, 127206 (2005)
- [48] S.V. Isakov, S. Wessel, R.G. Melko, K. Sengupta, and Yong Baek Kim, Phys. Rev. Lett. **97**, 147202 (2006)
- [49] R.G. Melko, J. Phys. Condens. Matter **19**, 145203 (2007)
- [50] S. Biswas, G. Rakala, and K. Damle, Phys. Rev. B **93**, 235103 (2016)
- [51] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, 3rd edition (Cambridge University Press, 2007)
- [52] H. Shao, Y.Q. Qin, S. Capponi, S. Chesi, Z.Y. Meng, and A.W. Sandvik, Phys. Rev. X **7**, 041072 (2017)
- [53] E.L. Pollock and D.M. Ceperley, Phys. Rev. B **36**, 8343 (1987)
- [54] A.W. Sandvik, Phys. Rev. B **56**, 11678 (1997)
- [55] L. Wang, Y.-H. Liu, J. Imriska, P.N. Ma, and M. Troyer, Phys. Rev. X **5**, 031007 (2015)
- [56] D. Schwandt, F. Alet, and S. Capponi, Phys. Rev. Lett. **103**, 170501 (2009)
- [57] M.B. Hastings, I. González, A.B. Kallin, and R.G. Melko, Phys. Rev. Lett. **104**, 157201 (2010)
- [58] R.G. Melko, A.B. Kallin, and M.B. Hastings, Phys. Rev. B **82**, 100409(R) (2010)
- [59] A. Dorneich and M. Troyer, Phys. Rev. E **64**, 066701 (2001)
- [60] R. Kaul, R.G. Melko, and A.W. Sandvik, Annu. Rev. Condens. Matter Phys. **4**, 179 (2013)
- [61] J.D. Reger and A.P. Young, Phys. Rev. B **37**, 5978 (1988)
- [62] A.W. Sandvik and D.J. Scalapino, Phys. Rev. Lett. **72**, 2777 (1994)
- [63] L. Wang, K.S.D. Beach, and A.W. Sandvik, Phys. Rev. B **73**, 014431 (2006)
- [64] S. Wenzel, L. Bogacz, and W. Janke, Phys. Rev. Lett. **101**, 127202 (2008)
- [65] N. Ma, P. Weinberg, H. Shao, W. Guo, D.-X. Yao, and A.W. Sandvik, Phys. Rev. Lett. **121**, 117202 (2018)
- [66] A.V. Chubukov, S. Sachdev, and J. Ye, Phys. Rev. B **49**, 11919 (1994)
- [67] P. Merchant, B. Normand, K.W. Krämer, M. Boehm, D.F. McMorrow, and Ch. Rüegg, Nat. Phys. **10**, 373 (2014)

- [68] S. Jin and A.W. Sandvik, Phys. Rev. B **85**, 020409(R) (2012)
- [69] D.-R. Tan, C.-D. Li, and F.-J. Jiang, Phys. Rev. B **97**, 094405 (2018)
- [70] Y.Q. Qin, B. Normand, A.W. Sandvik, and Z.Y. Meng, Phys. Rev. B **92**, 214401 (2015)
- [71] A.W. Sandvik, S. Daul, R.R.P. Singh, and D.J. Scalapino, Phys. Rev. Lett. **89**, 247201 (2002)
- [72] T. Senthil, A. Vishwanath, L. Balents, S. Sachdev, and M.P.A. Fisher, Science **303**, 1490 (2004)
- [73] A.W. Sandvik, Phys. Rev. Lett. **98**, 227202 (2007)
- [74] R.G. Melko and R.K. Kaul, Phys. Rev. Lett. **100**, 017203 (2008)
- [75] A.W. Sandvik, Phys. Rev. Lett. **104**, 177201 (2010)
- [76] R.K. Kaul, Phys. Rev. B **84**, 054407 (2011)
- [77] M.S. Block, R.G. Melko, and R.K. Kaul, Phys. Rev. Lett. **111**, 137202 (2013)
- [78] S. Pujari, F. Alet, and K. Damle, Phys. Rev. B **91**, 104411 (2015)
- [79] H. Shao, W. Guo, and A.W. Sandvik, Science **352**, 213 (2016)
- [80] C. Wang, A. Nahum, M.A. Metlitski, C. Xu, and T. Senthil, Phys. Rev. X **7**, 031051 (2017)
- [81] B. Zhao, P. Weinberg, and A.W. Sandvik, Nat. Phys. **15**, 678 (2019)
- [82] J. Wildeboer, J. D’Emidio, and R.K. Kaul, arXiv:1808.04731.
- [83] M. Zayed, M. et al., Nat. Phys. **13**, 962 (2017)
- [84] J. Guo, G. Sun, B. Zhao, L. Wang, W. Hong, V.A. Sidorov, N. Ma, Q. Wu, S. Li, Z.Y. Meng, A.W. Sandvik, and L. Sun, arXiv:1904.09927
- [85] R.K. Kaul, Phys. Rev. Lett. **115**, 157202 (2015)
- [86] R.K. Kaul, Phys. Rev. B **91**, 054413 (2015)
- [87] J. D’Emidio and R.K. Kaul, Phys. Rev. Lett. **118**, 187202 (2017)
- [88] N. Desai and R.K. Kaul, arXiv:1904.09629.
- [89] M.S. Block, J. D’Emidio, and R.K. Kaul, arXiv:1906.06246.
- [90] L. Dang, S. Inglis, and R.G. Melko, Phys. Rev. B **84**, 132409 (2011)

- [91] S.V. Isakov, R.G. Melko, and M.B. Hastings, *Science* **335**, 193 (2012)
- [92] J. Becker and S. Wessel, *Phys. Rev. Lett.* **121**, 077202 (2018)
- [93] M. Jarrell and J.E. Gubernatis, *Phys. Rep.* **269**, 133 (1996)
- [94] A.W. Sandvik, *Phys. Rev. E* **94**, 063308 (2016)
- [95] T. Ying, K.P. Schmidt, and S. Wessel, *Phys. Rev. Lett.* **122**, 127201 (2019)
- [96] N. Ma, G.-Y. Sun, Y.-Z. You, C. Xu, A. Vishwanath, A.W. Sandvik, and Z.Y. Meng, *Phys. Rev. B* **98**, 174421 (2018)
- [97] M. Lohöfer, T. Coletta, D.G. Joshi, F.F. Assaad, M. Vojta, S. Wessel, and F. Mila, *Phys. Rev. B* **92**, 245137 (2015)
- [98] M. Lohöfer and S. Wessel, *Phys. Rev. Lett.* **118**, 147206 (2017)
- [99] Y.Q. Qin, B. Normand, A.W. Sandvik, and Z.Y. Meng, *Phys. Rev. Lett.* **118**, 147207 (2017)
- [100] C.-J. Huang, Y. Deng, Y. Wan, Z.Y. Meng, *Phys. Rev. Lett.* **120**, 167202 (2018)
- [101] N. Laflorencie, S. Wessel, A. Läuchli, and H. Rieger, *Phys. Rev. B* **73**, 060403(R) (2006)
- [102] R. Yu, T. Roscilde, and S. Haas, *Phys. Rev. Lett.* **94**, 197204 (2005)
- [103] A.W. Sandvik, *Phys. Rev. Lett.* **90**, 207201 (2006)
- [104] D.-R. Tan and F.-J. Jiang, *Phys. Rev. B* **95**, 054435 (2017)
- [105] N. Ma, A.W. Sandvik, and D.-X. Yao, *Phys. Rev. B* **90**, 104425 (2014)
- [106] A.W. Sandvik, *Phys. Rev. Lett.* **89**, 177201 (2002)
- [107] A.W. Sandvik, *Phys. Rev. B* **66**, 024418 (2002)
- [108] L. Wang and A.W. Sandvik, *Phys. Rev. B* **81**, 054417 (2010)
- [109] L. Liu, H. Shao, Y.-C. Lin, W. Guo, and A.W. Sandvik, *Phys. Rev. X* **8**, 041040 (2018)
- [110] Y.-R. Shu, D.-X. Yao, C.-W. Ke, Y.-C. Lin, and A.W. Sandvik, *Phys. Rev. B* **94**, 174442 (2016)
- [111] Y.-R. Shu, M. Dupont, D.-X. Yao, S. Capponi, and A.W. Sandvik, *Phys. Rev. B* **97**, 104424 (2018)
- [112] Y. Wang, W. Guo, and A.W. Sandvik, *Phys. Rev. Lett.* **114**, 105303 (2015)
- [113] E. Farhi, D. Gosset, I. Hen, A.W. Sandvik, P. Shor, A.P. Young, and F. Zamponi, *Phys. Rev. A* **86**, 052334 (2012)

17 Algebraic Methods in Many-Body Physics

Gerardo Ortiz

Department of Physics

Indiana University, Bloomington, USA

Contents

1	The many languages of nature	2
2	Algebraic approach to interacting quantum systems	3
3	Bosonic and hierarchical languages	5
3.1	Languages and dictionaries: a fundamental theorem	6
3.2	Hierarchical language	7
4	Transmutation of statistics: fermionic (anyonic) languages	11
4.1	Local and non-local transmutation of statistics	11
4.2	Transmutation of statistics in higher dimensions	14
4.3	Jordan-Wigner transformation for $S = 1$	15
5	Some applications	18
5.1	Generalized Jordan-Wigner transformations	18
5.2	Connecting seemingly unrelated phenomena: Haldane gap systems	20
5.3	Unveiling hidden symmetries and exactly-solvable points	23
5.4	Identifying order parameters	25
5.5	Hierarchical mean-field theories	30
5.6	Quantum simulations	33
6	Concluding remarks	35

1 The many languages of nature

One of the most challenging and interesting phenomena in modern condensed matter physics is the one emerging from competing interactions in strongly correlated systems. The multiplicity of distinct and exotic quantum phases observed experimentally confronts us with new paradigms that question our understanding of the fundamental organizing principles behind such emergent phenomena [1]. The notion of symmetry has shaped our current conception of nature; however, nature is also full of broken symmetries. Thus, understanding the idea of invariance and its corresponding conservation laws is as fundamental as determining the causes that prevent such harmony, and leads to more complex behavior. While group theory and geometry have been fundamental to the physics of the Twentieth Century, only recently has topology become central to the nascent field of topological quantum matter [2–4]. Our ultimate goal is to exploit and extend tools borrowed from these fields of mathematics to unveil and master those underlying organizing principles [5]. These principles may lead the way to designing new materials and devices with specific functionalities and unprecedented technological applications.

On general grounds, the nature and degree of entanglement of the quantum state, characterizing the various thermodynamic phases of matter, is at the core of such complex behavior [6]. For instance, the last three decades have witnessed the discovery of fractional charges and skyrmion excitations in quantum Hall liquids [7] (i.e., electrons confined to two space dimensions in the presence of strong external magnetic fields). Another example of current interest is provided by unconventional high-temperature superconductivity and its startling strange metallic behavior [8]. From the theoretical viewpoint the hurdle is in the presence of non-linear couplings, non-perturbative phenomena, and a panoply of competing quantum orders. These systems happen to be strongly correlated because no obvious small coupling constant exists, and consequently exhibit high sensitivity to small parameter changes. The importance of developing a methodology based on qualitatively new concepts going beyond traditional mean-field (MF) and semi-classical approximations, that treats all possible competing orders on an equal footing with no privileged fixed-point phenomenon, becomes manifest. Despite great advances, there is a lack of a systematic and reliable methodology to study and predict the behavior of these complex systems. It is the purpose of these lecture notes to present some steps in that direction. Describing the structure and behavior of matter entails studying systems of interacting quantum constituents (bosons, fermions, gauge fields, spins). In the quantum-mechanical description of matter, each physical system is naturally associated with a *language* of operators (for example, quantum spin-1/2 operators) and thus to an algebra realizing this language (e.g., the Pauli spin algebra generated by a family of commuting quantum spin-1/2 operators). It is our point of view that crucial to the successful understanding of the mechanisms driving complexity is the realization of *dictionaries* (isomorphisms) connecting the different languages of nature and therefore linking seemingly unrelated physical phenomena [9, 10]. The existence of dictionaries provides not only a common ground to explore complexity but leads naturally to the fundamental concept of *universality*, meaning that different physical systems show the same behavior. In this way, there is a concept of physical equivalence hidden in these dictionaries.

On the other hand, the notion of algebra and its homomorphisms have also been essential to unravel hidden structures in theoretical physics: Internal symmetries which are hidden in a given algebraic representation of a model become manifest in another representation. In 1928 Jordan and Wigner [11] made a first step relating quantum spin $S = 1/2$ degrees of freedom to particles with fermion statistics. A simple application of their contribution is the mapping between the isotropic XY chain describing quantum magnets and the tight-binding spinless fermion model, which can be exactly solved in one spatial dimension. From the group theoretical viewpoint, an internal $U(1)$ continuous symmetry (related to particle number conservation), for instance, is evidenced in the fermion representation of the XY model which was hidden in the spin representation. Overall, what Jordan and Wigner established was an isomorphism of $*$ -algebras, i.e., an isomorphism between the Pauli and fermion algebras [12].

In this chapter we present a unifying algebraic framework for interacting quantum systems. We show that exact algebraic and group theoretic methods are one of the most elegant and promising approaches towards a complete understanding of quantum phases of matter and their corresponding phase transitions. Can we connect the different (spin-particle-gauge) languages of nature within a single algebraic framework? We will present a fundamental theorem which connects operators generating different algebras (e.g., $su(D)$ spin-particle connections), unifying the different languages known so far in the quantum-mechanical description of matter. We will illustrate the methodology with various examples borrowed from strongly correlated physics, including quantum magnets and superfluids. Applications aim at:

- Illustrating connections between seemingly unrelated physical phenomena
- Unveiling hidden symmetries to identify new states of matter
- Identifying order parameters (OPs) in phase transitions [13, 14]
- Establishing exact (algebraic) solutions and developing better approximation schemes
- Explaining the use of languages in quantum simulations

The chapter has been written with the intention of providing the reader with the most fundamental concepts involved in our algebraic framework and how they apply to study complex phenomena. Much more details and examples can be found in the original manuscripts [9, 10, 12, 15].

2 Algebraic approach to interacting quantum systems

The theory of operator algebras on Hilbert spaces was initiated by Murray and von Neumann [16] as a tool to study unitary representations of groups, and as a framework for a reformulation of quantum mechanics. This area of research continued its development independently in the realm of mathematical physics, and therefore knowledge of those investigations remained bounded to specialists. For use of C^* and W^* algebras as a framework for quantum statistical mechanics one can look at the books of Bratteli and Robinson [17]. For the purposes of our presentation one only needs to have an elementary background in basic algebra [18], and specially group theory [19], in particular, Lie algebras and groups.

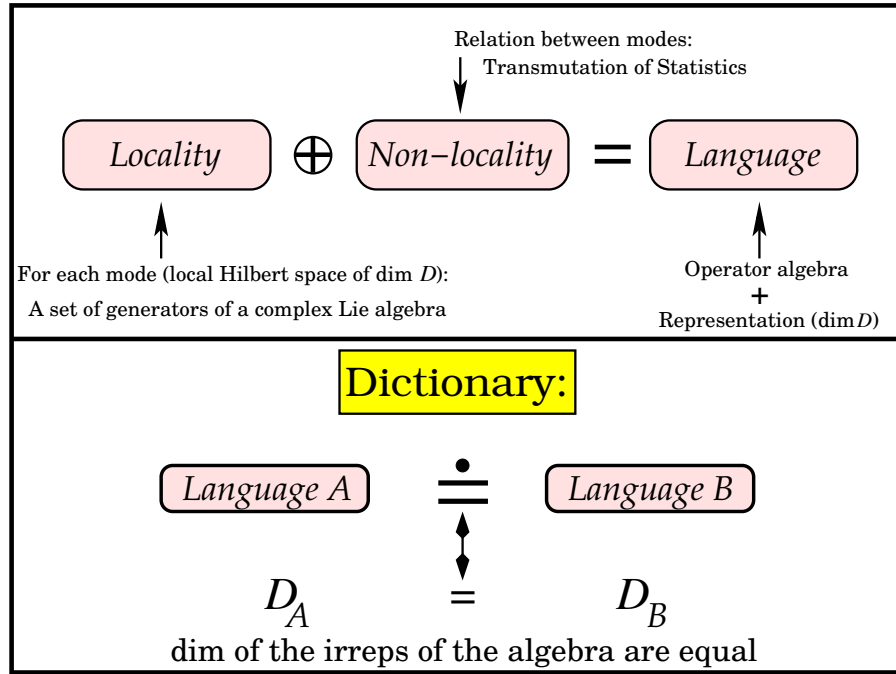


Fig. 1: Definition of a language and fundamental theorem [12, 9] behind the construction of the dictionaries of nature. In the upper panel we show schematically what elements define a language $= \mathcal{A} \wedge \Gamma_A$, where $\mathcal{A}$ is the algebra and Γ_A a particular irrep. In the lower panel we establish the criteria to build a dictionary, given two languages A and B . This criteria is based upon Burnside's theorem of algebra [18].

Here we are concerned with quantum lattice systems. A quantum lattice is identified with $\mathbb{Z}^{N_s}$, where N_s is the total number of lattice sites (or modes). Associated to each lattice site $\mathbf{i} \in \mathbb{Z}^{N_s}$ there is a Hilbert space $\mathcal{H}_{\mathbf{i}}$ of finite dimension D describing the “local” modes. The total Hilbert space is $\mathcal{H} = \bigotimes_{\mathbf{i}} \mathcal{H}_{\mathbf{i}}$ (we will also consider its symmetric and antisymmetric subspaces). Since we are mostly interested in zero temperature properties, a state of the system is simply a vector $|\Psi\rangle$ in $\mathcal{H}$, and an observable is a self-adjoint operator $\mathcal{O} : \mathcal{H} \rightarrow \mathcal{H}$. The dynamical evolution of the system is determined by its Hamiltonian H . The topology of the lattice, dictated by the connectivity and range of the interactions in H , is an important element in establishing complexity. In the case of quantum continuous systems we can still use the present formalism after discretizing space. Going beyond this approach is outside the scope of these notes.

As mentioned above, each physical system is naturally associated with a language of operators, and thus to an algebra realizing this language. Formally, a *language* is defined by an operator algebra and a specific representation of the algebra. We use the following notation: $language = \mathcal{A} \wedge \Gamma_A$, where $\mathcal{A}$ is the operator algebra and Γ_A is a particular irreducible representation (irrep) of the local algebra $\mathcal{A}_{\mathbf{j}}$ associated to $\mathcal{A}$, of dimension $\dim \Gamma_A = D$ (see Fig. 1).

For the sake of clarity, let us choose the phenomenon of magnetism to illustrate the key ideas. This is one of the most intriguing and not fully understood problems in condensed matter physics where strong correlations between electrons (of electrostatic origin) are believed to be the essence of the problem. To describe the phenomenon using a minimal model (i.e., a

model that only includes the relevant degrees of freedom) distinct approaches can be advocated depending upon the itineracy of the electrons that participate in the magnetic processes. In one extreme (e.g., insulators) a description in terms of localized quantum spins is appropriate, while in the other (e.g., metals) delocalization of the electrons is decisive and cannot be ignored. We immediately identify the languages associated to each description: quantum spins (e.g., Pauli algebra) and fermions (spin-1/2 Fermi algebra). Are these really different descriptions? Is there a dictionary that may connect the two languages? Let's assume that we decide to use the quantum spins language. What other seemingly unrelated phenomena are connected to magnetism? Can we relate phases of matter corresponding to dissimilar phenomena? Can an arbitrary physical system be mapped, for instance, onto a pure magnetic system (an array of quantum spins)? In the following we will answer these questions by examples. As mentioned above, a fundamental concept of universality, complementary to the one used in critical phenomena, emerges as a consequence of unveiling the hidden unity in the quantum-mechanical description of matter.

3 Bosonic and hierarchical languages

A *bosonic language* is a set of operators grouped in subsets S_i (associated to each mode) that satisfy the conditions

- Each element b_i^μ of S_i ($\mu \in [1, N_g]$) belongs to the algebra of endomorphisms for the vector space $\mathcal{H}_i$ over the field of complex numbers $\mathbb{C}$, $b_i^\mu : \mathcal{H}_i \rightarrow \mathcal{H}_i$, and these elements are linearly independent.
- The elements of S_i generate a monoid [20] of linear transformations under the associative product in the algebra which acts irreducibly on $\mathcal{H}_i$ in the sense that the only subspaces stabilized by S_i are $\mathcal{H}_i$ and 0 (0 is the null vector).
- If b_i^μ and b_j^ν are elements of different subsets S_i and S_j , then $b_i^\mu b_j^\nu = \square(b_i^\mu, b_j^\nu) = \square(b_j^\nu, b_i^\mu)$.

Combining the associative product $\square$ and the additive operations, we can define the non-associative Lie product $[,]$, which is called commutator

$$[b_i^\mu, b_j^\nu] = b_i^\mu b_j^\nu - b_j^\nu b_i^\mu. \quad (1)$$

Using this product the last condition can be reformulated

$$[b_i^\mu, b_j^\nu] = 0, \quad \text{if } i \neq j. \quad (2)$$

The set S_i is not necessarily closed under the regular product (composition) or the Lie product (commutator). If the set S_i is closed under the Lie product

$$[b_i^\mu, b_i^\nu] = \sum_{\mu'=1}^{N_g} \lambda_{\mu'} b_i^{\mu'}, \quad \lambda_{\mu'} \in \mathbb{C}, \quad (3)$$

the elements of $\mathcal{S}_i$ generate a Lie algebra $\mathcal{S}_i$. In addition, since each generator is represented by an endomorphism of $\mathcal{H}_i$ there is a particular representation $\Gamma_{\mathcal{S}}$ of $\mathcal{S}_i$ associated to the bosonic language. The second condition for a bosonic language implies that $\Gamma_{\mathcal{S}}$ is irreducible. The third condition implies that the global Lie algebra associated to the complete set of generators is the direct sum of local algebras $\mathcal{S}_i$, $\mathcal{S} = \bigoplus_i \mathcal{S}_i$. Therefore, if the set $\mathcal{S}_i$ is closed under the Lie product, we can represent the bosonic language by the conjunction of the Lie algebra $\mathcal{S}$, and the irreducible representation $\Gamma_{\mathcal{S}}$: $\mathcal{S} \wedge \Gamma_{\mathcal{S}}$. The dimension of $\Gamma_{\mathcal{S}}$ is equal to the dimension of the local Hilbert space $\mathcal{H}_i$: $\dim \Gamma_{\mathcal{S}} = D$.

3.1 Languages and dictionaries: a fundamental theorem

The demonstration of the fundamental theorem of this section is a direct consequence of the classical theorem of Burnside [18]

Burnside's theorem. Let G be a monoid of linear transformations in a finite dimensional vector space $\mathbf{V}$ over an algebraically closed field F (in quantum mechanics, the complex numbers $\mathbb{C}$), that acts irreducibly on $\mathbf{V}$ in the sense that the only subspaces stabilized by G are $\mathbf{V}$ and 0 . Then G contains a base for $\text{End}_F \mathbf{V}$ (ring of endomorphisms of $\mathbf{V}$ over F).

The following theorem shows that two languages are *equivalent* if they have in common the dimension D of their local Hilbert space $\mathcal{H}_i$, and D is finite.

Fundamental Theorem [12, 9]: Given two bosonic languages having the same finite dimension D of their local Hilbert spaces $\mathcal{H}_i$, the generators of one of them can be written as a polynomial function of the generators of the other language and vice versa.

This theorem provides the necessary and sufficient conditions to connect two bosonic languages. What is the unifying concept behind the construction of the dictionaries (isomorphisms) of nature? When is it possible to build a dictionary between two arbitrary (bosonic, fermionic or anyonic) languages? The answers lie in the application of the fundamental theorem together with the transmutation of statistics [9] (see Fig. 1). This will be done in Section 4.

The fundamental theorem establishes the existence of dictionaries connecting languages within the same class. As a consequence, we can use any bosonic language in the class to describe a given physical phenomena. The natural question which emerges from this result is: What is the most appropriate language in a given class for describing our particular problem? There is no generic answer to this question. Nonetheless, the following two corollaries give an important hint for problems which are invariant under particular transformations, because they relate the notion of language to the generators of symmetry groups.

- **Corollary I:** In each class of bosonic languages there is at least one which is the conjunction of a Lie algebra $\mathcal{S}$ and an irreducible representation $\Gamma_{\mathcal{S}}$ ($\mathcal{S} \wedge \Gamma_{\mathcal{S}}$), i.e., the generators of the bosonic language are generators of the Lie algebra $\mathcal{S}_i$ in the rep. $\Gamma_{\mathcal{S}}$.

Consider the Lie algebra $\mathcal{L}_i = u(1) \oplus su(2)$ ($\mathcal{S} = \bigoplus_i \mathcal{L}_i$) whose generators are $\{I_i, S_i^x, S_i^y, S_i^z\}$. The three basis elements S_i^α (linear and Hermitian operators) of the Lie algebra $su(2)$ for each

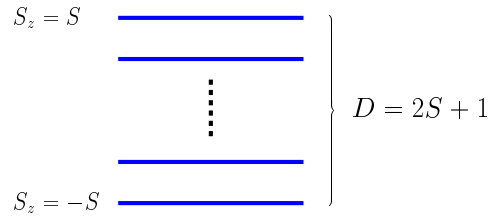


Fig. 2: Spin S representation of $su(2)$ of dimensionality $D = 2S + 1$.

lattice site (mode) $\mathbf{i}$ ($\mathbf{i} = 1, \dots, N_s$) satisfy the equal-time commutation relations

$$[S_{\mathbf{i}}^a, S_{\mathbf{j}}^b] = i\delta_{\mathbf{ij}} \epsilon_{abc} S_{\mathbf{i}}^c, \quad a, b, c = x, y, z, \quad (4)$$

with ϵ the totally antisymmetric Levi-Civita symbol. Equivalently, in terms of the ladder operators $S_{\mathbf{j}}^{\pm} = S_{\mathbf{j}}^x \pm iS_{\mathbf{j}}^y$

$$[S_{\mathbf{j}}^+, S_{\mathbf{j}}^-] = 2S_{\mathbf{j}}^z, \quad [S_{\mathbf{j}}^z, S_{\mathbf{j}}^{\pm}] = \pm S_{\mathbf{j}}^{\pm}, \quad \{S_{\mathbf{j}}^+, S_{\mathbf{j}}^-\} = 2(S(S+1) - (S_{\mathbf{j}}^z)^2), \quad (5)$$

where $D = 2S + 1$ is the dimension of the irreducible representation Γ_S^D (Fig. 2). The demonstration of corollary I shows that any bosonic problem with a local Hilbert space of dimension D can be described with $su(2)$ -spins of magnitude (representation) $S = (D - 1)/2$.

3.2 Hierarchical language

A given bosonic language will be called *hierarchical* if any local physical operator $\hat{\mathcal{O}}_{\mathbf{i}} : \mathcal{H}_{\mathbf{i}} \rightarrow \mathcal{H}_{\mathbf{i}}$, can be written as a linear combination of the generators of the language

$$\hat{\mathcal{O}}_{\mathbf{i}} = \sum_{\mu=1}^{N_g} \lambda_{\mu} \mathfrak{b}_{\mathbf{i}}^{\mu}, \quad \lambda_{\mu} \in \mathbb{C}. \quad (6)$$

- **Corollary II:** In each class of bosonic languages there is at least one which is hierarchical, and its generators are the identity and the generators of $su(N=D)$ in the fundamental representation.

Proof: The Lie algebra generated by the identity $I_{\mathbf{i}}$ and the generators of $su(N)$ is $\mathcal{L}_{\mathbf{i}} = u(1) \oplus su(N)$. Since $\dim \mathcal{L}_{\mathbf{i}} = N^2$, the generators of $\mathcal{L}_{\mathbf{i}}$ are also a base for $\text{End}_{\mathbb{C}} \mathcal{H}_{\mathbf{i}}$ if $D = \dim \mathcal{H}_{\mathbf{i}} = N$.

A consequence of corollary II is that the generators of any language can be expressed as a linear combination of generators of a hierarchical language (HL) in the same class. The most trivial example is given by the class of bosonic languages with $D = 2$. The generators of any language (any two level system) in the same class can be written as a linear combination of the identity and the Pauli matrices. We will see later that corollary II is the key to get a hierarchical classification of the possible broken symmetries of a given physical Hamiltonian.

3.2.1 Example: different classes of bosonic particles

Canonical bosons with N_f different flavors α satisfy the standard canonical commutation relations

$$\begin{cases} [b_{i\alpha}, b_{j\beta}] = [b_{i\alpha}^\dagger, b_{j\beta}^\dagger] = 0, \\ [b_{i\beta}, b_{j\alpha}^\dagger] = \delta_{ij} \delta_{\alpha\beta}, [n_{i\alpha}, b_{j\beta}^\dagger] = \delta_{ij} \delta_{\alpha\beta} b_{i\alpha}^\dagger. \end{cases} \quad (7)$$

The local Hilbert space is infinite dimensional. Consider now two additional examples that illustrate in detail the contents of the fundamental theorem and the subsequent corollaries. The first example corresponds to hard-core (HC) bosons with N_f different flavors α . Since they are HC, only single occupancy is allowed, i.e., the eigenvalues of $\bar{n}_i = \sum_\alpha \bar{n}_{i\alpha}$ are either 0 or 1 ($(\bar{b}_{i\alpha}^\dagger)^2 = 0$, and $\bar{n}_{i\alpha} = \bar{b}_{i\alpha}^\dagger \bar{b}_{i\alpha}$ is the number operator for the flavor α at the site i). The minimal set S'_i of operators that we can use to generate a bosonic language which is appropriate for HC bosons is: $S'_i = \{\mathbf{1}_i, \bar{b}_{i\alpha}^\dagger, \bar{b}_{i\alpha}\}$ with $1 \leq \alpha \leq N_f$. It can be shown that this set satisfies the three requirements for a bosonic language. The dimension of the local Hilbert space for these endomorphisms is $D = N_f + 1$. Then by varying the total number of flavors we can generate all possible values of D . Since each class of bosonic languages is characterized by the value of D , these HC bosons provide an example of a bosonic language in each class. It is clear that the set S'_i is not closed under the Lie product. Therefore, we cannot associate a Lie algebra to this minimal bosonic language. However, if we increase the number of generators in the set S'_i by including bilinear forms of the type $\bar{b}_{i\alpha}^\dagger \bar{b}_{i\beta}$, then the new set $S_i = \{\mathbf{1}_i, \bar{b}_{i\alpha}^\dagger, \bar{b}_{i\alpha}, \bar{b}_{i\alpha}^\dagger \bar{b}_{i\beta}\}$, with $1 \leq \alpha, \beta \leq N_f$, becomes closed under the Lie product

$$\begin{cases} [\bar{b}_{i\alpha}, \bar{b}_{j\beta}] = [\bar{b}_{i\alpha}^\dagger, \bar{b}_{j\beta}^\dagger] = 0, \\ [\bar{b}_{i\beta}, \bar{b}_{j\alpha}^\dagger] = \delta_{ij}(\delta_{\alpha\beta} - \bar{n}_i \delta_{\alpha\beta} - \bar{b}_{i\alpha}^\dagger \bar{b}_{i\beta}), \\ [\bar{b}_{i\alpha}^\dagger \bar{b}_{i\beta}, \bar{b}_{j\gamma}^\dagger] = \delta_{ij} \delta_{\beta\gamma} \bar{b}_{i\alpha}^\dagger. \end{cases} \quad (8)$$

This means that the extended set S_i is now a set of generators for a Lie algebra in a particular representation. From the commutation relations (Eq. (8)) we can conclude that S_i is the direct sum of an $u(1)$ algebra, generated by the identity $\mathbf{1}_i$, and an $su(N)$ ($N = D = N_f + 1$) algebra generated by $\{\bar{b}_{i\alpha}^\dagger, \bar{b}_{i\alpha}, \bar{b}_{i\alpha}^\dagger \bar{b}_{i\beta}\}$: $S_i = u(1) \oplus su(N)$. The representation Γ_S is the fundamental representation of $su(N)$ ($\dim \Gamma_S = N$). Therefore, the new language is a hierarchical one.

We will see in later sections that *canonical fermions* can be transformed continuously into bosons which are HC in each flavor

$$\begin{cases} [\tilde{b}_{i\alpha}, \tilde{b}_{j\beta}] = [\tilde{b}_{i\alpha}^\dagger, \tilde{b}_{j\beta}^\dagger] = 0, \\ [\tilde{b}_{i\alpha}, \tilde{b}_{j\beta}^\dagger] = \delta_{ij} \delta_{\alpha\beta} (1 - 2\tilde{n}_\alpha), [\tilde{n}_{i\alpha}, \tilde{b}_{j\beta}^\dagger] = \delta_{ij} \delta_{\alpha\beta} \tilde{b}_{i\alpha}^\dagger, \end{cases} \quad (9)$$

which implies $\{\tilde{b}_{i\alpha}, \tilde{b}_{i\alpha}^\dagger\} = 1$. The Lie algebra generated by these bosonic operators is $\mathcal{L} = \bigoplus_{\alpha,i} su(2)$, i.e. each set $\{\tilde{b}_{i\alpha}^\dagger, \tilde{b}_{i\alpha}, \tilde{n}_{i\alpha} - 1/2\}$ generates an $su(2)$ algebra.

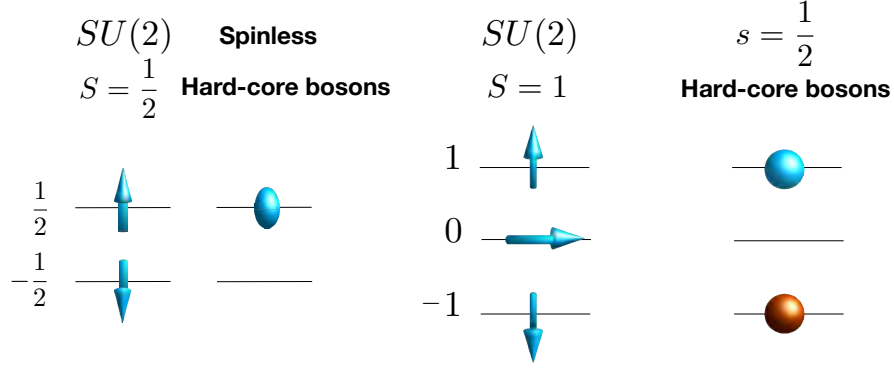


Fig. 3: Local Hilbert spaces $\mathcal{H}_i$, for the spin and HC bosonic worlds, at site (mode) i : (a) $S = \frac{1}{2}$ (dimension $D = 2$), (b) $S = 1$ (dimension $D = 3$).

3.2.2 Example: Matsubara-Matsuda transformation and its generalization to $S = 1$

To explain the peculiar properties of liquid Helium II, Matsubara and Matsuda [21] introduced a lattice model of atoms and holes. Think of a Helium-4 atom as a spinless HC boson ($N_f = 1$). On each lattice site, one can accommodate at most one Helium atom ($(\bar{b}_i^\dagger)^2 = 0$). Then, the simplest model of a liquid one can think of is described by a Hamiltonian with a kinetic energy term, with hopping amplitude t , and a nearest-neighbor density-density interaction of magnitude V . The expression for the Hamiltonian in terms of the first language defined by the sets S_i is

$$H_{\text{xxz}} = t \sum_{\langle i,j \rangle} (\bar{b}_i^\dagger \bar{b}_j + \bar{b}_j^\dagger \bar{b}_i) + V \sum_{\langle i,j \rangle} (\bar{n}_i - \frac{1}{2}) (\bar{n}_j - \frac{1}{2}), \quad (10)$$

where $\langle i,j \rangle$, refers to nearest-neighbors in an otherwise regular d -dimensional lattice. Since $\bar{b}_i^\dagger$, creating an atom at lattice point i , and $\bar{b}_i$, annihilating an atom at lattice point i , are not generators of a Lie algebra, the eventual global symmetries of H_{xxz} remain hidden in this particular language. However, if we *translate* H_{xxz} to the second $SU(2)$ -spin language using the dictionary provided by Matsubara and Matsuda [21]

$$S_i^+ = \bar{b}_i^\dagger, \quad S_i^- = \bar{b}_i, \quad S_i^z = \bar{b}_i^\dagger \bar{b}_i - \frac{1}{2} = \bar{n}_i - \frac{1}{2}, \quad (11)$$

we can immediately unveil the hidden symmetries of H_{xxz} . This dictionary maps single-site (mode) states in the following way (see Fig. 3)

$$|-\frac{1}{2}\rangle \longleftrightarrow |0\rangle, \quad |\frac{1}{2}\rangle \longleftrightarrow \bar{b}_i^\dagger |0\rangle, \quad (12)$$

where $|0\rangle$ is the vacuum of HC bosons. The well-known expression for H_{xxz} in terms of the $su(2)$ generators (i.e., the equivalent spin Hamiltonian) is

$$H_{\text{xxz}} = \sum_{\langle i,j \rangle} J_z S_i^z S_j^z + \frac{J_\perp}{2} (S_i^+ S_j^- + S_i^- S_j^+). \quad (13)$$

The magnetic couplings, J_z and $J_\perp$, are related to the original parameters, t and V , by the relations: $J_z = V$ and $J_\perp = 2t$. It is clear from Eq. (13) that the original model has a global

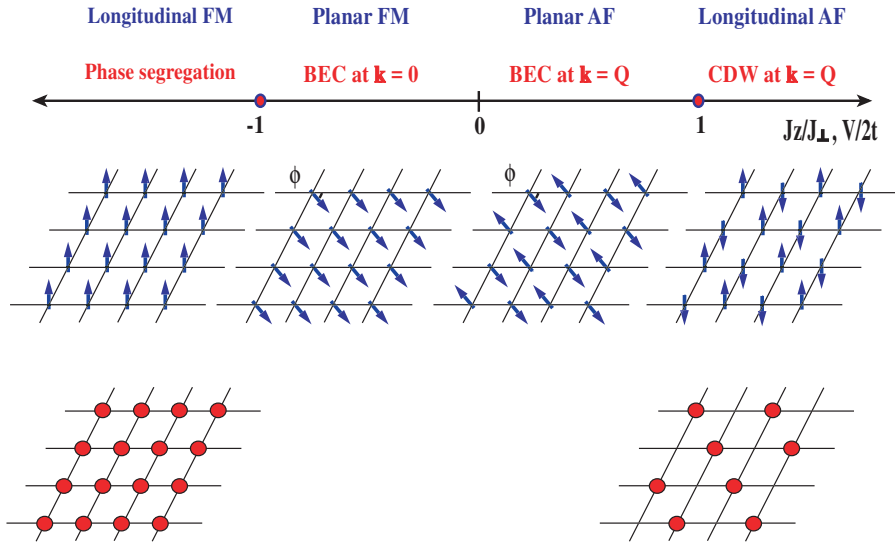


Fig. 4: Quantum phase diagram of the H_{xxz} model Hamiltonian. $\mathbf{Q}$ stands for the AF wave vector, and FM refers to the ferromagnetic phase.

$SU(2)$ invariance if $V = 2t$, i.e., it is in the isotropic Heisenberg point. The existence of this $SU(2)$ -symmetric point has a very important consequence for the quantum phase diagram (Fig. 4) of the bosonic model of Eq. (10): If there is a charge density wave (CDW) instability at that point, the $SU(2)$ invariance implies that *there is also* a Bose-Einstein (BE) condensation and vice versa. The order parameters (OPs) of both phases are different components of a unique OP in the spin language, i.e., the staggered magnetization of the antiferromagnetic (AF) phase ($t > 0$). The z -component of the staggered magnetization is mapped onto the CDW OP for the bosonic gas, while the transverse component is equivalent to the OP for the BE condensation. Only one of these two phases, which are coexisting at the $SU(2)$ invariant point, is stable when we depart from the symmetric point in any of both directions in parameter space (BE condensation if $V < 2t$ and CDW if $V > 2t$). In this very simple example we can see the advantages of using a HL ($su(2)$ in this case). In the first place, we can immediately recognize the high symmetry points. Secondly, we can describe an eventual broken symmetry state at those points in terms of a unified OP. If we were to use a non-HL to describe the same problem, we would find coexistence of more than one phase at the high symmetry points. The OPs of each of those phases are different components of the unified OP that we simply found with the HL. These ideas are developed in more detail in Section 5.4.

One can generalize [12] this idea to the case of HC bosons with two flavors, i.e., $N_f = 2$, that we can call spin $\sigma = \uparrow, \downarrow$, with the property that $\bar{b}_{i\sigma}^\dagger \bar{b}_{i\sigma'}^\dagger = 0$. In this case, the dictionary (isomorphic map) is given by (see Fig. 3, where $\uparrow(\downarrow)$ is represented by a blue(red) ball)

$$S_i^+ = \sqrt{2} (\bar{b}_{i\uparrow}^\dagger + \bar{b}_{i\downarrow}^\dagger), \quad S_i^- = \sqrt{2} (\bar{b}_{i\uparrow} + \bar{b}_{i\downarrow}), \quad S_i^z = \bar{n}_{i\uparrow} - \bar{n}_{i\downarrow}, \quad (14)$$

and single-site (mode) states map as

$$|-1\rangle \longleftrightarrow \bar{b}_{i\downarrow}^\dagger |0\rangle, \quad |0\rangle \longleftrightarrow |0\rangle, \quad |1\rangle \longleftrightarrow \bar{b}_{i\uparrow}^\dagger |0\rangle. \quad (15)$$

We will utilize this dictionary in later sections.

4 Transmutation of statistics: fermionic (anyonic) languages

We know that bosonic languages do not exhaust all possible languages of quantum mechanics. We have seen that the notion of bosonic languages is closely related to the concept of Lie algebras, since for each class of bosonic languages there is at least one language whose elements are generators of a Lie algebra. However, the same cannot be done for fermions. This can be easily seen by noticing that the condition for a bosonic language, Eq. (2), is not valid for fermions. In addition, the main consequence of Eq. (2) for a bosonic language is that the global Lie algebra is the direct sum of the local ones associated to each subset S_i . The generalization of these concepts to fermionic languages can be done by introducing the notion of Lie superalgebras (see for instance Ref. [19]), albeit we will not need this mathematical construct. The fermionic languages are associated to Lie superalgebras in the same way the bosonic languages are associated to Lie algebras. We will only consider the *fermionic language* generated by the *canonical* creation and annihilation operators

$$\begin{cases} \{c_{i\alpha}, c_{j\beta}\} = \{c_{i\alpha}^\dagger, c_{j\beta}^\dagger\} = 0, \\ \{c_{i\alpha}, c_{j\beta}^\dagger\} = \delta_{ij} \delta_{\alpha\beta}, \end{cases} \quad (16)$$

and other languages obtained by imposing particular local constraints on the canonical fermions. These generators, together with the identity, generate the Heisenberg Lie superalgebra. In analogy to the bosonic languages (see Eq. (2)), the Lie product (in the superalgebra) of two elements acting in different sites (different subsets S_i, S_j) is zero. Thus, instead of having a direct sum of local Lie algebras as in the bosonic case, we have a direct sum of local Lie superalgebras. In the case of canonical fermions the local Lie superalgebras are Heisenberg superalgebras.

4.1 Local and non-local transmutation of statistics

We will start by considering multiframe fermions $c_{i\alpha}^\dagger$ ($\alpha \in [1, N_f]$) which satisfy the canonical anticommutation relations (Eq. (16)). Other types of fermions, usually considered in physics, can be derived from the canonical ones by imposing particular constraints.

The canonical fermions can be transformed into bosons $\tilde{b}_{i\alpha}^\dagger$ which are hard-core in each flavor (the eigenvalues of $\tilde{n}_{i\alpha} = \tilde{b}_{i\alpha}^\dagger \tilde{b}_{i\alpha}$ are either 0 or 1, see Eq. (9)), i.e., two or more bosons of the same flavor are not allowed on the same site, but two of different flavors are permitted. In previous sections we have shown that a physical theory for objects obeying commutation relations (Lie brackets) can be formulated in terms of a bosonic language. By the present connection we will be able to extend this statement to fermions (or anyons, in general) through a transmutation of statistics. The local anticommutation relations are transmuted into anti-commutation relations (and vice versa) when the creation and annihilation operators are multiplied by the “local transmutator” $\hat{\mathcal{T}}_{j\alpha}^\dagger$

$$c_{j\alpha}^\dagger = \tilde{b}_{j\alpha}^\dagger \hat{\mathcal{T}}_{j\alpha}^\dagger, \quad (17)$$

where $\hat{\mathcal{T}}_{j\alpha} = \exp(i\pi \sum_{\beta < \alpha} \tilde{n}_{j\beta})$ is the “local transmutator,” and we are assuming a particular

ordering for the flavor indices α . From the expression for $\hat{\mathcal{T}}_{j\alpha}$ it is clear that

$$\hat{\mathcal{T}}_{j\alpha}^2 = I, \quad \hat{\mathcal{T}}_{j\alpha}^\dagger = \hat{\mathcal{T}}_{j\alpha}. \quad (18)$$

In this way we have established a mapping between fermions and bosons which are operating locally (on a given orbital or mode j)

$$\mathcal{S}_j = \{\tilde{b}_{j\alpha}^\dagger, \tilde{b}_{j\alpha}, \tilde{n}_{j\alpha} - \frac{1}{2}\} \longleftrightarrow \hat{\mathcal{S}}_j = \{c_{j\alpha}^\dagger, c_{j\alpha}, \hat{n}_{j\alpha} - \frac{1}{2}\} \quad (19)$$

So far, we have only transmuted the commutation relations between generators which belong to the same site or subset $\mathcal{S}_i$. For commutation relations of two generators at different sites we need to introduce a non-local operator K_j . For spinless fermions, in one spatial dimension, Jordan and Wigner [11] introduced the so-called kink-operator

$$K_j^{1d} = \exp\left(i\pi \sum_{l < j} \hat{n}_l\right), \quad (20)$$

to establish a map between quantum $S = 1/2$ spins and spinless fermions

$$S_j^+ = c_j^\dagger K_j, \quad S_j^- = K_j^\dagger c_j, \quad S_j^z = \hat{n}_j - \frac{1}{2}, \quad (21)$$

where $K_j = K_j^{1d}$ for a one-dimensional lattice. The generalization of K_j to multiflavored canonical fermions is straightforward

$$K_j = \exp\left(i\pi \sum_{l < j, \alpha} \hat{n}_{l\alpha}\right). \quad (22)$$

The complete transmutator for canonical fermions $\mathcal{K}_{j\alpha}^\dagger$ is the product of the local and the non-local components

$$c_{j\alpha}^\dagger = \tilde{b}_{j\alpha}^\dagger \hat{\mathcal{T}}_{j\alpha}^\dagger K_j^\dagger = \tilde{b}_{j\alpha}^\dagger \mathcal{K}_{j\alpha}^\dagger. \quad (23)$$

Similarly, one can extend this idea of transmutation of statistics to particles satisfying general equal-time anyonic canonical commutation relations defined by an angle θ . To this end we need to generalize the transmutators to any statistical angle $0 \leq \theta \leq \pi$. By replacing the phase π by θ , $\mathcal{K}_{j\alpha}^\dagger(\theta)$ transmutes the bosons (or fermions) into (Abelian) anyons

$$\tilde{a}_{j\alpha}^\dagger = \tilde{b}_{j\alpha}^\dagger \hat{\mathcal{T}}_{j\alpha}^\dagger(\theta) K_j^\dagger(\theta) = \tilde{b}_{j\alpha}^\dagger \mathcal{K}_{j\alpha}^\dagger(\theta). \quad (24)$$

Like in the previous example, the local commutation relations are preserved ($\tilde{n}_{j\alpha} = \tilde{a}_{j\alpha}^\dagger \tilde{a}_{j\alpha}$, $\tilde{n}_j = \sum_{\alpha=1}^{N_f} \tilde{n}_{j\alpha}$)

$$\begin{cases} [\tilde{a}_{j\alpha}, \tilde{a}_{j\alpha}] = [\tilde{a}_{j\alpha}^\dagger, \tilde{a}_{j\alpha}^\dagger] = 0, \\ [\tilde{a}_{j\alpha}, \tilde{a}_{j\alpha}^\dagger] = 1 - 2\tilde{n}_{j\alpha}. \end{cases} \quad (25)$$

In this particular case, since there is a hard-core condition $\tilde{a}_{j\alpha}^\dagger \tilde{a}_{j\alpha}^\dagger = 0$, the operators also satisfy the following local anticommutation relations

$$\begin{cases} \{\tilde{a}_{j\alpha}, \tilde{a}_{j\alpha}\} = \{\tilde{a}_{j\alpha}^\dagger, \tilde{a}_{j\alpha}^\dagger\} = 0, \\ \{\tilde{a}_{j\alpha}, \tilde{a}_{j\alpha}^\dagger\} = 1. \end{cases} \quad (26)$$

Thus, the local anticommutation relations are also preserved under statistical transmutation. Clearly, Eq. (26) are the local anticommutation relations for canonical fermions. This is not surprising since the multiflavored hard-core bosons defined by Eq. (9) can be transmuted into canonical fermions (see Eq. (23)). For operators involving different indices we have to define an index ordering. For $(\mathbf{j}, \beta) > (\mathbf{i}, \alpha)$

$$\begin{cases} [\tilde{a}_{\mathbf{i}\alpha}, \tilde{a}_{\mathbf{j}\beta}]_\theta = [\tilde{a}_{\mathbf{i}\alpha}^\dagger, \tilde{a}_{\mathbf{j}\beta}^\dagger]_\theta = 0, \\ [\tilde{a}_{\mathbf{j}\beta}^\dagger, \tilde{a}_{\mathbf{i}\alpha}]_\theta = 0, \end{cases} \quad (27)$$

where $[A, B]_\theta = AB - \exp[i\theta]BA$. Even though $\mathcal{K}_{\mathbf{j}\alpha}^\dagger(\theta)$ is a non-local operator, it does not introduce long-range interactions if the model has only nearest-neighbor-range terms. The only consequence of changing the statistics of the particles is a change of the short-range interactions in the original basis. Therefore, the concept of particle statistics in one dimension becomes irrelevant since any physical system can be described with a bosonic language without changing the short-range character of the interactions.

There is another type of HC anyon that is relevant for the generalized Jordan-Wigner transformations [12] of Sections 4.3 and 5.1. They are defined as

$$\bar{a}_{\mathbf{j}\alpha}^\dagger = \bar{b}_{\mathbf{j}\alpha}^\dagger \hat{\mathcal{T}}_{\mathbf{j}\alpha}^\dagger(\theta) K_{\mathbf{j}}^\dagger(\theta) = \bar{b}_{\mathbf{j}\alpha}^\dagger \mathcal{K}_{\mathbf{j}\alpha}^\dagger(\theta), \quad (28)$$

with

$$\hat{\mathcal{T}}_{\mathbf{j}\alpha}(\theta) = \exp\left(i\theta \sum_{\beta < \alpha} \bar{n}_{\mathbf{j}\beta}\right) \quad \text{and} \quad K_{\mathbf{j}}(\theta) = \exp\left(i\theta \sum_{1 < \mathbf{j}, \alpha} \bar{n}_{1, \alpha}\right). \quad (29)$$

Since the local commutation relations are preserved, we have $(\bar{n}_{\mathbf{j}\alpha} = \bar{a}_{\mathbf{j}\alpha}^\dagger \bar{a}_{\mathbf{j}\alpha}, \bar{n}_{\mathbf{j}} = \sum_{\alpha=1}^{N_f} \bar{n}_{\mathbf{j}\alpha})$

$$\begin{cases} [\bar{a}_{\mathbf{j}\alpha}, \bar{a}_{\mathbf{j}\alpha}] = [\bar{a}_{\mathbf{j}\alpha}^\dagger, \bar{a}_{\mathbf{j}\alpha}^\dagger] = 0, \\ [\bar{a}_{\mathbf{j}\alpha}, \bar{a}_{\mathbf{j}\alpha}^\dagger] = 1 - \bar{n}_{\mathbf{j}\alpha} - \bar{n}_{\mathbf{j}}. \end{cases} \quad (30)$$

Again, we need to define an index ordering for the deformed commutation relations involving operators with different indices. For $(\mathbf{j}, \beta) > (\mathbf{i}, \alpha)$

$$\begin{cases} [\bar{a}_{\mathbf{i}\alpha}, \bar{a}_{\mathbf{j}\beta}]_\theta = [\bar{a}_{\mathbf{i}\alpha}^\dagger, \bar{a}_{\mathbf{j}\beta}^\dagger]_\theta = 0, \\ [\bar{a}_{\mathbf{j}\beta}^\dagger, \bar{a}_{\mathbf{i}\alpha}]_\theta = 0. \end{cases} \quad (31)$$

In closing this section we would like to emphasize that these lecture notes deals with Abelian anyons only. The case of non-Abelian anyons is more subtle and beyond the scope of these lecture notes. In general, it is not clear how to realize non-Abelian representations of the braid group in terms of ‘‘Fock-space particles’’. We only know examples in terms of Weyl parafermions [22].

4.1.1 Example: Spin-1/2 – anyon mapping

The simplest case of a spin-anyon mapping that one can imagine is realized in the case $D = 2$. Consider the map ($\bar{n}_j = \bar{a}_j^\dagger \bar{a}_j$ and $0 \leq \theta < 2\pi$)

$$S_j^+ = \bar{a}_j^\dagger K_j(\theta), \quad S_j^- = K_j^\dagger(\theta) \bar{a}_j, \quad S_j^z = \bar{n}_j - \frac{1}{2}, \quad (32)$$

where the non-local statistical operator or transmutator $K_j(\theta)$ is given by

$$K_j(\theta) = \exp\left(i\theta \sum_{i < j} \bar{n}_i\right) = \prod_{i < j} (1 + (e^{i\theta} - 1) \bar{n}_i) \quad (33)$$

since $\bar{n}_j^2 = \bar{n}_j$ (for any θ), and satisfy $K_j(\theta) K_j^\dagger(\theta) = K_j^\dagger(\theta) K_j(\theta) = \mathbb{1}$. In this way we transformed the original localized spin $S = 1/2$ problem into an itinerant gas of (anyon) particles obeying the operator algebra (for $i \leq j$)

$$\begin{cases} [\bar{a}_i, \bar{a}_j]_\theta = [\bar{a}_i^\dagger, \bar{a}_j^\dagger]_\theta = 0, \\ [\bar{a}_i, \bar{a}_j^\dagger]_{-\theta} = \delta_{ij}(1 - (\exp[-i\theta] + 1)\bar{n}_j), \quad [\bar{n}_i, \bar{a}_j^\dagger] = \delta_{ij} \bar{a}_j^\dagger. \end{cases} \quad (34)$$

Each statistical angle θ provides a different particle language and defines the exchange statistics of the particles. The case $\theta = \pi$ corresponds to canonical spinless fermions [11] while $\theta = 0$ represents HC bosons [21]. In all cases one can accommodate up to a single particle ($p = 1$) per quantum state, $(\bar{a}_j^\dagger)^{p+1} = 0$ (i.e, the particles are HC). In order to construct a dictionary one also needs the inverse mapping

$$\begin{cases} \bar{a}_j^\dagger = \prod_{i < j} \left(\frac{e^{-i\theta} + 1}{2} + (e^{-i\theta} - 1) S_i^z \right) S_j^+ \\ \bar{a}_j = \prod_{i < j} \left(\frac{e^{+i\theta} + 1}{2} + (e^{+i\theta} - 1) S_i^z \right) S_j^- \\ \bar{n}_j = S_j^z + \frac{1}{2} \end{cases} \quad (35)$$

4.2 Transmutation of statistics in higher dimensions

We propose the following general expression for K_j

$$K_j = \exp\left(i \sum_l \omega(l, j) \bar{n}_l\right), \quad (36)$$

where $\omega(l, j)$ is a function to be determined by imposing the transmutation of statistics. This is equivalent to the antisymmetric condition

$$e^{i\omega(l, j)} = -e^{i\omega(j, l)}, \text{ if } l \neq j, \quad \omega(l, l) = 0. \quad (37)$$

One possible solution is based on a generalized Chern-Simons construction for a $2d$ lattice: $\omega(l, j) = a(l, j)$ (see Fig. 5). However, this is not the only possible solution. Another solution

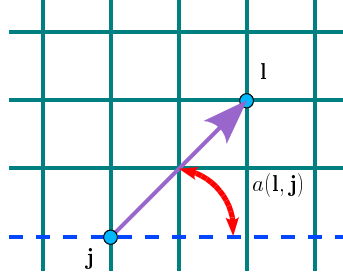


Fig. 5: One of the possible ways of defining the statistical transmutator K_j^{2d} in 2d. The dotted line represents a fixed direction on the lattice.

in 2d is

$$\omega(\mathbf{l}, \mathbf{j}) = \pi \left(\Theta(j_1 - l_1)(1 - \delta_{l_1 j_1}) + \Theta(j_2 - l_2) \delta_{l_1 j_1} \right), \quad (38)$$

with $\mathbf{l} = l_1 \mathbf{e}_1 + l_2 \mathbf{e}_2$ and $\mathbf{j} = j_1 \mathbf{e}_1 + j_2 \mathbf{e}_2$. The advantage of this solution is that its generalization to higher dimensions is straightforward ($\mathbf{l} = \sum_a l_a \mathbf{e}_a$, and $a \in [1, d]$). For instance, in 3d we have

$$\omega(\mathbf{l}, \mathbf{j}) = \pi \left(\Theta(j_1 - l_1)(1 - \delta_{l_1 j_1}) + \Theta(j_2 - l_2) \delta_{l_1 j_1} (1 - \delta_{l_2 j_2}) + \Theta(j_3 - l_3) \delta_{l_1 j_1} \delta_{l_2 j_2} \right). \quad (39)$$

4.3 Jordan-Wigner transformation for $S = 1$

Mathematically, the Jordan-Wigner transformation [11] involves the $S = 1/2$ irreducible representation of the Lie group $SU(2)$. In this section we generalize this transformation to $S = 1$. As in the case of $S = 1/2$ HC bosons (Fig. 3) the dimension of the local Hilbert space is $D = 3$. In this case, the particle Hilbert space corresponds to spin- $\frac{1}{2}$ (two flavors) fermions with the constraint of no double occupancy. This constraint can be taken into account by introducing the Hubbard operators $\bar{c}_{j\alpha}^\dagger = c_{j\alpha}^\dagger (1 - n_{j\bar{\alpha}})$ and $\bar{c}_{j\alpha} = (1 - n_{j\bar{\alpha}}) c_{j\alpha}$ ($\alpha = 1, -1; \bar{\alpha} = -\alpha$), which form a subalgebra of the so-called double graded algebra $\text{Spl}(1,2)$ [19]. From Fig. 3 one realizes that S_j^z is the difference between the occupation numbers of the two different fermion flavors. S_j^+ must be a linear combination of annihilation and creation operators since we need to annihilate one fermion to go from $S_j^z = -1$ to the $S_j^z = 0$ state and to create the other fermion to go from $S_j^z = 0$ to $S_j^z = 1$. To simplify notation we introduce the following composite operators

$$f_j^\dagger = \bar{c}_{j1}^\dagger + \bar{c}_{j\bar{1}}, \quad f_j = \bar{c}_{j1} + \bar{c}_{j\bar{1}}^\dagger. \quad (40)$$

For spins on a lattice we again fermionize the spins and reproduce the correct spin algebra with the following transformation

$$S_j^+ = \sqrt{2} (\bar{c}_{j1}^\dagger K_j + K_j^\dagger \bar{c}_{j\bar{1}}), \quad S_j^- = \sqrt{2} (K_j^\dagger \bar{c}_{j1} + \bar{c}_{j\bar{1}}^\dagger K_j), \quad S_j^z = \bar{n}_{j1} - \bar{n}_{j\bar{1}}, \quad (41)$$

whose inverse manifests the nonlocal character of the mapping

$$f_j^\dagger = \frac{1}{\sqrt{2}} \exp \left(i\pi \sum_{i < j} (S_i^z)^2 \right) S_j^+, \quad f_j = \frac{1}{\sqrt{2}} \exp \left(-i\pi \sum_{i < j} (S_i^z)^2 \right) S_j^-, \quad (42)$$

$$\bar{c}_{j1}^\dagger = S_j^z f_j^\dagger, \quad \bar{c}_{j1} = f_j S_j^z, \quad \bar{c}_{j\bar{1}}^\dagger = -S_j^z f_j, \quad \bar{c}_{j\bar{1}} = -f_j^\dagger S_j^z, \quad (43)$$

where the string operators $K_j = \exp(i\pi \sum_{i < j} \bar{n}_i) = \prod_{i < j} \prod_{\alpha} (1 - 2\bar{n}_{i\alpha})$ are the natural generalizations of the ones introduced before. The number operators are defined as $\bar{n}_i = \bar{n}_{i1} + \bar{n}_{i\bar{1}}$ ($\bar{n}_{i\alpha} = \bar{c}_{i\alpha}^\dagger \bar{c}_{i\alpha}$). These f -operators have the remarkable property that

$$\{f_j^\dagger, f_j\} = \{S_j^+, S_j^-\}, \quad (44)$$

which suggests an analogy between spin operators and “constrained” fermions.

The generalization to two-flavor (or $s = 1/2$) HC anyons is achieved by the isomorphic map

$$\begin{cases} S_j^+ = \sqrt{2} (\bar{a}_{j1}^\dagger K_j(\theta) + K_j^\dagger(\theta) \bar{a}_{j\bar{1}}) \\ S_j^- = \sqrt{2} (K_j^\dagger(\theta) \bar{a}_{j1} + \bar{a}_{j\bar{1}}^\dagger K_j(\theta)) \\ S_j^z = \bar{n}_{j1} - \bar{n}_{j\bar{1}} \end{cases} \quad (45)$$

where the non-local transmutator ($\bar{n}_{j\alpha} = \bar{a}_{j\alpha}^\dagger \bar{a}_{j\alpha}$, $\bar{n}_j = \bar{n}_{j1} + \bar{n}_{j\bar{1}}$)

$$K_j(\theta) = \exp\left(i\theta \sum_{i < j} \bar{n}_i\right) = \prod_{i < j} \left(1 + (\exp[i\theta] - 1) \bar{n}_i\right) \quad (46)$$

($\bar{n}_{j\alpha} \bar{n}_{j\beta} = \delta_{\alpha\beta} \bar{n}_{j\alpha}$) allows rotation of the statistics of the particles whose algebra is determined by ($i \leq j$)

$$\begin{cases} [\bar{a}_{i\alpha}, \bar{a}_{j\beta}]_\theta = [\bar{a}_{i\alpha}^\dagger, \bar{a}_{j\beta}^\dagger]_\theta = 0, \\ [\bar{a}_{i\alpha}, \bar{a}_{j\beta}^\dagger]_{-\theta} = \delta_{ij} \begin{cases} 1 - e^{-i\theta} \bar{n}_{j\alpha} - \bar{n}_j & \text{if } \alpha = \beta, \\ -e^{-i\theta} \bar{a}_{j\beta}^\dagger \bar{a}_{j\alpha} & \text{if } \alpha \neq \beta, \end{cases} \end{cases} \quad (47)$$

In this case a more restrictive version of the Pauli exclusion principle applies where one can accommodate no more than a single particle per site regardless of α , i.e., $\bar{a}_{j\alpha}^\dagger \bar{a}_{j\beta}^\dagger = 0$, $\forall(\alpha, \beta)$.

4.3.1 Example: the one-dimensional fermionic Hubbard model

The Hubbard model [23] is the most popular model of a strongly interacting system in condensed matter physics. It contains a kinetic energy term represented by a hopping integral t plus a local on-site Coulomb repulsion U . The single-band Hubbard Hamiltonian is ($\sigma = \uparrow, \downarrow$)

$$H_{\text{Hubb}}^{1d} = t \sum_{j,\sigma} (c_{j\sigma}^\dagger c_{j+1\sigma} + c_{j+1\sigma}^\dagger c_{j\sigma}) + U \sum_{j=1}^{N_s} \left(\hat{n}_{j\uparrow} - \frac{1}{2}\right) \left(\hat{n}_{j\downarrow} - \frac{1}{2}\right), \quad (48)$$

Let us introduce a new transformation that is not a generalized Jordan-Wigner mapping but that works in the case of Hamiltonian H_{Hubb}^{1d} . The map is defined by

$$\begin{cases} S_{j1}^+ = c_{j\uparrow}^\dagger \bar{K}_{j\uparrow}, \\ S_{j1}^z = \hat{n}_{j\uparrow} - \frac{1}{2}, \\ S_{j2}^+ = c_{j\downarrow}^\dagger \bar{K}_{j\downarrow}, \\ S_{j2}^z = \hat{n}_{j\downarrow} - \frac{1}{2}, \end{cases} \quad (49)$$

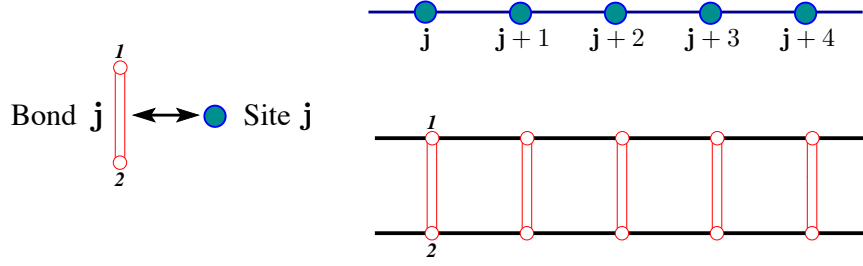


Fig. 6: Fermionic Hubbard chain maps onto a spin $S = 1/2$ ladder.

where the nonlocal operator $\bar{K}_{j\sigma}$ is defined as

$$\bar{K}_{j\uparrow} = \exp \left(i\pi \left(\sum_1 \hat{n}_{1\downarrow} + \sum_{1 < j} \hat{n}_{1\uparrow} \right) \right), \quad \bar{K}_{j\downarrow} = \exp \left(i\pi \sum_{1 < j} \hat{n}_{1\downarrow} \right). \quad (50)$$

Using this transformation, the one-dimensional Hubbard Hamiltonian becomes [9] ($\nu = 1, 2$)

$$H_{\text{Hubb}}^{1d} = 2t \sum_{j,\nu}^{N_s-1} (S_{j\nu}^x S_{j+1\nu}^x + S_{j\nu}^y S_{j+1\nu}^y) + U \sum_{j=1}^{N_s} S_{j1}^z S_{j2}^z, \quad (51)$$

which represents a two-leg ladder made out of two XY-chains coupled by an Ising interaction. The beauty of this map is that the fermionic Hubbard model and its dynamics can, in principle, be simulated with ultracold atoms in optical lattices or other quantum simulators.

4.3.2 Example: $su(N)$ spin-particle mappings

The fundamental (quark) representation of $su(N)$ can be mapped onto an algebra of constrained fermions ($\bar{c}_{j\alpha}^\dagger = \bar{a}_{j\alpha}^\dagger(\theta = \pi)$) or HC bosons ($\bar{b}_{j\alpha}^\dagger = \bar{a}_{j\alpha}^\dagger(\theta = 0)$) with $N_f = N-1$ flavors

$$\begin{aligned} \mathcal{S}^{\alpha\beta}(\mathbf{j}) &= \bar{a}_{j\alpha}^\dagger \bar{a}_{j\beta} - \delta_{\alpha\beta}/N \\ \mathcal{S}^{\alpha 0}(\mathbf{j}) &= \bar{a}_{j\alpha}^\dagger K_j^\theta, \quad \mathcal{S}^{0\beta}(\mathbf{j}) = (K_j^\theta)^\dagger \bar{a}_{j\beta} \\ \mathcal{S}^{00}(\mathbf{j}) &= \frac{N_f}{N} - \sum_{\alpha=1}^{N_f} \bar{n}_{j\alpha} = - \sum_{\alpha=1}^{N_f} \mathcal{S}^{\alpha\alpha}(\mathbf{j}), \end{aligned} \quad (52)$$

where $1 \leq \alpha, \beta \leq N_f$ runs over the set of particle flavors, and $\bar{a}_{j\alpha}^\dagger = \bar{a}_{j\alpha}^\dagger \prod_{\beta=1}^{N_f} (1 - \bar{n}_{j\beta})$. $\mathcal{S}^{\alpha_1\beta_1}(\mathbf{j})$ ($0 \leq \alpha_1, \beta_1 \leq N_f$) are the components of the $SU(N)$ -spin (i.e., there are N^2-1 linear independent components). It is easy to verify that these are generators of an $su(N)$ Lie algebra satisfying the commutation relations

$$[\mathcal{S}^{\alpha_1\alpha'_1}(\mathbf{j}), \mathcal{S}^{\beta_1\beta'_1}(\mathbf{j})] = \delta_{\alpha'_1\beta_1} \mathcal{S}^{\alpha_1\beta'_1}(\mathbf{j}) - \delta_{\alpha_1\beta'_1} \mathcal{S}^{\beta_1\alpha'_1}(\mathbf{j}). \quad (53)$$

For instance, for $N = 3$ we have ($\alpha = 1, 2$)

$$\mathcal{S}(\mathbf{j}) = \begin{pmatrix} \frac{2}{3} - \bar{n}_j & (K_j^\theta)^\dagger \bar{a}_{j1} & (K_j^\theta)^\dagger \bar{a}_{j2} \\ \bar{a}_{j1}^\dagger K_j^\theta & \bar{n}_{j1} - \frac{1}{3} & \bar{a}_{j1}^\dagger \bar{a}_{j2} \\ \bar{a}_{j2}^\dagger K_j^\theta & \bar{a}_{j2}^\dagger \bar{a}_{j1} & \bar{n}_{j2} - \frac{1}{3} \end{pmatrix}. \quad (54)$$

We can immediately see that the 2×2 block matrix $\mathcal{S}^{\alpha\beta}(\mathbf{j})$ ($1 \leq \alpha, \beta \leq 2$) contains the generators of $su(2)$. In general, from the commutation relations (53), we can verify that if $\mathcal{S}^{\alpha_1\beta_1}(\mathbf{j})$ are generators of $su(N)$, then $\mathcal{S}^{\alpha\beta}(\mathbf{j})$ are the generators of the subalgebra $su(N-1)$. This will be useful in Section 5.4.

5 Some applications

We have so far developed an algebraic framework for interacting quantum systems that will allow us to study complex phenomena characterized by the coexistence and competition of various broken symmetry states [3-5]. We have also proved a theorem that allowed us to connect all possible languages used in the quantum description of matter. Connecting the various languages through isomorphic mappings enables us to relate seemingly different physical phenomena, unveil hidden symmetries (i.e., uncover the *accidental degeneracies* of the original physical system), and, in some limiting cases, obtain the exact spectrum of the problem (or of a set of orthogonal subspaces). The ultimate goal is to use that framework to explore those unconventional complex states of matter from a unified perspective.

5.1 Generalized Jordan-Wigner transformations

In a similar fashion, one could continue for higher spin S irreps and would find that HC particles have $N_f = 2S$ flavors (we call these generalized Jordan-Wigner particles) [12]. Of course, this is not the only way to proceed. For example, for half-odd integer cases where $2S + 1 = 2^{\bar{N}_f}$ a simple transformation in terms of standard canonical multiflavor fermions is possible [12, 9].

We next generalize the JW spin-fermion mapping to any irreducible representation S [12]. Our mappings are valid for regular lattices in any spatial dimension d and particle statistics. These generalized JW mappings constitute a quantum version of the well-known classical spin-lattice-gas transformations. The significance of these transformations is that they help us understand various aspects of the same physical system by transforming intricate interaction terms in one representation into simpler ones in the other. Problems which seem intractable can even be exactly solved after the mapping. In other cases, new and better approximations can, in principle, be realized since fundamental symmetries which are hidden in one representation are manifest in the other. From a physical viewpoint, what our spin-particle transformations achieve is an exact connection between models of localized quantum spins S to models of itinerant particles with $(2S = N_f)$ color degrees of freedom or “effective” spin $s = S - \frac{1}{2}$.

We will consider now a type of fermions which naturally emerges from the strong coupling limit of models for interacting electrons. If the short range component of the Coulomb repulsion is much larger than the kinetic energy, the repulsion can be effectively replaced by a constraint of no double occupancy. This perturbative approach is usually implemented by a canonical transformation, which leads to an effective Hamiltonian acting on the subspace of states with no double occupancy. The fermionic subalgebra used to describe this effective model is generated by the so-called constrained fermions. Therefore, the constrained fermions are obtained by

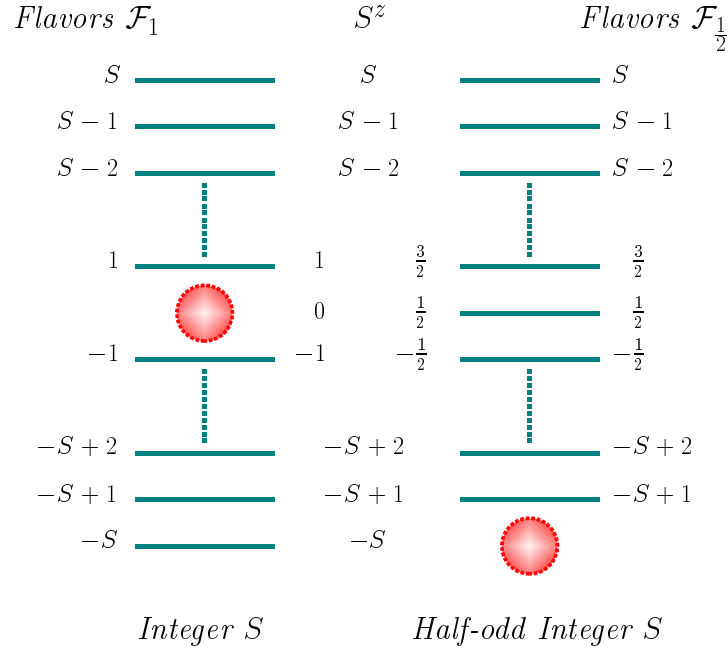


Fig. 7: Constrained fermion states per site for integer and half-odd integer spin S . There are $2S$ flavors and the corresponding $2S + 1$ values of S^z are shown in the middle column.

imposing to the canonical fermions a local constraint of no more than one particle per orbital (or site). This constraint may be incorporated into the fermionic algebra by defining the following creation and annihilation operators for the constrained fields

$$\bar{c}_{j\sigma}^\dagger = c_{j\sigma}^\dagger \prod_{\tau \in \mathcal{F}_\eta} (1 - \hat{n}_{j\tau}), \quad \bar{c}_{j\sigma} = \prod_{\tau \in \mathcal{F}_\eta} (1 - \hat{n}_{j\tau}) c_{j\sigma}, \quad (55)$$

where $\mathcal{F}_\eta$ is the set of flavors, with $\eta = \frac{1}{2}$ or 1 depending upon the spin character of the irreducible representation. It is easy to check that the particles generated by this fermionic algebra satisfy the constraint of single occupancy, i.e., the eigenvalues of $\bar{n}_j = \sum_{\sigma=1}^{N_f} \bar{n}_{j\sigma}$ are either 0 or 1. The most well-known context where these fermions appear in condensed matter physics is the strong coupling limit of the Hubbard model, which leads to the t - J Hamiltonian [23]. The set of commutation relations those constrained fermions satisfy is

$$\begin{aligned} \{\bar{c}_{i\sigma}, \bar{c}_{j\sigma'}\} &= \{\bar{c}_{i\sigma}^\dagger, \bar{c}_{j\sigma'}^\dagger\} = 0, \\ \{\bar{c}_{i\sigma}, \bar{c}_{j\sigma'}^\dagger\} &= \delta_{ij} \begin{cases} 1 + \bar{n}_{j\sigma} - \bar{n}_j & \text{if } \sigma = \sigma', \\ \bar{c}_{j\sigma'}^\dagger \bar{c}_{j\sigma} & \text{if } \sigma \neq \sigma'. \end{cases} \end{aligned} \quad (56)$$

Notice that $\prod_{\tau \in \mathcal{F}_\eta}^{\tau \neq \sigma} (1 - \bar{n}_{j\tau}) = 1 + \bar{n}_{j\sigma} - \bar{n}_j$ with number operators satisfying $\bar{n}_{j\sigma} \bar{n}_{j\sigma'} = \delta_{\sigma\sigma'} \bar{n}_{j\sigma}$. The explicit form of the generalization is

Half-odd integer spin S : $\sigma \in \mathcal{F}_{\frac{1}{2}} = \{-S + 1, \dots, S\}$

$$\begin{aligned}
S_j^+ &= \eta_{\bar{S}} \bar{c}_{j\bar{S}+1}^\dagger K_j + \sum_{\substack{\sigma \in \mathcal{F}_{\frac{1}{2}} \\ \sigma \neq \bar{S}}} \eta_\sigma \bar{c}_{j\sigma+1}^\dagger \bar{c}_{j\sigma} , \\
S_j^- &= \eta_{\bar{S}} K_j^\dagger \bar{c}_{j\bar{S}+1} + \sum_{\substack{\sigma \in \mathcal{F}_{\frac{1}{2}} \\ \sigma \neq \bar{S}}} \eta_\sigma \bar{c}_{j\sigma}^\dagger \bar{c}_{j\sigma+1} , \\
S_j^z &= -S + \sum_{\sigma \in \mathcal{F}_{\frac{1}{2}}} (S + \sigma) \bar{n}_{j\sigma} , \\
\bar{c}_{j\sigma}^\dagger &= K_j^\dagger \prod_{\tau=-S}^{\sigma-1} \eta_\tau^{-1} (S_j^+)^{\sigma+S} \prod_{\tau \in \mathcal{F}_{\frac{1}{2}}} \frac{\tau - S_j^z}{\tau + S} .
\end{aligned} \tag{57}$$

Integer spin S : $\sigma \in \mathcal{F}_1 = \{-S, \dots, -1, 1, \dots, S\}$

$$\begin{aligned}
S_j^+ &= \eta_0 (\bar{c}_{j1}^\dagger K_j + K_j^\dagger \bar{c}_{j\bar{1}}) + \sum_{\substack{\sigma \in \mathcal{F}_1 \\ \sigma \neq -1, S}} \eta_\sigma \bar{c}_{j\sigma+1}^\dagger \bar{c}_{j\sigma} , \\
S_j^- &= \eta_0 (K_j^\dagger \bar{c}_{j1} + \bar{c}_{j\bar{1}}^\dagger K_j) + \sum_{\substack{\sigma \in \mathcal{F}_1 \\ \sigma \neq -1, S}} \eta_\sigma \bar{c}_{j\sigma}^\dagger \bar{c}_{j\sigma+1} , \\
S_j^z &= \sum_{\sigma \in \mathcal{F}_1} \sigma \bar{n}_{j\sigma} , \\
\bar{c}_{j\sigma}^\dagger &= K_j^\dagger \prod_{\tau=0}^{|\sigma|-1} \eta_\tau^{-1} \begin{cases} (S_j^+)^{\sigma} \prod_{\tau \in \mathcal{F}_1} \frac{\tau - S_j^z}{\tau} & \text{if } \sigma > 0 , \\ (S_j^-)^{\sigma} \prod_{\tau \in \mathcal{F}_1} \frac{\tau - S_j^z}{\tau} & \text{if } \sigma < 0 , \end{cases}
\end{aligned} \tag{58}$$

with $\eta_\sigma = \sqrt{(S - \sigma)(S + \sigma + 1)}$ (see Fig. 7). [A bar in a subindex means the negative of that number (e.g., $\bar{\sigma} = -\sigma$).] These mappings enforce the condition on the Casimir operator $S_j^2 = S(S + 1)$. The generalized JW spin-fermion mapping can be easily extended to include a spin-anyon mapping simply by using the anyonic particles generated by $\bar{a}_{j\sigma}^\dagger$ and $\bar{a}_{j\sigma}$ [9].

5.2 Connecting seemingly unrelated phenomena: Haldane gap systems

Generically, half-odd integer spin chains have a qualitatively different excitation spectrum than integer spin chains. The Lieb, Schultz, Mattis, and Affleck theorem [24] establishes that the half-odd integer AF bilinear nearest-neighbors Heisenberg chain is gapless if the ground state (GS) is non-degenerate. The same model with integer spins is conjectured to have a Haldane gap [25]. To understand the origin of the Haldane gap we analyze the $S = 1$ XXZ Hamiltonian chain (an overall omitted constant $J > 0$ determines the energy scale)

$$H_{\text{xxz}} = \sum_j S_j^z S_{j+1}^z + \Delta (S_j^x S_{j+1}^x + S_j^y S_{j+1}^y) = \sum_j H_j^z + H_j^{\text{xx}} . \tag{59}$$

It is easy to show that the constrained fermion version of this Hamiltonian is a ($S = 1/2$) t - J_z model [26] plus particle non-conserving terms which break the $U(1)$ symmetry ($\sigma = \uparrow, \downarrow$)

$$H_{\text{xxz}} = \sum_{\mathbf{j}} (\bar{n}_{\mathbf{j}\uparrow} - \bar{n}_{\mathbf{j}\downarrow})(\bar{n}_{\mathbf{j}+1\uparrow} - \bar{n}_{\mathbf{j}+1\downarrow}) + \Delta \sum_{\mathbf{j}\sigma} \left(\bar{c}_{\mathbf{j}\sigma}^\dagger \bar{c}_{\mathbf{j}+1\sigma} + \bar{c}_{\mathbf{j}\sigma}^\dagger \bar{c}_{\mathbf{j}+1\bar{\sigma}} + \text{H.c.} \right). \quad (60)$$

The charge spectrum of the ($S = 1/2$) t - J_z model is gapless but the spin spectrum is gapped due to the explicitly broken $SU(2)$ symmetry (Luther-Emery liquid) [26]. Therefore, the spectrum of the $S = 1$ Hamiltonian associated with the t - J_z model, with $t = -\Delta$ and $J_z = 4$, (which has only spin excitations) is gapless. Hence the term which explicitly breaks $U(1)$ must be responsible for the opening of the Haldane gap. We can prove this by considering the perturbative effect that the pairing interaction $\eta \sum_{\mathbf{j}\sigma} (\bar{c}_{\mathbf{j}\sigma}^\dagger \bar{c}_{\mathbf{j}+1\bar{\sigma}} + \text{H.c.})$ has on the t - J_z Hamiltonian. To linear order in η (> 0), Eq. (60) maps onto the ($S = 1/2$) XYZ model with $\mathcal{J}_x = 2(\eta + \Delta)$, $\mathcal{J}_y = -2(\eta - \Delta)$, and $\mathcal{J}_z = -1$. To prove this statement we need to explain first how the lowest energy subspace of the t - J_z Hamiltonian can be mapped onto a spinless t - V model (the complete demonstration is presented in Ref. [26]).

The t - J_z Hamiltonian represents a hole-doped Ising model

$$H_{t-J_z} = \hat{H}_{J_z} + \hat{T} = J_z \sum_{\mathbf{j}} S_{\mathbf{j}}^z S_{\mathbf{j}+1}^z - t \sum_{\mathbf{j}\sigma} \left(\bar{c}_{\mathbf{j}\sigma}^\dagger \bar{c}_{\mathbf{j}+1\sigma} + \text{H.c.} \right). \quad (61)$$

Consider the set of parent states with M holes and $N_s - M = N_\uparrow + N_\downarrow$ quantum particles, $|\Phi_0(N_{\uparrow\uparrow}, N_{\uparrow\downarrow})\rangle$, defined as

$$|\Phi_0(N_{\uparrow\uparrow}, N_{\uparrow\downarrow})\rangle = |\underbrace{\uparrow\uparrow\downarrow\uparrow\downarrow\cdots}_{N_s-M} \underbrace{\circ\circ\circ\circ\cdots}_M\rangle, \quad (62)$$

where $N_{\uparrow\uparrow}$ ($N_{\uparrow\downarrow}$) is the number of ferro (antiferro)-magnetic links ($N_{\uparrow\uparrow} + N_{\uparrow\downarrow} = N_s - M - 1$). These states are eigenstates of the magnetic part of H_{t-J_z} , $\hat{H}_{J_z}$, with energy $E_M(N_{\uparrow\uparrow}, N_{\uparrow\downarrow}) = J_z(N_{\uparrow\uparrow} - N_{\uparrow\downarrow})/4$, and z -component of the total spin $(N_\uparrow - N_\downarrow)/2$.

From a given parent state one can generate a subspace of the Hilbert space, $\mathcal{M}(N_\uparrow, N_{\uparrow\uparrow}, N_{\uparrow\downarrow})$, by applying the hopping operators $\hat{T}_{\mathbf{j},\sigma} = c_{\mathbf{j}\sigma}^\dagger c_{\mathbf{j}+1\sigma} + \text{H.c.}$ ($\mathbf{j} = 1, \dots, N_s - 1$) to the parent state and its descendants

$$|\Phi_1(N_{\uparrow\uparrow}, N_{\uparrow\downarrow})\rangle = \hat{T}_{N_s-M,\sigma} |\Phi_0(N_{\uparrow\uparrow}, N_{\uparrow\downarrow})\rangle \quad (63)$$

or, in general,

$$|\Phi_n(N_{\uparrow\uparrow}, N_{\uparrow\downarrow})\rangle = \hat{T}_{\mathbf{j},\sigma} |\Phi_m(N_{\uparrow\uparrow}, N_{\uparrow\downarrow})\rangle. \quad (64)$$

The dimension $\mathcal{D}$ of the subspace $\mathcal{M}(N_\uparrow, N_{\uparrow\uparrow}, N_{\uparrow\downarrow})$ is $\binom{N_s}{M}$. Moreover, these different subspaces are orthogonal and not mixed by the Hamiltonian H_{t-J_z} .

We want to show now that, for a given number of holes M , the subspace generated by the Néel parent state, $\mathcal{M}(N_\uparrow, 0, N_{\uparrow\downarrow}) \equiv \mathcal{M}_0$, contains the GS. To this end, one has to note that the matrix elements $\langle \Phi_n(N_{\uparrow\uparrow}, N_{\uparrow\downarrow}) | \hat{T} | \Phi_m(N_{\uparrow\uparrow}, N_{\uparrow\downarrow}) \rangle$ are the same for the different subspaces $\mathcal{M}$. Nonetheless, the magnetic matrix elements $\langle \Phi_n(N_{\uparrow\uparrow}, N_{\uparrow\downarrow}) | \hat{H}_{J_z} | \Phi_m(N_{\uparrow\uparrow}, N_{\uparrow\downarrow}) \rangle = \delta_{nm} A(\mathcal{N}_{\uparrow\uparrow}, \mathcal{N}_{\uparrow\downarrow})$ are

different for the different subspaces, with $A(0, \mathcal{N}_{\uparrow\downarrow}) \leq A(\bar{\mathcal{N}}_{\uparrow\uparrow}, \bar{\mathcal{N}}_{\uparrow\downarrow})$, $\mathcal{N}_{\uparrow\downarrow} = \bar{\mathcal{N}}_{\uparrow\uparrow} + \bar{\mathcal{N}}_{\uparrow\downarrow}$. [Notice that, for a generic state of a given subspace, $\mathcal{N}_{\uparrow\uparrow} + \mathcal{N}_{\uparrow\downarrow} \leq N_s - M - 1$ with the equality satisfied by the parent state only, where $\mathcal{N}_{\uparrow\uparrow} = N_{\uparrow\uparrow}$ and $\mathcal{N}_{\uparrow\downarrow} = N_{\uparrow\downarrow}$.] Therefore, the Hamiltonian matrices $H_{n,m}^{\mathcal{M}}$ (of dimension $\mathcal{D} \times \mathcal{D}$) in each subspace $\mathcal{M}$, consists of identical off-diagonal matrix elements ($H_{n,m}^{\mathcal{M}} = H_{n,m}^{\mathcal{M}'}$, $n \neq m$) and different diagonal ones. These Hermitian matrices can be ordered according to the increasing value of the energy E_M of their parent states, which is equivalent (for fixed N_s and M) to ordering by the increasing number of ferromagnetic (FM) links $N_{\uparrow\uparrow}$, $H^{\mathcal{M}} \equiv H^{N_{\uparrow\uparrow}}$. For any $N_{\uparrow\uparrow} < N'_{\uparrow\uparrow}$, $H^{N'_{\uparrow\uparrow}} = H^{N_{\uparrow\uparrow}} + B$, where B is a positive semidefinite matrix. Then, the monotonicity theorem tells us that

$$E_k(N_{\uparrow\uparrow}) \leq E_k(N'_{\uparrow\uparrow}) \quad \forall k = 1, \dots, \mathcal{D}, \quad (65)$$

where $E_k(N_{\uparrow\uparrow})$ are the eigenvalues of $H^{N_{\uparrow\uparrow}}$ arranged in increasing order. Therefore, we conclude that the lowest eigenvalue of H_{t-J_z} must be in $\mathcal{M}_0$ and is $E_1(0)$.

The next step consists in showing that, within the GS subspace $\mathcal{M}_0$, the Hamiltonian H_{t-J_z} maps into an attractive spinless fermion model. If one makes the following identification

$$|\underbrace{\uparrow\downarrow\uparrow\downarrow\cdots}_{N_s-M} \underbrace{\circ\circ\circ\circ\cdots}_M\rangle \rightarrow |\underbrace{\bullet\bullet\bullet\bullet\cdots}_{N_s-M} \underbrace{\circ\circ\circ\circ\cdots}_M\rangle, \quad (66)$$

i.e., any spin particle ($\bar{c}_{j\sigma}^\dagger$, independently of the value of σ) maps into a single spinless fermion ($d_j^\dagger$) in $\mathcal{M}_0$, it is straightforward to realize that all matrix elements of H^0 are identical to the matrix elements of

$$H_0 = -t \sum_j (d_j^\dagger d_{j+1} + \text{H.c.}) - \frac{J_z}{4} \sum_j \hat{n}_j \hat{n}_{j+1} \quad (67)$$

in the corresponding new basis, with $\hat{n}_j = d_j^\dagger d_j$.

The addition of the pairing term $\eta \sum_{j\sigma} (\bar{c}_{j\sigma}^\dagger \bar{c}_{j+1\bar{\sigma}}^\dagger + \text{H.c.})$ to the $t-J_z$ Hamiltonian has two different effects in the lowest energy subspace $\mathcal{M}_0$. The process where the pair of up and down particles created preserves the Néel ordering of the spins can be mapped into the creation of two spinless particles in the effective spinless model. The other possible process creates at least one FM link in the parent state and connects $\mathcal{M}_0$ with the subspace containing one FM link, $N_{\uparrow\uparrow} = 1$, (the lowest spin excitation) $\mathcal{M}_1$. This means that the subspaces $\mathcal{M}$ are no longer invariant under the application of the Hamiltonian. However this second process contributes to second order in η (η^2/Δ_s) due to the existence of a spin gap between the GSs of $\mathcal{M}_0$ and $\mathcal{M}_1$. Therefore to first order in η , $\mathcal{M}_0$ is still an invariant subspace and the reduced Hamiltonian is a spinless model with a pairing (superconducting) term

$$H'_0 = H_0 + \eta \sum_j (d_j^\dagger d_{j+1}^\dagger + d_j d_{j+1}). \quad (68)$$

For arbitrary values of J_z , t , and hole density ν , H'_0 is equivalent (via the traditional Jordan-Wigner transformation) to the spin-1/2 XYZ chain Hamiltonian (up to an irrelevant constant)

$$H'_0 = \sum_j (\mathcal{J}_x s_j^x s_{j+1}^x + \mathcal{J}_y s_j^y s_{j+1}^y + \mathcal{J}_z (s_j^z s_{j+1}^z + s_j^z)), \quad (69)$$

and $\mathcal{J}_x = 2(\eta - t)$, $\mathcal{J}_y = -2(\eta + t)$, and $\mathcal{J}_z = -\frac{J_z}{4}$. In the language of our original Hamiltonian H_{xxz} , Eq. (60), $\mathcal{J}_x = 2(\eta + \Delta)$, $\mathcal{J}_y = -2(\eta - \Delta)$, and $\mathcal{J}_z = -1$. From exact solution of this model [27], it is seen that the system is critical only when $\eta = 0$ while for $\eta \neq 0$ a gap to all excitations opens.

It is important to note that the GS of the t - J_z model, $|\Psi_0^0\rangle$, has the same topological long-range order as the valence-bond-solid (Haldane state) [28, 26, 3], i.e., the correlation function [29] $\langle \Psi_0^0 | S_j^z \exp(i\pi \sum_{i=j+1}^{j+r-1} S_i^z) S_{j+r}^z | \Psi_0^0 \rangle = -\langle \Psi_0^0 | n_j n_{j+r} | \Psi_0^0 \rangle$ has a power law decay as a function of distance r to a constant value at the t - J_z point. This means that the superconducting term in H_{xxz} , although it is opening a gap (Haldane gap), does not change the topological order characterizing the GS. Therefore, the GS is in the same Haldane phase for $0 \leq \eta \leq \Delta$. In the particle language, the Haldane gap is a superconducting gap. Since each hole ($S^z = 0$ state in the spin language) is an anti-phase boundary (soliton) for the Néel ordering, the AF correlation function is short ranged for the GS of the t - J_z model. As demonstrated above, these solitonic excitations are massless at the t - J_z point, but become massive (gapped) as soon as the superconducting term is turned on ($\eta \neq 0$). As the superconducting term is derived from the transverse part of the Heisenberg interaction, it will not restore the AF ordering along the z direction. In this way it is easy to understand why the spin-spin correlation function, of the $S = 1$ AF Heisenberg chain is short ranged: $\langle S_j \cdot S_{j+r} \rangle \sim e^{-r/\xi}$ with ξ the correlation length.

5.3 Unveiling hidden symmetries and exactly-solvable points

We next illustrate the power of these transformations by showing exact solutions to lattice models previously unsolved by standard techniques [15]. The key is the existence of a general set of $SU(N)$ spin-particle transformations, fundamental to understanding the order hidden in complex behavior. More simply put, through the generalized spin-particle transformations, features that are subtle and hard to identify in one representation (hidden symmetries) can become prominent and easy to analyze in another (explicit symmetries). Indeed, these mappings *connect* seemingly unrelated physical phenomena, establishing equivalence relations among them. In the hierarchical group, all elements of the operator basis are symmetry generators. This allows one to study the coexistence and competition of phases, like ferromagnetism and BE condensation [15], with the corresponding OPs derived from the subgroup generators embedded in the largest global symmetry group of the problem.

Consider the $SU(2)$ -invariant model Hamiltonian

$$H_\phi = J\sqrt{2} \sum_{\langle \mathbf{i}, \mathbf{j} \rangle} (\cos \phi \mathbf{S}_i \cdot \mathbf{S}_j + \sin \phi (\mathbf{S}_i \cdot \mathbf{S}_j)^2), \quad (70)$$

where $J > 0$ and summation is over bonds $\langle \mathbf{i}, \mathbf{j} \rangle$ of a regular d -dimensional hypercubic lattice with N_s sites and coordination z . A spin $S = 1$ operator $\mathbf{S}_i$ is associated with lattice site $\mathbf{i}$ and locally satisfies the $su(2)$ Lie algebra. The case $H_1 = H_{\phi=\frac{5\pi}{4}}$, which as we will see displays a global $SU(3)$ symmetry, can be conveniently written in a ($s = \frac{1}{2}$) HC boson representation

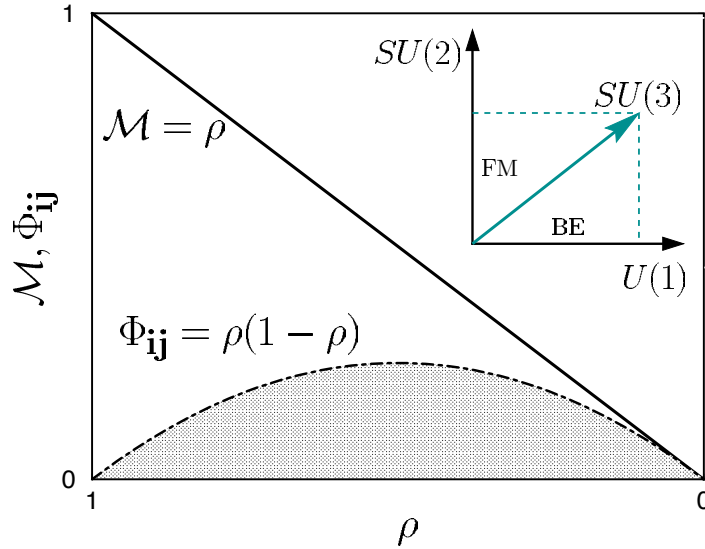


Fig. 8: Coexistence of FM and BE condensation. The correlation function $\Phi_{\uparrow\uparrow}(\mathbf{i}, \mathbf{j}) = \Phi_{\mathbf{i}\mathbf{j}}$ is expressed in the thermodynamic limit. The inset schematically displays the $SU(3)$ OP living in an eight-dimensional space with projections onto the $SU(2)$ (FM) and $U(1)$ (BE) axes.

($\bar{n}_{\mathbf{j}\sigma}\bar{n}_{\mathbf{j}\sigma'} = \delta_{\sigma\sigma'}\bar{n}_{\mathbf{j}\sigma}$), Eq. (14), as an extended t - J like Hamiltonian,

$$H_1 = -J \sum_{\langle \mathbf{i}, \mathbf{j} \rangle, \sigma} \left(\bar{b}_{\mathbf{i}\sigma}^\dagger \bar{b}_{\mathbf{j}\sigma} + \text{H.c.} \right) - 2J \sum_{\langle \mathbf{i}, \mathbf{j} \rangle} \mathbf{s}_{\mathbf{i}} \cdot \mathbf{s}_{\mathbf{j}} - 2J \sum_{\langle \mathbf{i}, \mathbf{j} \rangle} \left(1 - \frac{\bar{n}_{\mathbf{i}} + \bar{n}_{\mathbf{j}}}{2} + \frac{3}{4} \bar{n}_{\mathbf{i}} \bar{n}_{\mathbf{j}} \right), \quad (71)$$

with $\mathbf{s}_{\mathbf{j}} = \frac{1}{2} \bar{b}_{\mathbf{j}\alpha}^\dagger \boldsymbol{\sigma}_{\alpha\beta} \bar{b}_{\mathbf{j}\beta}$ an $s = \frac{1}{2}$ operator ($\boldsymbol{\sigma}$ denoting Pauli matrices), and $\bar{n}_{\mathbf{j}} = \bar{b}_{\mathbf{j}\uparrow}^\dagger \bar{b}_{\mathbf{j}\uparrow} + \bar{b}_{\mathbf{j}\downarrow}^\dagger \bar{b}_{\mathbf{j}\downarrow}$. This last form in turn can be rewritten as

$$H_1 = -2J \sum_{\langle \mathbf{i}, \mathbf{j} \rangle} P_s(\mathbf{i}, \mathbf{j}), \quad (72)$$

where $P_s(\mathbf{i}, \mathbf{j}) = P_s^2(\mathbf{i}, \mathbf{j})$ is the projector onto the symmetric subspace ($S = 0, 2$) corresponding to the bond $\langle \mathbf{i}, \mathbf{j} \rangle$ which indicates that if one finds a state that is symmetric under the permutation of nearest neighbors $\mathbf{r}_{\mathbf{i}}$ and $\mathbf{r}_{\mathbf{j}}$, then that state is the GS.

For a system of $\mathcal{N}$ HC bosons the GS is

$$|\Psi_0(\mathcal{N}, S_z)\rangle = (\hat{b}_{\mathbf{0}\uparrow}^\dagger)^{\mathcal{N}_\uparrow} (\hat{b}_{\mathbf{0}\downarrow}^\dagger)^{\mathcal{N}_\downarrow} |0\rangle, \quad (73)$$

($\mathcal{N} = \mathcal{N}_\uparrow + \mathcal{N}_\downarrow \leq N_s$) with an energy $E_0 = -JN_s z$ and a total $S_z = \frac{\mathcal{N}_\uparrow - \mathcal{N}_\downarrow}{2}$. The operator $\hat{b}_{\mathbf{0}\sigma}^\dagger$ is the $\mathbf{k} = \mathbf{0}$ component of $\bar{b}_{\mathbf{j}\sigma}^\dagger$, i.e., $\hat{b}_{\mathbf{k}\sigma}^\dagger = \frac{1}{\sqrt{N_s}} \sum_{\mathbf{j}} e^{i\mathbf{k} \cdot \mathbf{r}_{\mathbf{j}}} \bar{b}_{\mathbf{j}\sigma}^\dagger$. The quasihole and quasiparticle excited states are

$$\begin{cases} |\Psi_{\mathbf{k}}^h(\mathcal{N}, S_z)\rangle = \hat{b}_{\mathbf{k}\sigma} |\Psi_0(\mathcal{N}, S_z)\rangle & \text{quasihole,} \\ |\Psi_{\mathbf{k}}^p(\mathcal{N}, S_z)\rangle = \hat{b}_{\mathbf{k}\sigma}^\dagger |\Psi_0(\mathcal{N}, S_z)\rangle & \text{quasiparticle,} \end{cases} \quad (74)$$

with the excitation energy of each being $\omega_{\mathbf{k}} = Jz(\frac{1}{z} \sum_{\nu} e^{i\mathbf{k} \cdot \mathbf{e}_\nu} - 1)$ where the sum runs over the vectors $\mathbf{e}_\nu$ which connect a given site to its z nearest neighbors. In the $|\mathbf{k}| \rightarrow 0$ limit, $\omega_{\mathbf{k}} \rightarrow 0$.

Clearly the GS in Eq. (73) is a FM BE condensate with any partial spin polarization, and the form of the result is *independent of the spatial dimensionality of the lattice*. We note that different values of S_z correspond to the different orientations of the magnetization $\mathcal{M}$ associated to the broken $SU(2)$ spin rotational symmetry of the GS. We also note that the degeneracy of states with different number of particles $\mathcal{N}$ indicates a broken $U(1)$ charge symmetry (conservation of the number of particles) associated to the BE condensate. A signature of BE condensation is the existence of off-diagonal long-range order (ODLRO) in the correlation function $\Phi_{\sigma\sigma'}(\mathbf{ij}) = \langle \bar{b}_{i\sigma}^\dagger \bar{b}_{j\sigma'} \rangle$. When $\mathcal{N}_\uparrow$ and $\mathcal{N}_\downarrow$ are both of order N_s , there are two eigenvectors with eigenvalues of order N_s and the condensate is thus a mixture.

We can easily compute the magnetization $\mathcal{M}$ and phase coherence of these various (non-normalized) degenerate GSs for a given density $\rho = \frac{\mathcal{N}}{N_s}$. For example, in the fully polarized case, $\mathcal{N} = \mathcal{N}_\uparrow$, $\mathcal{M} = \langle S_j^z \rangle = \rho$, and the ODLRO $\Phi_{\uparrow\uparrow}(\mathbf{ij}) = \frac{\rho(1-\rho)}{1-\epsilon} (\mathbf{r}_i \neq \mathbf{r}_j)$, where $\epsilon = 1/N_s$. Similarly, the two-particle correlation function $\langle \Delta_i^\dagger \Delta_j \rangle = \Phi_{\uparrow\uparrow}(\mathbf{ij}) \frac{(\rho-\epsilon)(1-\rho-\epsilon)}{(1-2\epsilon)(1-3\epsilon)}$, where $\Delta_i^\dagger = \bar{b}_{i\uparrow}^\dagger \bar{b}_{i+\delta\uparrow}^\dagger$. Therefore the exact GS has two spontaneously broken continuous symmetries (see Fig. 8).

The exact solution defines the features of the phase diagram that our proposed framework must qualitatively admit. We will see below that both OPs (magnetization and phase), as promised, are embedded in an $SU(3)$ order parameter. We remark that the phase coexistence in the boson representation maps back to a $S = 1$ FM phase coexisting with another spin phase. To see this consider the state $|\Psi_0(\mathcal{N}, 0)\rangle$ for which $\langle S_i^z \rangle = \langle S_i^x \rangle = \langle S_i^y \rangle = 0$, which implies that it is a singlet state in the $S = 1$ representation. We will show below that this other phase has a pure spin-nematic ordering.

5.4 Identifying order parameters

The theory of phase transitions starts with Landau's pioneering work in 1937 [13,14]. One of his achievements was the realization of the fundamental relation between spontaneous symmetry breaking and the OP that measures this violation, thus giving simple prescriptions to describe order in terms of irreducible representations of the symmetry group involved. Another was the development of a phenomenological scheme to study the behavior of systems near a phase transition. Landau's theory has been successfully applied to study phase transitions where thermal fluctuations are most relevant. The field of quantum phase transitions studies the changes that can occur in the macroscopic properties of matter at zero temperature due to changes in the parameters characterizing the system. While one generally knows what to do if the OP is known, Landau's postulate gives no procedure for finding it. In this section we describe a simple algebraic framework for identifying OPs.

We have seen that the local OP acquires its simplest form when it is expressed in terms of the HL. In addition, the generators of this language exhaust all possible local OPs which may result from the solution of the problem under consideration. In other words, any local OP can be written as a linear combination of generators of the HL. Consider the bilinear-biquadratic $S = 1$ model of Eq. (70). The parameter ϕ sets the relative strength between the bilinear and bi-

quadratic terms. As shown below, there are four isolated values of $\phi = \{\pi/4, \pi/2, 5\pi/4, 3\pi/2\}$ for which H_ϕ is $SU(3)$ invariant. The only symmetry which is present for any value of ϕ is the global $SU(2)$ invariance since H_ϕ is a function of the scalar products $\mathbf{S}_i \cdot \mathbf{S}_j$.

The Hamiltonian H_ϕ has been the subject of several studies in the last two decades [30–35], nevertheless, the complete characterization of the different phases is not completely solved. A semiclassical treatment for $d > 1$ [30] indicates that there are four different phases: the usual FM ($\pi/2 < \phi < 5\pi/4$) and AF ($3\pi/2 < \phi < \pi/4$) phases are separated on both sides by collinear- ($5\pi/4 < \phi < 3\pi/2$) and orthogonal-nematic ($\pi/4 < \phi < \pi/2$) orderings. We will show below that the collinear- and orthogonal-nematic phases obtained with the semiclassical approximation are replaced by uniform- and staggered-nematic orderings, respectively.

As we have seen in previous section, the $SU(3)$ spins in the fundamental representation and the $S = 1$ $SU(2)$ spins are two equivalent languages. In addition, we have shown in Sections 4.3.2 and 4.3 that the $SU(3)$ spins and the $S = 1$ $SU(2)$ spins can be respectively mapped onto $s = 1/2$ HC bosons. We will use now these transformations to map the spin one Hamiltonian H_ϕ onto its $SU(3)$ spin version. For pedagogical reasons, it is convenient to use the $s = 1/2$ HC bosons as an intermediate language.

In Section 4.3.2, we introduced a spin-particle transformation connecting $SU(N)$ spins and multiflavored HC bosons (JW particles, in general). In particular, the fundamental (*quark*) representations of $su(N)$ were mapped onto an algebra of HC bosons with $N_f = N - 1$ flavors (see Eq. (52)). For $N = 3$ the HC bosons have two flavors ($\alpha = \uparrow, \downarrow$) which can be associated to an internal spin $s = 1/2$ degree of freedom. A compact way of writing the $SU(3)$ spin in terms of HC bosons is

$$\mathcal{S}(\mathbf{j}) = \begin{pmatrix} \frac{2}{3} - \bar{n}_{\mathbf{j}} & \bar{b}_{\mathbf{j}\uparrow} & \bar{b}_{\mathbf{j}\downarrow} \\ \bar{b}_{\mathbf{j}\uparrow}^\dagger & \bar{n}_{\mathbf{j}\uparrow} - \frac{1}{3} & \bar{b}_{\mathbf{j}\uparrow}^\dagger \bar{b}_{\mathbf{j}\downarrow} \\ \bar{b}_{\mathbf{j}\downarrow}^\dagger & \bar{b}_{\mathbf{j}\downarrow}^\dagger \bar{b}_{\mathbf{j}\uparrow} & \bar{n}_{\mathbf{j}\downarrow} - \frac{1}{3} \end{pmatrix}. \quad (75)$$

In the same way we wrote in Eq. (75) the generators of $SU(3)$ in the fundamental representation, we can write down the corresponding expressions for the generators in the conjugate representation

$$\tilde{\mathcal{S}}(\mathbf{j}) = \begin{pmatrix} \frac{2}{3} - \bar{n}_{\mathbf{j}} & -\bar{b}_{\mathbf{j}\downarrow}^\dagger & -\bar{b}_{\mathbf{j}\uparrow}^\dagger \\ -\bar{b}_{\mathbf{j}\downarrow} & \bar{n}_{\mathbf{j}\downarrow} - \frac{1}{3} & \bar{b}_{\mathbf{j}\uparrow}^\dagger \bar{b}_{\mathbf{j}\downarrow} \\ -\bar{b}_{\mathbf{j}\uparrow} & \bar{b}_{\mathbf{j}\downarrow}^\dagger \bar{b}_{\mathbf{j}\uparrow} & \bar{n}_{\mathbf{j}\uparrow} - \frac{1}{3} \end{pmatrix}. \quad (76)$$

When the $S=1$ operators are replaced by the corresponding functions of $SU(3)$ generators in the fundamental and the conjugate representations, it turns out that H_ϕ , up to an irrelevant constant, is a linear combination of the FM and the AF $SU(3)$ Heisenberg models ($0 \leq \alpha_1, \beta_1 \leq 2$)

$$H_\phi = J\sqrt{2} \sum_{\langle \mathbf{i}, \mathbf{j} \rangle} \left(\cos \phi \mathcal{S}^{\alpha_1 \beta_1}(\mathbf{i}) \mathcal{S}^{\beta_1 \alpha_1}(\mathbf{j}) + (\sin \phi - \cos \phi) \mathcal{S}^{\alpha_1 \beta_1}(\mathbf{i}) \tilde{\mathcal{S}}^{\beta_1 \alpha_1}(\mathbf{j}) \right). \quad (77)$$

Repeated Greek superindices are summed and the site index $\mathbf{i}$ runs over one of the two sublattices. This expression for H_ϕ illustrates the very important result that any nonlinear interaction in the original representation is simply a bilinear term in the new representation when mapped

onto the highest rank algebra [15]. In particular, as mentioned above, there are certain special points in parameter space where the Hamiltonian is highly symmetric. For example, for $\phi = \frac{\pi}{4}$ and $\frac{5\pi}{4}$, H_ϕ is explicitly invariant under uniform $SU(3)$ transformations on the spins [36], while for $\phi = \frac{\pi}{2}, \frac{3\pi}{2}$, H_ϕ is explicitly invariant under staggered conjugate rotations of the two sublattices. These symmetries are hard to identify in the original spin representation but are manifest in the $SU(3)$ representation.

In the following we will concentrate on the determination of the quantum phase diagram of H_ϕ for spatial dimensions $d > 1$ to avoid the strong effects of quantum fluctuations which can restore the continuous symmetry when $d = 1$. In the previous section we have analyzed the high symmetry point $\phi = \frac{5\pi}{4}$. We found that the GS has a non-zero OP

$$\mathcal{S} = \sum_{\mathbf{j}} \mathcal{S}(\mathbf{j}) , \quad (78)$$

associated to a broken continuous $SU(3)$ symmetry. This order parameter is the uniform $SU(3)$ magnetization and corresponds to the coexistence of a FM and a uniform spin-nematic ordering (see Table 1). This indicates that $\phi = \frac{5\pi}{4}$ is a quantum phase transition point separating a FM phase from a uniform spin-nematic one. Let us consider now the related point $\phi = \frac{\pi}{4}$ which differs in an overall sign from the previous case. This sign changes the interaction from FM to AF. Therefore, for this new high symmetry point we expect to get a GS characterized by the staggered order parameter

$$\mathcal{S}_{\text{ST}} = \sum_{\mathbf{j}} \exp(i\mathbf{Q} \cdot \mathbf{j}) \mathcal{S}(\mathbf{j}) , \quad (79)$$

where $\mathbf{Q}$ is the AF wave vector. It is clear from the left column of Table 1, that this staggered $SU(3)$ OP corresponds to the coexistence of the staggered $SU(2)$ magnetization

$$\mathbf{M}_{\text{ST}} = \sum_{\mathbf{j}} \exp(i\mathbf{Q} \cdot \mathbf{j}) \mathbf{S}_{\mathbf{j}} , \quad (80)$$

and the staggered nematic OP

$$\mathbf{N}_{\text{ST}} = \sum_{\mathbf{j}} \exp(i\mathbf{Q} \cdot \mathbf{j}) \mathbf{N}_{\mathbf{j}} . \quad (81)$$

$\mathbf{N}_{\mathbf{j}}$ is the symmetric and traceless component of the tensor obtained from the tensorial product of two vectors $\mathbf{S}_{\mathbf{j}}$. Hence, $\phi = \frac{\pi}{4}$ is a transition point separating the usual AF ordering from a staggered spin-nematic phase characterized by the OP of Eq. (81).

We will consider now the other two high-symmetry points, $\phi = \frac{\pi}{2}, \frac{3\pi}{2}$. For $\phi = \frac{3\pi}{2}$, the $SU(3)$ symmetry is generated by the staggered operator

$$\mathcal{S}_+ = \sum_{\mathbf{j} \in A} \mathcal{S}(\mathbf{j}) + \sum_{\mathbf{j} \in B} \tilde{\mathcal{S}}(\mathbf{j}) , \quad (82)$$

where A and B denote the two different sublattices of a hypercubic lattice. In this case, we have a FM interaction between $\mathcal{S}(\mathbf{i})$ and $\tilde{\mathcal{S}}(\mathbf{j})$, and then $\mathcal{S}_+$ is the OP characterizing the broken

Table 1: Generators of OPs and its relations for three different languages $\mathcal{A} \wedge \Gamma_A$. Each column represents a language, in this case $\dim \Gamma_A = D = 3$. *M* stands for magnetism, *SN* spin-nematic, *BE* Bose-Einstein condensation, and *CDW* charge-density wave. $su(3) \wedge FR$ is the hierarchical language with *FR* meaning fundamental representation.

$su(2) \wedge S = 1$		HC bosons $\wedge \alpha = 2$	$su(3) \wedge \text{FR}$	
M	$\begin{cases} S^x = \frac{1}{\sqrt{2}}(\mathcal{S}^{01} + \mathcal{S}^{20} + \mathcal{S}^{02} + \mathcal{S}^{10}) \\ S^y = \frac{-1}{\sqrt{2}i}(\mathcal{S}^{01} + \mathcal{S}^{20} - \mathcal{S}^{02} - \mathcal{S}^{10}) \\ S^z = \mathcal{S}^{11} - \mathcal{S}^{22} \end{cases}$	M	$\begin{cases} s^x = \frac{1}{2}(\mathcal{S}^{12} + \mathcal{S}^{21}) \\ s^y = \frac{1}{2i}(\mathcal{S}^{12} - \mathcal{S}^{21}) \\ s^z = \frac{1}{2}(\mathcal{S}^{11} - \mathcal{S}^{22}) \end{cases}$	$\mathcal{S}^{\alpha\beta}$
SN	$\begin{cases} (S^x)^2 = \frac{2}{3} + \frac{1}{2}(\mathcal{S}^{12} + \mathcal{S}^{21} + \mathcal{S}^{00}) \\ (S^z)^2 = \frac{2}{3} - \mathcal{S}^{00} \\ \{S^x, S^y\} = i(\mathcal{S}^{21} - \mathcal{S}^{12}) \\ \{S^x, S^z\} = \frac{1}{\sqrt{2}}(\mathcal{S}^{01} - \mathcal{S}^{20} - \mathcal{S}^{02} + \mathcal{S}^{10}) \\ \{S^y, S^z\} = \frac{-1}{\sqrt{2}i}(\mathcal{S}^{01} - \mathcal{S}^{20} + \mathcal{S}^{02} - \mathcal{S}^{10}) \end{cases}$	BE	$\begin{cases} \bar{b}_\uparrow^\dagger = \mathcal{S}^{10} \\ \bar{b}_\downarrow^\dagger = \mathcal{S}^{20} \\ \bar{b}_\uparrow = \mathcal{S}^{01} \\ \bar{b}_\downarrow = \mathcal{S}^{02} \end{cases}$	$\alpha, \beta \in [0, 2]$
		CDW	$\bar{n} = \frac{2}{3} - \mathcal{S}^{00}$	$\text{Tr } \mathcal{S} = 0$

$SU(3)$ symmetry of the GS. It is interesting to note that when the $SU(3)$ OP $\mathcal{S}_+$ is reduced with respect to the $SU(2)$ group, the two coexisting OPs are the staggered magnetization (see Eq. (80)) and the *uniform* nematic OP

$$\mathbf{N} = \sum_{\mathbf{j}} \mathbf{N}_{\mathbf{j}}. \quad (83)$$

In other words, if we apply an $SU(3)$ rotation generated by $\mathcal{S}_+$ to the staggered magnetization we get the uniform nematic order parameter, and vice versa, the uniform nematic OP is rotated into the staggered magnetization. This can be immediately seen by writing down the components of the local $SU(2)$ magnetization and the nematic OP as a function of the local generators of $su(3)$ in the conjugate representation $\tilde{\mathcal{S}}(\mathbf{j})$

$$\begin{aligned} S^x &= \frac{-1}{\sqrt{2}}(\tilde{\mathcal{S}}^{01} + \tilde{\mathcal{S}}^{20} + \tilde{\mathcal{S}}^{02} + \tilde{\mathcal{S}}^{10}), \\ S^y &= \frac{1}{\sqrt{2}i}(\tilde{\mathcal{S}}^{01} + \tilde{\mathcal{S}}^{20} - \tilde{\mathcal{S}}^{02} - \tilde{\mathcal{S}}^{10}), \\ S^z &= \tilde{\mathcal{S}}^{22} - \tilde{\mathcal{S}}^{11}, \\ (S^x)^2 - \frac{2}{3} &= \frac{1}{2}(\tilde{\mathcal{S}}^{12} + \tilde{\mathcal{S}}^{21} + \tilde{\mathcal{S}}^{00}), \quad (S^z)^2 - \frac{2}{3} = -\tilde{\mathcal{S}}^{00}, \\ \{S^x, S^y\} &= i(\tilde{\mathcal{S}}^{21} - \tilde{\mathcal{S}}^{12}), \\ \{S^x, S^z\} &= \frac{1}{\sqrt{2}}(\tilde{\mathcal{S}}^{01} - \tilde{\mathcal{S}}^{20} - \tilde{\mathcal{S}}^{02} + \tilde{\mathcal{S}}^{10}), \\ \{S^y, S^z\} &= \frac{-1}{\sqrt{2}i}(\tilde{\mathcal{S}}^{01} - \tilde{\mathcal{S}}^{20} + \tilde{\mathcal{S}}^{02} - \tilde{\mathcal{S}}^{10}). \end{aligned} \quad (84)$$

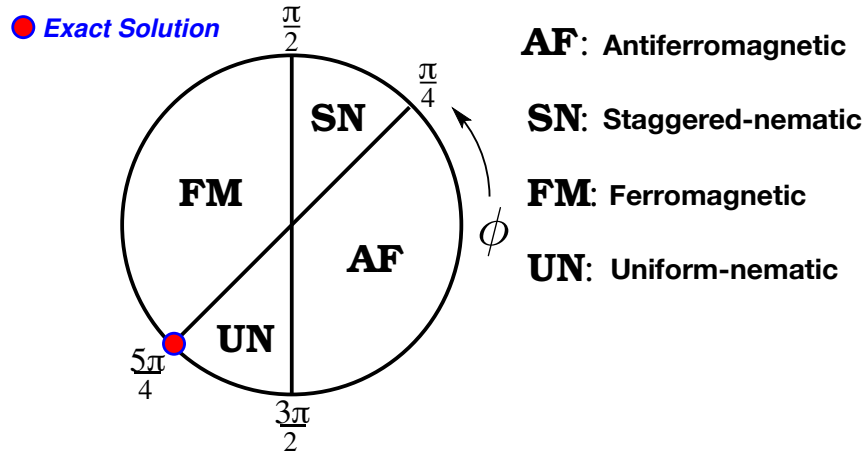


Fig. 9: Quantum phase diagram of the bilinear-biquadratic $S = 1$ model for $d \geq 2$.

Comparing these expressions to the ones in Table 1, we see that when we change from $\mathcal{S}(\mathbf{j})$ to $\tilde{\mathcal{S}}(\mathbf{j})$, there is a change in sign for the three components associated to the magnetization, while the five components corresponding to the nematic parameter remain the same. Then, it is clear that $\mathcal{S}_+$ describes the coexistence of a staggered magnetization and a uniform nematic ordering. Therefore, the conclusion is that $\phi = \frac{3\pi}{2}$ separates an ordinary AF phase from the uniform nematic ordering.

The last high symmetry point to be considered is $\phi = \frac{\pi}{2}$. In this case the coupling between $\mathcal{S}(\mathbf{i})$ and $\tilde{\mathcal{S}}(\mathbf{j})$ turns out to be positive, i.e., AF, and therefore we expect to get a broken continuous symmetry characterized by the OP

$$\mathcal{S}_- = \sum_{\mathbf{j} \in A} \mathcal{S}(\mathbf{j}) - \sum_{\mathbf{j} \in B} \tilde{\mathcal{S}}(\mathbf{j}) . \quad (85)$$

From the considerations above, it is clear that $\mathcal{S}_-$ describes the coexistence of ferromagnetism (uniform magnetization) and staggered nematic order. Hence, $\phi = \frac{\pi}{2}$ is a transition point separating these two phases.

ϕ	Global $SU(3)$ OP	OP 1	OP 2
$5\pi/4$ (FM-UN)	$\mathcal{S} = \sum_{\mathbf{j}} \mathcal{S}(\mathbf{j})$	M	N
$\pi/4$ (AF-SN)	$\mathcal{S}_{\text{ST}} = \sum_{\mathbf{j}} e^{i\mathbf{Q} \cdot \mathbf{j}} \mathcal{S}(\mathbf{j})$	$\mathbf{M}_{\text{ST}}$	$\mathbf{N}_{\text{ST}}$
$3\pi/2$ (AF-UN)	$\mathcal{S}_+ = \sum_{\mathbf{j} \in A} \mathcal{S}(\mathbf{j}) + \sum_{\mathbf{j} \in B} \tilde{\mathcal{S}}(\mathbf{j})$	$\mathbf{M}_{\text{ST}}$	N
$\pi/2$ (FM-SN)	$\mathcal{S}_- = \sum_{\mathbf{j} \in A} \mathcal{S}(\mathbf{j}) - \sum_{\mathbf{j} \in B} \tilde{\mathcal{S}}(\mathbf{j})$	M	$\mathbf{N}_{\text{ST}}$

Table 2: Order parameters describing the different phases of the bilinear-biquadratic $S = 1$ Heisenberg model for $d > 1$. ϕ indicates the phase boundary where the two phases in parentheses coexist.

In this way, by identifying the high-symmetry points of H_ϕ we have determined the quantum phase diagram of this model (see Fig. 9). In addition to the transition points, we have obtained explicit expressions for the OPs associated to each phase for any $d > 1$; they are summarized in Table 2. We can also predict from this analysis that the four transition points (high-symmetry points) correspond to first-order quantum phase transitions. In each phase, the corresponding OP has a finite value and they coexist pairwise at the high-symmetry points. However, as soon as we depart from this point in one or the other direction in ϕ , the $SU(3)$ symmetry is removed and one of the OPs goes discontinuously to zero. In other words, the states with pure magnetic (FM or AF) and nematic orderings belong to different representations of $SU(2)$ (the remaining symmetry) so only one of them remains as the GS when the $SU(3)$ symmetry is lifted. In closing this section let us summarize the main steps to follow in order to obtain and classify the local OPs.

- Identify the group $\mathcal{G}_{\text{HL}} = SU(D)$ associated to the HL whose fundamental representation has the same dimension D as the local Hilbert space of the problem. The generators of this language exhaust all possible local OPs.
- Identify the group of global symmetries of the Hamiltonian $\mathcal{G}$ which are direct products of local transformations.
- Given that $\mathcal{G} \subseteq \mathcal{G}_{\text{HL}}$, one can classify the generators of $\mathcal{G}_{\text{HL}}$ in the fundamental representation according to the irreps of $\mathcal{G}$. Each irrep leads to a different broken symmetry OP.
- Key: existence of a general set of $SU(D)$ transformations.

5.5 Hierarchical mean-field theories

In this section we are interested in using our algebraic framework to develop new approximation schemes whenever exact solutions are not available. In previous sections we outlined a framework to identify OPs based upon isomorphic mappings to a HL defined by the set of operators which in the fundamental representation (of dimension D) has the largest number of symmetry generators of the group. Any local operator can be expressed as a *linear* combination of the generators of the HL. The building of the HL depends upon the dimension D of the local Hilbert space, $\mathcal{H}_j$, modeling the physical phenomena. For instance, if one is modeling a doped AF insulator with a t - J Hamiltonian [23], then $D = 3$ (i.e., there are three possible states per site) and a HL is generated by a basis of the Lie algebra $su(3)$ in the fundamental representation [15]. As explained and proved in Refs. [15, 9], there is always a HL associated to each physical problem. These ideas complement Landau's concept of an OP providing a mechanism to reveal them, something that is outside the groundwork of his theory. Indeed, Landau's theory does not say what the OPs should be in a general situation.

As mentioned above, these isomorphic mappings not only unveil hidden symmetries of the original physical system but also manifestly establish equivalences between seemingly unrelated physical phenomena. Nonetheless, this is not sufficient to determine the *exact* phase diagram of

the problem: One has to resort to either numerical simulations with their well-known limitations or, as will be shown in here, to a *guided* approximation which at least preserves the qualitative nature of the possible thermodynamic states. A key observation in this regard is the fact that typical model Hamiltonian operators written in the HL become quadratic in the symmetry generators of the hierarchical group, and this result is independent of the group of symmetries of the Hamiltonian. This suggests a simple approximation, based upon group theoretical grounds, which deals with competing orders on an equal footing and will be termed *hierarchical mean-field theory* (HMFT) [37, 38]. In a sense, that will become clear below, HMFT constitutes the *optimum* MF or saddle-point solution that approximates the energy and correlation functions of the original problem. The HMFT is distinctly suitable when the various phases displayed by a system are the result of competing interactions and non-linear couplings of their constituents matter fields.

Since the $su(N)$ languages provide a complete set of HLs [9], any model Hamiltonian can be written in a similar fashion once we identify the appropriate HL and apply the corresponding SW mapping in the *fundamental representation* (the ordering operators will, of course, have a different meaning and algebraic expressions). The key point is that the Hamiltonian operator in the HL becomes quadratic in the symmetry generators of the hierarchical group ($SU(3)$ in the present case).

The idea behind any MF approximation is to disentangle interaction terms into quadratic ones replacing some of the elementary mode operators by their mean value. The crux of our HMFT is that the approximation is done in the HL where all possible local OPs are treated on an equal footing and the number of operators replaced by their mean value is minimized since the Hamiltonian is quadratic in the symmetry generators. In this way, the information required is minimal. In mathematical terms, given $\mathcal{O}_{ij}^\dagger \mathcal{O}_{ij} = \langle \mathcal{O}_{ij}^\dagger \rangle \mathcal{O}_{ij} + \mathcal{O}_{ij}^\dagger \langle \mathcal{O}_{ij} \rangle - \langle \mathcal{O}_{ij}^\dagger \rangle \langle \mathcal{O}_{ij} \rangle + (\mathcal{O}_{ij}^\dagger - \langle \mathcal{O}_{ij}^\dagger \rangle)(\mathcal{O}_{ij} - \langle \mathcal{O}_{ij} \rangle)$, for an arbitrary bond-operator $\mathcal{O}_{ij}$, the approximation amounts to neglecting the latter fluctuations, i.e., $\mathcal{O}_{ij}^\dagger \mathcal{O}_{ij} \approx \langle \mathcal{O}_{ij}^\dagger \rangle \mathcal{O}_{ij} + \mathcal{O}_{ij}^\dagger \langle \mathcal{O}_{ij} \rangle - \langle \mathcal{O}_{ij}^\dagger \rangle \langle \mathcal{O}_{ij} \rangle$. An important result is that all local OPs are treated equally and, moreover, symmetries of the original Hamiltonian related to the OPs are not broken explicitly in certain limits. In a sense, this is the *best* MF approximation that can be performed, i.e., the best *non-interacting* Hamiltonian that approximates the energy and correlation functions of the original problem.

We study now a simple model which displays coexistence and competition between antiferromagnetism and BE condensation (superfluidity). The model represents a gas of interacting spin-1/2 HC bosons with Hamiltonian ($t > 0$)

$$H = t \sum_{\langle i,j \rangle, \sigma} \left(\bar{b}_{i\sigma}^\dagger \bar{b}_{j\sigma} + \text{H.c.} \right) + J \sum_{\langle i,j \rangle} (\mathbf{s}_i \cdot \mathbf{s}_j - \frac{\bar{n}_i \bar{n}_j}{4}) + V \sum_{\langle i,j \rangle} \bar{n}_i \bar{n}_j - \bar{\mu} \sum_j \bar{n}_j, \quad (86)$$

where $\mathbf{s}_j = \frac{1}{2} \bar{b}_{j\alpha}^\dagger \vec{\sigma}_{\alpha\beta} \bar{b}_{j\beta}$ is an $s = \frac{1}{2}$ operator ($\vec{\sigma}$ denoting Pauli matrices). Notice that H is an extended t - J -like model of HC bosons instead of constrained fermions. These HC bosons could represent three-state atoms, like the ones used in trapped BE condensates (BECs), in an optical lattice potential. For the sake of simplicity we will only consider the AF, $J > 0$, case.

In the HL, H represents a Heisenberg-like Hamiltonian [12] in the presence of an external magnetic field μ' ($J_{\alpha_1\beta_1} = J_{\beta_1\alpha_1}$ with $0 \leq \alpha_1, \beta_1 \leq 2$)

$$H = \sum_{\langle \mathbf{i}, \mathbf{j} \rangle} J_{\alpha_1\beta_1} \mathcal{S}^{\alpha_1\beta_1}(\mathbf{i}) \mathcal{S}^{\beta_1\alpha_1}(\mathbf{j}) - \mu' \sum_{\mathbf{j}} \mathcal{S}^{00}(\mathbf{j}), \quad (87)$$

with $J_{00} = V - J/2$, $J_{01} = J_{02} = t$, $J_{11} = J_{12} = J_{22} = J/2$, and $\mu' = \frac{z}{3}(2V - J/2) - \bar{\mu}$. This HL furnishes the natural framework to analyze the symmetries of the Hamiltonian H . There is always an $SU(2)$ spin symmetry generated by $\mathcal{S}^{11} - \mathcal{S}^{22}$, $\mathcal{S}^{12}$, and $\mathcal{S}^{21}$. When $\mu' = 0$ and $V = 2t$, there are five additional generators of symmetries related to the charge degrees of freedom. Moreover, if $J = V = 2t$ there is full $SU(3)$ symmetry. For $\mu' \neq 0$, the only charge symmetry that remains is a $U(1)$ symmetry generated by $\mathcal{S}^{00}$ (conservation of the total charge). In this way the HL, leading to a unique OP from which all possible embedded orderings are derived, provides a unified description of the possible thermodynamic states of the system. Yet, it remains to establish the orderings that survive as a result of tuning the parameters of the Hamiltonian or external variables such as temperature and particle filling.

For arbitrary values of the parameters J/t and V/t we do not know a priori how to determine exactly the phase diagram of H (we know that for $J = V = 2t = \frac{2\bar{\mu}}{z} < 0$, we can find the exact GS and lowest energy states [15]). The resulting Hamiltonian ($V = 2t$ with no loss of generality) is up to irrelevant constant terms

$$H = - \sum_{\langle \mathbf{i}, \mathbf{j} \rangle} \left(\frac{J}{2} A_{\mathbf{ij}}^\dagger A_{\mathbf{ij}} + t \sum_{\sigma=\uparrow, \downarrow} B_{\sigma\mathbf{ij}}^\dagger B_{\sigma\mathbf{ij}} \right) - \mu \sum_{\mathbf{j}} n_{\mathbf{j}0}, \quad (88)$$

where $\mu = zt - \bar{\mu}$ and the ordering operators

$$\begin{cases} A_{\mathbf{ij}}^\dagger = b_{\mathbf{i}\uparrow}^\dagger b_{\mathbf{j}\downarrow}^\dagger - b_{\mathbf{i}\downarrow}^\dagger b_{\mathbf{j}\uparrow}^\dagger \\ B_{\sigma\mathbf{ij}}^\dagger = b_{\mathbf{i}\sigma}^\dagger b_{\mathbf{j}0}^\dagger - b_{\mathbf{i}0}^\dagger b_{\mathbf{j}\sigma}^\dagger \end{cases} \quad (89)$$

which transform as singlets with respect to the generators of $SU(2)$ spin and charge symmetries, respectively: $[A_{\mathbf{ij}}^\dagger, \mathcal{S}^{12(21)}(\mathbf{i}) + \mathcal{S}^{12(21)}(\mathbf{j})] = 0 = [B_{\uparrow(\downarrow)\mathbf{ij}}^\dagger, \mathcal{S}^{10(20)}(\mathbf{i}) + \mathcal{S}^{10(20)}(\mathbf{j})]$.

The resulting MF Hamiltonian reads

$$\tilde{H} = - \sum_{\langle \mathbf{i}, \mathbf{j} \rangle} \left(\frac{JA}{2} (A_{\mathbf{ij}}^\dagger + A_{\mathbf{ij}}) + tB \sum_{\sigma=\uparrow, \downarrow} (B_{\sigma\mathbf{ij}}^\dagger + B_{\sigma\mathbf{ij}}) \right) - \mu \sum_{\mathbf{j}} n_{\mathbf{j}0} + \lambda \sum_{\mathbf{j}, \alpha} n_{\mathbf{j}\alpha} \quad (90)$$

$$= \sum_{\mathbf{k} \in \text{RBZ}} \left(\Lambda_A b_{\mathbf{k}\uparrow}^\dagger b_{-\mathbf{k}+\mathbf{Q}\downarrow}^\dagger + \Lambda_B \sum_{\sigma=\uparrow, \downarrow} b_{\mathbf{k}\sigma}^\dagger b_{-\mathbf{k}+\mathbf{Q}0}^\dagger + \text{H.c.} + (\lambda - \mu) n_{\mathbf{k}0} + \lambda \sum_{\sigma=\uparrow, \downarrow} n_{\mathbf{k}\sigma} \right), \quad (91)$$

where the sum of momenta $\mathbf{k}$ is performed over the reduced Brillouin zone (RBZ) with AF ordering wave vector $\mathbf{Q}$, with $\Lambda_A = -2JA\gamma_{\mathbf{k}}$, $\Lambda_B = -4tB\gamma_{\mathbf{k}}$, with $\gamma_{\mathbf{k}} = \frac{1}{z} \sum_{\vec{\delta}} e^{i\mathbf{k} \cdot \vec{\delta}}$ ($\vec{\delta}$ are nearest-neighbor vectors). Note that when $B = 0$, the $SU(2)$ spin and $U(1)$, $\mathcal{S}^{00}$, symmetries are conserved; the opposite case $A = 0$ preserves $\mathcal{S}^{10(01)} + \mathcal{S}^{20(02)}$ and $\mathcal{S}^{11} + \mathcal{S}^{22} - \mathcal{S}^{00}$ symmetries. In Eq. (91) we have only considered homogeneous solutions.

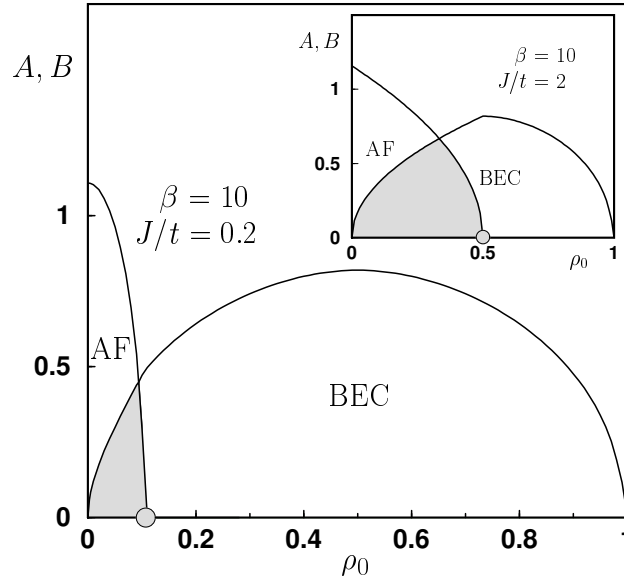


Fig. 10: Order fields A and B as a function of the density ρ_0 for different values of J/t and inverse temperature $\beta = 10$ (in units of t^{-1}). The filled circle on the density axis indicates a quantum critical point.

Generalization of these ideas to include clusters of many elementary degrees of freedom, e.g., quantum spins, can be found in [38]. In practice, we tile the original many-body lattice system into clusters preserving most of the symmetries of the Hamiltonian and represent each many-body state by the action of a composite operator over the vacuum of a new enlarged Fock space. The mapping that relates the original set of operators and the new composite ones is canonical if a physical constraint is implemented. As a consequence, the Hamiltonian of study can be exactly re-expressed in the new language of composite operators and treated by standard many-body techniques, with the advantage that the intra-cluster quantum correlations are computed exactly while the inter-cluster are dealt with in a MF way. The HMFT approach has been implemented successfully in a variety of frustrated strongly correlated system [39–43].

5.6 Quantum simulations

A new challenge in information theory and computer science has recently emerged as the result of exploiting the fundamental laws of quantum mechanics. This new set of ideas comprise what is known as “Theory of Quantum Computation and Quantum Information” and has as a major objective to process information in a way that exceeds the capabilities of classical information [44]. The device that performs the manipulation of information is named quantum computer and the standard unit of information is the *qubit* (i.e., a two-level system). The close relationship between information processing and the physical phenomena leading to it is perhaps the most remarkable aspect of this new paradigm. Since information can be represented in many different physical forms, and easily converted from one form to another without changing its meaning, quantum information represents a new abstract archetype for information processing independent of the precise implementation of the quantum computer, only requiring at least one physical representation to be useful.

A key fundamental concept in information theory is the realization [45, 46] that a model of computation is intimately connected to a physical system through a closed operator algebra. In other words, each physical system is associated to a certain language (e.g., spin $S = 1/2$) and thus to an algebra realizing it (e.g., Pauli algebra), and that particular algebra may become a possible model of computation. An immediate consequence is that an arbitrary physical system can be simulated by another physical system (e.g., a quantum computer) whenever there exists an isomorphic mapping between the different operator algebras representing the systems [45]. Simple examples are provided in Refs. [45], [46], and [47], where it is shown how to simulate a $1d$ impurity Anderson model using a quantum computer based on an array of spins $S=1/2$ [48]. It is very clear the power that our fundamental theorem has by providing the formal connections (isomorphisms) between the different languages of nature. Therefore, the implications for quantum information and computation are rather obvious, namely that one can identify quantum resources and define convenient models of computation, or imitate an arbitrary quantum phenomena with a given quantum computer given the appropriate dictionaries to translate nature's language to the *machine language*. In this way, one can recognize the subject of quantum simulations as one of those areas where the concepts of language and dictionaries developed in the present lecture are of particular relevance.

Physical phenomena can be simulated or *imitated* by a quantum network [45, 47, 49] with the help of a quantum computer. Imitation is realized through a quantum algorithm which consists of a quantum network with a means to repeat blocks of instructions. A quantum network is defined by a sequence of universal gates (unitary operations), applied to the system for the purpose of information processing, and measurements in a fixed temporal order. The measurement operation is mostly needed to classically access information about the state of the system. Every matrix which represents a reversible operation on quantum states can be expressed as a product of the one and two-qubit gates, and the minimum set needed to represent any such matrices is called a universal set of gates.

When trying to simulate a problem using quantum information processing, an important issue is to determine how many physical resources are needed for the solution. The main resources are *quantum space*, the number of qubits needed, and *quantum time*, the number of quantum gates required. The accounting of algorithmic resources forms the foundations of quantum complexity theory. One of the objectives in quantum information theory is to accomplish imitation efficiently, i.e., with polynomial complexity, and the hope is that quantum imitation is more efficient (i.e., needs less resources) than classical imitation. There are examples that support such hope (e.g., fermion simulations with polynomially bounded statistical errors [45, 47]), although there is no general proof that indicates the superiority of quantum over classical imitations, regarding efficiency. Indeed, there is, so far, no efficient quantum algorithm that can determine the GS (or, in general, the spectrum) of a given Hermitian operator [47], despite occasional claims. It is known that the ability to resolve this question leads to efficient algorithms for NP-hard problems like the traveling salesman conundrum.

A very important observation, in connection with the notion of efficiency, is a corollary of our fundamental theorem: Given two languages, the generators of one of them can be written as a

polynomial function, with *polynomial complexity in the number of modes or resources*, of the generators of the other and vice versa. This result implies that the important algorithmic step of *translation* from the language of the system to be imitated to the *machine language* does not change the complexity of the quantum space and time.

Certainly, a general purpose quantum computer is not the only device that allows simulation of physical phenomena in nature (with its many languages). Imitation can also be achieved in a conceptually different manner using a quantum simulator. The main distinction is the lack of universality of the latter. An example of a quantum simulator is an optical lattice [50] which is specifically designed to imitate a given physical Hamiltonian and where there is limited quantum control. The possibility of control and tunability of the interactions of the elementary constituents offers the potential to design new states of matter. This is of particular relevance in strongly correlated matter where these quantum simulators furnish the benchmark to test theories and approximations. Again, the importance of the languages and dictionaries developed in this manuscript is clear and concrete.

6 Concluding remarks

The development of exact algebraic methods is one of the most elegant and promising tools towards the complete understanding of quantum phases of matter and their corresponding phase transitions. We presented an algebraic framework aimed at uncovering the order behind the potential multiplicity of complex phases in interacting quantum systems, a paradigm at the frontiers of condensed matter physics. We argued that symmetry, and topology (not so much explained in this lecture notes), are key guiding principles behind such complex emergent behavior. Emphasis has been made in developing a systematic mathematical structure that allows one to attack these problems within a single unifying approach.

A key result, from which all other results follow, is the proof of a fundamental theorem that permits to connect the various operator languages used in the description of the properties of physical systems. This theorem together with the notion of transmutation of statistics provides the tools necessary to unify the quantum description of matter. To formalize this unification we needed to rigorously define the concepts of language and dictionary (isomorphism): To model a particular physical phenomena we commonly identify the main degrees of freedom of the problem and associate to them certain operators. One can furnish the resulting set of operators (that we call language) with an algebraic structure and ask whether two different languages have something in common. The fundamental theorem tells us that two languages can be connected whenever the dimension of their local Hilbert spaces are equal. We expanded the notion of local Hilbert space to embrace different Hilbert space decompositions (we saw, for instance, how to map the Hilbert space of a bond to a site). The resulting one-to-one language mappings we named dictionaries (a traditional example of which is the Jordan-Wigner mapping).

In the course of the presentation we showed, through example, many different dictionaries relating diverse operator languages. In this way we defined universality of behavior as an equivalence relation between seemingly different physical phenomena which share exactly the same

underlying mathematical structure as a result of one-to-one language mappings (for example, the spin nematic order and Bose-Einstein condensation of flavored hard-core bosons). If it is the whole system Hamiltonian that maps onto another in a different language (like the example we described above), the universality applies to all length and time scales. However, sometimes only particular invariant subspaces of the original Hamiltonian map onto another system Hamiltonian. In this case, universality is only manifested at certain energy scales. The t - J_z chain model provides a beautiful example of the latter situation: the low-energy manifold of states maps onto an XXZ model Hamiltonian, which can be exactly solved using the Bethe ansatz [26] (the t - J_z model is quasi-exactly solvable). Out of the many languages one can use to describe a given physical problem there is a class, we named hierarchical language, which has the advantage that any local operator can be expressed as a linear combination of its generators. In this way, hierarchical languages provide the tools necessary to classify order parameters.

There are several reasons why our algebraic framework constitutes a powerful method to unveil complex phenomena in interacting quantum systems. Most importantly: To connect seemingly unrelated physical phenomena (e.g., models for high-temperature superconductors or heavy-fermion systems and quantum spin theories); to identify general symmetry principles behind complex phase diagrams; to unveil hidden symmetries (and associated order parameters) to explore new states of matter with internal orders not contemplated before; to obtain exact solutions of relevant physical models that display complex ordering at certain points in Hamiltonian space; and to find new approximations which do not privilege any of the competing interactions. For instance, in the *hierarchical mean-field theory* approach, we approximated the dynamics (and thermodynamics) treating all possible local order parameters on an equal footing. One may say that this procedure follows the guiding principles of *maximum symmetry* and *minimum information*. This allowed us to obtain in a simple manner the phase diagram of a problem exhibiting coexistence and competition between antiferromagnetism and superfluidity. Combined with an analysis of fluctuations (to analyze the stability of the mean-field) one now has a simple machinery to design phase diagrams.

Several important concepts have been left out of this lecture notes. For instance, the notion of *emergent symmetry* [9], i.e., the fact that new symmetries not realized in the Hamiltonian describing the system can emerge at low energies [51]. There is one concept, in particular, that deserves special attention. This is the notion of a *duality transformation*. Dualities appear in nearly all disciplines of physics and play a central role in statistical mechanics and field theory [52, 53]. When available, these mathematical transformations provide an elegant and efficient way to obtain information about models that need not be exactly solvable. Most notably, dualities may be used to determine features of phase diagrams such as boundaries between phases, and the exact location of some critical/multicritical points. Historically, dualities were introduced in classical statistical mechanics by Kramers and Wannier [54] as a relation between the partition function of one system at high temperature (or weak coupling) to the partition function of another (dual) system at low temperatures (or strong coupling). This relation allowed for a determination of the exact critical temperature of the two-dimensional Ising model on a square lattice before the exact solution of the model was available. Later on, it was noticed that,

due to the connection between quantum theories in d space dimensions and classical statistical systems in $d + 1$ dimensions, dualities can provide relations between quantum theories in the strong coupling and weak coupling regimes [52]. Our work is motivated by a quest to realize a simple unifying framework for the detection and treatment of dualities.

An algebraic approach to dualities and self-dualities for systems of *arbitrary spatial dimensionality* d has been developed in Refs. [55, 10]. This theory of dualities is based on the notion of bond algebras [56, 10]. It deals with classical and quantum dualities in a unified fashion explaining the precise connection between quantum dualities and the low temperature (strong-coupling)/high temperature (weak-coupling) dualities of classical statistical mechanics (or (Euclidean) path integrals) [57]. Its range of applications includes discrete lattice, continuum field, and gauge theories. Dualities are revealed to be local, structure-preserving mappings between model-specific bond algebras that can be implemented as unitary transformations, or partial isometries if gauge symmetries are involved. This characterization permits to search systematically for dualities and self-dualities in quantum models of arbitrary system size, dimensionality and complexity, and any classical model admitting a transfer matrix or operator representation. In particular, special dualities like exact dimensional reduction, emergent, and gauge-reducing dualities that solve gauge constraints can be easily understood in terms of mappings of bond algebras. The transformations are, in general, quite non-local in the mapped degrees of freedom. Non-local transformations like dual variables and Jordan-Wigner dictionaries are algorithmically derived from the local mappings of bond algebras. This permits to establish a precise connection between quantum dual and classical disorder variables. Our bond-algebraic approach goes beyond the standard (Fourier transformation) approach to classical dualities (see, e.g., Appendix 1 of Ref. [10]), and could help resolve the long-standing problem of non-Abelian duality transformations [58]. Several interesting examples and applications, including location of phase boundaries, spectral behavior and, notably, how bond-algebraic dualities help constrain and realize fermionization in an arbitrary number of spatial dimensions, can be found in Refs. [10] and [59].

References

- [1] P.W. Anderson, *Science* **177**, 393 (1972)
- [2] B. Zeng, X. Chen, D.-L. Zhou, and X.-G. Wen:
Quantum Information Meets Quantum Matter (Springer, New York, 2019)
- [3] Z. Nussinov and G. Ortiz, *Ann. Phys. (NY)* **324**, 977 (2009)
- [4] Z. Nussinov and G. Ortiz, *Proc. Nat. Ac. Sc.* **106**, 16944 (2009)
- [5] P.W. Anderson: *Basic Notions of Condensed Matter Physics*
(Addison-Wesley, Redwood City, 1992)
- [6] H. Barnum, E. Knill, G. Ortiz, R. Somma, and L. Viola,
Phys. Rev. Lett. **92**, 107902 (2004)
- [7] J.K. Jain: *Composite Fermions* (Cambridge University Press, 2007)
- [8] J. Zaanen, *SciPost Physics* **6**, 061 (2019)
- [9] C.D. Batista and G. Ortiz, *Adv. in Phys.* **53**, 1 (2004)
- [10] E. Cobanera, G. Ortiz, and Z. Nussinov, *Adv. in Phys.* **60**, 679 (2011)
- [11] P. Jordan and E. Wigner, *Z. Phys.* **47**, 631 (1928)
- [12] C.D. Batista and G. Ortiz, *Phys. Rev. Lett.* **86**, 1082 (2001)
- [13] L.D. Landau and E.M. Lifshitz: *Statistical Physics*
(Butterworth-Heinemann, Oxford, 1980)
- [14] H. Nishimori and G. Ortiz: *Elements of Phase Transitions and Critical Phenomena*
(Oxford University Press, 2011)
- [15] C.D. Batista, G. Ortiz, and J.E. Gubernatis, *Phys. Rev. B* **65**, 180402 (2002)
- [16] F.J. Murray and J. von Neumann, *Ann. Math.* **37**, 116 (1936)
- [17] O. Bratteli and D.W. Robinson: *Operator Algebras and Quantum Statistical Mechanics*
Vols. I and II (Springer, New York, 1987-97)
- [18] N. Jacobson: *Basic Algebra I* (W.H. Freeman and Company, New York, 1985), and
Basic Algebra II (W.H. Freeman and Company, New York, 1989)
- [19] J.F. Cornwell: *Group Theory in Physics* (Academic Press, San Diego, 1997)

- [20] A monoid is a triple $(\mathfrak{M}, \square, 1)$ in which $\mathfrak{M}$ is a non-empty set, $\square$ is an associative product in $\mathfrak{M}$, and 1 is an element of $\mathfrak{M}$ such that $\square(1, \alpha) = \alpha = \square(\alpha, 1)$ for all $\alpha \in \mathfrak{M}$. In this way we see that the concept of monoid generalizes the notion of group; a group is a monoid all of whose elements are invertible. A monoid can also be defined as a semigroup $(\mathfrak{M}, \square)$ with an element that is the unit for $\square$, i.e., 1 .
- [21] T. Matsubara and H. Matsuda, Prog. Theor. Phys. **16**, 569 (1956)
- [22] E. Cobanera and G. Ortiz, Phys. Rev. A **89**, 012328 (2014);
Erratum, Phys. Rev. A **91**, 059901 (2015)
- [23] A. Auerbach: *Interacting Electrons and Quantum Magnetism* (Springer, New York, 1994)
- [24] E.H. Lieb, T.D. Schultz, and D.C. Mattis, Ann. Phys. **16**, 407 (1961);
I. Affleck and E.H. Lieb, Lett. Math. Phys. **12**, 57 (1986)
- [25] F.D.M. Haldane, Phys. Lett. A **93**, 464 (1983)
- [26] C.D. Batista and G. Ortiz, Phys. Rev. Lett. **85**, 4755 (2000)
- [27] R.J. Baxter: *Exactly Solved Models in Statistical Mechanics* (Academic Press, London, 1990)
- [28] I. Affleck, T. Kennedy, E.H. Lieb, and H. Tasaki, Phys. Rev. Lett. **59**, 799 (1987)
- [29] M. Den Nijs and K. Rommelse, Phys. Rev. B **40**, 4709 (1989)
- [30] N. Papanicolau, Phys. Lett. A **116**, 89 (1986); Nucl. Phys. **305**, 367 (1988)
- [31] J. Sólyom, Phys. Rev. B **36**, 8642 (1987)
- [32] J.B. Parkinson, J. Phys. C **21**, 3793 (1988)
- [33] M.N. Barber and M.T. Batchelor, Phys. Rev. B **40**, 4621 (1989)
- [34] A. Klumper, Europhys Lett. **9**, 815 (1989); J. Phys. A **23**, 809 (1990)
- [35] G. Fáth and J. Sólyom, Phys. Rev. B **51**, 3620 (1995)
- [36] C.K. Lai, J. Math. Phys. **15**, 1675 (1974); B. Sutherland, Phys. Rev. B **12**, 3795 (1975)
- [37] G. Ortiz and C.D. Batista, Phys. Rev. B **67**, 134301 (2003)
- [38] L. Isaev, G. Ortiz, and J. Dukelsky, Phys. Rev. B **79**, 024409 (2009)
- [39] L. Isaev, G. Ortiz, and J. Dukelsky, Phys. Rev. Lett **103**, 177201 (2009)
- [40] L. Isaev, G. Ortiz, and J. Dukelsky, J. Phys.: Condens. Matter **22**, 016006 (2010)
- [41] L. Isaev and G. Ortiz, Phys. Rev. B **86**, 100402(R) (2012)

- [42] D. Huerga, J. Dukelsky, N. Laflorencie, and G. Ortiz, Phys. Rev. B **89**, 094401 (2014)
- [43] D. Huerga, S. Capponi, J. Dukelsky, and G. Ortiz, Phys. Rev. B **94**, 165124 (2016)
- [44] See for example M.A. Nielsen and I.L. Chuang: *Quantum Computation and Information* (Cambridge University Press, 2000)
- [45] G. Ortiz, J.E. Gubernatis, R. Laflamme, and E. Knill, Phys. Rev. A **64**, 22319 (2001)
- [46] G. Ortiz, E. Knill, and J.E. Gubernatis, Nucl. Phys. B **106**, 151 (2002)
- [47] R. Somma, G. Ortiz, J.E. Gubernatis, R. Laflamme, and E. Knill, Phys. Rev. A **65**, 42323 (2002)
- [48] C. Negreverne, R. Somma, G. Ortiz, E. Knill, and R. Laflamme, Phys. Rev. A **71**, 032344 (2005)
- [49] R. Somma, G. Ortiz, E. Knill, and J.E. Gubernatis, Int. J. of Quant. Inf. **1**, 189 (2003)
- [50] M. Greiner *et al.*, Nature **415**, 39 (2002)
- [51] G. Ortiz, E. Cobanera, Z. Nussinov, Nucl. Phys. B **854**, 780 (2012)
- [52] R. Savit, Rev. Mod. Phys. **52**, 453 (1980)
- [53] E. Witten, Phys. Today **50**, 28 (1997)
- [54] H.A. Kramers and G.H. Wannier, Phys. Rev. **60**, 252 (1941)
- [55] E. Cobanera, G. Ortiz, and Z. Nussinov, Phys. Rev. Lett. **104**, 020402 (2010)
- [56] Z. Nussinov and G. Ortiz, Phys. Rev. B **79**, 214440 (2009)
- [57] Z. Nussinov, G. Ortiz, and M.-S. Vaezi, Nucl. Phys. B **892**, 132 (2015)
- [58] E. Cobanera, G. Ortiz, and E. Knill, Nucl. Phys. B **877**, 574 (2013)
- [59] Z. Nussinov, G. Ortiz, and E. Cobanera, Phys. Rev. B **86**, 085415 (2012)

Index

A

acceptance ratio, 15.3
adiabatic connection (AC), 2.4
 fluctuation-dissipation theorem (ACFDT),
 2.3
Anderson molecule, 11.7, 11.40
anticommutation relations, 7.6
antisymmetrized geminal power (AGP),
 6.19
anyons, 17.14
atomic-sphere truncation, 11.30
auxiliary-field quantum Monte Carlo (AFQMC),
 6.3

B

backflow correlations, 5.13, 5.21
band gap problem, 1.18
basis set, 8.10
BCS, 6.19
 Hamiltonian, 5.17
Bethe lattice, 10.10
Bloch oscillations, 13.21
branching random walks, 3.12

C

Ceperley and Alder QMC data, 1.10
cluster update, 16.16
complex auxiliary-fields, 6.15
computational complexity, 15.23
configuration interaction, 7.3, 7.14
configuration space, 15.2
conserving approximations, 12.2
continued fraction, 7.24
continuous-time quantum Monte Carlo,
 11.13
Coulomb matrix elements, 6.4
coupled cluster theory, 2.3, 2.8
current, 13.21
current operator, 13.10

D

deconfined quantum-critical point, 16.24
delta function contour, 13.14, 13.16
density matrix, 1.5, 1.21, 7.4, 9.3, 14.4
 coherence, 14.7

 eigenvalues, 9.4
density of states, 13.3
 equilibrium, 13.20
 nonequilibrium, 13.20
density response function, 2.3, 2.7
density-functional theory (DFT), 1.5,
 11.2
density-matrix renormalization group (DMRG),
 8.2, 9.2
 algorithm, 8.5, 9.5
 DMFT impurity solver, 9.9
 dynamical properties, 9.7
 Lanczos, 9.8
 symmetries, 9.6
designer Hamiltonian, 16.23
detailed balance, 3.6, 5.8
 in DMC, 3.15
DFT+DMFT, 11.25
diagrammatic Monte Carlo, 15.2
 bold (BDMC), 15.20
diffusion Monte Carlo, 3.13, 4.3
 fixed-node approximation, 4.3
dimerized Heisenberg model, 16.24
directed loop update, 16.18
double-counting correction, 11.3, 11.27,
 11.35
downfolding, 7.28, 11.27
drift-diffusion process, 4.3
dynamic structure factor, 16.26
dynamical mean-field theory (DMFT),
 9.9, 10.6, 11.22
 impurity solver, 11.13
 multi-site and multi-orbital problems,
 11.25
 Bethe lattice, 10.10
 cavity Green function, 10.9
 impurity solver, 13.12
 iterative algorithm, 13.11
 multi-site and multi-orbital problems,
 9.10
 nonequilibrium, 13.11
 self-consistency, 10.8
Dyson equation, 11.6, 12.6

E

electron density, 1.3
 operator, 1.4
 electronic transport, 14.4
 energy derivatives, 5.22
 energy fluctuation potential method, 4.11
 energy minimization, 4.7
 entanglement, 9.5, 14.5
 entropy, 16.23
 entropy, von Neumann, 9.5
 equation of motion, 13.7
 Euler-Lagrange equation, 1.7
 exchange, 14.3, 14.19
 exchange-correlation
 functionals, 1.10, 2.2
 hole, 1.4
 sum rule, 1.5
 extended-cell expansion, 16.19

F

Falicov-Kimball model, 13.9
 fermionic wave function
 fast update of determinants, 5.17
 Feynman diagrams, 12.4, 15.2
 fidelity susceptibility, 16.23
 field operators, 7.6
 fixed-node approximation, 4.4
 fixed-sample optimization, 4.5
 fluctuation-dissipation theorem, 13.5
 Fock space, 7.6
 force bias, 6.16
 fractional occupation, 1.21
 frontier orbital, 14.8
 f_{xc} kernel, 2.7

G

gauge invariance, 13.9
 Gaussian integration, 13.17
 generalized gradient approximations, 1.15
 generating functional, 12.15
 ghost states, 7.23
 grand canonical potential, 12.2
 Green function, 12.4, 13.2
 advanced, 13.3
 boundary condition, 13.14
 contour-ordered, 13.7
 equilibrium, 13.3
 gauge-invariant, 13.9

greater, 13.3
 history, 13.2
 imaginary time, 12.3
 lesser, 13.3
 local, 13.16
 momentum-dependent, 13.3
 nonequilibrium
 noninteracting, 13.9
 retarded, 13.3
 time representation, 13.5
 ground-state projector method, 16.8
 Gutzwiller
 projector, 5.11, 5.15
 wavefunction, 5.10

H

hard-core bosons, 17.8
 Hartree-Fock, 2.10, 2.12, 11.17, 11.23
 potential, 12.5
 wavefunctions, 5.9
 Heisenberg antiferromagnet, 16.11, 16.24
 Heisenberg representation, 13.3
 Hohenberg-Kohn
 theorem, 1.5
 universal functional, 1.7
 Hubbard chain, 17.16
 Hubbard dimer, 11.5, 11.38
 Hubbard model, 5.2, 9.13, 11.22
 Hubbard models for t_{2g} and e_g systems,
 11.31
 Hubbard operators, 17.15
 Hubbard-Stratonovich transformation,
 6.5
 hybrid formalism in AFQMC, 6.17
 hybrid functionals, 1.15
 hybridization function, 11.6
 hypercubic lattice, 13.11

I

imaginary time Green function, 12.3
 importance sampling, 3.11
 indistinguishability, 7.2
 iterative perturbation theory (IPT), 10.14,
 10.29

J

J - Q model, 16.25
 Jacob's ladder, 2.2

Janak theorem, 1.21
Jastrow wavefunction, 5.11
Jordan-Wigner transformation, 17.15, 17.18

K

Kadanoff-Baym-Keldysh contour, 13.5, 13.14
Kohn-Sham DFT formulation, 1.9, 2.5, 2.8
Kondo effect, 10.4
Krylov space, 7.20

L

Löwdin orthonormalization, 7.26
Lanczos method, 7.20
LDA+DMFT, 11.22, 11.25
Lehmann representation, 11.42, 12.4, 13.3
Levy constrained search, 1.7
linear optimization method, 3.28
Liouville
 equation, 14.4, 14.23
 superoperator, 14.4
local density approximation (LDA), 1.10
local Dyson equation, 11.6
local energy, 4.3
localization of the basis, 11.30
long-range interactions, 16.15
loop update, 16.16
Luttinger-Ward functional, 12.2, 12.10

M

machine learning, 15.23
many-body
 interference, 14.7
 rate, 14.7
many-body perturbation theory, 2.8
many-body wavefunctions, 3.19
Markov chain, 5.7
 master equation, 5.8
master equation, 14.5, 14.11, 14.23
matrix product operator (MPO), 8.6
matrix product state (MPS), 8.3
mean-field theory, 17.30
metal-insulator transition, 10.16
Metropolis-Hastings algorithm, 3.5, 4.3, 5.8
 acceptance probability, 3.6, 5.8

 proposal probability, 3.7, 5.8
model building, 11.27
molecular orbital, 14.4
moment expansion, 7.25
Mott insulator, 5.2
Mott transition, 10.16, 10.22, 11.22, 13.20
multi-orbital Hubbard models, 11.25
mutual information, 16.23
Møller-Plesset perturbation theory (MP2), 2.9

N

natural orbitals, 1.21
neural networks, 5.30
non-spherical Coulomb terms, 11.35
normalization integrals, 15.5

O

$O(3)$ transition, 16.24
occupation number representation, 7.15
one-body operators, 1.4, 7.10
orbital reversal, 14.15
order parameter, 17.25
orthonormalization, 7.26

P

pair correlation function, 1.4
pair density, 1.4
pair hopping, 14.19
particle-hole transformation, 5.17
path integral, 16.2
path integral ground state method, 16.9
Peierls' substitution, 13.8
phase problem, 6.18
phaseless AFQMC, 6.15
phonon, 15.7
photoemission, 1.17
plaquette-singlet state, 16.25
polaron, 15.4
population inversion, 14.14
probability density, 15.3
projector Monte Carlo, 3.10

Q

quantum annealing, 16.10, 16.27
quantum computing, 17.33
quantum criticality, 16.24
quantum impurity, 10.13

- quantum information, 9.5
- quasiparticle, 1.5, 10.19
 - wave function, 14.3
- quenched disorder, 16.26
- R
- random phase approximation (RPA), 2.2
- range-separated hybrids, 1.16
- reduced density matrix, 14.5
 - coherence, 14.7
 - population, 14.7
- renormalization, 9.2
- renormalized
 - second-order perturbation theory (rPT2), 2.11
 - singles excitations (rSE), 2.11
- resonating valence bond, 5.11
- resummation techniques, 15.17
- ring diagrams, 2.9
- RPA+rSE, 2.16
- S
- scanning tunneling microscopy, 14.15
- screened Coulomb potential, 1.16
- second quantization, 7.5
- second-order screened exchange (SOSEX), 2.9
- self-consistency, 11.22
- self-energy, 12.8, 13.7
 - Hubbard dimer, 11.6
 - local, 11.6, 11.8
- self-interaction problem, 1.16, 1.19
- shifted action, 15.19
- sign problem, 3.15, 6.11, 15.7
- single molecule junction, 14.2
- single particle rate matrix, 14.7
- single-particle Green function, 6.24
- skeleton diagrams, 12.12, 15.17
- Slater determinant, 5.15, 7.3
- spectral function, 7.24, 13.5
- spin crossover, 14.8
- spin stiffness, 16.22
- spin-orbit interaction, 11.32
- spin-statistics connection, 7.2
- stability of homogeneous electron gas, 1.11
- steepest descent, 5.24, 7.18
- superoperator, 14.11
- susceptibility, 16.20
- sweeping cluster update, 16.18
- symmetry factor of a diagram, 12.13
- T
- t - J model, 17.19
- target states, 9.3
- TDDFT, 6.7
- tensor diagrams, 8.4
- tensor networks, 9.2
- Thomas-Fermi kinetic energy, 1.8
- time-step error, 3.14
- transverse-field Ising model, 16.11
- two-body operators, 7.11
- U
- updates, 15.3
- V
- vacuum state, 7.5, 7.6
- valence-bond solid, 16.24
- van der Waals (vdW), 2.6
- variance minimization, 4.6
- variational Monte Carlo, 3.5, 4.2, 5.5
 - local energy, 4.3
 - zero-variance property, 5.6
- variational principle, 5.4
- von Weizsäcker functional, 1.8
- W
- Wannier functions, 11.28
- Wannier-Stark ladder, 13.21
- wavefunction optimization, 3.21, 4.5
 - linear method, 3.28, 4.9
 - Newton method, 3.23, 4.7
 - perturbative method, 4.11
 - stochastic reconfiguration, 3.28, 4.11, 5.26
- Weiss field, 11.23
- Z
- zero-variance estimator, 4.5

1. **The LDA+DMFT approach to strongly correlated materials**
Lecture Notes of the Autumn School 2011 Hands-on LDA+DMFT
edited by E. Pavarini, E. Koch, D. Vollhardt, A. Lichtenstein (2011), 420 pages
ISBN 978-3-89336-734-4
2. **Correlated Electrons: From Models to Materials**
Lecture Notes of the Autumn School on Correlated Electrons 2012
edited by E. Pavarini, E. Koch, F. Anders, and M. Jarrell (2012), 450 pages
ISBN 978-3-89336-796-2
3. **Emergent Phenomena in Correlated Matter**
Lecture Notes of the Autumn School on Correlated Electrons 2013
edited by E. Pavarini, E. Koch, and U. Schollwöck (2013), 520 pages
ISBN 978-3-89336-884-6
4. **DMFT at 25: Infinite Dimensions**
Lecture Notes of the Autumn School on Correlated Electrons 2014
edited by E. Pavarini, E. Koch, D. Vollhardt, A. Lichtenstein (2014), 450 pages
ISBN 978-3-89336-953-9
5. **Many-Body Physics: From Kondo to Hubbard**
Lecture Notes of the Autumn School on Correlated Electrons 2015
edited by E. Pavarini, E. Koch, and P. Coleman (2015), 500 pages
ISBN 978-3-95806-074-6
6. **Quantum Materials: Experiments and Theory**
Lecture Notes of the Autumn School on Correlated Electrons 2016
edited by E. Pavarini, E. Koch, J. van den Brink, G. Sawatzky (2016), 420 pages
ISBN 978-3-95806-159-0
7. **The Physics of Correlated Insulators, Metals, and Superconductors**
Lecture Notes of the Autumn School on Correlated Electrons 2017
edited by E. Pavarini, E. Koch, R. Scalettar, and R. Martin (2017), 450 pages
ISBN 978-3-95806-224-5
8. **DMFT: From Infinite Dimensions to Real Materials**
Lecture Notes of the Autumn School on Correlated Electrons 2018
edited by E. Pavarini, E. Koch, A. Lichtenstein, D. Vollhardt (2018), 480 pages
ISBN 978-3-95806-313-6
9. **Many-Body Methods for Real Materials**
Lecture Notes of the Autumn School on Correlated Electrons 2019
edited by E. Pavarini, E. Koch, and S. Zhang (2019), 520 pages
ISBN 978-3-95806-400-3