Reducing the Sign Problem with simple Contour Deformation

Christoph Gäntgen,[a,b,c,*] **Evan Berkowitz,**[a,b,d] **Thomas Luu,**[a,c,] **Johann Ostmeyer**[e] **and Marcel Rodekamp**[a,b,c,d]

[a]*Institute for Advanced Simulation, Forschungszentrum Jülich, 54245 Jülich, Germany*

[b]*Center for Advanced Simulation and Analytics (CASA), Forschungszentrum Jülich, 52425 Jülich, Germany*

[c]*Helmholtz-Institut für Strahlen- und Kernphysik, Rheinische Friedrich-Wilhelms-Universität Bonn, 53115 Bonn, Germany*

[d]*JARA & Jülich Supercomputing Center, Forschungszentrum Jülich, 54245 Jülich, Germany*

[e]*Department of Mathematical Sciences, University of Liverpool, Liverpool, L69 7ZL, United Kingdom*
E-mail: c.gaentgen@fz-juelich.de, e.berkowitz@fz-juelich.de,
t.luu@fz-juelich.de, m.rodekamp@fz-juelich.de, j.ostmeyer@liverpool.ac.uk

We apply constant imaginary offsets to the path integral for a reduction of the sign problem in the Hubbard model. These simple transformations enhance the quality of results from HMC calculations without compromising the speed of the algorithm. This method enables us to efficiently calculate systems that are otherwise inaccessible due to a severe sign problem. To support this claim, we present observables of the C20 and C60 fullerenes. Furthermore, we demonstrate that at a certain offset, the sign problem is completely lifted in the limit of large chemical potential.

---

*Speaker

1. Introduction

The Hubbard model named after John Hubbard originates from solid state physics. It is particularly well known for its ability to describe the transition from conducting to insulating metals. It describes a physical system as a spatial lattice with particles occupying the sites. These particles can *hop* from one site to another, which is described by a kinetic term and usually restricted to neighboring sites. Multiple particles sharing the same site interact with each other via a localized potential that can be repulsive or attractive. Since the system is fermionic the Pauli exclusion principle holds. Optionally a chemical potential can be added to influence the number of charged particles within the system. Here we only look at the fermionic Hubbard model with repelling potential which describes electrons on a lattice of ions

$$H = - \sum_{x,y} \kappa_{x,y} \left(a_{x\uparrow}^{\dagger} a_{y\uparrow} + a_{x\downarrow}^{\dagger} a_{y\downarrow} \right) - \frac{U}{2} \sum_{x} \left(n_{x\uparrow} - n_{x\downarrow} \right)^2 - \mu \sum_{x} \left(n_{x\uparrow} + n_{x\downarrow} \right). \tag{1}$$

Here κ is the hopping matrix describing the hopping amplitudes, U is the interaction potential and μ the chemical potential [4]. By applying a Hubbard-Stratonovich transformation on Equation 1 we can formulate an action with continuous auxiliary fields ϕ,

$$S = \sum_{x,t} \frac{\phi_{x,t}^2}{2\tilde{U}} - \log \det(M[\phi, \tilde{\kappa}, \tilde{\mu}] M[-\phi, -\tilde{\kappa}, -\tilde{\mu}]) \in \mathbb{C} \tag{2}$$

which allow lattice field calculations [8]. A tilde indicates parameters in lattice units, e.g. $\tilde{U} = U\delta$ where $\delta = \beta/N_t$ is the lattice spacing. The fermionic properties are encoded in the fermion matrices M, one for spin up and one for spin down (or equivalently, one for particles and one for holes). At non-zero μ there is no ergodicity problem with the *exponential* discretization that we use [11].

The term involving these matrices in Equation 2 can be complex for non-zero chemical potential or non-bipartite lattices. This property makes the Hubbard model susceptible to the *sign problem* that diminishes the convergence of Markov chain Monte Carlo calculations, making estimates of observables,

$$\langle \hat{O} \rangle = \frac{1}{\mathcal{Z}} \int \mathcal{D}\phi \, \hat{O}[\phi] \, e^{-S[\phi_n]} \approx \frac{1}{N} \sum_{n=0}^{N} \hat{O}[\phi_n] \; , \tag{3}$$

difficult, if not impossible, to obtain. Reweighting is thus required to handle the otherwise complex probability distribution $e^{-S[\phi_n]}/\mathcal{Z}$ appearing in every expectation value,

$$\langle \hat{O} \rangle = \frac{\langle \hat{O} e^{-\mathrm{i}S_I} \rangle_R}{\langle e^{-\mathrm{i}S_I} \rangle_R} \approx \frac{\sum_{n=0}^{N} \hat{O}[\phi_n] \, e^{-\mathrm{i}S_I[\phi_n]}}{\sum_{n=0}^{N} e^{-\mathrm{i}S_I[\phi_n]}} \; . \tag{4}$$

While including the imaginary phase in the observable and sampling according to the real part of the action is theoretically exact, the statistical convergence is exponentially slowed down by the average phase in the denominator of Equation 4 due to large phase oscillations. We call the absolute value of this phase,

$$\Sigma \equiv \left| \langle e^{-\mathrm{i}S_I} \rangle_R \right|, \tag{5}$$

the *statistical power* and use it to quantify the sign problem. A sign problem-*free* calculation has $\Sigma = 1$, whereas the worst sign problem occurs when $\Sigma \sim 0$.

Note that for this expectation value the real part of the action is used for the probability distribution, which is a non-holomorphic function. This means that its value depends on the chosen contour while the physical observables do not. Contour deformation according to Cauchy's theorem is one of many approaches to relieve the sign problem. We have observed that simple shifts into the complex space can greatly impact the sign problem. It does not have the same capabilities as more intricate transformations such as neural networks, but is simple to apply, does not affect the Hybrid Monte Carlo (HMC) and requires no Jacobian determinant which might blow up the volume scaling [6, 9, 10]. For these reasons we investigate the limits of constant imaginary shifts and develop methods for choosing one.

2. Advanced Offsets

When talking about an offset we are referring to an imaginary constant that is added to all components of the initially real auxiliary field $\phi \to \phi + i\phi_c$. The most basic, but also most reliable of those offsets occurs at the *tangent plane*. As the name suggests, it is the plane that is tangent to the critical point of the main Lefschetz thimble [1]. It is found by following holomorphic flow equations starting from $\phi = 0$ or solving the transcendental equation Equation 6, which we derived for the Hubbard model in [3], at finite imaginary time discretisation $\delta = \beta/N_t$,

$$\phi_0/\delta = -\frac{U}{N_x} \sum_k \tanh\left(\frac{\beta}{2}\left[\epsilon_k + \mu + \phi_0/\delta\right]\right) . \tag{6}$$

From this equation we see that the tangent plane scales inversely with the number of time steps N_t and is restricted to the range $[-\tilde{U}, +\tilde{U}]$ converging to its boundaries in the limit of $\mu \to \pm\infty$. We find the tangent plane consistently has a milder sign problem than the unmodified *real plane* calculation. Furthermore we proved that the sign problem vanishes on the tangent plane for $\mu\beta \to \pm\infty$ when μ is beyond a certain value.

The tangent plane is already known in the lattice community, but varying the offset quickly reveals that even greater improvements can be achieved. The next step is a next to leading order correction (NLO) that takes into account thermal fluctuations in the field ϕ. For this we expand the action around the main critical point and formulate an effective action Equation 7 [3]

$$S_{\text{eff}}[\phi_1] = S[\phi_1] + \frac{1}{2} \log \det \mathbb{H}_{S[\phi_1]} \tag{7}$$

which can be minimized to find ϕ_1. This offset leads to a further reduction in the sign problem most of the time.

A further reduction in the sign problem cannot rely on (quasi-)analytical methods alone and requires numerical iterative calculations. Fortunately the derivatives of the statistical power can be easily computed from an existing Markov chain. The derivatives derived in our paper are correct, but [2] provides a more elegant solution excluding terms that vanish analytically.

$$\frac{\mathrm{d}}{\mathrm{d}\phi_0} \left\langle e^{-iS_{I,\phi_0}} \right\rangle_{R,\phi_0} = \left\langle e^{-iS_{I,\phi_0}} \right\rangle_{R,\phi_0} \left\langle \frac{\mathrm{d}S_{R,\phi_0}}{\mathrm{d}\phi_0} \right\rangle_{R,\phi_0} \tag{8}$$

$$\frac{\mathrm{d}^2}{\mathrm{d}\phi_0{}^2} \left\langle e^{-iS_{I,\phi_0}} \right\rangle_{R,\phi_0} = \left\langle e^{-iS_{I,\phi_0}} \right\rangle_{R,\phi_0} \left(2\left\langle \frac{\mathrm{d}S_{R,\phi_0}}{\mathrm{d}\phi_0} \right\rangle_{R,\phi_0}^2 + \left\langle \frac{\mathrm{d}^2S_{R,\phi_0}}{\mathrm{d}\phi_0{}^2} - \frac{\mathrm{d}S_{R,\phi_0}}{\mathrm{d}\phi_0}^2 \right\rangle_{R,\phi_0}\right) \tag{9}$$
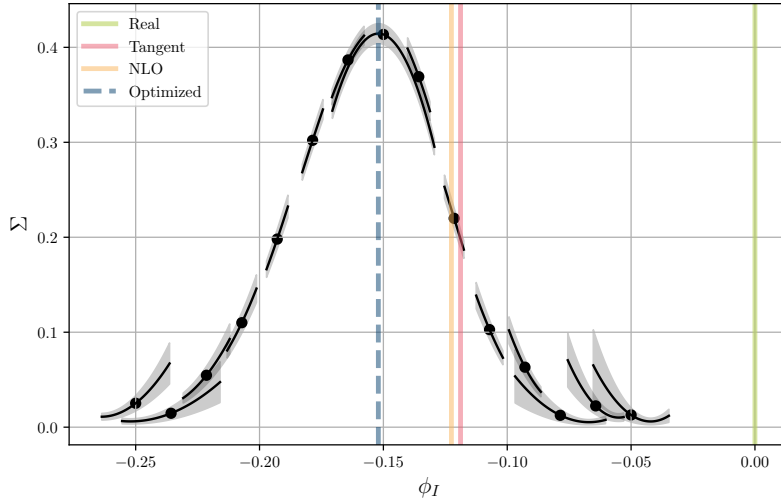
Figure 1: Statistical power with Taylor expansion to second order depending on imaginary offset ϕ_I. This is the C_{20} fullerene with $N_t = 16$, $\beta = 6$, $U = 2$ and $\mu = 1$. The vertical lines mark the named offsets in this example.

With those and a reasonable initial guess from NLO or tangent plane, this optimal offset that maximizes the statistical power can be approached iteratively. To estimate this *optimized plane* we simultaneously fit a Gaussian and its first two derivatives to Equation 5, Equation 8 and Equation 9 for all previously measured offsets. We have no formal proof for the statistical power forming a Gaussian around one offset, but judging from our observations it is a sufficiently accurate model for this optimization. This method converges more reliably and slightly quicker than the Newton-Rhapson routine we describe in [3]. Figure 1 illustrates the statistical power with derivatives for a range of offsets showing a clear peak structure motivating this research. The difference in statistical power between the marked offsets emphasizes the importance of the chosen contour. As the effective number of configurations scales with Σ^2 these offsets are orders of magnitude apart when it comes stochastic uncertainty.

3. Results

In this section we apply the developed techniques to a tetrahedron, C_{20}, C_{60} and a 72-site honeycomb lattice. The tetrahedron is small enough for exact diagonalization, allowing us to compare our stochastic algorithm analytic results.

First, the introduced offsets are quantitatively compared with each other. Figure 2 shows side by side the chosen shift at each μ and the resulting statistical power. It confirms our previous statement about the sign problem vanishing for $\mu \to \infty$. But in the interesting range we see that the improvement is greatly dependent on the parameters. At some μ the tangent plane is already the best we can find, at other μ the optimized plane surpasses it by magnitudes. The sign problem also gets worse for increasing U and β, for more information we refer to [3]. The sign problem also scales with volume, to compensate this effect we calculate larger lattices at smaller β [7].
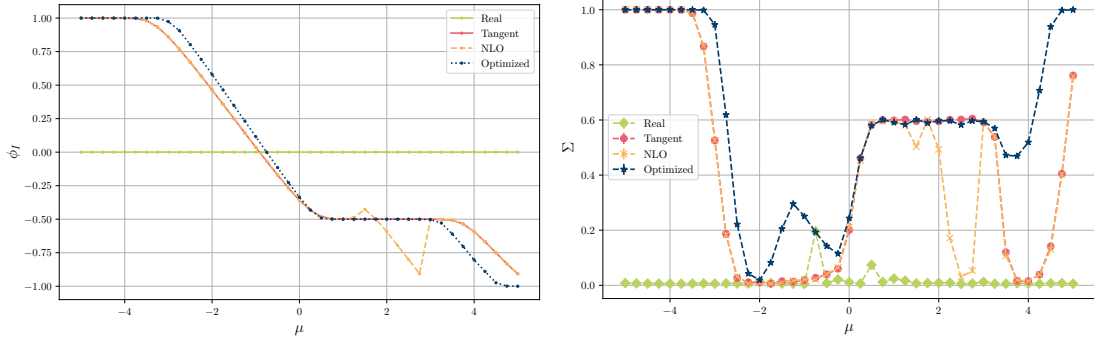
Figure 2: Comparing the introduced offsets and their corresponding average phase for the tetrahedron with $N_t = 16$, $\beta = 8$, $U = 2$.
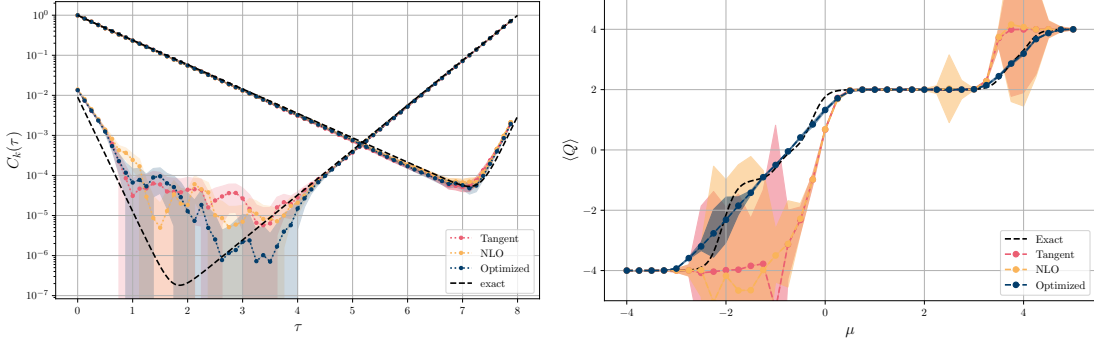

Figure 3: Observables of tetrahedron with $\beta = 8$, $U = 2$. Left: single particle correlator $C_k(\tau)$ at $\mu = 1$. Right: charge Q.

Furthermore, we show exemplary observables to highlight the quality difference that the choice of contour can cause even in this most basic form. The first observable is the single particle [1] correlation function given by Equation 10. It can be used to extract the low lying spectrum relative to the ground state, because at large β the smallest transition energies will dominate the slope at $\tau \to 0$ or $\tau \to \beta$:

$$C_k(\tau) = \left\langle a_k(\tau) a_k^\dagger(0) \right\rangle = \sum_{a,b} \left| \langle \Psi_a | a_k^\dagger | \Psi_b \rangle \right|^2 e^{-\tau(E_b - E_a)} e^{-\beta E_a} \tag{10}$$

The k-index labels the momenta rather the positions on the spatial lattice. For more details consider [5]. The systems we work with have degenerate energy states. We average the corresponding correlation functions and show in our figures only the resulting selection of unique momenta.

The second observable is the total charge expectation value Equation 11 which is an experimentally measurable quantity, depending on the chemical potential μ,

$$\langle Q \rangle = \left\langle \sum_x q_x \right\rangle = N_x - 2 \sum_k C_k(\tau = 0) \, . \tag{11}$$

---

[1]Here we use the Hubbard model in particle-hole basis where a particle are annihilated with $a_x = a_{x,\uparrow}$ and holes with $b_x = a_{x,\downarrow}^\dagger$. The creation operators are similar.
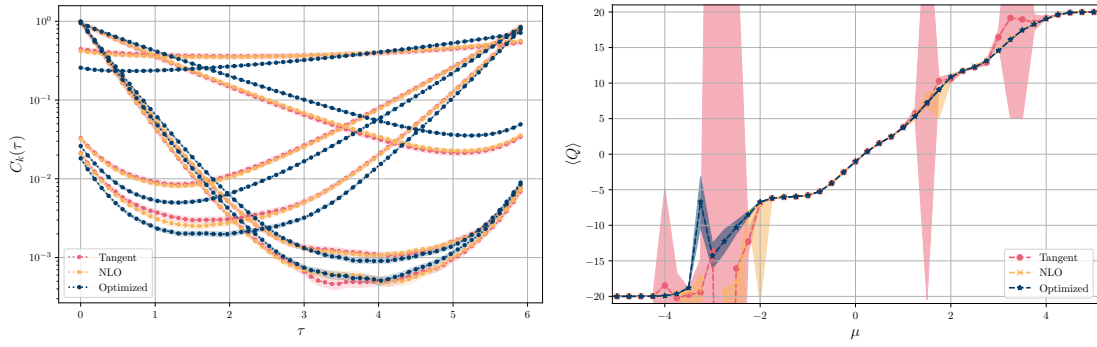
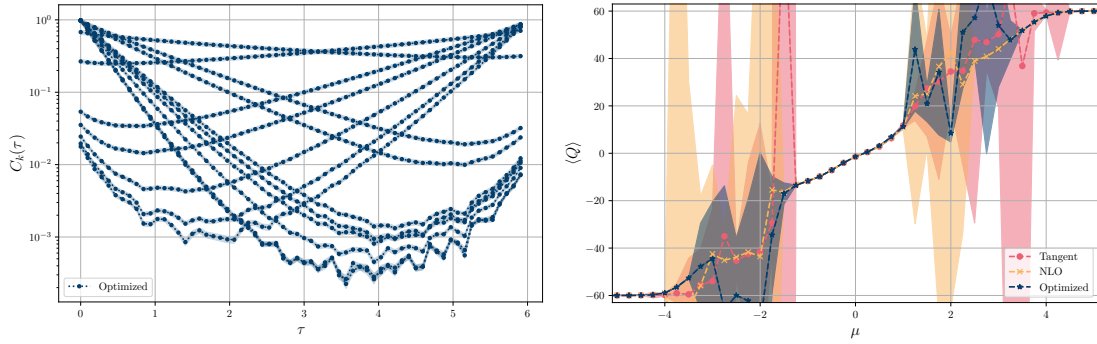**Figure 4:** Observables as in fig. 3 for C_{20} with $\beta = 6$, $U = 2$ and $\mu = 1$.

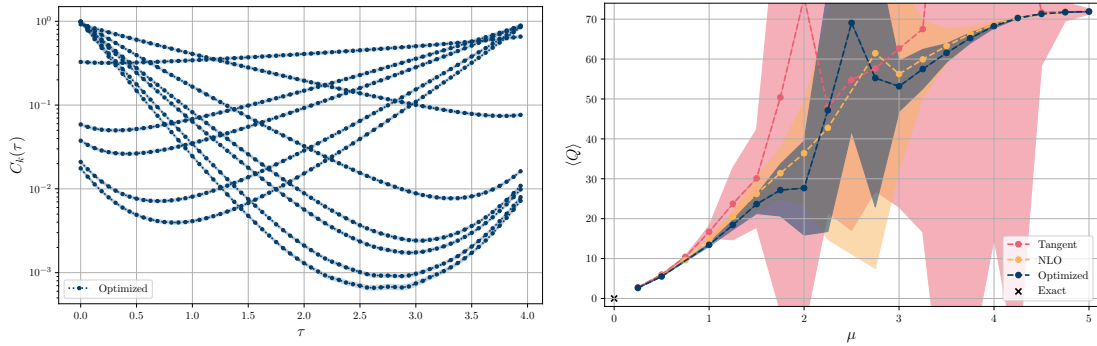**Figure 5:** Observables as in fig. 3 for C_{60} with $\beta = 6$, $U = 2$ and $\mu = 1$.

Figure 6: Observables as in fig. 3 for 72-site honeycomb lattice (6×6 unit cells) with $\beta = 4$, $U = 2$ and $\mu = 1$.

Comparing these charge expectation values with the ones from [3] shows that the afore mentioned Gaussian fits improved the stability of our optimization routine in areas where the sign problem is especially strong.

Energy Fits

It is clear from Equation 10 that the ground state has the greatest contribution, especially at large β. We fit an exponential to each correlation function shown in figures 3, 4, 5 and 6 to determine the particle excitation energy.

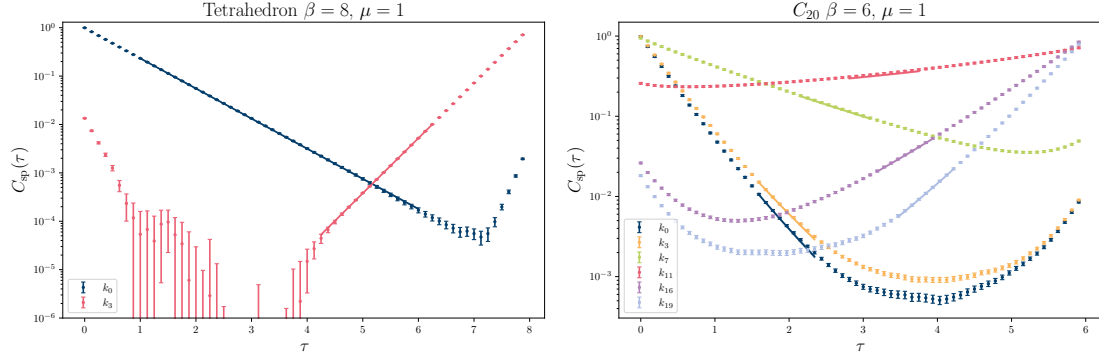Figure 7 illustrates our fits for tetrahedron and C_{20}. We identify negative slopes (i.e. positive

Figure 7: Single particle correlation functions with exponential fit. Both lattices have $U = 2$ and $N_t = 64$.

energies) as $\Delta N = +1$ transitions and positive slopes as $\Delta N = -1$ transitions, i.e. when the ground state has particle number N the excited state associated with a single particle correlation function has particle number $N + \Delta N$. In Figure 8 we show the resulting energies for all four lattices. Again we provide exact values for the tetrahedron that we calculated from exact diagonalization accounting for discretized time. Keep in mind that the spectrum is shifted due to the chemical potential and the ground state might not be at half filling.
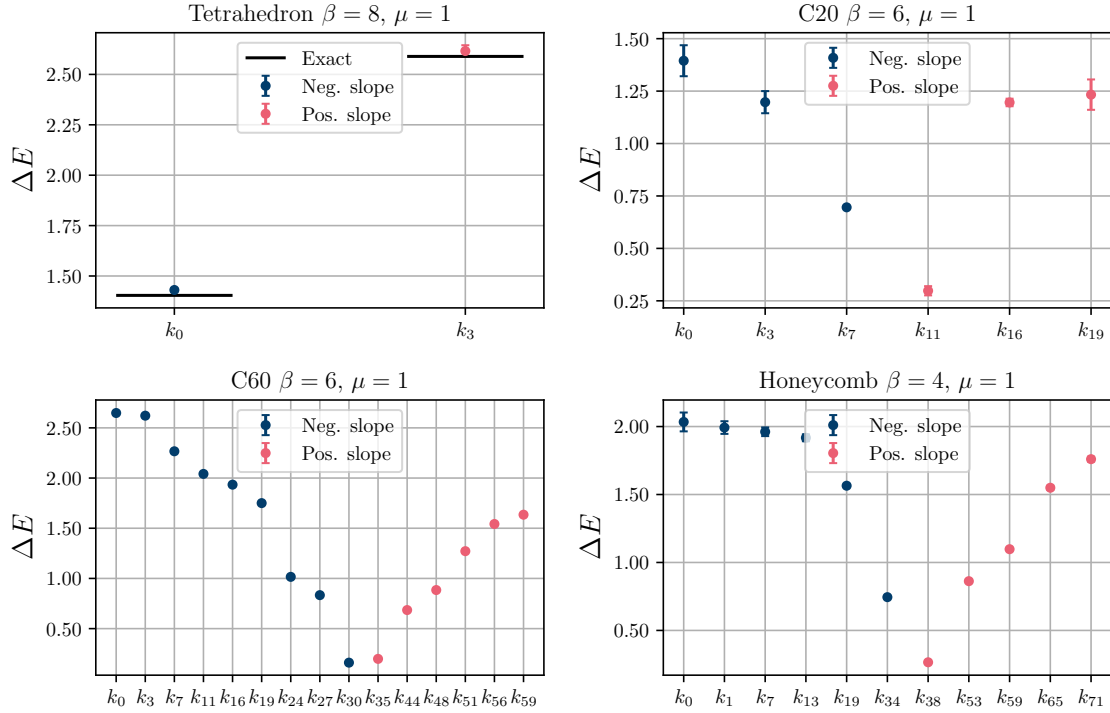

Figure 8: Energy states reachable by particle creation/annihilation from ground state. All systems have $U = 2$ and $N_t = 64$.

4. Summary & Outlook

We demonstrate that a smart choice of the integration contour can drastically reduce the sign problem in the Hubbard model without compromising numerical efficacy. Even a basic transformation like a constant shift can improve the quality of HMC calculations by orders of magnitude while entailing only negligible computational costs and small human effort. These methods cannot solve the exponentially hard sign problem, but they can expand the explorable parameter space at a given budget.

We present our methods of finding such a favorable offset and compare them with each other. All of them perform much better than the default calculation on the real plane. The recently developed optimization routine, utilizing fits to the statistical power and its derivatives, provides more stability than the initial version relying on the Newton-Rhapson method. Also we show the extraction of energy levels from the correlation functions and confirm that they align with the exact results on a small system.

In the future we intend to apply these methods to more carbon nano-systems to provide a detailed analysis of their electronic properties. We will also continue the exploration of simple contour deformations that do not induce a Jacobian with limiting volume scaling. Some current ideas that come into question are individual offsets for each component of ϕ and offsets combined with a Gaussian around 0. We hope that these investigations will deepen our understanding of sign optimizing manifolds and provide better starting points for the training of neural networks.

Acknowledgments

We thank Neill Warrington for many helpful discussions related to this work as well as Timo Lähde for his valuable comments. This work was funded in part by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through the funds provided to the Sino-German Collaborative Research Center "Symmetries and the Emergence of Structure in QCD" (NSFC Grant No. 12070131001, DFG Project-ID 196253076 – TRR110) as well as the STFC Consolidated Grant ST/T000988/1. This work is supported by the MKW NRW under the funding code NW21-024-A. We gratefully acknowledge the computing time granted by the JARA Vergabegremium and provided on the JARA Partition part of the supercomputer JURECA at Forschungszentrum Jülich.

References

[1] A. Alexandru, G. Basar, P. F. Bedaque, G. W. Ridgway, and N. C. Warrington. Monte carlo calculations of the finite density thirring model. *Physical Review D*, 95(1), jan 2017.

[2] A. Alexandru, P. F. Bedaque, H. Lamm, and S. Lawrence. Finite-Density Monte Carlo Calculations on Sign-Optimized Manifolds. *Phys. Rev. D*, 97(9):094510, 2018.

[3] C. Gäntgen, E. Berkowitz, T. Luu, J. Ostmeyer, and M. Rodekamp. Fermionic sign problem minimization by constant path integral contour shifts. 2023.

[4] J. Hubbard. Calculation of partition functions. *Physical Review Letters*, 3(2):77–78, 1959.

[5] J. Ostmeyer, E. Berkowitz, S. Krieg, T. A. Lähde, T. Luu, and C. Urbach. Semimetal–Mott insulator quantum phase transition of the Hubbard model on the honeycomb lattice. *Phys. Rev. B*, 102(24):245105, 2020.

[6] M. Rodekamp, E. Berkowitz, C. Gäntgen, S. Krieg, T. Luu, and J. Ostmeyer. Mitigating the hubbard sign problem with complex-valued neural networks. *Phys. Rev. B*, 106:125139, Sep 2022.

[7] K. Splittorff and J. Verbaarschot. Phase of the Fermion Determinant at Nonzero Chemical Potential. *Phys. Rev. Lett.*, 98:031601, 2007.

[8] R. L. Stratonovich. On a Method of Calculating Quantum Distribution Functions. *Soviet Physics Doklady*, 2:416, July 1957.

[9] M. Ulybyshev, C. Winterowd, and S. Zafeiropoulos. Lefschetz thimbles decomposition for the Hubbard model on the hexagonal lattice. *Phys. Rev. D*, 101(1):014508, 2020.

[10] J.-L. Wynen, E. Berkowitz, S. Krieg, T. Luu, and J. Ostmeyer. Leveraging machine learning to alleviate hubbard model sign problems. *Phys. Rev. B*, 103:125153, Mar 2021.

[11] J.-L. Wynen, E. Berkowitz, C. Körber, T. A. Lähde, and T. Luu. Avoiding ergodicity problems in lattice discretizations of the hubbard model. *Physical Review B*, 100(7), aug 2019.