

51st IFF Spring School 2020

Quantum Technology

Lecture Notes

Schlüsseltechnologien / Key Technologies

Band / Volume 210

ISBN 978-3-95806-449-2

Forschungszentrum Jülich GmbH
Peter Grünberg Institute (PGI);
Institute for Advanced Simulation (IAS);
Institute of Complex Systems (ICS);
Jülich Centre for Neutron Science (JCNS)

Lecture Notes of the
51st IFF Spring School 2020

Hendrik Bluhm, Tommaso Calarco,
David DiVincenzo (Eds.)

Quantum Technology

This Spring School was organized
by the Institutes PGI and IAS
of the Forschungszentrum Jülich
on 23 March until 3 April 2020.

In collaboration with Universities
and research institutions.

Schriften des Forschungszentrums Jülich
Reihe Schlüsseltechnologien / Key Technologies

Band / Volume 210

ISSN 1866-1807

ISBN 978-3-95806-449-2

Bibliographic information published by the Deutsche Nationalbibliothek.
The 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
PGI, IAS, JCNS and ICS
52425 Jülich
Tel.: +49 2461 61-6048
Fax: +49 2461 61-2410

Cover Design: Grafische Medien, Forschungszentrum Jülich GmbH

Source Cover Image: Quantum Flagship/H.Ritsch

Printer: Schloemer + Partner GmbH, Düren

Copyright: Forschungszentrum Jülich 2020

Distributor: Forschungszentrum Jülich GmbH
Zentralbibliothek, Verlag
52425 Jülich
Tel.: +49 2461 61-5368
Fax: +49 2461 61-6103
zb-publikation@fz-juelich.de
www.fz-juelich.de/zb

Schriften des Forschungszentrums Jülich
Reihe Schlüsseltechnologien / Key Technologies, Band / Volume 210

ISSN 1866-1807
ISBN 978-3-95806-449-2

The complete volume is freely available on the Internet on the Jülicher Open Access Server (JuSER)
at 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

- A1** Origins of Quantum Information Science
D. DiVincenzo, Forschungszentrum Jülich, Germany
- A2** Quantum Simulation
A. Browaeys, Université Paris-Saclay, France
- A3** Quantum Supremacy Using a Programmable Superconducting Processor
H. Neven and collaborators, Google Inc., USA
- A4** Quantum Optics and Quantum Information
J.-W. Pan, University of Science and Technology, China
- A5** Statistics for Quantum Estimation
D. Gross, University of Cologne, Germany
- A6** Quantum Measurement
D. DiVincenzo, Forschungszentrum Jülich, Germany
- A7** Quantum Networks with Superconducting Circuits
P. Kurpiers, P. Magnard, B. Royer, M. Pechal, S. Storz, J.-C. Besse, S. Gasparinetti, J. Heinsoo, Y. Salathé, A. Akin, A. Blais, A. Wallraff, ETH Zurich, Switzerland
- A8** Quantum Communication
C. Salomon, Ecole Normale Supérieure, France
- B1** Quantum Computing with Semiconductor Spins
*L. M. K. Vandersypen, Delft University of Technology, the Netherlands;
M. A. Eriksson, University of Wisconsin, USA*
- B2** Topological Quantum Computers
F. Hassler, RWTH Aachen University, Germany
- B3** Majorana Qubits
C. Marcus, Niels Bohr Institute, Denmark
- B4** An Introduction into Optimal Control for Quantum Technologies
*F. Wilhelm, S. Kirchhoff, S. Machnes, N. Wittler, Saarland University, Germany;
D. Sugny, CNRS-Université Bourgogne, France*
- B5** Programming a Quantum Computer: the Difficulty in Designing Optimal Quantum Algorithms
H. Buhrman, S. Patro, F. Speelman, University of Amsterdam, the Netherlands

Appendix

Preface

This is the first time in the 51-year history of the IFF Spring School that the subject has been a “Technology”. Can there be two weeks of intensive lectures by top scientists on a mere technology?

The decision just a few years ago, to designate **Quantum Technology** as a coherent societal endeavour, was taken after much deliberation within a large circle of working scientists. This endeavour is not a technology in a traditional sense, but is rather a unique intermingling of basic scientific insights, new capabilities demonstrated in laboratories, and an ambition to turn these unique capabilities into applications for the further advancement of technical capability on a number of fronts.

The “quantum revolution” has been declared multiple times over the last century. The first uncovering of quantum mechanics at the beginning of the 20th century was an intellectual revolution, albeit a small one confined to the circle of modern physicists. But by mid-century, quantum knowledge was power: first with the quantum properties of the nucleus, but much more extensively with the quantum behaviour of light and of electrons, entirely new capabilities arose. This first (quantum) technological revolution produced the information processing world of today.

But our second quantum revolution, which has prompted the birth of our new quantum-technology era, was waiting to happen because only a subset of the phenomena that are possible in the quantum world were harnessed in the first edition. We have simple, if inscrutable, names for some of these phenomena – “quantum entanglement”, “spooky action at a distance”, “quantum logic gates”. But it takes more than a glance to perceive what are the new things that are happening that produce our new quantum-technological era.

One helpful guide has been provided, and will be followed in the lecture scheme of our two weeks together. We speak of quantum technology as consisting of four pillars, which define the new application areas that are foreseen as a result of the fuller exploitation of the phenomenology of the quantum world:

Quantum Sensing and Metrology: Quantum mechanics defines the smallest detectable influence; here the aim is to produce detection devices working at this limit.

Quantum Communication: We can work towards networks that communicate not bits, but two-level quantum systems (qubits); these can yield absolute advantages in the security, privacy, and authenticatability of transmissions, and are important for interconnecting quantum computers:

Quantum Computing: If bits are replaced by qubits in processors, a new style of computing machine can come into being. For some problems it will have unrivalled algorithmic power.

Quantum Simulation: When appropriately specialized, quantum processors can efficiently mimic the dynamics of natural objects obeying the laws of quantum mechanics. The quantum simulator can sharpen our modelling abilities for complex molecular or solid-state quantum systems.

The 2020 IFF Spring School will cover all the main areas of quantum technology. Each lecturer, each a renowned research expert in the area, will give an in-depth presentation, with three hours devoted to each of the four pillars of quantum technology. These main “pillar” lectures, by Browaeys, Wallraff, Pan, and Salomon, will provide the core concepts of quantum technology, and the main current set of laboratory achievements. In the case of Pan, the “laboratory” is a full satellite-plus-earth stations quantum communications system!

Other lectures, extending into the second week, will expose the fundamental-science flavour that is still very present in the field. Gross, DiVincenzo, and Wilhelm-Mauch will give several current perspectives on how the basic physical quantum theory continues to be refined to deal with new problems. Buhrman will give the pure-theory perspective from the computer-science and mathematics direction. Hassler will discuss how the mathematical subject of topology has provided new concepts for condensed matter physics and, at the same time, for the theory of quantum computing.

The real technological aspect of quantum technology is also a key component of our lectures. The forward focus of the Jülich Research Center is seen in the several perspectives that are given on the progress towards building real solid-state quantum computers: Vandersypen will touch on what has become possible with semiconducting qubits, Marcus will give the Microsoft view of developments in topological-qubit systems, and, following on Wallraff’s lectures, Neven will show Google’s active program, which currently leads the world in producing the most capable superconducting qubit processor.

While these devices are still some way from changing the world of information processing, they have produced an enormous leap in the degree of commercial interest in quantum technologies. As a first for the IFF Spring School, we will devote a whole day in the second week to industry, with presentations from active contributors from small and large companies. Students will hear some frank views about where all this is heading, will have a chance to ask questions directly to the experts, and can think about career moves into quantum technology!

We would like to close with some thanks: First, thank you to all the lecturers who have worked long in advance to provide lectures notes for the school (in traditional and non-traditional formats), and who have agreed to, in most cases, give in-depth, three-hour presentations of their subject. We are very grateful to the board of directors of the Forschungszentrum Jülich for its continuous organizational and financial support of this school and of this book of lecture notes. Finally, our special thanks go to Ms. Barbara Daegener and Ms. L. Snyders for the general management and the help in compiling this book.

David DiVincenzo
Hendrik Bluhm
Tommaso Calarco

January, 2020

A 1 **Origins of Quantum Information Science¹**

David P. DiVincenzo

Peter Grünberg Institut

Forschungszentrum Jülich GmbH

Contents

1	What is Quantum Information Science?	2
2	Pioneers of Quantum Information Science	4
3	Postscript: the 2012 Nobel Prize in Physics	11

Lecture Notes of the 51st IFF Spring School “Quantum Technology” This is an Open Access publication distributed under the terms of the Creative Commons Attribution License 4.0, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. (Forschungszentrum Jülich, 2020)

These notes are reprinted from: Lecture Notes of the 44th IFF Spring School "Quantum Information Processing", lecture A1, copyright Forschungszentrum Jülich 2013. Used by permission.

1 What is Quantum Information Science?

In this volume you will learn a great deal about current developments in Quantum Information Science. I would like first to define the term, since it didn't exist, with its current meaning, fifteen years ago. The term tries to convey the idea that an entirely new field has been created (or, more accurately, is in the making) that, while it has strong affinities to Physics, Informatics, Computer Engineering (and, less obviously, to pure Mathematics, to Electrical Engineering, Materials Science, and Chemistry), it can stand as its own subject. As the title of this school we use the related phrase *Quantum Information Processing* (QIP), because there is a strong emphasis here on the devices that will manipulate quantum information. Beware: QIP is also the title of one of the longest-running series of international conferences in this field, but the QIP conferences have evolved to mainly cover the specialty of theoretical quantum computer science.

I have already used a few additional phrases worth noting. *Quantum information* is, one might say, the descriptor of the raw material of the subject, the stuff that is actually processed or communicated. *Theoretical quantum computer science* concerns itself with rigorous formal properties of quantum information; but, the acronym QCS is in use in the USA as the name of a major national research program that is oriented towards rather more practical computer science issues. *Quantum Computing* is sometimes used as a generic phrase for the whole field, but is more properly reserved for the part of the field concerned with the implementation of *Quantum Algorithms*, a central but in fact quite difficult specialty in the field. *Quantum Communications* and *Quantum Cryptography* are two more terms that are more-or-less self defining.

Returning to Quantum Information Science itself, it is worth quoting in full a carefully thought out definition that was created in a workshop with this name sponsored by the National Science Foundation of the United States in 1999 [1]:

Quantum information science (QIS) is a new field of science and technology, combining and drawing on the disciplines of physical science, mathematics, computer science, and engineering. Its aim is to understand how certain fundamental laws of physics discovered earlier in this century [20th] can be harnessed to dramatically improve the acquisition, transmission, and processing of information.

You will hear in this school our progress in achieving these dramatic improvements. Indeed, it must be admitted that we still have not built the quantum computer, or anything other than the most rudimentary quantum communications system. The things we have achieved in this new science are less tangible, but arguably more profound. Here is my personal list of present-day successes of QIS:

New ways have arisen of teaching and learning the quantum theory. Many computer science curricula include lectures in quantum theory today, something that was almost unheard of and unthinkable fifteen years ago. This shows that quantum mechanics is more than what its originators thought it was: it is a theory of knowledge, a theory of the transformations of information, in addition to its being a theory of matter and energy. Lectures in theoretical physics are also

gradually adopting both these points of view. On the research side, it has enlivened discussions of what the quantum world is (does quantum mechanics require many worlds) and what it could be (there is a theory called quantum boxes, which is like quantum theory but is even more nonlocal in the correlations are possible between separated systems).

A mathematically sophisticated and rigorous information theory has been developed around quantum mechanics. This refers to specific technical developments that have been decisively influenced by QIS concepts. I have two rather different things in mind here: first, the highly developed entropic characterizations of quantum properties and protocols that have arisen in a large literature. They are natural outgrowths of the reasoning introduced by Shannon in communication theory; a great variety of “Shannon entropies” are used to quantify a large number of things, for example, the carrying capacity of a noisy qubit channel. Second, there is a surge of interest in QIS in the areas of quantum gravity and astrophysics. I am not competent to comment in any detail about these developments, but evidently quantum entanglement has been a very useful concept in making more sense of Hawking radiation, and about the question of whether information is truly lost in a black hole. This seems to be still going strong, as you could read in a current semi-popular account of “firewalls” in black holes [2].

We have a new mindset for discussing quantum experiments. QIS has defined a set of metrics that we apply confidently and profitably to experiments from many different fields of experimental quantum physics and chemistry. We can talk about process fidelities, leakage rates, and decoherence times for experiments from a vast range of fields – nuclear magnetic resonance, quantum dots, ion traps, or optical lattices. We have both a good grasp of the fundamentals of this work, and we have a reasonably accurate assessment of how these different systems progress towards achieving real quantum information processing. In some fields, QIS has completely turned around the idea of what is interesting. In the 1980s, it was fashionable in quantum dot research to design experiments to see how *fast* the relaxation times could be made. Since the advent of QIS, we have recognized that it is also scientifically interesting for these experiments to investigate how *slow* relaxation can be. Slow is good, if what is slow is decoherence.

Quantum experiments have gone in entirely new directions. Building a quantum computer is really hard, in the 1990s we realized that, despite a century of remarkable advances, we had only taken the first baby steps towards such a machine. This has motivated experiments to tackle quantum phenomena in a remarkably ambitious new way. We had to think about ion traps, not just for studying single ions, but for large arrays of ions that could be shuttled around, recooled, and coupled at will. This exciting science would not have happened without QIS. In quantum dot research, we have achieved not only one-by-one control of trapped electrons, not only precise temporal control of exchange couplings, not only exquisite control of the nuclear spins in the crystal lattice, not only intricate coherence-recovery protocols, not only precise and accurate in-place quantum measurements, but all of these things together in a single complex device. In superconducting devices, we have gone in ten years from the barest hint of quantum coherence to fidelities that rival those in atomic systems, with designed-to-order qubits with tailored coupling networks on the way. Experiments in this area are now driven explicitly by architectural ideas for a fault-tolerant quantum computer.

In the remainder of this introductory chapter, I will tell the story, as I have seen it and reconstructed it, of how we got to where we are. Of course, a tale of personalities, as I will tell it, is intrinsically limited. It does not by itself convey the detailed implications of the discoveries, and it is certainly very unfair to the large number of other persons that took part in these



(a) Erwin Schrödinger.



(b) Rolf Landauer.

Fig. 1: Schrodinger clarified the ideas of correlations as introduced, by Einstein, Podolsky, and Rosen, introducing the notion of quantum entanglement. Landauer articulated ideas of the information-carrying capacity of physical systems.

developments and played many different roles, major and minor, in moving them along.

2 Pioneers of Quantum Information Science

How did QIS happen?

The inventors of modern quantum theory in the 1920s certainly did not think of a quantum computer. It would have been anachronistic for them to have done so: it should be recalled that the whole notion of an automatic computing device barely existed at the time. Also, the concept of the "bit" was hardly in use – such mechanical calculators as existed in those times implemented base-10 arithmetic in their gearworks, rather absurdly imitating human calculating hardware. But there was a further reason why the first generation of quantum theorists did not think about quantum computers: being just barely understood for the first time, quantum theory was only comprehended as a special theory of the very small – nuclei, electrons, molecules, the smallest quanta of light. The idea that one could manufacture something that intentionally followed chosen features of the quantum theory was utterly foreign to them. Thus, even when there arose a next wave of scientists like von Neumann and Turing, both at the forefront of new developments in computing machines and quite aware of current developments in quantum mechanics, the connection of information processing and quantum theory was barely made.

I say "barely" because there was a significant intellectual stir triggered in 1935 by the publication of the paper of Einstein, Podolsky, and Rosen [3] (EPR), which put into print the misgivings

that Einstein had had for many years about quantum theory. This paper featured the strangeness of correlations between two observers sharing parts of a quantum state, which convinced Einstein that quantum theory was incorrect or incomplete. Others were not convinced. Of all the popular and scholarly reaction in 1935 to the EPR paper, in my opinion the most relevant and interesting from our current perspective is that of Schrödinger (Fig. 1(a)), in a set of immortal papers in 1935-6 [4]. Schrödinger gave us the name for correlations in quantum states that we still use in English and in German (entanglement/Verschränkung). Schrödinger accepted quantum theory, and accepted that it was very weird – he famously gave us his Cat in these papers. Famous as they are, I think that these papers deserve to be more famous. In the technical side, he develops the relation of mixed quantum states and the pure states from which they must arise – the work presages by many decades the Uhlmann theorem that you will learn in quantum information theory. He rather poetically expresses himself about quantum measurement, thus: *Es ist ein Unterschied zwischen einer verwackelten oder unscharf eingestellten Photographie und einer Aufnahme von Wolken und Nebelschaden./ There is a difference between a shaky or out-of-focus photograph and a snapshot of clouds and fog banks.* A modern, technical translation would run: there is a difference between a weak quantum measurement on a pure quantum state and a strong quantum measurement on a mixed state. This insight is very central to what you will learn in quantum information theory about measurement.

Schrödinger, in his discussion of entanglement and how it is created by interaction and then separation of two subsystems, in fact gives us the rudiments of a theory of communication with quantum systems. There was perhaps only one other instance of an information-theoretic argument in physics before this, in the discussion of the paradox of the Maxwell demon, for example, in the Habilitationsschrift of Szilard [5]. In fact, both Schrödinger's entanglement and Szilard's engine played some role in the next few steps in the development of quantum information theory, as we will now see.

These developments were, however, a long time in coming. From the 1930s to the 1960s, practitioners of quantum science were solving momentous practical problems at a furious pace, and there was not very much room for "philosophy". Computer science underwent tremendous development in those times, the classification of the hardness of mathematical problems from an algorithmic perspective was well underway; and, of course, we became comfortable with the concept of the "bit". Two contributors in quantum physics in those times are worth mentioning: David Bohm, in the midst of many practical contributions to quantum mechanics, pointed out the canonical example of entanglement of two qubits that we use today, states like $(|00\rangle + |11\rangle)/\sqrt{2}$, loosely called "EPR pairs". The role of the other worker, John Wheeler, is a little harder to pin down. It is clear that in this period, Wheeler became convinced that quantum theory was some sort of information science. We have from this period his catchphrase, "it from bit" [6]. His influence on QIS becomes much clearer in a later period, in the group of other scientists with whom he surrounded himself.

In 1961 Rolf Landauer (Fig. 1(b)) ushered in the contemporary period of the physics of information with his article [7] on heat generation in computing. His $k_B T \ln 2$ formula for dissipated energy per logical operation, an important milestone today in digital computing, was understood to come from the *discarding* of logical information during the performance of boolean logic operations. The refinement of this understanding in the following decade [8] was crucial in the acceptance of quantum mechanics as a valid platform for a computing theory.

The discovery by John Bell of his inequalities in 1964 [9] gave an entirely new motivation



(a) Alexander Holevo.



(b) Yuri Manin.

Fig. 2: Earliest contributors in the U.S.S.R. to quantum information.

for doing quantum experiments. He showed that classes of non-quantum physical theories could be straightforwardly ruled out by the measurement of correlations of entangled, separated quantum particles. The experiments that resulted (e.g., [10], and especially later [11]) were clear forerunners of experiments done today in quantum communication; they evolved towards much more reliable and copious sources of EPR quantum states, in the form of photon pairs created by spontaneous downconversion [12]. A very readable, semi-technical account of these developments may be found in [13]. In fact, even more powerful physical resources for the implementation of quantum computation were gestating in these times: see my Postscript to this chapter.

While much had happened up to the year 1970, no one at that time recognized any trend in the making. The developments of the next decade were still isolated; but we see in them the seeds from which the current trends clearly sprang up.

A singular individual in the Soviet Union invented the modern, rigorous information-theoretic treatment of quantum channels. This was Alexander Holevo (Fig. 2(a)), whose “Bounds for the Quantity of Information Transmitted by a Quantum Communication Channel” [14] in 1973 ushered in the application of Shannon entropies to the quantum questions. First Holevo, and very gradually others, added to this corpus of ideas, which is a large and flourishing area of analysis today.

Even more isolated was the advent of Steven Wiesner’s observations (Fig. 3(a)) on “conjugate coding” – the potential applications of preparing and measuring a two-state quantum system in one of two bases defined by non-commuting operators. Wiesner was quite serious and explicit in using this feature of quantum mechanics for cryptographic purposes, he immediately conceived of the possibility of money that could not be counterfeited. It would perhaps not be useful to date his contribution to around 1970, as it was known to hardly any other person at that time, appearing in publication only in 1983 [15]. But, crucially, Wiesner’s ideas were well known and appreciated by his University classmate Charles Bennett (Fig. 3(b)), who will figure prominently in the remainder of this history. In the following 25 years Bennett worked very effectively to make Wiesner’s concepts known, and to greatly extend them to produce some of



(a) Steven Wiesner.



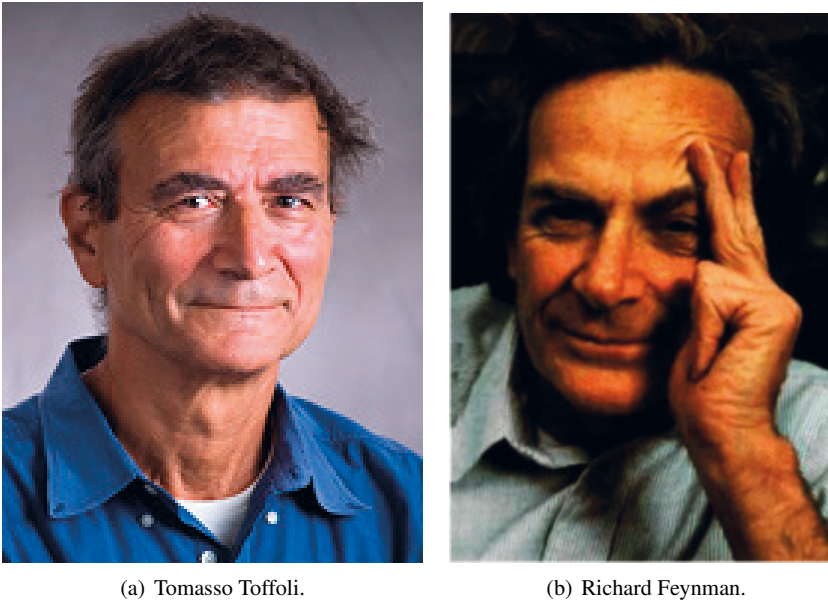
(b) Charles H. Bennett.

Fig. 3: Seminal American contributors to quantum information.

the best-known discoveries in QIS.

Independent of Wiesner's ideas, Bennett in 1973 contributed his revolutionary correction [16] to Landauer's arguments about the energy consumption in computing, by showing that there is *no* lower limit on how little heat needs to be dissipated during the computational process; Landauer's $k_B T \ln 2$ is avoided by efficiently eliminating the need to erase workspaces. This discovery of *reversible computation* initially had no impact on quantum research – but by the time of Bennett's review of the subject in 1982 [17] it had begun to. In the meantime, Bennett's result, which had been formulated for the Turing machine model of computation, was extended to the logic-gate model of computation in a series of papers by Toffoli (see Fig. 4(a)) and Fredkin [18]. Toffoli's gate was his concept for an all-purpose Boolean logic gate that could perform the logical OR or the logical AND as desired, and would be reversible – the inputs could be reconstructed from the outputs. In its minimal form, as everybody today knows, it involves three bits. But it is only known to us as the “Toffoli gate” on account of Richard Feynman (Fig. 4(b)).

Feynman had been speculating about the physics of the very small for a long time; his 1959 speech “There's Plenty of Room at the Bottom” [19] is considering a founding manifesto of the field of nanoscience. He connected his own thoughts about quantum physics with his recent awareness of the work of Landauer, Bennett and Toffoli to put in print in 1982 [20], for (almost) the first time, the phrase *quantum computer*, and some (but not all) of the thoughts that we associate today with this phrase. I say “almost” because our Russian friends tell us that the well known mathematician Yuri Manin (Fig. 2(b)) made some reasonable speculations about the idea of a quantum computer in 1980 [21]. I have never seen this reference, and I don't believe it



(a) Tommaso Toffoli.

(b) Richard Feynman.

Fig. 4: These two contributed essential ideas to quantum computing in the early 1980s.

has been read outside the Russian community. Feynman's vision of a quantum computer arose from his understanding that determining the properties of many-particle quantum systems is computationally intractable; thus he introduced the quantum computer as a computer simulator. In his only follow-up to his 1982 work [22], Feynman groped towards a more definite model for the quantum computer, which he built as an extension of the idea of reversible computing, quantum evolution being a more general version of reversible dynamics. It was in this work that he gave the world Toffoli's name for the three-bit logic gate. But Feynman's final work was also notable for the crucial element that he (and others like Benioff) missed: he failed to recognize that *entangled* quantum states, and entangling logic gates, would be crucial for the real quantum aspects of this new computer. This lack was made up in the next seminal work towards quantum computation.

Before getting to this, I might mention that the set of papers I have been citing from 1982, in the International Journal of Theoretical Physics, were all written for a conference, on "The Physics of Computation" held at MIT in May, 1981. There were about 60 participants, including Bennett, Feynman, Fredkin, Landauer, and Toffoli, as well as others like Dyson, Büttiker, Greenberger (of the Greenberger, Horne, Zeilinger (GHZ) state), and prominent computer architects Hillison and Cocke. One should not get the impression that quantum computing then became a large-scale enterprise; the conference was not repeated until about a dozen years later. Important developments began to occur with some regularity, but I can think of no quantum information papers in 1983 or in 1987. But these were perhaps the last years for which this could ever be said.

Not present at the 1981 conference, but soon to contribute decisively to further developments, was the active group around Wheeler at that time; Fig. 5 shows part of his group then. Two of those pictured, Wootters and Zurek, gave the essential formal underpinning for the Wiesner



Fig. 5: The group of John A. Wheeler attending the NATO Advanced Study Institute on Quantum Optics and Experimental General Relativity, held August 16-29, 1981, in Bad Windsheim, Bavaria, Federal Republic of Germany. Left to right are William Wootters (see Fig. 7), Kip Thorne, Wheeler, Wojciech Zurek, and William Unruh.

ideas, which were just beginning to be known to a wider circle. They proved (as Dieks did independently at the same time) that a single, unknown quantum state cannot be copied, or “cloned” [23]. This work provided a simple physical insight into why quantum systems could be interesting for cryptography, where privacy is of paramount importance. Bennett had been continuing to follow up on Wiesner’s concepts, culminating in the announcement in 1984, with cryptographer Gilles Brassard (see Fig. 7), of a full protocol for distributing keys for secret message transmissions, as well as secure, remote coin tossing [24]. Even at this juncture, it was still some years before “BB84” was a widely known concept.

The next, and perhaps *the*, decisive step in the further articulation of the quantum computer came also indirectly out of the Wheeler group. Another member of the group in Texas shortly after the time of Fig. 5 was David Deutsch (Fig. 6(a)), who came on a postdoctoral fellowship from Oxford. Deutsch’s 1985 paper [25] took the crucial step of generalizing Bennett’s reversible computing to the quantum case, taking the step which had eluded Feynman. Working with Bennett’s original reversible Turing machine approach, Deutsch makes a small change to its definition; this change permits the bits of the machine to evolve into entangled states. He exhibits a Turing machine program for the “EPR experiment”. Further, he points out that the equivalence concept that had developed in computer science at that time, that all reasonable physical implementations of a computing device had the same scaling of efficiency up to “small” polynomial factors, would in all likelihood not apply to his simple quantum generalization. The quantum computer would be in a different complexity class. In this paper Deutsch is silent on the connection that his discovery had to ideas in the Texas group, but one tiny detail gives it away: he acknowledges grant PHY8205717, which a short investigation reveals to be Wheeler’s National Science Foundation grant in the early 1980s.



(a) David Deutsch.



(b) Peter Shor.

Fig. 6: The workers who began and completed the revolution in quantum algorithms.

Deutsch's follow up in 1989 [26] took several more giant steps, and put the field irrevocably on that track that it has gone on ever since. This paper introduced the circuit model of quantum computing that we are most familiar with today, provided us with the notion of a universal set of quantum gates (he gave a simple generalization of the three-bit Toffoli gate – simpler universal sets came later), discussed the controlled NOT gate and its roles in producing entanglement and performing noninvasive quantum measurements. Finally, it reinforced the notion of his 1985 paper of quantum computing being in a different class from ordinary computation, by introducing a problem for which a quantum algorithm provably took fewer steps than on any classical machine! This algorithm was “useless” in itself, but was revolutionary in what it made possible.

Indeed, the timeline from Deutsch's toy algorithm to Peter Shor's (Fig. 6(b)) renowned factoring algorithm was, by the standards of the rest of this history, very fast and sure. (I will not give a reference, since the student can find this in any textbook.) Also, a much wider circle of workers soon came into the picture. The Deutsch-Jozsa algorithm (1992) [27] was followed by the Bernstein-Vazirani algorithm from Berkeley workers (1993) [28], followed within months by the Simon algorithm (Montreal, 1994) [29]. Shor's work at Bell labs came a few months afterwards. I will not take my story into this era; the prehistory of QIS, at this point, was complete. To say that the discovery of the factoring algorithm was inevitable once Deutsch's algorithm came to light perhaps has some truth to it, but it unfairly minimizes the great innovations of these five years, and Shor's remarkable synthesis of the ideas of others with sophisticated knowledge of the structure of integer factoring.

It should also not be understood that the only activity of the time was the advancement of quantum algorithms. In fact, the founding ideas of QIS were being developed on many fronts by the early 1990s. I will not tell the story of the considerable insights of quantum cryptography, of the further innovations in quantum communication theory, in the developing concepts in quantum measurement and in quantum decoherence, and the fortuitous dramatic improvements in the manipulation of single quanta in the lab (but see Postscript), that occurred in these few years. Wootters and Schumacher coined the term *qubit* in this time period. I must say a word, however,



Fig. 7: The discoverers of quantum teleportation; photo taken c. 2004, 11 years after the discovery. Left to right: Gilles Brassard, Richard Jozsa, William Wootters (see Fig. 5), Claude Crépeau, Charles Bennett (see Fig. 3), and Asher Peres.

about quantum teleportation [30] (Fig. 7), which was a synthesis of many ideas that I have been mentioning: EPR states as a kind of communication channel, the conjugate bases of Wiesner, quantum logic circuits, and entangling quantum measurements. Within a few years quantum teleportation entered quantum computer theory in a big way, as a predecessor of quantum error correcting codes and fault tolerant quantum computing; of course, in many ways, this 1993 paper has become an iconic representative of our entire field.

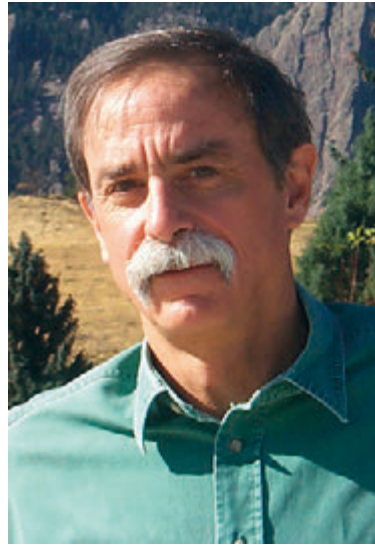
After 1994 the field took on a different character, with the involvement, in succession, of many additional communities in theoretical and experimental physics. Much of our school will concentrate on a rather later wave, the advent of solid state devices as high-quality qubits. With scarcely an hour going by without a new paper appearing on some aspect of QIS, it is clear that the telling of the post-1994 history is a very much harder job than the one I have done in my story of the pioneers of the field.

3 Postscript: the 2012 Nobel Prize in Physics

After the construction of the above history, the 2012 Nobel Prizes for Physics were announced. While the laureates Serge Haroche and David Wineland (Fig. 8) did not figure in my discussion, they are in fact the outstanding representatives of one of the important specialties that got pulled into the QIS orbit almost immediately after the advent of Shor's factoring algorithm. While Bell's work in the 1960 had already caught the attention of the quantum optics community, and downconverted photon pairs became a readily available resource for simple experiments on



(a) Serge Haroche.



(b) David Wineland.

Fig. 8: Nobel laureates in physics for 2012: they are experimentaists who brought single quanta under control in atomic physics in the 1980s.

entanglement, workers like Haroche and Wineland progressed to qualitatively more powerful control of single quantum systems. As we will hear in this school, they worked on controlled entanglement and measurement, Haroche using Rydberg states (highly excited states) of atoms passing through microwave cavities, Wineland using vibrational and other states of individual ions in Paul-style traps.

I am delighted and proud that the Nobel Committee had honored their outstanding work, and that it has seen fit to present their work to the public in the language of its contributions to the further development of Quantum Information Science.

Haroche and Wineland did not work on qubits or quantum gates – the terms had not been invented in the 1980s when they began their work. But by the mid-90s, their systems became prime testing ground for the new ideas from QIS. Haroche, especially, was not an unabashed supporter of quantum information concepts at that time; he wrote a commentary entitled “Quantum Computing: Dream or Nightmare” [31]. But he engaged in serious and fair debates on the future of our subject at that time, being an especially spirited participant in the discussions at the end of our Program on Quantum Computing at the Institute for Theoretical Physics (now Kavli) at Santa Barbara in 1996. I would like to end by quoting Preskill’s final remarks during that debate, as set down in his written lecture notes of the time [32]:

Serge Haroche, while a leader at the frontier of experimental quantum computing, continues to deride the vision of practical quantum computers as an impossible dream that can come to fruition only in the wake of some as yet unglimped revolution in physics. As everyone at this meeting knows well, building a quantum computer will be an enormous technical challenge, and perhaps the naysayers will be vindicated in the end. Surely, their skepticism is reasonable.

But to me, quantum computing is not an impossible dream; it is a possible dream. It is a dream that can be held without flouting the laws of physics as currently understood. It is a dream that can stimulate an enormously productive collaboration of experimenters and theorists seeking deep insights into the nature of decoherence. It is a dream that can be pursued by responsible scientists determined to explore, without prejudice, the potential of a fascinating and powerful new idea. It is a dream that could change the world. So let us dream.

References

- [1] “Quantum Information Science: An Emerging Field of Interdisciplinary Research and Education in Science and Engineering,” Report of the NSF Workshop, October 28-29, 1999, Arlington, Virginia, <http://www.nsf.gov/pubs/2000/nsf00101/nsf00101.htm>; An update of this can be found in “Quantum Information Science,” Report of the United States National Science and Technology Council Workshop, 23-25 April 2009, Tysons Corner, Virginia, <http://calyptus.caltech.edu/qis2009/QIS-Workshop-Report-1-July-2009.pdf>.
- [2] J. Polchinski, “Black Holes, Complementarity, and Firewalls,” <http://blogs.discovermagazine.com/cosmicvariance/2012/09/27/>.
- [3] A. Einstein, B. Podolsky, N. Rosen, “Can quantum-mechanical description of physical reality be considered complete?,” *Phys. Rev.* **47** (1935), 777 - 780.
- [4] Erwin Schrödinger, “Die gegenwärtige Situation in der Quantenmechanik,” *Naturwissenschaften* **48**, pp. 807-812; **49**, pp. 823-828; **50**, 844-849 (1935); “Discussion of probability relations between separated systems,” *Proceedings of the Cambridge Philosophical Society*, **31**, (1935), 555-563; “Probability relations between separated systems,” *ibid.* **32**, (1936), 446-452.
- [5] Leo Szilárd: “Über die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen,” *Zeitschrift für Physik* **53** (1929), pp. 840 - 856.
- [6] This may be seen in much later writing, e.g., John A. Wheeler, “Hermann Weyl and the Unity of Knowledge,” *American Scientist*, Vol. 74, July-August 1986, pp. 366-375.
- [7] Rolf Landauer, “Irreversibility and heat generation in the computing process,” *IBM Journal of Research and Development*, **5**, pp. 183-191 (1961).
- [8] see Rolf Landauer, “Information is Physical”, *Physics Today*, **44**, no. 5, p. 23 (May 1991).
- [9] J. S. Bell, “On the Einstein-Podolsky-Rosen paradox,” *Physics* (Long Is. City, N. Y.) **1**, 195-200 (1964).
- [10] Stuart J. Freedman and John F. Clauser, “Experimental Test of Local Hidden-Variable Theories”, *Phys. Rev. Lett.* **28**, 938-941 (1972).
- [11] Alain Aspect, Phillipe Grangier, and Gérard Roger, “Experimental Tests of Realistic Local Theories via Bell’s Theorem,” *Phys. Rev. Lett.* **47**, 460 (1981).

- [12] Paul G. Kwiat, Klaus Mattle, Harald Weinfurter, Anton Zeilinger, Alexander V. Sergienko and Yanhua Shih, "New High-Intensity Source of Polarization-Entangled Photon Pairs," *Phys. Rev. Lett.* **75**, 4337–4341 (1995).
- [13] Louisa Gilder, *The Age of Entanglement: When Quantum Physics Was Reborn* (A. A. Knopf, New York, 2008).
- [14] A. S. Holevo, "Bounds for the Quantity of Information Transmitted by a Quantum Communication Channel", *Probl. Peredachi Inf.*, **9**:3 (1973), 3–11; "On the capacity of quantum communication channel," *Probl. Inform. Transm.* **15**, 3-11 (1979).
- [15] S.J. Wiesner, "Conjugate Coding", *SIGACT News* **15**:1, pp. 78–88, 1983.
- [16] C. H. Bennett, "Logical reversibility of computation," *IBM Journal of Research and Development*, **17**, no. 6, pp. 525-532 (1973).
- [17] Bennett, C. H. (1982). "The thermodynamics of computation – a review". *International Journal of Theoretical Physics* **21** (12): 905–940.
- [18] T. Toffoli, "Computation and Construction Universality of Reversible Cellular Automata," *J. Computer and System Sciences*, **15**:2 213-231 (1977); T. Toffoli, "Reversible Computing," *Laboratory of Computer Science Report MIT/LCS/TM-151* (MIT, February 1980); T. Toffoli: "Bicontinuous Extensions of Invertible Combinatorial Functions." *Mathematical Systems Theory*, **14**, 13-23 (1981); E. Fredkin and T. Toffoli, "Conservative Logic", *International Journal of Theoretical Physics* **21**, Numbers 3-4 (1982), 219-253.
- [19] R. P. Feynman, "There's plenty of room at the bottom," *Caltech Engineering and Science*, **23**:5, February 1960, pp 22-36.
- [20] Feynman, Richard (1982). "Simulating Physics with Computers". *International Journal of Theoretical Physics* **21** (6–7): 467–488.
- [21] Yu. Manin, "Computable and Uncomputable" (in Russian), *Sovetskoye Radio, Moscow* (1980).
- [22] Richard P. Feynman, "Quantum Mechanical Computers," *Optics News* **11**(2), 11-20 (1985).
- [23] W.K. Wootters and W.H. Zurek, "A Single Quantum Cannot be Cloned," *Nature* **299** (1982), pp. 802–803; D. Dieks, "Communication by EPR devices," *Physics Letters A* **92** (1982) 271–272.
- [24] C. H. Bennett and G. Brassard, "Quantum Cryptography: Public key distribution and coin tossing", in *Proceedings of the IEEE International Conference on Computers, Systems, and Signal Processing, Bangalore*, p. 175 (1984).
- [25] Deutsch, David (July 1985). "Quantum theory, the Church-Turing principle and the universal quantum computer". *Proceedings of the Royal Society of London A* **400**: pp. 97–117.
- [26] Deutsch, D.: "Quantum computational networks." *Proceedings of the Royal Society of London A* **425**, 73–90 (1989).

- [27] David Deutsch and Richard Jozsa (1992). "Rapid solutions of problems by quantum computation." *Proceedings of the Royal Society of London A* **439**: 553.
- [28] E. Bernstein and U. Vazirani, "Quantum complexity theory," in *Proc. 25th Annual ACM Symposium on Theory of Computing*, ACM, New York, 1993, pp. 11-20.
- [29] D. Simon, "On the power of quantum computation," in *Proc. 35th Annual IEEE Symposium on Foundations of Computer Science*, IEEE Press, Piscataway, NJ, 1994, pp. 116-123.
- [30] Bennett C H, Brassard G, Crepeau C, Jozsa R, Peres A and W K Wootters 1993 "Teleporting an unknown quantum state via dual classical and EPR channels," *Phys. Rev. Lett.* **70** 1895-1899.
- [31] Serge Haroche and Jean-Michel Raymond, "Quantum Computing: Dream or Nightmare?" *Physics Today* **49**(8), p. 51 (1996).
- [32] John Preskill, "Quantum Computing: Pro and Con," *Proc. Roy. Soc. Lond.* **A454** (1998) 469-486 (quant-ph/9705032).

A 2 Quantum Simulation

A. Browaeys
Université Paris-Saclay,
Institut d'Optique Graduate School, CNRS,
Laboratoire Charles Fabry.
F-91127 Palaiseau Cedex, France

Contents

1	What is quantum simulation?	2
2	The approaches to quantum simulation	3
3	Experimental platforms	5
3.1	Requirements for experimental implementations	5
3.2	Trapped ions [6, 7]	6
3.3	Neutral atoms [13, 14]	8
3.4	Quantum circuits [17]	14
3.5	Quantum fluids of light with polaritons in semiconductor cavities [20]	16
3.6	Other platforms	18
4	Examples of quantum simulations	18
4.1	Out-of-equilibrium dynamics	18
4.2	Adiabatic preparation of ground states and quantum phase transition	23
4.3	Quantum chemistry and the hybrid approach	27
4.4	Other examples	30
5	Challenges and outlook	32
	References	33

This chapter introduces the basic ideas of quantum simulation, presents a few experimental platforms used so-far and illustrates the concepts on a selection of recent experiments. Several reviews, giving more details or emphasizing different aspects, have been published over the last few years (e.g. [1]) and will be cited throughout the text¹.

1 What is quantum simulation?

Quantum simulation can be defined as the process by which one performs the simulation of a many-body problem using a well-controlled synthetic quantum system, called the *quantum simulator*. Many-body problems deal with situations where ensembles of quantum particles interact. They are thus encountered in many areas of physics, such as condensed matter, nuclear and high-energy physics. Usually one has in mind that with the knowledge of the microscopic laws ruling the particles and their interaction, one could in principle calculate *ab-initio* the macroscopic properties of the ensemble. However, the size of the Hilbert space scales exponentially with the number of particles involved, making any *ab-initio* calculations impossible beyond a few tens of particles. For example, for the simplest case of an ensemble of N spin-1/2 particles, the size of the Hilbert space is 2^N . Today, the most advanced computers can calculate up to $N \approx 50$ interacting spins. To investigate systems involving larger number of particles, one must therefore rely on approximations. Many approximate methods have been devised such as mean-field theory, density functional theory, density matrix renormalization group, quantum Monte-Carlo, etc., that allow calculating ground state properties of quantum systems or their dynamics in low dimension. However they all have limitations: for example finding the ground state of an ensemble of interacting fermions is hard as one must keep track of a large number of minus signs when anti-symmetrizing the many-body wave functions. This is outside the validity range of the approximate methods that quantum simulation becomes really relevant. In some cases, the range of applicability of the methods is not even known and one may want to compare their predictions to the “exact” solution provided by a quantum simulator.

The idea of quantum simulation was introduced by Richard Feynman in 1982 [2]. He stated that “with a suitable class of quantum machine you could imitate any quantum system, including the physical world”. Let us illustrate the idea on a particular example. Assume you want to understand a phenomenon observed in an experiment, for example high- T_c superconductivity. You know experimentally that below a certain temperature the resistivity of certain alloys vanishes (macroscopic property), but you want to understand it from a *microscopic* perspective. You guess that the electrons need to be involved and that a minimal model should include some kind of transport of the electrons and the fact that they interact. The simplest model with all the relevant ingredients was written 1963 by J. Hubbard [3]. The corresponding Hamiltonian is

$$H_{\text{FH}} = -J \sum_{\langle i,j \rangle, \sigma} (\hat{c}_{i\sigma}^\dagger \hat{c}_{j\sigma} + \hat{c}_{i\sigma} \hat{c}_{j\sigma}^\dagger) + U \sum_i \hat{n}_{i\downarrow} \hat{n}_{i\uparrow}, \quad (1)$$

with $\hat{c}_{i\sigma}$ ($\hat{c}_{i\sigma}^\dagger$) the annihilation (creation) operator of an electron on site i with a spin $\sigma = \uparrow, \downarrow$. This Fermi-Hubbard Hamiltonian describes the hopping of electrons from site i to j with amplitude J in the tight-binding limit and with an on-site interaction U . This model is only a very crude

¹ I thank my colleague Thierry Lahaye for his careful reading of these notes.

approximation of a real metal or of a superconducting material. However it remains very hard to solve with the arsenal of theoretical physics and numerical calculations are restricted to small numbers of fermions. There is therefore no point in proposing more elaborate models, if this one is not already solved. In particular, one does not know whether the ground state of the model exhibits superconductivity when the number of spins \uparrow and \downarrow is not equal. It turns out however that the Hamiltonian (1) can be readily implemented using a gas of ultracold fermionic atoms placed in optical lattices (see Sec. 3.3). Hence, by measuring the ground state of this *synthetic atomic system*, one can hope to answer the question of the existence of superconductivity for the Fermi-Hubbard model.

This example illustrates one first important idea about quantum simulation: one has to find a mapping between the system we are interested in (here the electrons of a metal in a periodic lattice) and the one that can be implemented in the lab. Other examples are the study of the motion of electrons in a strong magnetic field emulated by an ensemble of atoms placed in a suitable laser field, or the study of the magnetic properties of electrons in a lattice that are mapped onto an ensemble of interacting two-level atoms or ions.

The quantum simulation approach therefore has several appealing features. First, as explained above, it helps answering questions about real-life materials or situations. Second, relying on a synthetic system allows one to vary the parameters of the model in a range inaccessible otherwise, thus providing a way to better understand their respective influence. For example mimicking a magnetic field seen by an electron by a laser field applied on atoms may allow reaching values of the magnetic field impossible to produce in a condensed matter laboratory. Also, if one is interested in the influence of interatomic interactions on the phase diagram of a given compound, synthetic systems are valuable as they allow varying their strength in a way which is usually impossible in real materials. Third, synthetic systems usually provide new types of probes, such as the direct measurement of the correlation functions between the particles, or the direct observation of the wave functions. Finally, it allows implementing models that are purely mathematical, and hence not even an idealization of real materials. In this case, their implementation leads to the fabrication of an artificial many-body system, which becomes an object of study in its own. From this perspective, quantum simulation can be viewed as exploring many-body physics with synthetic systems: in the same way chemists design new materials exhibiting interesting properties (such as magnetism, superconductivity, etc.), physicists assemble artificial systems and study their properties, with the hope to observe new phenomena.

2 The approaches to quantum simulation

Two main approaches to quantum simulation have been proposed.

In the first approach, called *analog quantum simulation*, the model Hamiltonian H_{model} that one is interested in is directly implemented on the experimental platform (see examples in Sec. 3). Performing a quantum simulation of the model means evolving the initial state $|\psi(0)\rangle$ of the quantum simulator, assumed to be relatively easy to prepare, under the action of H_{model} , to reach a final state of interest:

$$|\psi(t)\rangle = U(t) |\psi(0)\rangle \quad \text{with} \quad U(t) = \exp \left(-\frac{i}{\hbar} \int_0^t H_{\text{model}}(t') dt' \right). \quad (2)$$

Following this evolution, the state of the simulator is read out. Usually, this does not mean measuring exactly $|\psi(t)\rangle$, as this is too difficult (the number of components of the wave function scales exponentially with the particle number), but to measure the average value $\langle\psi(t)|\hat{O}|\psi(t)\rangle$ of an observable \hat{O} , such as a magnetization, the mean occupancy on a given site of an array, or a pair correlation function. This time evolution of the state of the quantum simulator allows studying two kinds of situations: (i) the out-of-equilibrium dynamics of the system following the sudden variation of one parameter of the Hamiltonian; (ii) the adiabatic preparation of the ground state of the synthetic system following the slow variation of the parameters of the Hamiltonian. We will illustrate this in more details in Sec. 4. This first, analog approach is the easiest to implement experimentally. It is problem-specific, in the sense that for any problem of interest, one has to build a dedicated synthetic quantum system to study it. By engineering the system in a way that all parameters of the Hamiltonian are tunable, the machine can be thought as an analog *programmable quantum simulator*. Analog quantum simulation relies on the idea of mapping between two systems: the one we are interested in (e.g. the motion of an electron in an ultra-strong magnetic field, transport of electrons in metal, etc.) and the one that emulates it, which is implemented in the lab (e.g. atoms subjected to shaped laser fields, interacting atoms in optical lattices, etc.).

An interesting aspect of the analog approach is that it is, to a certain extent, relatively robust to imperfections and errors in the implementation. Although this statement would need to be carefully defined (which is still the topics of current research), one naive way to understand it is to realize that some interesting features could emerge even in the presence of errors: for example, a magnetic ordering of interest may develop on top of a background resulting from the imperfections. If one is only interested in the possibility of such ordering, one gets already a partial answer to the question. This idea that imperfect quantum machines can already be useful has been conceptualized recently by J. Preskill under the name Noisy Intermediate Scale Quantum technology [4].

The second approach, named *digital quantum simulation*, allows synthetizing any arbitrary Hamiltonian, including ones that can not be directly realized experimentally. The idea was introduced by Seth Lloyd in 1996 [5]. Let us first assume that we want to implement a Hamiltonian, which is the sum of terms describing local interactions (i.e. single and two-body): $H = \sum_{\ell} H_{\ell}$. As an example, think of the Ising Hamiltonian for spins in a transverse magnetic field: $H_{\text{Ising}} = J \sum_{i,j} \sigma_i^z \sigma_j^z + B \sum_i \sigma_i^x$. In general, the various terms H_{ℓ} in the Hamiltonian do not commute, and therefore the evolution operator $U(t) = \exp[-iHt/\hbar]$ (here H is assumed time-independent) can not be decomposed into a product of terms $\exp[-iH_{\ell}t/\hbar]$. However, an approximate factorization is possible when breaking the evolution time t into n time steps, using the first-order Trotter decomposition:

$$U(t) = \left[e^{-i\frac{H}{\hbar}\frac{t}{n}} \right]^n \approx \left[\prod_{\ell} e^{-i\frac{H_{\ell}}{\hbar}\frac{t}{n}} \right]^n + \mathcal{O}\left(\frac{1}{n^2}\right). \quad (3)$$

When $n \rightarrow \infty$, the evolution operator is a product of elementary evolutions that can be efficiently calculated. This has a price, as the number of time steps can be large to reach a given precision. It was however shown in [5] that, for a fixed precision, the complexity of the simulation grows only polynomially with the number of particles involved in the simulation.

More generally, any unitary evolution operator $U(t) = \exp[-iHt/\hbar]$ can be decomposed as the product of universal gates, such as the ones operating on one or two quantum bits. It is

therefore possible to synthesise any Hamiltonian with this digital approach, including the ones that do not involve interactions between particles allowed by Nature. As an example, consider the Hamiltonian describing the interaction between three spins

$$H = \sigma_1^z \otimes \sigma_2^z \otimes \sigma_3^z. \quad (4)$$

This 3-body Hamiltonian does not correspond to any physical interaction, as only two-body interactions are found in Nature. However, it can be digitally synthesized using sets of C-NOT gates and the unitary evolution of a fourth ancillary quantum bit, as shown in Fig. 1. As can be seen on this example, the implementation of an arbitrary Hamiltonian may require auxiliary quantum bits and therefore may not necessarily scale favorably with the number of particles to consider.

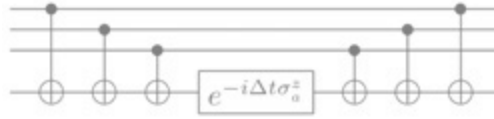


Fig. 1: Quantum circuit to synthesize the 3-body spin interaction described by Eq. (4). Figure from [1].

Using the digital approach, it becomes possible to perform a *universal* quantum simulation: the quantum simulator does not need to be rebuilt for each Hamiltonian H to be studied, but just reprogrammed for a specific problem. It should also be clear that the digital approach to quantum simulation actually amounts to performing a quantum computation, the quantum simulator being now a full fledged quantum computer. It is therefore much more demanding experimentally than the analog approach: the fidelity of each gate must be as high as possible for the final result to be meaningful.

3 Experimental platforms

Although the idea of quantum simulation using synthetic quantum systems is now rather old, it was for a very long time essentially theoretical. This was due to the fact that the control over quantum systems was not advanced enough. The situation changed radically in the last 20 years with the development of experimental techniques to control the quantum state of individual quantum objects, be they atoms, molecules, ions, photons, or even artificial atoms such as quantum dots, superconducting circuits or excitons in semi-conductors, to name a few.

In this Section, we will describe some of the platforms that have already demonstrated their potential for quantum simulation. It is not meant to be exhaustive, but rather to illustrate the idea of mapping introduced in Sec. 1. More detailed reviews on each platform have been published recently: they are mentioned in the titles of the sections below.

3.1 Requirements for experimental implementations

A key question for the experimentalist is how faithfully one can implement the model Hamiltonian, and how well one can initialize the quantum simulator in a given state $|\psi(0)\rangle$. For the

digital approach, this is critical as the highest fidelity is required for the results to be meaningful. This is also true for analogous quantum simulators, although here the constraints on fidelity of the mapping are less severe.

The programmability of the simulators, i.e. the ability to tune the parameters of the Hamiltonian at will over large ranges, is another important key feature. This includes the control of the external parameters of the simulators acting at the single particle level (ex: magnetic field, laser light, trapping potential, etc.), but also the tunability of the interactions between the constituent particles. Importantly, the energy scale U associated to the interactions sets the timescale \hbar/U for a dynamics induced by them. It should of course be much smaller than the decoherence time of the platform.

When considering programmable and digital quantum simulation, one usually adds the requirement of individual particle manipulations and detection. In particular, the ability to address any individual constituent allows for local variations of the parameters (e.g. placing one impurity somewhere and following its evolution, local gates, or gates acting on two particular particles, etc.). It also makes it possible to measure the state of each particle individually, and in this way allow reconstructing important quantities characterizing the many-body properties of the system: this is in particular the case of the n -order correlation functions $\langle \psi | \hat{O}_1 \hat{O}_2 \dots \hat{O}_n | \psi \rangle$ of a relevant observable \hat{O}_i for the particle i . One can thus link macroscopic variables to the details of the microscopic configuration.

3.2 Trapped ions [6,7]

In this first platform, ions are trapped and laser-cooled in electromagnetic traps, such as Paul or Penning traps. There, the interplay between their electrostatic repulsion and the harmonic confinement provided by the traps leads to a “Coulomb crystal” for sufficiently low temperatures.

A popular configuration for the trap is the so-called linear Paul trap (Fig. 2a) where a combination of DC and RF electric fields results into a highly anisotropic, nearly harmonic potential for an ion (mass M , charge q), of the form $V_{\text{trap}}(x, y, z) = M\omega_{\perp}^2(x^2 + y^2)/2 + M\omega_z^2 z^2/2$ with $\omega_{\perp} \gg \omega_z$. In practice: $\omega_{\perp}/(2\pi) \approx 5$ MHz and $\omega_z/(2\pi) \approx 100$ kHz. The distances between the ions are set by balancing the confinement energy with the electrostatic repulsion between ions.

For two ions, this gives $M\omega_z^2 \Delta z/4 = q^2/(4\pi\epsilon_0 \Delta z^2)$, yielding $\Delta z = [q^2/(\pi\epsilon_0 M\omega_z^2)]^{1/3}$. For two $^{40}\text{Ca}^+$ ions, $\Delta z \approx 3.5 \mu\text{m}$. For larger numbers N of ions, no analytic expression exist for the inter-ion spacings, which are numerically found to scale as $[q^2/(4\pi\epsilon_0 M\omega_z^2)]^{1/3} N^{-0.59}$. The largest linear ion crystal produced contains around 100 ions separated by a few micrometers (see Fig. 2a). This large spacing between ions makes each of them addressable (manipulation by focused lasers and individual readout). In the crystal, the ions can oscillate around their equilibrium positions. As the N ions are all coupled by the Coulomb interaction, their motion is collective and the crystal features $3N$ mechanical eigen-modes of vibration. These modes are quantized, and when the crystal is laser-cooled it is even possible to reach the ground state of the collective vibrational modes.

Popular choices for the ions used in the experiments are Ca^+ , Yb^+ , Be^+ , Cd^+ , Mg^+ , etc. When considering their internal degrees of freedom, it is possible to isolate a two-level structure and to map the two states onto a spin-1/2 with spin states $|\uparrow\rangle$ and $|\downarrow\rangle$ (Fig. 2b). Depending on the ions, they are coupled by a microwave transition (e.g. Yb^+ , Be^+) or by an optical transition

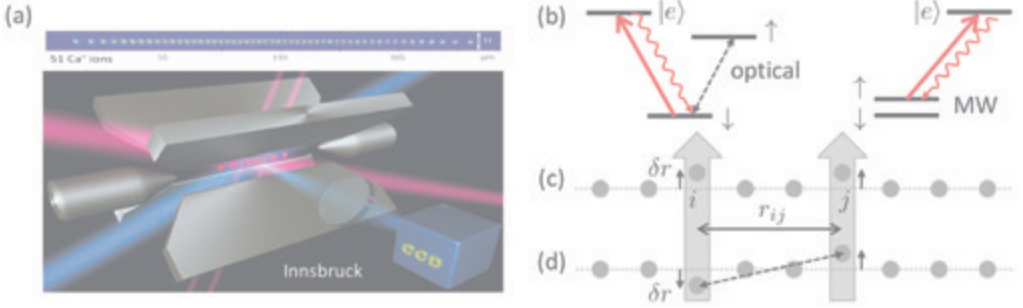


Fig. 2: Trapped ion platform. (a) Schematic view of a linear Paul trap and fluorescence image of a chain of 51 Ca⁺ ions (figures from University of Innsbruck). (b) Two examples of encoding of spin states on the atomic structure of the ion using optical or microwave (MW) transitions. The extra level $|e\rangle$ coupled by a laser is used for fluorescence state detection. (c) Action of two lasers on an ion chain, when ions i and j are both in state $|\uparrow\rangle$. (d) Same as (c) for the two ions in different spin states.

(e.g. Ca⁺). Microwave or laser pulses allow for the coherent manipulation of any superposition of the two spin states.

To make the spins interact, one uses laser beams that induce a spin-dependent force. To see how this works, consider two laser beams propagating perpendicularly to the chain and focused on ions i and j (Fig. 2c,d). With a suitable choice of parameters, the force resulting from the lasers pushes the ion in one direction or another depending on its state $|\uparrow\rangle$ or $|\downarrow\rangle$. As a consequence, if the two ions are in the same state, i.e. in $|\uparrow, \uparrow\rangle$ or $|\downarrow, \downarrow\rangle$, they are pulled or pushed in the same direction and their Coulomb interaction q^2/r_{ij} will remain nearly unchanged (Fig. 2c). On the contrary, if they are in different spin states, i.e. $|\uparrow, \downarrow\rangle$ or $|\downarrow, \uparrow\rangle$, they are pushed in two opposite directions by a quantity δr (Fig. 2d). Their Coulomb interaction energy is now changed by an amount:

$$\Delta V_{C,ij} = \frac{q^2}{4\pi\epsilon_0} \left(\frac{1}{\sqrt{r_{ij}^2 + 4\delta r^2}} - \frac{1}{r_{ij}} \right) \approx -2 \frac{(q\delta r)^2}{4\pi\epsilon_0 r_{ij}^3}. \quad (5)$$

Hence the energy shift induced by the light depends on the relative configuration of the two spins: it realizes an Ising-type of interaction $\sim J_{ij} \hat{\sigma}_i^z \hat{\sigma}_j^z$, with $J_{ij} = \Delta V_{C,ij}/4$. Here $\hat{\sigma}^z = |\uparrow\rangle\langle\uparrow| - |\downarrow\rangle\langle\downarrow|$ is the usual Pauli operator.

In fact, the exact derivation of the light-induced force is more involved than the argument presented above, but the idea of a state-dependent force remains valid. The laser illuminates the whole chain perpendicularly to it. Its central frequency ω is detuned with respect to the one of the $(|\uparrow\rangle, |\downarrow\rangle)$ transition, and it contains two frequencies ω and $\omega + \delta$, leading to an amplitude modulation. Calculating the action of the force requires considering the quantized mechanical modes of the ion chain. The result of the derivation (see for example [8]) is the following:

- For $\delta = 0$, the Hamiltonian is

$$H_{XX} = \sum_{i,j} J_{ij} \hat{\sigma}_i^x \hat{\sigma}_j^x. \quad (6)$$

- For $\delta \gg J_{ij}$, one gets the XY (or flip-flop, or exchange) Hamiltonian, with $\hat{\sigma}^+ = |\uparrow\rangle\langle\downarrow|$ and $\hat{\sigma}^- = |\downarrow\rangle\langle\uparrow|$:

$$H_{\text{XY}} = \sum_{i,j} J_{ij} (\hat{\sigma}_i^+ \hat{\sigma}_j^- + \hat{\sigma}_i^- \hat{\sigma}_j^+) . \quad (7)$$

- Finally, $\delta \approx J_{ij}$ leads to an Ising-like Hamiltonian:

$$H_{\text{Ising}} = \sum_{i,j} J_{ij} \hat{\sigma}_i^x \hat{\sigma}_j^x + \frac{\delta}{2} \sum_i \hat{\sigma}_i^z . \quad (8)$$

Numerical calculations show that $J_{ij} \approx J_0/|i-j|^\alpha$, with α varying in practice between 0.5 and 2.5 depending on the frequency detuning δ . The coupling is thus long-range and allows studying many-body physics in unusual regimes. The constant J_0 depends on the laser power, the frequency ω , and involves the frequencies of the mechanical modes. It can be set positive or negative. Typically, in the experiments $J_0/(2\pi) \sim 0.5 - 1$ kHz, leading to a timescale for the interactions in the millisecond range.

Finally, an alternative approach to prepare Coulomb crystal relies on Penning traps, which combine electrostatic and magnetic fields [9]. The crystal is then a plane of ions, with a separations also around several micrometers. Although such crystal can contain up to a few hundred ions (thus more than the linear Paul trap), it has the drawback that it rotates at high frequency (typically 50 kHz) around an axis perpendicular to the ion plane, making local manipulations difficult. Quantum simulations of the Ising [10] and Dicke [11] models have however already been performed in this platform.

As a conclusion on trapped ions, they implement naturally spin models and realize the quantum Ising or XY Hamiltonians encountered in quantum magnetism. Importantly, the trapped ion platform is probably the best one in terms of fidelity of single and many-ion manipulations. Thanks to this good level of control, this platform has also attempted quantum simulation following the digital approach (see Sec. 4.3 and [12]).

3.3 Neutral atoms [13,14]

Neutral atom-based platforms are probably the richest from the point of view of quantum simulation. This comes from the possibility to change the quantum statistic of the particles (fermions or bosons), as well as the ability to map them onto spin systems. Besides, this platform is so-far easier to scale up than many others.

When speaking about neutral atom-based platforms, it is usually understood cold (i.e. laser-cooled) or ultra-cold atoms prepared in quantum degenerate gases (Bose-Einstein condensate or Fermi sea). In many situations, the temperature associated to the atomic motion can be ignored and the dominant energy scale is associated to the interactions U between the atoms. The atom-based platforms exist in several implementations that we now briefly describe.

- **Ultra-cold atoms in quantum degenerate gases.** By combining laser and evaporative cooling, a gas of atoms (alkali: Li, Na, K, Rb, Cs; alkaline-earth: Sr ; Yb, Cr, Er, Dy; He*, etc.) is cooled below typically 100 nK where it reaches the quantum degenerate

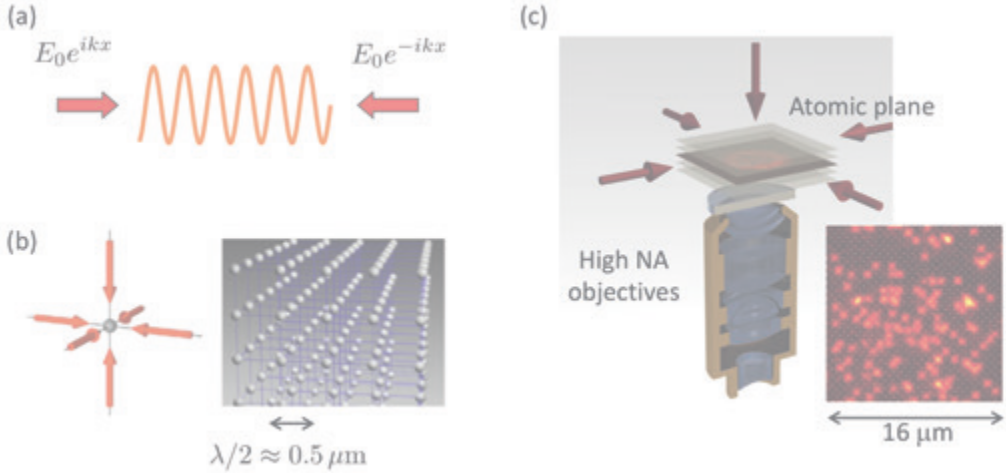


Fig. 3: Neutral atom platform: optical lattices. (a) One dimensional optical lattice. (b) Three dimensional cubic lattices obtained by interfering three pairs of one-dimensional lattices (figure from M. Greiner's thesis). (c) Quantum gas microscope: a high NA lens (here $NA= 0.68$) monitors a plane of atoms using fluorescence imaging. Each dot in the lower right corner corresponds to one atom. Adapted from [14].

regime. This corresponds to a phase-space density $n\lambda_{\text{dB}}^3 \sim 1$, with n the density of the vapor and $\lambda_{\text{dB}} = h/\sqrt{2\pi M k_B T}$ the thermal de Broglie wavelength at a temperature T . If the atoms are bosons, the gas forms a Bose-Einstein condensate, where all the atoms occupy the same wave function. If the atoms are fermions, one gets a Fermi sea. In this implementation, the positions of the atoms are randomly distributed. The vapors are confined in traps produced by static magnetic fields, by the dipole force of a laser or a combination of both. The typical number of atoms ranges between 10^3 to 10^6 .

- Ultra-cold atoms in optical lattices and quantum gas microscope.** Here ultracold atoms are held in periodic light structures, called optical lattices [14]. These lattices rely on the dipole force exerted by a laser on an atom: a laser with intensity I and frequency ω produces an energy shift of the atomic ground state by an amount $\propto I/(\omega - \omega_0)$, with ω_0 the resonance frequency of an optical transition. If the detuning of the laser $|\omega - \omega_0|$ is larger than the linewidth of the transition Γ , the action of the laser is nearly conservative, as spontaneous emission is suppressed as I/Δ^2 . When the intensity $I(\mathbf{r})$ varies in space, the atom is attracted towards the region of high or low intensity depending on the sign of $\omega - \omega_0$. To create periodic structures, one interferes several beams. For example, in one dimension, two counter-propagating lasers with wave vectors $k = 2\pi/\lambda$ produce a one-dimensional lattice leading to a trapping potential $V_L(x) \propto |e^{ikx} + e^{-ikx}|^2 = V_0 \cos^2(kx)$, as shown in Fig. 3(a). The spacing between the nodes or antinodes is $\lambda/2$ but it can also be increased by changing the geometry of the interfering beams. One, two and three dimensional lattices are commonly used (Fig. 3b). The elementary cell can be square in 2D, cubic in 3D, but also triangular, honeycomb, etc.

It is possible to load atoms in 2D and 3D lattices with exactly one atom per node or antinode. The most common technique consists in starting from a Bose-Einstein condensate

or a Fermi degenerate gas and to adiabatically switch on the lattice [14] (see Sec. 4.2). As the atoms can tunnel from one site to another, this platform realizes situations where electrons move in crystals of various geometries. It is thus ideal to study condensed matter situations (see below).

An important experimental breakthrough of the past decade was the development of “quantum gas microscopes”, which combine optical lattices with large numerical aperture objectives [14] (see Fig. 3c). The resolution of the objectives being on the order of $\lambda/2$, also the typical spacing between atoms in the lattices, this tool allows for the in-situ observation of the atoms and their local manipulations. One can thus study transport, measure correlation functions, etc., locally. Typically, a few hundred atoms are routinely manipulated in 2D optical lattices under a quantum gas microscope.

- **Cold atoms in arrays of optical tweezers [15].** More recently, a platform complementary to optical lattices has emerged. It relies on laser beams tightly focused by a large numerical aperture optical system. The spot size is on the order of $1\ \mu\text{m}$ and this single beam, red-detuned with respect to an optical transition of the atoms, ensures a 3D confinement around the focal point where the light intensity is maximum. This microscopic dipole trap is called an optical tweezers. By placing a diffractive element before the focusing lens (such as a liquid-crystal-based spatial light modulator), one obtains an array of spots that can be arranged in almost any arbitrary 2D and 3D geometry (see Fig. 4a,c,d). This again allows emulating condensed matter situations.

The appealing feature of the platform relies on the fact that a single tweezers can host at most one atom for a proper choice of the parameters. To load it, one uses a cloud of laser-cooled atoms and focuses the tweezers array inside this cloud. Atoms enter at random in the tweezers and get trapped. When two atoms are present at the same time, a light-induced collision expels both of them, and the process starts again. Although simpler to implement experimentally than the quantum gas microscope, the fact that the atoms enter at random times in the tweezers is a drawback: it results into arrays with $1/2$ filling fraction. Techniques have however been devised to move the atoms in the arrays and thus assemble atom-by-atom atomic arrays with unit filling (Fig. 4b). They operate up to 120 atoms in 2D and 72 in 3D. In this platform, the typical distances between the atoms are several micrometers, which ensures the ability of local addressing for manipulations and readout.

Interaction between ultra-cold atoms. Performing quantum simulation requires interactions between the atoms. For ultra-cold bosonic atoms (mass M), the interaction $V(\mathbf{r}_1, \mathbf{r}_2)$ between two of them is described to a good approximation by a contact potential governed by one parameter a called the scattering length:

$$V(\mathbf{r}_1, \mathbf{r}_2) = \frac{4\pi\hbar^2 a}{M} \delta(\mathbf{r}_1 - \mathbf{r}_2). \quad (9)$$

Hence, for 2 atoms in wave functions $\phi_{a,b}(\mathbf{r})$, the interaction energy is

$$U = \int |\phi_a(\mathbf{r}_1)|^2 V(\mathbf{r}_1, \mathbf{r}_2) |\phi_b(\mathbf{r}_1)|^2 d^3\mathbf{r}_1 d^3\mathbf{r}_2 = \frac{4\pi\hbar^2 a}{M} \int |\phi_a(\mathbf{r})|^2 |\phi_b(\mathbf{r})|^2 d^3\mathbf{r}. \quad (10)$$

For two fermions in the same wave function, the interaction vanishes due to the Pauli principle. Fermions in different spin states do however interact, with an interaction also characterized by

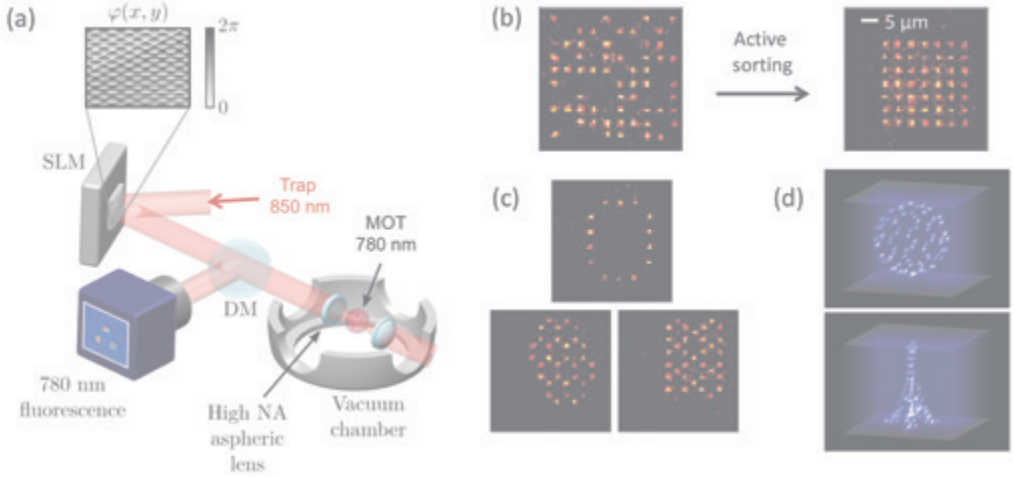


Fig. 4: Neutral atom platform: optical tweezers arrays. (a) Setup to create an array of optical tweezers. A spatial light modulator (SLM) imprints a phase pattern on the trapping beam, resulting in the focal plane of the high NA lens into an array of tweezers. The atoms are observed by fluorescence imaging on a sensitive camera. (b) Active sorting of the atoms in the array results into a sub-array with unit filling. (c) Examples of fluorescence images of 2D and 3D structures (d). Each bright spot corresponds to one atom. Figures from the Institut d’Optique.

a scattering length. Importantly, the scattering length is controllable by a magnetic field B , a phenomenon called a Feshbach resonance. The scattering length of two atoms placed in a magnetic field B varies around a resonance located at B_0 as $a(B) = a_{\text{bg}} + \Delta/(B - B_0)$. It is thus possible to vary the interaction between two atoms from perfect cancellation ($a = 0$) to a much larger than any lengthscale in the problem (the unitary regime).

Implementation of Bose- and Fermi-Hubbard models. Let us look at the consequence of this contact interaction for ultra-cold atoms held in optical lattices. Consider first the case of bosons. The Hamiltonian governing the behavior of N bosons in an optical lattice contains two terms:

$$H_{\text{BH}} = -J \sum_{\langle i,j \rangle} (\hat{a}_i^\dagger \hat{a}_j + \hat{a}_i \hat{a}_j^\dagger) + \frac{U}{2} \sum_{i=1} \hat{n}_i (\hat{n}_i - 1), \quad (11)$$

with \hat{a}_i (\hat{a}_i^\dagger) the annihilation (creation) operator of an atom at site i . The first term describes the tunneling of atoms between neighboring sites, in the tight-binding model. The second term describes the on-site interaction when two bosons are placed on the same site of the lattice. Here $U = \frac{4\pi\hbar^2 a}{M} \int |\phi(\mathbf{r})|^4 d^3\mathbf{r}$, with $\phi(\mathbf{r})$ the wave function of an atom in one lattice site. In practice $U/\hbar \approx 1$ kHz. This is the Bose-Hubbard model. In the same way, fermionic atoms held in an optical lattice and with two internal states mapped onto the spin states (\uparrow, \downarrow) directly implement the Fermi-Hubbard Hamiltonian of Eq. (1). Thus atoms in optical lattices naturally realize two of the basic Hamiltonians used to describe strongly correlated electronic systems in condensed matter physics, which are still actively studied.

Implementation of spin models in optical lattices. Ultra-cold atoms in optical lattices also

naturally implement spin models, such as the Heisenberg one described by the interaction Hamiltonian $H_{\text{Heisenberg}} = J \mathbf{S}_a \cdot \mathbf{S}_b$ for two spins \mathbf{S}_a and \mathbf{S}_b [16]. Consider two *bosonic* alkali atoms and two of their hyperfine ground states $|\uparrow\rangle$ and $|\downarrow\rangle$ (for example, for ^{87}Rb atoms, $|\downarrow\rangle = |5S_{1/2}, F=1, M=-1\rangle$ and $|\uparrow\rangle = |5S_{1/2}, F=1, M=1\rangle$). Place each of them on the two sites a and b of a double-well potential, coupled with a tunneling amplitude J . We assume that the on-site interaction U is spin-independent and that $U \gg J$. If the atoms are initially in opposite spin states, i.e. in $|\uparrow, \downarrow\rangle_{a,b}$ or $|\downarrow, \uparrow\rangle_{a,b}$, the tunneling couples them to the states $|\uparrow\downarrow, 0\rangle_{a,b}$ or $|0, \uparrow\downarrow\rangle_{a,b}$ with two atoms in one of the sites at an energy cost U . This off-resonant coupling induces an energy shift $-2J^2/U$ obtained from second-order perturbation theory. If now the atoms are in the two sites with the same spin state, e.g. $|\uparrow, \uparrow\rangle_{a,b}$, the tunneling couples them to the states $|\uparrow\uparrow, 0\rangle_{a,b}$ or $|0, \uparrow\uparrow\rangle_{a,b}$ with an amplitude $J\sqrt{2}$ due to the bosonic enhancement. The resulting energy shift is thus $-2(J\sqrt{2})^2/U$. The same argument holds for the two-atom state $|\downarrow, \downarrow\rangle_{a,b}$, and the diagonal, effective Hamiltonian describing these shifts is:

$$H_1 = -\frac{2J^2}{U}(|\uparrow, \downarrow\rangle \langle\uparrow, \downarrow| + |\downarrow, \uparrow\rangle \langle\downarrow, \uparrow|) - \frac{4J^2}{U}(|\uparrow, \uparrow\rangle \langle\uparrow, \uparrow| + |\downarrow, \downarrow\rangle \langle\downarrow, \downarrow|). \quad (12)$$

If the atoms are initially in $|\uparrow, \downarrow\rangle_{a,b}$, the tunneling also induces a spin-flip to the state $|\downarrow, \uparrow\rangle_{a,b}$ via the intermediate states $|\uparrow\downarrow, 0\rangle_{a,b}$ or $|0, \uparrow\downarrow\rangle_{a,b}$. In perturbation theory, the amplitude of the spin-flip is $-2J^2/U$, and the effective Hamiltonian is off-diagonal:

$$H_2 = -\frac{2J^2}{U}(|\uparrow, \downarrow\rangle \langle\downarrow, \uparrow| + |\downarrow, \uparrow\rangle \langle\uparrow, \downarrow|). \quad (13)$$

Introducing the spin operators $S_{a,b}^z = (|\uparrow\rangle \langle\uparrow| - |\downarrow\rangle \langle\downarrow|)_{a,b}/2$, $S_{a,b}^+ = |\uparrow\rangle \langle\downarrow|_{a,b}$ and $S_{a,b}^- = |\downarrow\rangle \langle\uparrow|_{a,b}$, the total effective Hamiltonian has the Heisenberg form, within an offset:

$$H_{\text{Heisenberg}} = H_1 + H_2 = -\frac{4J^2}{U} \left[S_a^z S_b^z + \frac{1}{2} (S_a^+ S_b^- + S_a^- S_b^+) \right] = 2J_{\text{ex}} \mathbf{S}_a \cdot \mathbf{S}_b, \quad (14)$$

with $J_{\text{ex}} = -2J^2/U$. This coupling favors a ferromagnetic ordering of the spins. The process by which the hopping occurs via an off-resonant state is called super-exchange and is used to describe the magnetism of ionic solids such as MnO and CuO. Starting from the Fermi-Hubbard Hamiltonian of Eq. (1), the same line of arguments for two spin states of *fermionic* atoms (e.g. ^6Li) also leads to a Heisenberg spin Hamiltonian, but the coupling is now $J_{\text{ex}} = 2J^2/U$, i.e. favors an anti-ferromagnetic order. This comes from: (i) the Pauli blockade that prevents the coupling of, e.g., $|\uparrow, \uparrow\rangle_{a,b}$ to $|\uparrow\uparrow, 0\rangle_{a,b}$ or $|0, \uparrow\uparrow\rangle_{a,b}$, and (ii) the fact that the spin-flip from $|\uparrow, \downarrow\rangle_{a,b}$ to $|\downarrow, \uparrow\rangle_{a,b}$ amounts to exchanging the two fermions, which leads to a sign change in the coupling.

Some atoms used in the experiments also possess a strong magnetic moment $\boldsymbol{\mu}$ in their ground state. This is the case of Cr ($\mu = 6\mu_B$), Er ($\mu = 7\mu_B$) and Dy ($\mu = 10\mu_B$). Consequently, on top of the contact interaction described by the scattering length, they also interact by magnetic dipole interaction:

$$H_{\text{mag}} = \frac{\mu_0}{4\pi r^3} [\boldsymbol{\mu}_1 \cdot \boldsymbol{\mu}_2 - 3(\boldsymbol{\mu}_1 \cdot \hat{\mathbf{r}})(\boldsymbol{\mu}_2 \cdot \hat{\mathbf{r}})], \quad (15)$$

for two atoms separated by a distance r . The energy scale $\mu_0 \mu^2 / (4\pi \hbar r^3)$ associated to this interaction is rather small, around 10 Hz for $r = 500$ nm, the typical spacing between atoms

in an optical lattice. To observe the influence of this magnetic interaction, one usually suppresses the contact interaction using a Feshbach resonance. Finally, as the magnetic moment is proportional to the spin operator, the Hamiltonian (15) also allows implementing spin models.

Rydberg quantum simulator. In platforms based on tweezers arrays, the separations between the atoms are on the order of a few micrometers. At these distances all the interactions between atoms mentioned above are negligible. To make them interact at such large distances, one uses Rydberg states, i.e. states with large principal quantum numbers n (in practice $50 \leq n \leq 100$) [15]. When in such a state, the atoms feature two important properties. First, their lifetime, scaling as n^3 , is much longer than for low lying transitions (typically in the $100 \mu\text{s}$ range for $n \approx 50$). Second, they exhibit large electric dipole moments between states n and $n - 1$ with opposite parity, scaling as n^2 . This leads to large interaction strengths V , corresponding to frequencies $V/h \gtrsim 1 \text{ MHz}$ for $n \approx 50$ at distances around $5 \mu\text{m}$. The timescale associated with this Rydberg-based interaction is thus sub-microsecond. Ensembles of atoms in arrays of tweezers and excited to Rydberg states naturally implement Ising and XY models, as for the ions, and constitutes a “Rydberg quantum simulator”, as we now describe.

Consider first the case where two atoms are excited to the *same* Rydberg state $|r\rangle$. There, the dipole-dipole interaction leads to the van der Waals interaction, which induces an energy shift of the pair state $|rr\rangle$ scaling as C_6/R^6 . The van der Waals C_6 coefficient varies as n^{11} , meaning that placing two atoms in, e.g., Rydberg state $n = 50$ enhances the interaction by 11 orders of magnitude. This shift occurs only when both atoms are excited to the Rydberg state. Mapping the ground and Rydberg states $|g\rangle$ and $|r\rangle$ of each atom onto a spin $1/2$ following $|\downarrow\rangle = |g\rangle$ and $|\uparrow\rangle = |r\rangle$, the Hamiltonian of an ensemble of atoms driven by a coherent laser (Rabi frequency Ω , frequency detuning Δ) is

$$H = \frac{\hbar\Omega}{2} \sum_i \hat{\sigma}_i^x - \hbar\Delta \sum_i \hat{n}_i + \sum_{i<j} V_{ij} \hat{n}_i \hat{n}_j, \text{ with } V_{ij} = \frac{C_6}{R_{ij}^6}. \quad (16)$$

Here \hat{n}_i is the operator counting the number of Rydberg excitation (0 or 1) at site i , related to the $\hat{\sigma}_i^z$ Pauli matrix by $\hat{n}_i = (\hat{\sigma}_i^z + 1)/2$. Hence, Eq. (16) has the form of the quantum Ising model, with a transverse field $B_\perp \propto \Omega$, a longitudinal field $B_\parallel \propto -\Delta$ and Ising couplings J_{ij} decaying as $1/R_{ij}^6$ with the distance. The Hamiltonian (16) assumes that the excitation laser covers uniformly the atomic array, but owing to the single-site addressability, the detunings and Rabi frequency can be made site dependent by adding local laser control.

Consider now the case where the atoms are prepared in two *different* Rydberg states that are dipole-coupled, such as $|nS\rangle$ and $|nP\rangle$, separated by a transition frequency typically in the 10 GHz range. There, the dipole-dipole interaction gives rise to a coherent exchange of the internal states of the atoms and the interaction potential scales as C_3/R^3 , with $C_3 \propto n^4$. The mapping onto a spin- $1/2$ model is then $|\downarrow\rangle = |nS\rangle$ and $|\uparrow\rangle = |nP\rangle$. Microwave radiation can be used to manipulate the spin and thus acts as an external magnetic field. The Hamiltonian for a system of atoms then reads:

$$H = \frac{\hbar\Omega_{\mu\text{w}}}{2} \sum_i \hat{\sigma}_i^x - \frac{\hbar\Delta_{\mu\text{w}}}{2} \sum_i \hat{\sigma}_i^z + \sum_{i \neq j} \frac{C_3}{R_{ij}^3} (\hat{\sigma}_i^+ \hat{\sigma}_j^- + \hat{\sigma}_i^- \hat{\sigma}_j^+), \quad (17)$$

which is the XY spin Hamiltonian with transverse and longitudinal fields given by the Rabi frequency $\Omega_{\mu\text{w}}$ and the detuning $\Delta_{\mu\text{w}}$ of the microwave field.

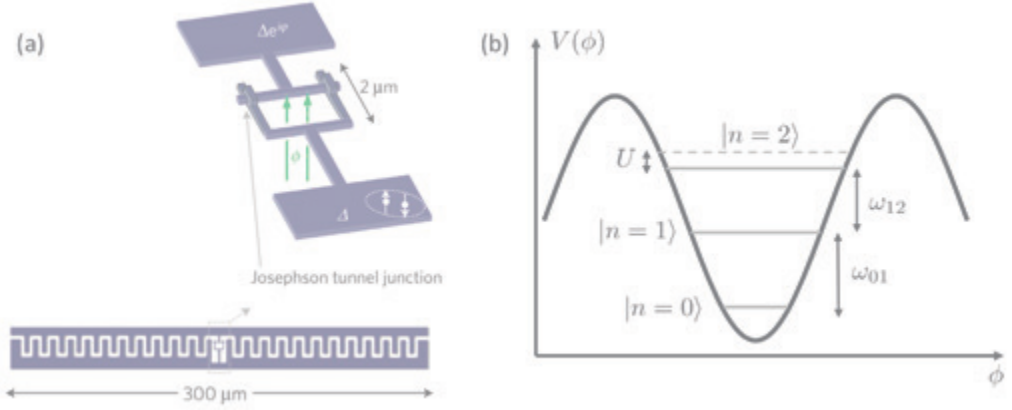


Fig. 5: Quantum circuit platform. (a) Example of a quantum circuit based on a Cooper-pair box. It consists of two Josephson junctions in a loop (SQUID) placed between two superconducting islands. This device is patterned on a chip and can be coupled to a microwave guide also patterned on the chip. Figures from [17]. (b) The non-linear inductance of the Josephson junction results into an effective sinusoidal potential as a function of the flux through the loop.

Ultra-cold molecule-based platform. To conclude this description of the atom-based platforms, one must cite a related platform manipulating ultra-cold polar molecules such as KRb, LiCs, CaF [14]. They possess electric dipole moments, which can be aligned and enhanced using DC electric fields. This results into a dipole-dipole interaction, which has the form given by Eq. (15), replacing the magnetic dipoles by the electric dipoles. The molecules can be trapped either in optical lattices or in arrays of tweezers. They also implement spin models with XY and Ising-like terms in the interaction Hamiltonian [14].

3.4 Quantum circuits [17]

Quantum circuit-based platforms have made impressive progress over the last decade, culminating with the recent demonstration of a quantum computation with 53 connected circuits [18]. This platform exists in various implementations (phase, charge, flux quantum bits). They all rely on three interconnected ingredients: (i) the pairing of electrons into Cooper pairs (charge $2e$) at low temperature in some metallic conductors; (ii) the superconductivity which allows the Cooper pairs to propagate almost without resistivity through the conductors; (iii) a Josephson Junction, which provides a non-linear inductance.

Consider one of the simplest case to understand, called a Cooper-pair box, represented in Fig. 5(a) (today, a popular implementation, the transmon qubit, is a variant of the Cooper-pair box). It consists of two Josephson junctions in a loop placed between two superconducting islands. The Cooper pairs have an amplitude E_J to tunnel through the junctions, and the flux through the loop ϕ is controllable by external means (e.g. external potentials). The junction also acts as a capacitor C . The Hamiltonian describing this situation can be written in the limit $(2e)^2/C \ll E_J$ [19]

$$H_J = \frac{Q^2}{2C} - E_J \cos\left(2\pi \frac{\phi}{\phi_0}\right). \quad (18)$$

with Q the charge difference between the two islands, and $\phi_0 = h/(2e)$. As the induction laws giving the voltage drop between the islands is $\dot{\phi} = -Q/C$, the Hamiltonian (18) is analogous to the one of a mechanical pendulum in a potential $V(\phi) = -E_J \cos\left(2\pi \frac{\phi}{\phi_0}\right)$ (Fig. 5b). It is non-harmonic for large fluxes.

When quantizing this non-linear oscillator, energy levels appear, with a frequency separation between the two lowest states $\omega_{01} = \sqrt{(E_J/\hbar^2)(2e)^2/C}$. For a typical experimental realizations using transmons, $\omega_{01}/(2\pi) \approx 5$ GHz. The non-linearity is weak, as the second energy spacing ω_{12} differs from ω_{01} by only about $2\pi \times 200$ MHz. For the quantum treatment to be relevant, one has to work at low temperature to avoid thermally populating the excited states. In practice, this means operating at a temperature lower than $\hbar\omega_{01}/k_B \approx 100$ mK. These experiments thus operate in a cryogenic environment.

Let us consider first the case where we isolate the two lowest levels, and let us map them onto a spin-1/2: $|\downarrow\rangle = |n=0\rangle$ and $|\uparrow\rangle = |n=1\rangle$. A microwave tuned near the frequency ω_{01} drives the transition and the corresponding Hamiltonian is $\hbar\Omega(\hat{\sigma}^+ + \hat{\sigma}^-)/2$, with Ω the Rabi frequency, $\hat{\sigma}^+ = |\uparrow\rangle\langle\downarrow|$ and $\hat{\sigma}^- = |\downarrow\rangle\langle\uparrow|$. Two such circuits can also be coupled together by connecting them by a microwave guide. Consider the mode of the waveguide with the frequency ω the closest to the resonance frequency ω_{01} of the two-level system. Its quantized electric field is $\hat{E} = \sqrt{\hbar\omega/(2\epsilon_0 V)}(\hat{a} + \hat{a}^\dagger)$ (V is the mode volume). The two coupled circuits thus implement the Jaynes-Cumming Hamiltonian:

$$H_{JC} = \hbar\omega_{01}(\hat{\sigma}_1^z + \hat{\sigma}_2^z) + \hbar\omega \hat{a}^\dagger \hat{a} + \hbar g(\hat{\sigma}_1^+ + \hat{\sigma}_1^- + \hat{\sigma}_2^+ + \hat{\sigma}_2^-)(\hat{a} + \hat{a}^\dagger), \quad (19)$$

with g the coupling constant between the circuits and the waveguide, which acts as a cavity. If the waveguide cavity is detuned by $\Delta = \omega - \omega_{01}$ so that $|\Delta| \gg g$, it is possible to eliminate the field and obtain an effective Hamiltonian containing only the circuits' degrees of freedom. Consider the two circuits in opposite state and an empty waveguide, i.e. the (circuits + field) state $|\uparrow, \downarrow, 0\rangle$. The first circuit can off-resonantly drop a photon in the cavity waveguide, which gets absorbed by the second circuit, leading to $|\downarrow, \uparrow, 0\rangle$ via the intermediate state $|\downarrow, \downarrow, 1\rangle$ shifted in energy by $\hbar\Delta$. As for the derivation of the Heisenberg model for atoms in optical lattices (Eq. 14), perturbation theory gives the amplitude of this spin-flip process: $\hbar g^2/\Delta$. The effective coupling Hamiltonian is now:

$$H_{\text{eff}} = \frac{\hbar g^2}{\Delta}(\hat{\sigma}_1^+ \hat{\sigma}_2^- + \hat{\sigma}_1^- \hat{\sigma}_2^+). \quad (20)$$

This is the coupling needed to implement the XY spin model introduced earlier.

Interestingly, the intrinsic non-linearity introduced by the Josephson junctions also allows a mapping onto the Bose-Hubbard model. To understand this, let us expand the cosine in Eq. (18) up to fourth order, and write the flux as a quantized operator $\hat{\phi} \propto \hat{b} + \hat{b}^\dagger$, with \hat{b} (\hat{b}^\dagger) the annihilation (creation) operator of a flux quantum. The expansion leads to

$$H_J \approx \frac{\hat{Q}^2}{2C} + \frac{4\pi^2 E_J}{2\phi_0^2} \hat{\phi}^2 - \frac{4\pi^4 E_J}{\phi_0^4} \hat{\phi}^4 = \hbar\omega_{01} \hat{b}^\dagger \hat{b} + U \hat{n}(\hat{n} - 1), \quad (21)$$

with $U = 4\pi^4 E_J$, $\hat{n} = \hat{b}^\dagger \hat{b}$, using $[\hat{b}, \hat{b}^\dagger] = 1$ and neglecting offset terms when restricting to up to $n = 2$ excitations. In the language of creation and annihilation operators, the effective

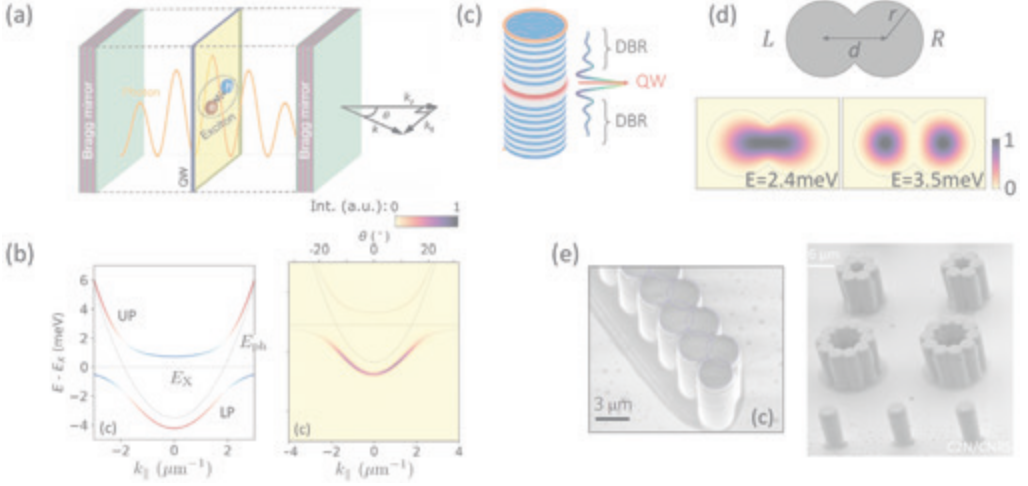


Fig. 6: Polaritonic platform. (a) Quantum well (QW) containing one exciton placed in the middle of a cavity made of Bragg mirrors. Figure from [20]. (b) (Left) Calculated energy spectrum $E_{L,U}(k)$ for the upper and lower polaritons resulting from the coupling between the excitons and the cavity field (parameters: $\Omega/[E_{ph}(0) - E_X(0)] = 1$). (Right) Light intensity measured on a sensitive camera placed in the far-field of a QW pumped by a laser: the spatial intensity distribution directly maps the relation dispersion $E_{L,U}(k)$. (c) Micropillar cavity containing a quantum well. (d) Radial intensity distribution for the two lowest modes of the electromagnetic field inside two coupled micropillars. (e) Examples of arrays of micropillars realizing a tight-binding model. Figures (b)-(e) from [21].

exchange Hamiltonian (20) takes the form $\frac{\hbar g^2}{\Delta}(\hat{b}_1^\dagger \hat{b}_2 + \hat{b}_1 \hat{b}_2^\dagger)$. If now N circuits are coupled by waveguides (nearest-neighbor couplings), one obtains the Bose-Hubbard Hamiltonian:

$$H_{BH} = \frac{\hbar g^2}{\Delta} \sum_{\langle i,j \rangle} (\hat{b}_i^\dagger \hat{b}_j + \hat{b}_i \hat{b}_j^\dagger) + U \sum_{i=1} \hat{n}_i (\hat{n}_i - 1). \quad (22)$$

For a long time, people were skeptical about the scalability of the quantum circuit-based approach. The recent demonstration of a quantum algorithm with up to 53 connected circuits [18] indicates that this platform now has the potential to scale up. The high degree of control achieved by the circuit-based platforms also allows for digital quantum simulation.

3.5 Quantum fluids of light with polaritons in semiconductor cavities [20]

This platform consists of hybrid particles, called polaritons, with photonic and matter components. It combines excitons in a semi-conductor material with photons confined in a high finesse cavity in the strong coupling regime. As the light leaks out of the cavity, the polaritons are very short-lived. Hence the system needs to be continuously pumped by a laser. Contrarily to the platforms described in the previous sections, this one is intrinsically dissipative and must operate under the drive of a laser. It is therefore an ideal platform to perform quantum simulation of driven dissipative systems.

The first ingredient of the platform is a cavity of length L made of two Bragg mirrors surrounding a semi-conductor material of refractive index n_{cav} . Along the cavity axis (Oz) the wavevector of the light is quantized: $k_z = p\pi/L$ (p is an integer). The total wavevector is thus $k = \sqrt{k_{\parallel}^2 + (p\pi/L)^2}$. Expanding around low k_{\parallel}/k_z yields a quadratic dispersion relation

$$E_{\text{ph}}(k_{\parallel}) = \hbar\omega(k_{\parallel}) = \frac{\hbar kc}{n_{\text{cav}}} = C + \frac{\hbar^2 k_{\parallel}^2}{2M_{\text{ph}}}, \quad (23)$$

with $M_{\text{ph}} = p\pi\hbar n_{\text{cav}}/(Lc)$ the effective mass of the photon in the plane parallel to the mirrors.

The second ingredient is an exciton. It is a bound state made of an electron excited to the conduction band and a hole in the valence band in a semiconductor material (e.g. in InGaAs, with gap energy 1.35 eV). If the semiconductor where the exciton is formed is placed between two different semi-conductors with larger gap (e.g. GaAs, with gap energy 1.45 eV) separated by a distance d , perpendicular to the (Oz) axis, the exciton is confined in the layer of thickness d in the z direction with a wave function $\phi_n(z)$. This arrangement forms a quantum well (QW), with the exciton free to move in the plane perpendicular to (Oz). The wave function of the external degree of freedom of the exciton in the QW is thus $\psi_X(\mathbf{r}, z) = \phi_n(z) \exp(i\mathbf{K}_{\parallel} \cdot \mathbf{r})$, with an energy $E_X = \hbar^2 K_{\parallel}^2/(2M_X)$ and the exciton mass $M_X \gg M_{\text{ph}}$.

Let us now place the quantum well inside the cavity. The electric dipole moment of the exciton couples to the cavity field, and the coupled exciton-cavity system is described by the Jaynes-Cumming Hamiltonian:

$$H_{X-\text{cav}} = E_{\text{ph}}(\mathbf{k}) \hat{a}_{\mathbf{k}}^{\dagger} \hat{a}_{\mathbf{k}} + E_X(\mathbf{k}) \hat{b}_{\mathbf{k}}^{\dagger} \hat{b}_{\mathbf{k}} + \frac{\hbar\Omega}{2} (\hat{a}_{\mathbf{k}}^{\dagger} \hat{b}_{\mathbf{k}} + \hat{b}_{\mathbf{k}}^{\dagger} \hat{a}_{\mathbf{k}}), \quad (24)$$

with \mathbf{k} the wavevector in the plane perpendicular to (Oz), Ω the Rabi frequency associated to the exciton-cavity coupling and $(\hat{a}_{\mathbf{k}}, \hat{a}_{\mathbf{k}}^{\dagger})$, $(\hat{b}_{\mathbf{k}}, \hat{b}_{\mathbf{k}}^{\dagger})$ the bosonic operators associated to the photons and excitons respectively. The lowest energy eigenstates of the Hamiltonian mix one photon and one exciton:

$$|\text{pol}_L\rangle = \cos(\theta_k) |X\rangle + \sin(\theta_k) |\text{ph}\rangle; |\text{pol}_U\rangle = \sin(\theta_k) |X\rangle - \cos(\theta_k) |\text{ph}\rangle, \quad (25)$$

with $\tan(2\theta_k) = \Omega/[E_X(\mathbf{k}) - E_{\text{ph}}(\mathbf{k})]$. The corresponding eigen-energies $E_{L,U}(k)$ are represented in Fig. 6(b). These two states, called Lower and Upper Polaritons, behave as hybrid particles with momentum \mathbf{k} . Their mass is adjustable by changing the detuning between the cavity and the exciton, e.g., $1/M_L = \cos^2(\theta_k)/M_X + \sin^2(\theta_k)/M_{\text{ph}}$. Through its photonic component, the particle emits light that exits the cavity. The k_{\parallel} -component of the light is conserved, meaning that light intensity pattern in the far-field directly maps onto the dispersion relation of the polariton $E_{L,U}(k_{\parallel})$, as shown in Fig. 6(b).

When the pumping intensity is large enough, several polaritons can be prepared in the system, that now interact via their exciton component. The interaction between two excitons results from their Coulomb interaction. It can be described by a contact interaction $V_{XX}\delta(\mathbf{r})$ with $V_{XX} = 3e^2 a_B/\epsilon$ for two excitons of the same spin (a_B is the 2D exciton Bohr radius, and ϵ is the dielectric constant of the QW material). Hence the interaction between, e.g., two lower polaritons has a contact constant $g = V_{XX} \cos^4(\theta_k)$. This interaction is usually quite weak and is treated in the mean-field approach. Therefore, an ensemble of interacting polaritons can also be

viewed as a fluid of interacting photons (especially in situation where they form a Bose-Einstein condensate), hence the name *quantum fluid of light*.

It is possible to confine the polaritons in 3D and to arrange them in arrays of various geometries in order to emulate electrons in lattices. To do so, one starts from micropillar cavities with the QW layer in the middle (Fig. 6c). The small radial size of the pillars now ensures a 3D confinement of the polaritons. Placing two micropillars close to each other, the solution of Maxwell's equations in the resulting cavity presents new eigenmodes that are *delocalized* on the two pillars as shown in Fig. 6(d), and with energy difference ΔE . Conversely, these eigenmodes can be seen as resulting from the symmetric and anti-symmetric combinations of two modes *localized* in each micropillar with a coupling amplitude $J = \Delta E/\hbar$. Hence this coupled micropillars cavity is mapped onto a two coupled-wells problem (with states $|L\rangle$ and $|R\rangle$) described by the Hamiltonian $H = -J(|L\rangle\langle R| + |R\rangle\langle L|)$ in the tight-binding limit. Nanotechnology techniques now allow fabricating arrays of micropillars, as illustrated by a few examples in Fig. 6(e): benzene-like structures, 1D chain, honeycomb lattice, etc. These 1D or 2D periodic structures are characterized by band-structures with dispersion relations $E(\mathbf{k})$, which can be directly visualized by monitoring the light emitted in the far-field, as in Fig. 6(b). If the number of excitons in the structure is large enough, they interact, although at the moment only in the mean-field. One important current goal for this platforms is to increase the interactions between polaritons to reach the correlated regime.

3.6 Other platforms

Many other platforms are considered for quantum simulation, such as free space photons [22] nuclear spins, semi-conductor quantum dots, NV-centers, etc., with the list growing regularly [1].

4 Examples of quantum simulations

Many problems can be studied using the quantum simulation approach. These include: quantum magnetism, quantum chemistry, quantum transport, topological matter, high energy physics, lattice gauge theories, cosmology, etc. [1]. New ranges of applicability appear regularly. In this Section, we illustrate the concepts introduced in the first two sections on a few recent experiments related to the applications mentioned above, using the platforms described in Sec. 3. Here again we will not attempt to be exhaustive. In a biased way, most of the examples will belong to atomic-based platforms and will illustrate the analog approach to quantum simulation as it is currently the most advanced.

4.1 Out-of-equilibrium dynamics

One of the simplest ways to study many-body physics questions with a quantum simulator consists in placing it out-of-equilibrium. This is usually achieved by varying abruptly one parameter of the Hamiltonian (a process called a “quench”), after its preparation in a state easy to access, and monitoring the ensuing dynamics. This dynamics is very challenging to calculate

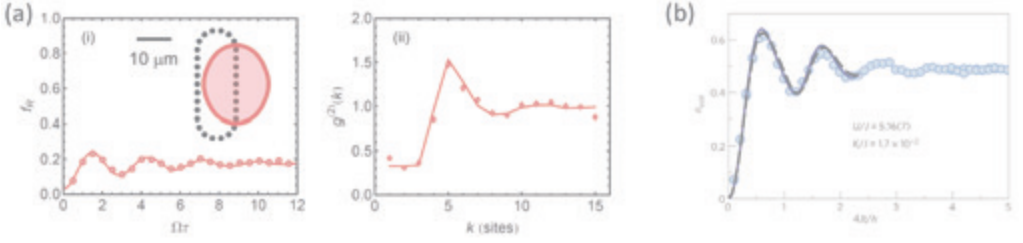


Fig. 7: Out-of-equilibrium experiments. (a) Dynamics in the Ising-like model using a Rydberg quantum simulator implementing a chain with periodic boundary conditions. (Left) Dynamics of the fraction f_R of atoms excited to a Rydberg state equivalent to the average magnetization $\langle \hat{\sigma}^z \rangle$ as a function of time τ in unit of Ω . (Right) Pair correlation function given by Eq. (27) for $\Omega\tau = 0.3$. For both figures the solid lines are the *ab-initio* solutions of the Schrödinger equation. Figures from [23]. (b) Dynamics of the Bose-Hubbard model after suddenly turning on the tunneling between lattices sites in a 1D array with only the even sites initially filled. Measured mean occupation $\langle n_{\text{odd}} \rangle$ of the odd sites as a function of the time in unit of the tunneling constant J/\hbar . The solid lines are the results of a *t*-DMRG calculation including the residual imperfections of the experiment. Solid line: nearest neighbor hoppings only. Dashed line: includes also next-nearest neighbor hoppings. Figure from [25].

by traditional methods, as it involves the knowledge of the full spectrum of the Hamiltonian (see below). Out-of-equilibrium situations are therefore ideal testbeds for a quantum simulation approach.

To see how the dynamics involves the spectrum, consider a many-body Hamiltonian H_{MB} and denote $|\psi_\alpha\rangle$ the eigenstates and E_α the corresponding eigen-energies. The initial state of the simulator $|\psi(0)\rangle$ right after the quench is $|\psi(0)\rangle = \sum_\alpha c_\alpha |\psi_\alpha\rangle$. After a time t , the system driven by H_{MB} evolves as $|\psi(t)\rangle = \sum_\alpha c_\alpha e^{-iE_\alpha t/\hbar} |\psi_\alpha\rangle$. Hence, the expectation value of an observable \hat{O} is

$$\langle \hat{O} \rangle = \langle \psi(t) | \hat{O} | \psi(t) \rangle = \sum_{\alpha, \alpha'} c_\alpha c_{\alpha'}^* \langle \psi_{\alpha'} | \hat{O} | \psi_\alpha \rangle e^{-\frac{i}{\hbar}(E_\alpha - E_{\alpha'})t}. \quad (26)$$

This shows that the evolution of the expectation value of the observable results from the interference of all the eigen-frequencies of the Hamiltonian.

Let us illustrate the out-of-equilibrium approach on a few examples.

Dynamics in the quantum Ising model. The first one uses the mapping of spin models onto ensembles of atoms interacting by the van der Waals interaction when excited to Rydberg states (Rydberg quantum simulator). The experiment, done at the Institut d’Optique, relies on a chain of atoms held in optical tweezers, wrapped on itself to realize periodic boundary conditions [23] (see Fig. 7a). We use Rydberg states to mediate the interaction and implement the Ising-like Hamiltonian of Eq. (16). The Rydberg excitation laser mimicking the magnetic field is applied at resonance ($\Delta = 0$), after having prepared the atoms in their ground state (i.e. spin $|\downarrow\rangle$). We then measure two quantities relevant to the study of spin systems. First, the average magnetization $\langle \sigma^z \rangle = \sum_i \langle \sigma_i^z \rangle / N$, related to the average number of atoms excited to the Rydberg states, or equivalently in the spin state $|\uparrow\rangle$. We observe an oscillatory evolution of the magnetization, as shown in Fig 7(a). In particular the system seems to equilibrate at long time. Second, we

use the ability to measure the state of each atom after an evolution time to reconstruct the spin-spin correlation function $g^{(2)}(k)$, which is the probability to find a Rydberg excitation at site j when one is already present at site i :

$$g^{(2)}(k) = \frac{\langle n_i n_{i+k} \rangle}{\langle n_i \rangle \langle n_{i+k} \rangle}. \quad (27)$$

This correlation function varies as a function of time. It is represented in Fig 7(a) for a particular evolution time. It exhibits a shape characteristic of the correlation function observed in a liquid, with a suppressed probability of excitation of two nearby atoms due to the strong energy cost associated to it. The solid lines in the figures are the results of *ab-initio* calculations using the Schrödinger equation with no adjustable parameters. The good agreement between theory and data obtained in a regime where benchmarking against exact diagonalization is possible validates the experimental platform. We also considered an array of 7×7 , for which an exact diagonalization is not possible [24]. In this regime, we cross validate the approximate theory with the accuracy of the experimental implementation of the model Hamiltonian.

Dynamics in the Bose-Hubbard model. As a second experiment, we discuss a quantum quench experiment performed on the 1D Bose-Hubbard model in the group of I. Bloch [25]. They initialized a one-dimensional lattice so that only the even sites are filled with exactly one atom. The lattice depth is initially large to suppress tunneling between the sites. They suddenly switch on the tunneling by reducing the depth and measure after an evolution time t the occupation number n_{odd} in the odd sites initially empty. The system is driven both by the tunneling J and the on-site interaction U . They observe also a damped oscillatory behavior (see Fig. 7b), that they compare with a state-of-the-art t -DMRG calculations including all the knowledge of the system's parameters. Here also the agreement between theory and experiment is very good at short time. Interestingly, for the largest number of atoms ($N \approx 43$), the theory can not calculate the evolution after a certain duration, showing that the quantum simulator gives access to timescales and regimes (here the steady-state) not accessible to the theory.

Equilibrium of closed systems and many-body localization. In the two experiments described above, the systems evolve towards a steady-state that can not be explained by imperfections of the experiment, but results from the beat-notes of the eigen-frequencies of the many-body Hamiltonian (Eq. 26). This equilibration is therefore a genuine many-body effect. The appearance of a steady-state at long time hence raises the question of how closed many-body systems (as are the two implementations described above) relax to equilibrium. The principle of maximum entropy specifies what should be the equilibrium state but not how the system reaches it. It was assumed for a long time that the equilibrium state of a many-body system obeys the Eigenstate Thermalization Hypothesis (ETH). The hypothesis states that when considering a subpart A of the system and tracing over the degrees of freedom of the complementary subpart, the resulting density matrix of A , ρ_A , should be thermal, i.e. $\rho_A \propto \exp[-\beta H_A]$. The system thus acts as his own bath. It was later realized that this may not always be true. For example, by adding some disorder in the system (in a sense described below), the many-body system may retain the memory of its initial state and never thermalize. This effect is called Many-Body Localization (MBL) and can be seen as the extension of Anderson localization for interacting systems. Quantum simulators are ideal platforms to investigate MBL, as they allow controlling the degree of disorder. Let us describe two recent experiments investigating MBL in one and two dimensions respectively.

In the first experiment, the group of C. Monroe in the US used a chain of 10 ions implementing

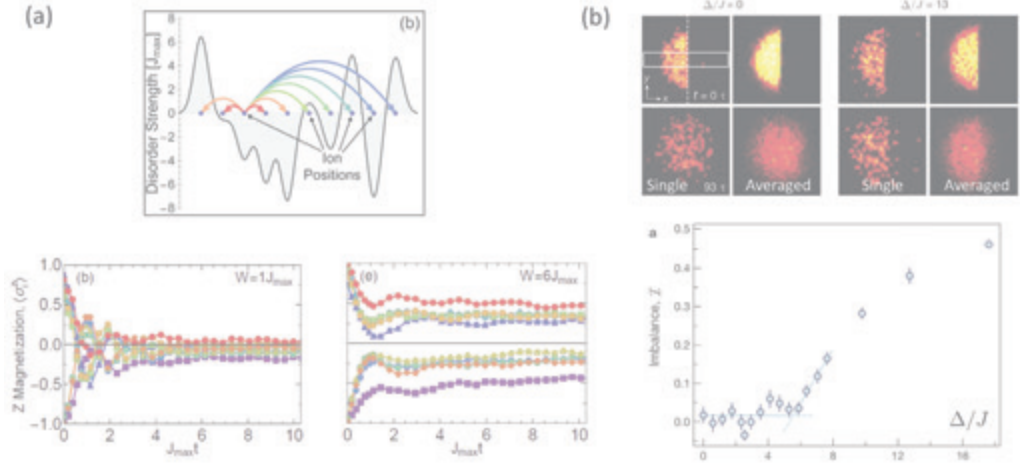


Fig. 8: Many-Body localization experiments. (a) (Upper) Chain of 10 ions with random shifts of the resonant frequency of the $(|\downarrow\rangle, |\uparrow\rangle)$ transition. (Lower) Expectation value of the magnetization $\langle\sigma_i^z\rangle$ for small (left) or large (right) disorder as a function of time. Figures from [26]. (b) MBL in a 2D square array with random disorder from site-to-site. (Upper) Fluorescence images obtained with a quantum gas microscope of the atomic distribution at the start ($t = 0$) and after an evolution time $t = 93\tau$ ($\tau = \hbar/J$), for small and large disorders. (Lower) Normalized imbalance $\mathcal{I} = (N_L - N_R)/(N_L + N_R)$ after a time $t = 187\tau$ as a function of the disorder amplitude Δ/J . Figures from [27].

the Ising model in a transverse magnetic field in the way described in Sec. 3.2 [26]. Using extra laser beams they could imprint a random dispersion on the resonance frequencies of each ion (see Fig. 8a), thus adding to the Hamiltonian a term of the form $\sum_i D_i \hat{\sigma}_i^z$. Here, the D_i 's are sampled from a uniform random distribution ranging from $-W$ to W . They initialized the ion chain in the Néel state $|\uparrow, \downarrow, \uparrow, \downarrow, \uparrow, \downarrow, \uparrow, \downarrow, \uparrow, \downarrow\rangle$ and let it evolve after turning on the coupling constants J_{ij} between the ions. They observed the subsequent evolution of the expectation values $\langle\hat{\sigma}_i^z\rangle$ of each ion for different degrees of disorder set by W . The results are presented in Fig. 8(a): for $W = 0$, the system does relax and they could check by measuring the density matrix locally that the ETH is valid in this case. On the contrary, for $W \neq 0$, the ions do not relax towards $\langle\hat{\sigma}_i^z\rangle = 0$ but rather keep track of their initial state. This is the signature of MBL and of the violation of the ETH. With 10 ions, the system is numerically easy to calculate. However, this experiment is a nice illustration of the power of a quantum simulator: it provides new tools (e.g. controllable disorder) and new probes (e.g. local tomography of the state) that allows one to think in a different way about those systems.

Many open questions remain about MBL. Among them was the one of the existence in 2D of a transition between localized and non-localized states as a function of the strength of the disorder. The answer was obtained experimentally in 2016 by using a quantum simulator based on a quantum gas microscope [27]. In this work, the group of I. Bloch and C. Gross demonstrated that a transition does occur, and more, they were able to measure its position in terms of disorder strength, in a regime inaccessible to the theory. Hence their experiment is in fact the quantum simulation of the MBL transition in 2D, with a machine that “calculates” some parameters of the transition. The experiment relied on a 2D optical lattices. They filled the lattice sites

with individual atoms in a semi-circular region, with a sharp boundary (see Fig. 8b). They could control the initial filling fraction. The lattice depth was initially large so that tunneling is frozen. On top of the lattice, they projected a random light intensity pattern, which adds to the Bose-Hubbard Hamiltonian of Eq. (11) a term $\sum_i \delta_i \hat{n}_i$, with the δ_i randomly distributed over a distribution of width Δ . At a given time they decreased the lattice depth to induce tunneling between the sites. After some time they measure the imbalance $\mathcal{I} = (N_L - N_R)/(N_L + N_R)$ between the left region of the circle initially filled and the right region initially empty. For low disorder $\Delta \ll J, U$, they found that the atoms occupy the full circle at long time and that the imbalance converges to zero. However, for a ratio Δ/J larger than a critical value Δ_c/J , the imbalance does not converge to zero any longer, but to a finite value, indicating that the systems retains the memory of its initial preparation. This demonstrates the existence of a transition between many-body localized and delocalized states in 2D. Moreover the value of the transition is “calculated” by the simulator and found to be $\Delta_c/J \approx 5.5$ for the largest initial filling. This regime, where $\Delta \sim U \sim J$ is inaccessible numerically for the number of atoms used in the experiment (125), thus showing the power of the quantum simulation approach. Many-body localization has also been recently investigated with up to 21 quantum circuits [28].

Propagation of correlations and Lieb-Robinson bound. To finish this section on out-of-equilibrium dynamics, let us discuss the question of the speed at which the correlations between the particles build up during the evolution. We will mainly discuss this problem in relation with the XY spin model, described by a flip-flop interaction $J_{ij}(\hat{\sigma}_i^+ \hat{\sigma}_j^- + \hat{\sigma}_i^- \hat{\sigma}_j^+)$. The first thing to realize is that one can think of this flip-flop term as leading to the transport of a spin excitation $|\uparrow\rangle$, which now acts as a quasi-particle. With this image in mind, the XY model is well adapted to study the transport properties of this quasi particle. Importantly, this spin excitation can also interact when placing several of them in the same chain. The question of the propagation speed of the correlations was first considered from a mathematical point of view by Lieb and Robinson [29]. They found that the connected correlation function for, e.g., the σ_i^z operators $C_{ij}(t) = \langle \sigma_i^z(t) \sigma_j^z(t) \rangle - \langle \sigma_i^z(t) \rangle \langle \sigma_j^z(t) \rangle$ was bounded:

$$|C_{ij}(t)| \leq A \exp[-(|i - j| - 2v_{\text{LR}}t)/\xi], \quad (28)$$

with v_{LR} the Lieb-Robinson velocity, related to the coupling constant between the sites. For example, for the Ising model with nearest neighbor interaction J , $v_{\text{LR}} = 6eJ$. This bound thus predicts the equivalent of a “light-cone”, but with the boundary of the cone being exponentially soft over a distance ξ . For short-range interactions, the Lieb-Robinson bound is well understood. However the case of long-range to infinite range interactions is much less clear. Although this may look like a mathematical problem, as no interaction is infinite range, this question can now be investigated experimentally by quantum simulators based on ions (Sec. 3.2). This points out again one of the virtue of the quantum simulation approach presented in the Introduction: although the problem of Lieb Robinson could be explored numerically to up ~ 30 ions, the existence of an experimental realization makes it look much less like an exotic mathematical problem. The simulator thus acts as a trigger for a more in-depth investigation prone to an experimental validation.

Two experiments using chains containing around 10 ions explored the Lieb-Robinson bound for long-range interaction [30, 31]. They placed the chain out of equilibrium by either a global quench induced by suddenly switching on the interactions between the spins with a laser [30], or a local quench by exciting one ion in state $|\uparrow\rangle$, all the others being in $|\downarrow\rangle$ [31]. After some evolution time, they measured both the magnetization $\sigma_i^z(t)$ for each ion and the connected

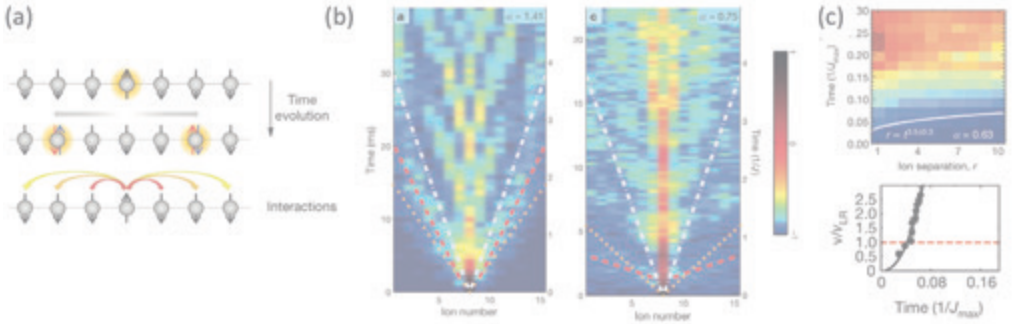


Fig. 9: Transport and Lieb-Robinson bound. (a) Principle of the experiment studying the transport of excitations and the buildup of correlations in a chain of 15 ions implementing the XY spin model with coupling constant $J_{ij} \propto 1/|i - j|^\alpha$. The spin of the central ion is flipped to $|\uparrow\rangle$, all the others being in $|\downarrow\rangle$. (b) After an evolution time t the magnetization of each ion is measured and plotted in a 2D map for two values of α . The case $\alpha \approx 1.4$ shows the appearance of a “light-cone” for the propagation of the magnetization (edge in red dashed lines). On the contrary, the case $\alpha = 0.75$ indicates the violation of the light-cone picture. Figures (a) and (b) from [30]. (c) (Upper) Propagation of the correlations $C_{i,i+r}(t)$ following a global quench ($\alpha = 0.63$). The white solid line is the light-cone boundary defined as the time it takes for a correlation of a fixed amplitude (here $C_{i,i+r} = 0.1C_{i,i+r}^{\text{max}}$) to reach a given distance r . (Lower) Velocity of the boundary dr/dt as a function of time. The Lieb-Robinson limit is rapidly violated. Figures from [31].

correlation function $C_{ij}(t)$ (see Fig. 9). They performed the experiment for various power law of the coupling constant: $J_{ij} \propto 1/|i - j|^\alpha$, with α ranging from 0.5 to 2. They could observe that for $\alpha \geq 1$, the Lieb-Robinson cone exist, while it disappears for $\alpha \leq 1$. In this last regime, the Lieb-Robinson bound (28) is violated, and the speed at which the correlations build up is larger than v_{LR} .

4.2 Adiabatic preparation of ground states and quantum phase transition

The second way to study many-body systems consists in preparing them in their ground state, and then study their properties. Contrarily to the out-of-equilibrium case, here, only the low part of the energy spectrum of the Hamiltonian is relevant. To prepare this ground state, one usually applies an adiabatic approach, which amounts to slowly varying the parameters of the Hamiltonian to drive the system in the many-body ground state. Let us assume that one wants to reach the ground state of a particular Hamiltonian H_{MB} , starting from an Hamiltonian H_0 whose ground state $|\psi_{\text{ini}}\rangle$ is easy to prepare. Varying a parameter $\lambda(t)$ between 0 and 1 allows changing the total Hamiltonian from H_0 to H_{MB} :

$$H(t) = (1 - \lambda(t))H_0 + \lambda(t)H_{\text{MB}}. \quad (29)$$

Provided that $\lambda(t)$ varies slowly enough, the state of the system continuously changes from $|\psi_{\text{ini}}\rangle$ to the desired $|\psi_{\text{MB}}\rangle$. This statement is valid if the ground state energy of $H(t)$ is at all time separated from the first excited state by a gap $\Delta E(t) = E_1(t) - E_0(t)$ to prevent the

excitation of the system during the preparation. The validity condition for the sweep to be adiabatic is (see [32], Eqs. from (35.21), p. 289):

$$|\langle \psi_1(t) | \frac{dH}{dt} | \psi_0(t) \rangle| \ll \frac{\Delta E(t)^2}{\hbar}, \quad (30)$$

with $|\psi_0(t)\rangle$ ($|\psi_1(t)\rangle$) the instantaneous ground and first excited states.

Preparation of a Mott insulator phase. This adiabatic procedure is in fact the one used to prepare one atom per site in an optical lattice starting from a Bose-Einstein condensate [13]. The Bose-Hubbard Hamiltonian features a quantum phase transition (called the superfluid-insulator Mott transition) at a critical value $(U/J)_c$ between a superfluid phase for $U/J < (U/J)_c$ and an insulating phase for $U/J > (U/J)_c$. This insulating phase corresponds to the situation where exactly one atom is placed at the sites of the lattices. Indeed, for $U \gg J$, emptying one site and placing two atoms on the neighboring one has an energy cost of U , and corresponds to the first excited state. Thus starting from a Bose-Einstein condensate and gradually increasing the lattice depth drives the system from the ground state at low U/J (atoms delocalized over the sites), to the ground state consisting of one atom per site (see however the discussion below about the phase transition).

Adiabatic preparation of an Ising anti-ferromagnet on a Rydberg simulator. As a second example, consider the Ising model in a transverse magnetic field emulated by an ensemble of atoms in arrays of tweezers driven by a laser, and in van der Waals interaction when in Rydberg states, as described in Sec. 3.3 and Eq. (16). Figure 10(a) presents the generic phase diagram for this Hamiltonian at zero temperature and for spins placed on a chain or on two-dimensional square arrays. In the case of nearest-neighbor couplings only ($V_{i,i+1} = V$ and 0 otherwise) and with $V > 0$, so that the interactions favor anti-ferromagnetic ordering, the phase diagram consists of two regions: a paramagnetic (PM) and an anti-ferromagnetic (AF) one, separated by a quantum phase transition (see below). Two limiting cases are easy to understand: for $\Omega, \Delta \gg V$, the ground state is paramagnetic, i.e. the spins align along the effective magnetic field; for $\Omega = 0$, the phases results from the minimization of the energy of the classical configuration. When we relax the constraint of nearest-neighbor couplings only (as is the case for a van der Waals interaction) the phase diagram exhibits several phases around the line separating the PM from the AF phases. For example, on a chain, if $V_{i,i+1}, V_{i,i+2} \gg \Delta \gg \Omega \gg V_{i,i+3}$, the ground state corresponds to one excitation separated by two ground state atoms (Z_3 symmetry). Similarly, $V_{i,i+1}, V_{i,i+2}, V_{i,i+3} \gg \Delta \gg \Omega \gg V_{i,i+4}$ leads to a phase with Z_4 -symmetry, and so on. By controlling the detuning Δ and Rabi frequency Ω , one can explore the phase diagram of this Ising model.

Several recent experiments explored this approach. For example, the group of M. Lukin in Harvard investigated the one-dimensional case [33] with up to 51 atoms. The idea of the experiment is to start from an initial state where all the atoms are prepared in their ground state ($|\downarrow\rangle$). Then, by sweeping the Rabi frequency $\Omega(t)$ and detuning $\Delta(t)$, the system is driven adiabatically in the ground state of the interacting ensemble for a given final value of Ω and Δ . They could access several Z_n phases by controlling the final ratio of $\Omega/V_{i,i+1}$, as shown in Fig. 10(b). Our group at the Institut d'Optique [34] and the group of W. Bakr in Princeton [35] explored the two-dimensional case using respectively atoms in arrays of tweezers (36 atoms) and in optical lattices (200 atoms) both with square unit cells. The procedure is the same as in the one-dimensional case. Both groups observed the appearance of antiferromagnetic correlations (see

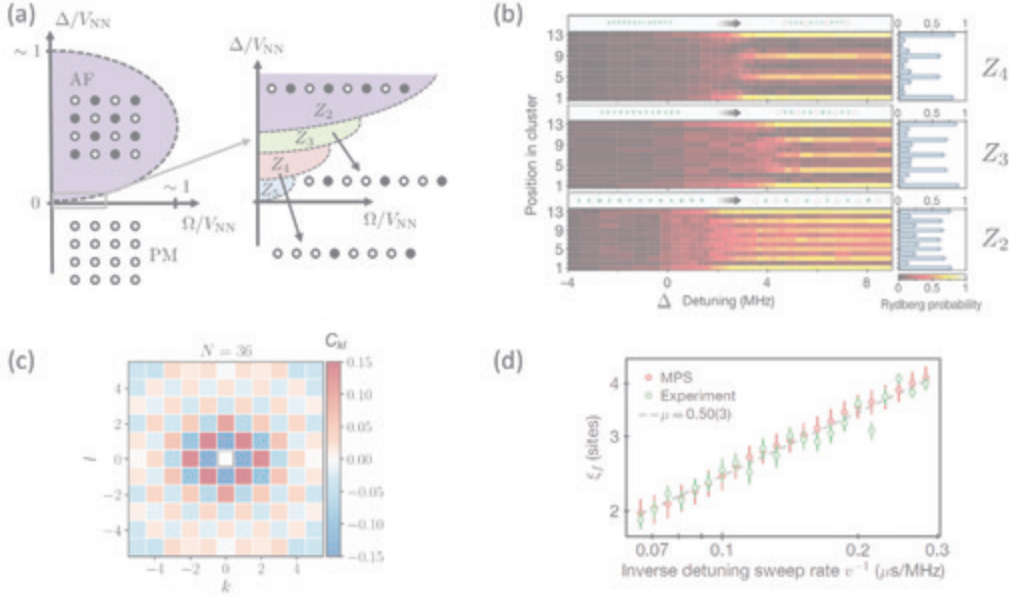


Fig. 10: Adiabatic preparation and quantum phase transition. (a) Schematic phase diagram of the quantum Ising model of Eq. (16) for a 2D square or a 1D chain of atoms interacting via the van der Waals interaction, showing the paramagnetic (PM) and antiferromagnetic (AF) phases. Here $V_{NN} = C_6/a^6$, with a the spacing between atoms (we take $C_6 > 0$). The right side is a zoom of the phase diagram around the critical point ($\Omega = 0, \Delta = 0$), for the 1D chain. (b) Adiabatic preparation of the ground state of a 13-atom chain in a Rydberg quantum simulator [33] for various values of Ω/V_{NN} : this gives an excitation every second site, every third site, or every fourth site, corresponding to different ordered phases Z_n . (c) Antiferromagnetic spin-spin correlation function in two dimensional square tweezers arrays of atoms after adiabatic preparation for the experiment of Ref. [34]. (d) Antiferromagnetic correlation length ξ obtained in a chain of 51 atoms of a Rydberg quantum simulator as function of the detuning sweep rate [36]. The slope gives access to critical exponents characterizing the quantum phase transition.

Fig. 10c) by measuring the connected spin-spin correlation function:

$$g^{(2)}(k, l) = \frac{1}{N_{kl}} \sum_{(i,j)} [\langle n_i n_j \rangle - \langle n_i \rangle \langle n_j \rangle]. \quad (31)$$

Fitting this correlation function by a decaying exponential yields the correlation length ξ , on the order of $1 - 1.5a$, with a the distance between the sites. This length was limited by experimental imperfections, but was in good agreement with a modeling of the system including them.

The group at the Institut d'Optique also started to explore the case of arrays arranged in a triangular configuration [34]. This case is particularly interesting as it exhibits geometrical frustration [37]. Consider three spins placed at the corner of an equilateral triangle with antiferromagnetic Ising interaction. The pair-interaction energy is minimized for two opposite spins. This means that the minimum energy of three atoms can not be minimized by minimizing the energy on each bond. This leads to a ground state featuring entanglement and large scale degeneracy. For large triangular arrays, this geometrical frustration is expected to lead, for some parameters (Ω, Δ), to an exponential scaling of the degeneracy with the number of spins. The nature of the ground state for frustrated geometries such as triangle, honeycomb, etc., is a topic of current study. In some cases, the ground state is expected to be a spin liquid featuring strong correlations but no long-range order. In condensed matter systems, some materials exhibit frustration, but no direct evidence of spin liquids has been obtained, mainly due to the difficulty to perform local measurements [38]. Quantum simulators as the one based on arrays of tweezers with tunable geometries could be a new way to investigate frustration and study spin liquids.

Quantum phase transition and adiabatic preparation. There is an obvious drawback to the procedure of adiabatic ground state preparation when a quantum phase transition is crossed. Quantum phase transitions (QPT) are zero temperature transitions where the symmetry of the system suddenly changes when varying a parameter in the Hamiltonian. The superfluid-to-insulator and para- to ferromagnetic transitions described above are important examples of quantum phase transitions. The main issue for an adiabatic preparation comes from the fact that the energy gap between the ground and first excited state closes at the transition, making the adiabaticity criteria (30) impossible to fulfill. For finite-size system, the gaps remain finite, but decrease with the number of atoms N . For example, for the Ising model described above, the gap at the transition scales like $\Delta E \sim 1/N$ in 1D, and $\Delta E \sim 1/\sqrt{N}$ in a 2D square array. This means that the time to reach adiabatically the ground state increases polynomially with the atom numbers. Conversely, for a fixed duration of the parameters sweep, an increasing number of excitations appears in the system as the number of atoms grows. This apparent drawback can actually be turned into a way to study the properties of the quantum phase transition in a quantum simulation approach, as we now illustrate by a recent experiment using a Rydberg quantum simulator [36].

In this experiment, the group of M. Lukin in Harvard implemented the quantum Ising model of Eq. (16) in a one dimensional chain with 51 atoms. As explained above the system exhibits a quantum phase transition between a para- and anti-ferromagnetic phase (Fig. 10a). To study the transition, the group swept the detuning of the laser Δ emulating the longitudinal magnetic field across the transition, for a given “transverse” magnetic field (the laser Rabi frequency). At the end of the ramp characterized by a sweep rate v , they measured the connected correlation function $G(r) = [\langle n_i n_{i+r} \rangle - \langle n_i \rangle \langle n_{i+r} \rangle]$ and fitted the envelope by an exponential decay $\exp[-r/\xi_f]$ to extract the correlation length ξ_f . The final correlation length is found to scale

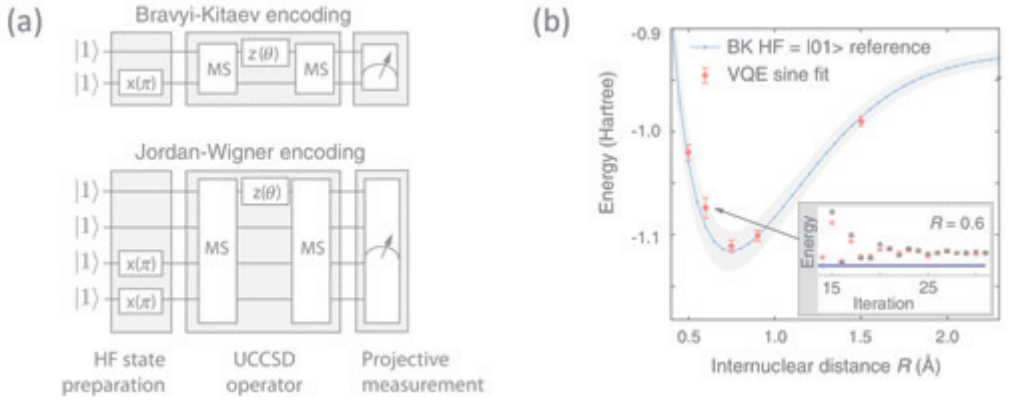


Fig. 11: Quantum chemistry calculation using a quantum-classical hybrid approach. (a) Circuit implementation of the unitary coupled-cluster operator generating the trial state $|\phi(\theta)\rangle$, for the Bravyi-Kitaev or the Jordan-Wigner mapping. (b) Results of the experiment performed on a two-ion quantum computer (Bravyi-Kitaev encoding): reconstructed energy of the ground state of the H_2 molecule as a function of the internuclear distance R . Figures from [41].

with the sweeping rate v as $\xi_f \propto v^{-\mu}$, with $\mu = 0.5$. To understand this scaling, the group used a model proposed by Kibble and Zurek in the 1980's [39, 40]. Initially introduced in the context of cosmology, this model relates the number of defects introduced in a system around a QPT to the sweep rate. In the context of the Ising model, these defects are excitation of two nearby Rydberg atoms, while the ground state in, e.g., a Z_2 phase should consist of Rydberg atoms every other sites. The characteristic distance between the defects are directly the correlation length and results from a compromise between two effects occurring near the QPT. The first one is the divergence of the correlation length ξ at the transition, characterized here by a critical detuning Δ_c and a critical exponent ν characteristic of the universality class of the model: $\xi \sim |\Delta - \Delta_c|^{-\nu}$. However building the correlations occurs on a typical response time τ that also diverges at the transition: $\tau \sim |\Delta - \Delta_c|^{-z\nu}$, with $z\nu$ the dynamical critical exponent. Hence when sweeping the detuning across the transition as $\Delta - \Delta_c = vt$, the correlation length is frozen at a value ξ_f when $\Delta^* - \Delta_c = v\tau$. This corresponds to $\Delta^* - \Delta_c \sim v^{\frac{1}{1+z\nu}}$ and $\xi_f \sim v^{-\mu}$ with $\mu = \nu/(1 + z\nu)$. Measuring μ thus amounts to measuring the critical exponents of the QPT. For the quantum Ising model one expects $\nu = z = 1$, leading to $\mu = 1/2$, in very good agreement with the data. In particular, a mean field treatment would yield $\mu = 1/3$, and is inconsistent with the experiment. This demonstration, which investigated the Kibble-Zurek mechanism in the quantum case, was made possible thanks to the very good decoupling of the platform from the environment. It was not the case in previous attempts. This nicely illustrates that synthetic quantum systems can be used to measure the properties of quantum phase transitions with high precision.

4.3 Quantum chemistry and the hybrid approach

One of the central goals of quantum chemistry is to calculate the energy levels of molecules, in particular their ground state. This task is one more example of a difficult many-body problem as,

apart for small molecules, it involve large number of interacting electrons and nuclei. Chemists have however developed an arsenal of sophisticated approximate methods, but they can become hard to implement when the size of the molecule grows. Quantum simulation may be able to address the question and proof-of-principle experiments have already investigated the potential of this approach (see [41] for a review). Here we will introduce the basic ideas of one method, called Variational Quantum Eigensolver (VQE) algorithm. It combines two concepts: (i) the mapping of the quantum chemistry problem onto a spin model naturally implemented by many quantum simulation platforms; (ii) an hybrid approach where the hardest task is performed by a quantum simulator (see below), while the others are done by a classical hardware.

Mapping of the chemistry problem onto a spin model. We want to find the ground state of the Coulomb Hamiltonian describing N_e electrons (positions \mathbf{r}_i , mass m) and N_n nuclei (position \mathbf{R}_j , charge $Z_j e$, mass M_j):

$$H = - \sum_{j=1}^{N_n} \frac{\hbar^2}{2M_j} \nabla_{\mathbf{R}_j}^2 - \sum_{i=1}^{N_e} \frac{\hbar^2}{2m} \nabla_{\mathbf{r}_i}^2 - \sum_{i,j} \frac{Z_j e^2}{|\mathbf{R}_j - \mathbf{r}_i|} + \sum_{i,j>i} \frac{Z_j Z_i e^2}{|\mathbf{R}_j - \mathbf{R}_i|} + \sum_{i,j>i} \frac{e^2}{|\mathbf{r}_j - \mathbf{r}_i|}. \quad (32)$$

The energy E_0 of the ground state is usually found by a variational method where one uses the fact that for any trial many-body normalized wave function ϕ , $\langle H \rangle = \langle \phi | H | \phi \rangle \geq E_0$. To be able to perform the minimization of the Hamiltonian on a quantum simulator, one writes first the electronic part of the Hamiltonian (32) in the second quantization form, after having fixed the positions of the nuclei (Born-Oppenheimer approximation):

$$H = \sum_{p,q} h_{pq} \hat{c}_p^\dagger \hat{c}_q + \frac{1}{2} \sum_{p,q,r,s} h_{pqrs} \hat{c}_p^\dagger \hat{c}_q^\dagger \hat{c}_r \hat{c}_s, \quad (33)$$

with \hat{c}_p^\dagger (\hat{c}_q) the fermionic creation and annihilation operators, and the coefficients:

$$\begin{aligned} h_{pq} &= \int d\sigma \phi_p^*(\sigma) \left[-\frac{\hbar^2}{2m} \nabla_{\mathbf{r}}^2 - \sum_i \frac{Z_i e^2}{|\mathbf{R}_i - \mathbf{r}|} \right] \phi_q(\sigma); \\ h_{pqrs} &= \int d\sigma_1 d\sigma_2 \phi_p^*(\sigma_1) \phi_q^*(\sigma_2) \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \phi_r(\sigma_1) \phi_s(\sigma_2). \end{aligned} \quad (34)$$

Here $\sigma_i = (\mathbf{r}_i, s_i)$ describes the spatial and spin variables of the electrons. Importantly, the single-electrons orbital $\phi(\sigma)_i$ are supposed to be known and calculable by classical methods. Often, they are obtained by a Linear Combination of Atomic Orbitals calculated by a mean-field, Hartree Fock model. They can also be ansatz functions. To write the Hamiltonian (33) as a spin-model, one uses the mapping of the creation (annihilation) operators onto the spin operators $\hat{\sigma}^\pm$. This can be done by the Jordan-Wigner or the Barvnyi-Kitaev transformations [42]. As a result of these pretty involved formal mappings, the Hamiltonian now has the form of a sum of tensor products of spin operators:

$$H = \sum_{\ell} c_{\ell} H_{\ell} \quad \text{with} \quad H_{\ell} = \hat{\sigma}_1^{\alpha} \hat{\sigma}_2^{\beta} \hat{\sigma}_3^{\gamma} \dots \quad (35)$$

The c_{ℓ} -coefficients depend on the h_{pq} and h_{pqrs} integrals and of the positions of the nuclei (see example below). One now has to minimize $\langle H \rangle$. However, the Hamiltonian (35) corresponds to a spin model involving unphysical spin-spin interactions that can not be directly implemented

on a quantum simulation platform. One option consists in synthesizing H by a digital approach as mentioned in Sec. 2 and illustrated in Fig. 1. At the present stage of the technology, this is challenging. Another approach relies on a hybrid Variational Quantum Eigensolver, where the calculation of the expectation value for a trial function is performed by a quantum simulator, which does not need to implement directly H .

The Variational Quantum Eigensolver algorithm consists of four steps, which combine quantum and classical resources:

1. Initialization of the simulator in the many-electron state $|\phi(0)\rangle$ encoded on spin degrees of freedom. This step is usually relatively easy, as the initial state is often a product state.
2. Construction of a trial state $|\phi(\theta)\rangle$ depending on a parameter θ . To do so, one evolves $|\phi(0)\rangle$ with a unitary operator $U(\theta)$ (called coupled-cluster): $|\phi(\theta)\rangle = U(\theta) |\phi(0)\rangle$. The operator $U(\theta)$ is chosen so that the trial state explores efficiently the parameter space when varying θ . This is where the quantum simulator operates, as calculating the evolution of $|\phi(0)\rangle$ under $U(\theta)$ is very demanding using classical methods. The operator $U(\theta)$ can either be directly implemented in the lab (analog approach) or is digitally synthesized by a set of qubit gates (case of the experiments done so far). Importantly, the resources required in the lab to realize $U(\theta)$ do not need to realize the full Hamiltonian H .
3. Readout of the simulator and repetition of the two first steps until obtaining enough statistics to estimate the correlation functions: $\langle \phi(\theta) | \hat{\sigma}_1^\alpha \hat{\sigma}_2^\beta \hat{\sigma}_3^\gamma \dots | \phi(\theta) \rangle = \langle \phi(\theta) | H_\ell | \phi(\theta) \rangle$. This allows estimating $\langle \phi(\theta) | H | \phi(\theta) \rangle$.
4. Finally, variation of θ and use of a *classical* algorithm (e.g. gradient-descent) to minimize the expectation value $\langle \phi(\theta) | H | \phi(\theta) \rangle$.

The third step shows that the non-physical many-spin interactions of Eq. (35) do not have to be implemented experimentally: one only needs as quantum resources as many particles as spin operators in the tensor products, and the ability to synthesize $U(\theta)$, which can be considerably simpler than H . Let us illustrate briefly the method on H_2 , the simplest molecule. In this case the Bravyi-Kitaev transformation leads to an Hamiltonian involving up to 4-body terms of the form $\hat{\sigma}_i^z, \hat{\sigma}_i^z \hat{\sigma}_j^z, \hat{\sigma}_i^x \hat{\sigma}_j^z \hat{\sigma}_k^x$ and $\hat{\sigma}_i^z \hat{\sigma}_j^y \hat{\sigma}_k^z \hat{\sigma}_l^z$ [41]. However, it turns out that the initial state $|\phi(0)\rangle$ resulting from a Hartree-Fock calculation is $|0001\rangle$, making the three- and four-body terms irrelevant. This fact allows restricting the Hamiltonian to:

$$H_{\text{BK}} = c_0 \mathcal{I} + c_1 \hat{\sigma}_0^z + c_2 \hat{\sigma}_1^z + c_3 \hat{\sigma}_0^z \hat{\sigma}_1^z + c_4 \hat{\sigma}_0^x \hat{\sigma}_1^x + c_5 \hat{\sigma}_0^y \hat{\sigma}_1^y. \quad (36)$$

As the new Hamiltonian involves only up to two-body terms, the experimental implementation requires a simulator with two quantum bits (e.g. two ions as in [41]). Using the Jordan-Wigner transformation does not lead to this simplification and the final Hamiltonian involves up to 4-body terms, thus requiring four qubits. The coupled-cluster operators are $U_{\text{BK}}(\theta) = \exp(-i\theta \hat{\sigma}_1^x \hat{\sigma}_0^y)$ and $U_{\text{JW}}(\theta) = \exp(-i\theta \hat{\sigma}_3^x \hat{\sigma}_2^x \hat{\sigma}_1^x \hat{\sigma}_0^y)$ for the Bravyi-Kitaev and Jordan-Wigner mappings respectively. Figure 11(a) shows the circuit implementations of these operators for the two mappings, as done on the experiment of Innsbruck using an ion quantum simulator in digital mode [41]. Measuring the correlation functions necessary to reconstruct $\langle \phi(\theta) | H | \phi(\theta) \rangle$ is relatively straightforward. The final result of the experiment is presented in Fig. 11(b), which

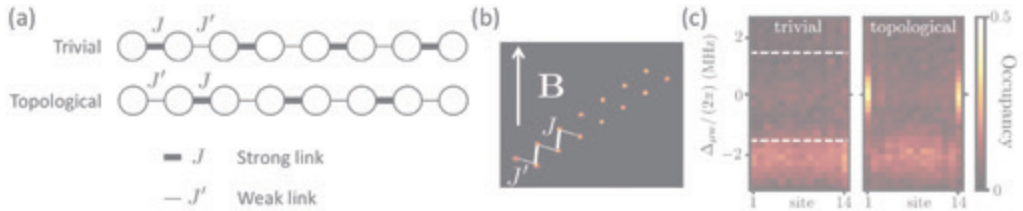


Fig. 12: Implementation of the Su-Schrieffer-Heeger model. (a) The two configurations of the model, in the tight-binding limit. (b) Fluorescence image of atoms in tweezers array implementing the topological configuration. (c) Microwave spectroscopy of the Su-Schrieffer-Heeger model implemented on a chain of Rydberg atoms: probability of exciting the atoms from Rydberg state nS to nP as a function of the frequency of the microwave, for different positions in the chain. The topological configuration features two states at zero energy, corresponding to atoms located at the two edges of the chain. Figure from [47].

shows the minimal energy obtained as a function of the internuclear distance R . The accuracy can be obtained by comparing to a quantum chemistry calculation, very accurate for H_2 . It is around 4×10^{-2} .

The method has been also tested on LiH , BeH_2 , and H_2O using circuit-based or ion-based quantum simulators. The precision achieved are not yet at the chemical accuracy ($\sim 10^{-5}$). We are therefore quite far from the simulation of molecules that challenges chemists. However these experiments validate the method and is a nice illustration of quantum simulation methods.

4.4 Other examples

There are many other examples of application of the quantum simulation approach that we will not cover in these notes (see [1] and the reviews cited below).

Among them, a very important line of research is the quantum simulation of topological matter. Many experimental platforms have been used, such as ultracold atoms (quantum degenerate gases or placed in optical lattices), superconducting circuits or photonic platforms (silicon waveguides, polaritons in semi-conductor cavities described in Sec. 3.5, etc.). For recent reviews on this subject, see [43, 44]. So far, the implementations mainly consisted in engineering topological situations, in particular by tailoring band-structures with non-trivial topological properties. However, for most of the experiments, the interactions between the particles were negligible and the observations could be explained at the single-particle level.

Interacting topological system. The first realization of a topological situation with strong interactions between the particles was performed in our group at the Institut d’Optique using a Rydberg quantum simulator [47]. We implemented the Su, Schrieffer and Heeger (SSH) model, initially developed in the late 1970’s to explain the conductivity of some organic polymers [45, 46]. In its simplest setting, it consists of a one-dimensional chain of sites that are coupled by an alternation of strong and weak links and where an excitation can hop in the chain. Since then, the SSH model has been recognized as one of the simplest examples of a system exhibiting topological properties. Consider the two configurations of a finite chain represented in Fig. 12(a): either the chain ends up with the strongest link J , or with the weakest

one J' . It turns out that in the first configuration, the single-particle spectrum consists of two bands with width J' , separated by an energy gap $|J - J'|$. On the contrary, in the second configuration, two states at zero energy appear in the middle of the gap, and correspond to states localized on the edges of the chain. The fact that they have zero energy is rather intuitive in the extreme case where $J' = 0$, as adding a particle on each edge does not cost energy. It turns out that this remains true even when $J' \neq 0$. The two configurations correspond to two different topological classes of the system: it is impossible to vary the ratio J/J' and continuously transform one configuration into the other without closing the gap. To implement this tight-binding model, we used the XY model of Eq. (17) and its mapping onto the propagation of spin excitations introduced at the end of Sec. 4.1: this situation is equivalent to a single particle (here the spin excitation) tunnelling between neighboring sites in a lattice. Also, the excitations behave as artificial particles with infinite on-site interactions, i.e. with a hard-core constraint. It turns out that the spin excitations have the same commutation relations as those of bosons. Therefore the problem of an ensemble of two-level Rydberg atoms interacting by the resonant dipole-dipole interaction can be equivalently mapped onto a spin-1/2 XY model, or onto a system of so called hard-core bosons [48]. But one should keep in mind that the resonant dipole interaction that drives the transport of an excitation leads to a *single-particle* problem when considering a single excitation. What makes the excitations interact is the fact that the atoms only have two levels, and not the fact that the atoms carrying the excitations interact. This mapping allowed us to study the genuinely many-body properties arising from a hard-core constraint [47]. We used two Rydberg states $|nS\rangle = |\downarrow\rangle$ and $|nP\rangle = |\uparrow\rangle$ separated by a transition with a frequency around 17 GHz. The anisotropy of the dipole interaction makes it possible to realize a situation where next nearest-neighbor interactions are negligible, hence implementing a 1D situation using a 2D arrangement of the atoms (see Fig. 12b). After preparing all the atoms in the state $|nS\rangle$, we used microwave spectroscopy to measure the single-particle energy spectrum (see Fig. 12c). We entered the many-body regime by preparing the ground state of the chain comprising $N/2$ excitations (where N is the number of sites), using an adiabatic preparation in the spirit of the one we used to prepare the antiferromagnetic correlations in the Ising model described in Sec. 4.2. We have characterized this many-body ground state in the topological configuration, and found that it displayed a characteristic robustness with respect to the breaking of certain symmetries of the Hamiltonian. The prepared state is the first experimental realization of a type of topological order for bosons introduced in 2012 based on formal considerations [49], called symmetry-protected topological phases. These phases are the only topological orders that can exist in one dimension. This experiment shows once again that synthetic quantum systems allow implementing situations originally devised in a mathematical physics context.

Simulation of open systems. Finally, it is important to mention that all quantum simulation platforms are coupled to an environment. For all the examples mentioned above, the goal was to keep this coupling to a minimum value. In some situations however, it is possible to engineer the environment and therefore quantum simulate driven, open many-body systems. As described in Sec. 3.5, the polariton platform described in 3.5 naturally operates in this regime. Quantum circuits are also an interesting platform to study open systems [17].

5 Challenges and outlook

In this last section, we first briefly discuss some of the challenges ahead for quantum simulators.

As seen in the examples discussed in the previous section, quantum simulation platforms are already at the “useful” stage: they allowed answering open questions (e.g. MBL in 2D) and they operate for some of them in a regime inaccessible to classical calculations. Also they provide new tools to address open questions, forcing one to think differently about the problems. This triggering aspect provided by the platforms is, in my opinion, one important aspect of the quantum simulation approach. This shows, that even plagued with imperfections and noise, we learn from quantum simulators. As mentioned in Sec. 2, this fact has been conceptualized: programmable quantum simulators can be seen as noisy, intermediate scale quantum computers [4] able to perform tasks too hard for classical hardware, despite the fact that they are not universal. There is also no doubt that as the quality of the simulators improve, thanks to technological developments, the range of physical effects accessible to them will enlarge, and the obvious tasks for the experimentalists is to work hard on this improvements.

It is also fair to say that for many problems, quantum simulators are still in the benchmarking phase, where the outcomes of the simulators can be compared to a classical simulation as the number of particles is small enough. This phase is necessary to build trust in the experimental implementation, and it is crucial for any new platform that is being developed. However this raises the question of the validation of the quantum simulation in a regime where no benchmarking against a classical solution is possible. The problem is quite different for quantum simulators and quantum computers. For computers, performing the calculation is hard but the solution is often easy to check (this is actually an important point as some quantum computing algorithms only give a probabilistic answer). Most of the times, quantum simulation does not allow for this possibility of directly confirming the solution. A current line of research is thus to investigate benchmarking and verification protocols. Among the identified directions:

- as already mentioned, benchmarking in regimes where we can compare the solution directly against (classically) solvable problems, or with formal generalisations of randomized benchmarking from digital quantum computers.
- benchmarking with problems for which we can efficiently check the final outcome, including e.g., classical optimization problems.
- comparing different platforms by performing the quantum simulation of a given model Hamiltonian (e.g. Heisenberg, Quantum Ising, Fermi Hubbard etc) on different implementations. One can probe the many-body state fidelity of the model Hamiltonians, as well as dynamical state evolution. This approach is similar to the one used for the most precise atomic clocks: to check that the time given by the clock is accurate, one builds a second clock and beat it with the first one to investigate systematic effects.
- when verifying the solution is computationally hard, we may be able to identify some observables or correlation function witnesses to check the solution.

To conclude, and as an outlook, one should not think about quantum simulators only as being able to address academic scientific questions. In fact these synthetic systems can also be viewed

as machines able to prepare quantum states useful for many applications. For example, they can generate large entangled states, whose correlations are useful to beat the standard quantum limit, hence leading to clocks or sensors with enhanced precision [50]. In the long-term, they could of course lead to quantum computers. Interestingly, machines able to implement spin models could be useful to answer computationally hard problems well beyond physics, such as combinatorial optimization problems (one prominent example being the traveling salesman’s problem). Many of these optimization problems can be cast as Ising models [51] that most quantum simulators implement naturally, as we have seen in these notes. By varying the parameters on the experiment, one could drive the system into a state encoding the solution of the problem. Therefore, quantum simulation using synthetic many-body systems appear as a wonderful playground for physicists, but also with many potential of applications in our everyday life. We are just at the beginning of exploring them.

References

- [1] I.M. Georgescu, S. Ashhab, F. Nori, “Quantum Simulation”, *Rev. Mod. Phys.* **86**, 153 (2014).
- [2] R. P. Feynman, “Simulating Physics with Computers”, *Int. J. Theor. Phys.*, **21**, 467 (1982).
- [3] J. Hubbard, “Electron Correlations in Narrow Energy Bands”, *Proceedings of the Royal Society of London*, **276** (1365), 238 (1963).
- [4] J. Preskill, “Quantum Computing in the NISQ era and beyond” *Quantum* **2**, 79 (2018).
- [5] S. Lloyd, “Universal Quantum Simulators”, *Science* **273**, 1073 (1996).
- [6] R. Blatt & C.F. Roos, “Quantum simulations with trapped ions”, *Nat. Phys.* **8**, 277 (2012).
- [7] C. Monroe *et al.*, “Programmable Quantum Simulations of Spin Systems with Trapped Ions”, arXiv:1912.07845
- [8] Petar Jurcevic, “Quantum Computation and Many-Body Physics with Trapped Ions”, PhD Thesis, University of Innsbruck (2017).
- [9] J. W. Britton *et al.*, “Engineered two-dimensional Ising interactions in a trapped-ion quantum simulator with hundreds of spins”, *Nature* **484**, 489 (2012).
- [10] M. Gärtner *et al.*, “Measuring out-of-time-order correlations and multiple quantum spectra in a trapped-ion quantum magnet” *Nat. Phys.* **13** 781 (2017).
- [11] A. Safavi-Naini *et al.*, “Verification of a Many-Ion Simulator of the Dicke Model Through Slow Quenches across a Phase Transition” *Phys. Rev. Lett.* **121**, 040503 (2018).
- [12] E.A. Martinez *et al.*, “Real-time dynamics of lattice gauge theories with a few-qubit quantum computer”, *Nature* **534**, 516 (2016).
- [13] I. Bloch, J. Dalibard, S. Nascimbène, “Quantum simulations with ultracold quantum gases”, *Nat. Phys.* **8**, 267 (2012).

- [14] C. Gross & I. Bloch, “Quantum simulation with ultra-cold atoms in optical lattices”, *Science* **357**, 995 (2017).
- [15] A. Browaeys and T. Lahaye, “Many-Body Physics with Individually-Controlled Rydberg Atoms”, *Nat. Phys.* in press (2020).
- [16] L.-M. Duan, E. Demler, and M. D. Lukin, “Controlling Spin Exchange Interactions of Ultracold Atoms in Optical Lattices”, *Phys. Rev. Lett.* **91**, 090402 (2003).
- [17] A.A. Houck, H.E. Türeci, J. Koch, “On-chip quantum simulation with superconducting circuits”, *Nat. Phys.* **8**, 292 (2012).
- [18] F. Arute, *et al.*, “Quantum supremacy using a programmable superconducting processor” *Nature* **574**, 505 (2019).
- [19] David Isaac Schuster, “Circuit Quantum Electrodynamics”, PhD Thesis, Yale University (2007).
- [20] I. Carusotto and C. Ciuti, “Quantum fluids of light” *Rev. Mod. Phys.* **85**, 299 (2013).
- [21] V. Goblot, “Polariton quantum fluids in one-dimensional synthetic lattices: localization, propagation and interactions”, PhD thesis, Université Paris-Saclay (2019).
- [22] A. Aspuru-Guzik & P. Walther, “Photonic quantum simulators”. *Nat. Phys.* **8**, 285 (2012).
- [23] H. Labuhn, D. Barredo, S. Ravets, S. de Léséleuc, T. Macrì, T. Lahaye, & A. Browaeys, “Tunable two-dimensional arrays of single Rydberg atoms for realizing quantum Ising models”, *Nature* **534**, 667 (2016).
- [24] S. de Léséleuc, S. Weber, V. Lienhard, D. Barredo, H. P. Büchler, T. Lahaye, & A. Browaeys, “Accurate mapping of multilevel Rydberg atoms on interacting spin-1/2 particles for the quantum simulation of Ising models”, *Phys. Rev. Lett.* **120**, 113602 (2018).
- [25] S. Trotzky *et al.*, Probing the relaxation towards equilibrium in an isolated strongly correlated one-dimensional Bose gas” *Nat. Phys.* **8**, 325 (2012).
- [26] J. Smith *et al.*, “Many-body localization in a quantum simulator with programmable random disorder”, *Nat. Phys.* **12**, 907 (2016).
- [27] J.-Y. Choi *et al.*, “Exploring the many-body localization transition in two dimensions”, *Science* **352**, 1547 (2016).
- [28] P. Roushan *et al.*, “Spectroscopic signatures of localization with interacting photons in superconducting qubits”, *Science* **358**, 1175 (2017).
- [29] E. H. Lieb and D. W. Robinson, “The Finite Group Velocity of Quantum Spin Systems”, *Commun. Math. Phys.* **28**, 251 (1972).
- [30] P. Richerme *et al.*, “Non-local propagation of correlations in quantum systems with long-range interactions” *Nature* **511**, 198 (2014).
- [31] P. Jurcevic *et al.*, “Quasiparticle engineering and entanglement propagation in a quantum many-body system” *Nature* **511**, 211 (2014).

- [32] L.I. Schiff, *Quantum Mechanics*, 3rd edition, McGraw-Hill (1968).
- [33] H. Bernien *et al.*, “Probing many-body dynamics on a 51-atom quantum simulator”, *Nature* **551**, 579 (2017).
- [34] V. Lienhard *et al.*, “Observing the space- and time-dependent growth of correlations in dynamically tuned synthetic Ising antiferromagnets”, *Phys. Rev. X* **8**, 021070 (2018).
- [35] E. Guardado-Sanchez *et al.*, “Probing the Quench Dynamics of Antiferromagnetic Correlations in a 2D Quantum Ising Spin System”, *Phys. Rev. X* **8**, 021069 (2018).
- [36] A. Keesling *et al.* “Probing quantum critical dynamics on a programmable Rydberg simulator”, *Nature* **568**, 207 (2019).
- [37] R. Moessner & A.P. Ramirez, “Geometrical Frustration”, *Physics Today*, Feb. 2006, p. 24.
- [38] J. Knolle, R. Moessner, “A field guide to spin liquids”, *Annu. Rev. Cond. Mat. Phys.* **10**, 451 (2019).
- [39] T.W.B. Kibble, “Topology of cosmic domains and strings”, *J. Phys. A* **9**, 1387 (1976).
- [40] W.H. Zurek, “Cosmological experiments in superfluid helium”, *Nature* **317**, 505 (1985).
- [41] C. Hempel *et al.*, “Quantum Chemistry Calculations on a Trapped-Ion Quantum Simulator”, *Phys. Rev. X* **8**, 031022 (2018).
- [42] J.T. Seeley, “The bravyi-kitaev transformation for quantum computation of electronic structure”, *The Journal of chemical physics* **137**, 224109 (2012).
- [43] T. Ozawa, H.M. Price, A. Amo, N. Goldman, M. Hafezi, L. Lu, M.C. Rechtsman, D. Schuster, “Topological photonics”, *Rev. Mod. Phys.* **91**, 015006 (2019).
- [44] N.R. Cooper, J. Dalibard, I.B. Spielman, “Topological bands for ultracold atoms”, *Rev. Mod. Phys.* **91**, 015005 (2019).
- [45] W.P. Su, J.R. Schrieffer, A.J. Heeger, “Solitons in Polyacetylene” *Phys. Rev. Lett.* **42**, 1698-1701 (1979).
- [46] A.J. Heeger, S. Kivelson, J.R. Schrieffer, W.P. Su, “Solitons in conducting polymers” *Rev. Mod. Phys.* **60**, 781 (1988).
- [47] S. de Léséleuc *et al.*, “Observation of a symmetry-protected topological phase of interacting bosons with Rydberg atoms”, *Science* **365**, 775 (2019).
- [48] T. Giamarchi, “Quantum Physics in one dimension”, Oxford University Press, Oxford (2004).
- [49] X. Chen, Z.-C. Gu, Z.-X. Liu, X.-G. Wen, “Symmetry-protected topological orders in interacting bosonic systems”, *Science* **338**, 1604 (2012).
- [50] V. Giovannetti, S. Lloyd, & L. Maccone, “Advances in quantum metrology”, *Nat. Phot.* **5**, 222 (2011).
- [51] A. Lucas, “Ising formulations of many NP problems”, *Front. Phys.* **2**, **5**, 1 (2014).

A 3 Quantum supremacy using a programmable superconducting processor

H. Neven and collaborators

Google Inc., Venice, California 90291, USA

nature<https://doi.org/10.1038/s41586-019-1666-5>**Supplementary information****Quantum supremacy using a programmable superconducting processor**

In the format provided by the
authors and unedited

Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C. Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando G. S. L. Brandao, David A. Buell, Brian Burkett, Yu Chen, Zijun Chen, Ben Chiaro, Roberto Collins, William Courtney, Andrew Dunsworth, Edward Farhi, Brooks Foxen, Austin Fowler, Craig Gidney, Marissa Giustina, Rob Graff, Keith Guerín, Steve Habegger, Matthew P. Harrigan, Michael J. Hartmann, Alan Ho, Markus Hoffmann, Trent Huang, Travis S. Humble, Sergei V. Isakov, Evan Jeffrey, Zhang Jiang, Dvir Kafri, Kostyantyn Kechedzhi, Julian Kelly, Paul V. Klimov, Sergey Knysh, Alexander Korotkov, Fedor Kostritsa, David Landhuis, Mike Lindmark, Erik Lucero, Dmitry Lyakh, Salvatore Mandrà, Jarrod R. McClean, Matthew McEwen, Anthony Megrant, Xiao Mi, Kristel Michielsen, Masoud Mohseni, Josh Mutus, Ofer Naaman, Matthew Neeley, Charles Neill, Murphy Yuezhen Niu, Eric Ostby, Andre Petukhov, John C. Platt, Chris Quintana, Eleanor G. Rieffel, Pedram Roushan, Nicholas C. Rubin, Daniel Sank, Kevin J. Satzinger, Vadim Smelyanskiy, Kevin J. Sung, Matthew D. Trevithick, Amit Vainsencher, Benjamin Villalonga, Theodore White, Z. Jamie Yao, Ping Yeh, Adam Zalcman, Hartmut Neven & John M. Martinis

Supplementary information for “Quantum supremacy using a programmable superconducting processor”

Google AI Quantum and collaborators[†]
(Dated: October 8, 2019)

CONTENTS

I. Device design and architecture	2	2. Universality for SU(2)	30
II. Fabrication and layout	2	G. Circuit variants	30
III. Qubit control and readout	3	1. Gate elision	31
A. Control	3	2. Wedge formation	31
B. Readout	3	VIII. Large scale XEB results	31
IV. XEB theory	5	A. Limitations of full circuits	32
A. XEB of a small number of qubits	5	B. Patch circuits: a quick performance indicator for large systems	33
B. XEB of a large number of qubits	7	C. Elided circuits: a more rigorous performance estimator for large systems	33
C. Two limiting cases	8	D. Choice of unitary model for two-qubit entangling gates	34
D. Measurement errors	9	E. Understanding system performance: error model prediction	35
V. Quantifying errors	9	F. Distribution of bitstring probabilities	36
VI. Metrology and calibration	11	G. Statistical uncertainties of XEB measurements	39
A. Calibration overview	11	H. System stability and systematic uncertainties	40
1. Device registry	11	I. The fidelity result and the null hypothesis on quantum supremacy	41
2. Scheduling calibrations: “Optimus”	12	IX. Sensitivity of XEB to errors	42
B. Calibration procedure	12	X. Classical simulations	44
1. Device configuration	12	A. Local Schrödinger and Schrödinger-Feynman simulators	44
2. Root config: procedure	12	B. Feynman simulator	45
3. Single-qubit config: procedure	13	C. Supercomputer Schrödinger simulator	49
4. Optimizing qubit operating frequencies	13	D. Simulation of random circuit sampling with a target fidelity	50
5. Grid config: procedure	14	1. Optimality of the Schmidt decomposition for gates embedded in a random circuit	50
C. Two-qubit gate metrology	15	2. Classical speedup for imbalanced gates	52
1. The natural two-qubit gate for transmon qubits	15	3. Verifiable and supremacy circuits	53
2. Using cross entropy to learn a unitary model	16	E. Treewidth upper bounds and variable elimination algorithms	54
3. Comparison with randomized benchmarking	16	F. Computational cost estimation for the sampling task	55
4. Speckle purity benchmarking (SPB)	18	G. Understanding the scaling with width and depth of the computational cost of verification	56
5. “Per-layer” parallel XEB	19	1. Runtime scaling formulas	56
D. Grid readout calibration	20	2. Assumptions and corrections	57
1. Choosing qubit frequencies for readout	20	3. Fitting constants	58
2. Single qubit calibration	20	4. Memory usage scaling	58
3. Characterizing multi-qubit readout	21	H. Energy advantage for quantum computing	59
E. Summary of system parameters	21	XI. Complexity-theoretic foundation of the experiment	59
VII. Quantum circuits	27		
A. Background	27		
B. Overview and technical requirements	27		
C. Circuit structure	27		
D. Randomness	27		
E. Quantum gates	28		
F. Programmability and universality	29		
1. Decomposition of CZ into fSim gates	29		

A. Error model	60
B. Definition of computational problem	60
C. Computational hardness of unbiased-noise sampling	61
D. Proof of Theorem 1	62
Acknowledgments	63
References	64

I. DEVICE DESIGN AND ARCHITECTURE

The Sycamore device was designed with both the quantum supremacy experiment and small noisy intermediate scale quantum (NISQ) applications in mind. The architecture is also suitable for initial experiments with quantum error correction based on the surface code. While we are targeting 0.1% error two-qubit gates for error correction, a quantum supremacy demonstration can be achieved with 0.3-0.6% error rates.

For decoherence-dominated errors, a 0.1% error means a factor of about 1000 between coherence and gate times. For example, a 25 μ s coherence time implies a 25 ns gate. A key design objective in our architecture is achieving short two-qubit gate time, leading to the choice of tunable transmon qubits with direct, tunable coupling.

A difficult challenge for achieving a high-performance two-qubit gate is designing a sufficiently strong coupling when the gate is active, which is needed for fast gates, while minimizing the coupling otherwise for low residual control errors. These two competing requirements are difficult to satisfy with a fixed-coupling architecture: our prior processors [1] used large qubit-qubit detuning (~ 1 GHz) to turn off the effective interaction, requiring relatively high-amplitude precise flux pulses to tune the qubit frequencies to implement a CZ gate. In the Sycamore device, we use adjustable couplers [2] as a natural solution to this control problem, albeit at the cost of more wiring and control signals. This means that the qubits can idle at much smaller relative detuning. We chose a capacitor-coupled design [2, 3], which is simpler to layout and scale, over the inductor-based coupler of previous gmon devices [4, 5]. In Sycamore, the coupling g is tunable from 5 MHz to -40 MHz. The experiment uses ‘on’ coupling of about -20 MHz.

By needing only small frequency excursions to perform a two-qubit gate, the tunable qubit can be operated much closer to its maximum frequency, thus greatly reducing flux sensitivity and dephasing from $1/f$ flux noise. Additionally, the coupling can be turned off during measurement, reducing the effect of measurement crosstalk, a phenomenon that has shown to be somewhat difficult to understand and minimize [6].

The interaction Hamiltonian of a system of on-resonance transmons with adjustable coupling (truncated

to the qubit levels) has the following approximate form,

$$H_{\text{int}}(t) \approx \sum_{(i,j)} g_{ij}(t) (\sigma_i^+ \sigma_j^- + \sigma_i^- \sigma_j^+) + \frac{g_{ij}^2(t)}{|\eta|} \sigma_i^z \sigma_j^z, \quad (1)$$

where g_{ij} is the nearest neighbor coupling, η is the non-linearity of the qubits (roughly constant), i and j index nearest-neighbor qubit pairs, and $\sigma^\pm = (\sigma^x \pm i\sigma^y)/2$. We pulse the coupling in time to create coupling gates.

Our two-qubit gate can be understood using Cartan decomposition [7], which enables an arbitrary two-qubit gate to be decomposed into four single-qubit gates around a central two-qubit gate that can be described by a unitary matrix describing only XX, YY and ZZ interactions, with 3 parameters indicating their strengths. For the physical interaction describing our hardware, we see a swapping interaction between the $|01\rangle$ and $|10\rangle$ qubits states, corresponding to an XX+YY interaction. Interaction of the qubit $|11\rangle$ state with the $|2\rangle$ states of the data transmons produce a phase shift of that state, corresponding to a ZZ interaction. By changing the qubit frequencies and coupling strength we can vary the magnitude of these interactions, giving net control of 2 out of the 3 possible parameters for an arbitrary gate.

II. FABRICATION AND LAYOUT

Our Sycamore quantum processor is configured as a diagonal array of qubits as seen in the schematic of Fig. 1 in the main text. The processor contains 142 transmon qubits, of which 54 qubits have individual microwave and frequency controls and are individually read out (referred to as qubits). The remaining 88 transmons are operated as adjustable couplers remaining in their ground state during the algorithms (referred to as couplers).

The qubits consist of a DC SQUID sandwiched between two metal islands, operating in the transmon regime. An on-chip bias line is inductively coupled to the DC SQUID, which allows us to tune qubit frequency by applying control fluxes into the SQUID loop. For regular operations, we tune qubits through a small frequency range (< 100 MHz). This corresponds to a relatively small control signal and makes qubit operation less sensitive to flux crosstalk.

Each pair of nearest-neighbor qubits are coupled through two parallel channels: direct capacitive coupling and indirect coupling mediated by coupler [2, 3, 8]. Both channels result in qubit-qubit coupling in the form of $\sigma_i^x \sigma_j^x + \sigma_i^y \sigma_j^y$ in the rotating frame, although with different signs. The indirect coupling is negative, given it is a second-order virtual process. The strength of the indirect coupling is adjusted by changing the coupler frequency with an additional on-chip bias line, giving a net zero qubit-qubit coupling at a specific flux bias.

The Sycamore processor consists of two die that we fabricated on separate high resistivity silicon wafers. The fabrication process, using aluminum on silicon, requires

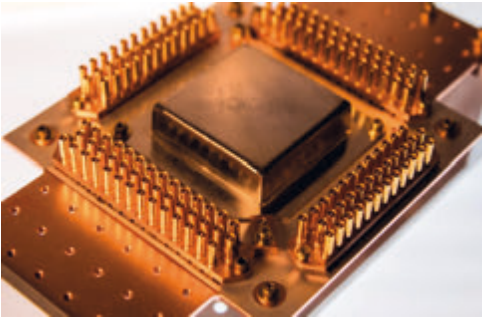


FIG. S1. **A photograph of a packaged Sycamore processor.** The processor is shielded from the electromagnetic environment by a mu-metal shield (middle) and a superconducting Aluminum cap, inside the mu-metal shield. The processor control wires are routed, through PCB circuit board, to coaxial connectors shown around the edge.

a total of 14 lithography layers utilizing both optical and electron beam lithography. Crosstalk and dissipation are mitigated through ground plane shielding [9]. After fabrication and die singulation, we use indium bump bonding [10, 11] of the two separate dies to form the Sycamore processor.

The Sycamore processor is connected to a 3-layer Al-plated circuit board with Al wirebonds [12]. Each line is routed through a microwave connector to an individual coax cable. We shield the processor from stray light using a superconducting Al lid with black coating, and from magnetic fields using a mu-metal shield as shown in Fig. S1.

III. QUBIT CONTROL AND READOUT

A. Control

Operating the device requires simultaneous synchronized control waveforms for each of the qubits and couplers. We use 54 coherent microwave control signals for qubit XY rotations, 54 fast flux bias lines for qubit frequency tuning, and 88 fast flux biases for the adjustable couplers. Dispersive readout requires an additional 9 microwave signals and phase sensitive receivers. A schematic of the room temperature electronics is shown in Fig. S2, and the cryogenic wiring is shown in Fig. S3.

Waveform generation is based on a custom-built multi-channel digital to analog converter (DAC) module. Each DAC module provides 8 DACs with 14-bit resolution and 1 GS/s sample rate. Each DAC sample clock is synchronized to a global 10 MHz reference oscillator, and their trigger is connected by a daisy chain to synchronize all modules used in the experiment. This set of DAC mod-

ules forms a >250-channel, phase-synchronous waveform generator. We have measured 20 ps of jitter between channels. The modules are mounted in 14-slot 6U rack-mount chassis. A single chassis, shown in FIG. S4, can control approximately 15 qubits including their associated couplers and readout signals. A total of 4 chassis are used to control the entire Sycamore chip.

The DAC outputs are used directly for fast flux biasing the qubits and couplers required for two-qubit gates. Microwave control for single-qubit XY rotations and dispersive readout combine two DAC channels and a mixer module to form a microwave arbitrary waveform generator (Microwave AWG) via single-sideband upconversion in an IQ mixer as shown in Figure S2 a. The microwave AWG provides signals with arbitrary spectral content within ± 350 MHz of the local oscillator (LO). A single LO signal is distributed to all IQ mixers so that all qubits' XY controls are phase coherent. The mixer modules are mounted in the same chassis as the DAC modules. Each mixer's I and Q port DC offsets are calibrated for minimum carrier leakage and the I and Q amplitudes and phases are calibrated to maximize image rejection.

Each DAC module contains an FPGA that provides a gigabit ethernet interface, SRAM to store waveform patterns, and sends the waveform data to the DAC module's 8 DACs. To optimize the use of SRAM, the FPGA implements a simple jump table to allow reusing or repeating waveform segments. A computer loads the desired waveforms and jump table onto each FPGA using a UDP-based protocol and then requests the first (master) FPGA to start. The start pulse is passed down the daisy chain causing the remainder (slave) DACs and ADCs to start.

B. Readout

Qubit state measurement and readout (hereafter “readout”) are done via the dispersive interaction between the qubit and a far-detuned harmonic resonator [13–15]. A change in the qubit state from $|0\rangle$ to $|1\rangle$ causes a frequency shift of the resonator from $\omega_{|0\rangle}$ to $\omega_{|1\rangle}$. A readout probe signal applied to the resonator at a frequency in between $\omega_{|0\rangle}$ and $\omega_{|1\rangle}$ reflects with a phase shift $\phi_{|0\rangle}$ or $\phi_{|1\rangle}$ that depends on the resonator frequency and therefore on the qubit state. By detecting the phase of the reflected probe signal we infer the qubit state. The readout probe signal is generated with the same microwave AWG as the XY control signals, but with a separate local oscillator, and is received and demodulated by the circuit shown in Figure S2 b.

The readout probe intensity is typically set to populate the readout resonator with only a few photons to avoid readout-induced transitions in the qubit [16]. Detecting this weak signal at room temperature with conventional electronics requires 100 dB of amplification. To limit the integration time to a small fraction of the qubit coherence time, the amplification chain must operate near the

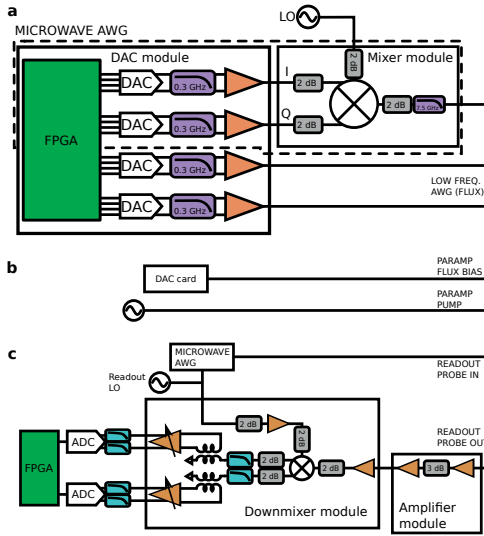


FIG. S2. **Control electronics.** **a**, The custom DAC module provides 8 DAC channels (4 shown). DACs are used individually for flux pulses or in pairs combined with a mixer module to comprise a microwave AWG channel (dashed box). **b**, A single DAC channel and a microwave source are used to bias and pump the parametric amplifier for readout. **c**, Readout pulses are generated by a microwave AWG. The reflected signal is amplified, mixed down to IF, and then digitized in a pair of ADCs. The digital samples are analyzed in the FPGA.

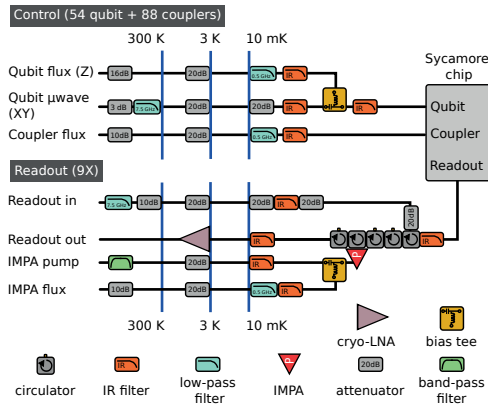


FIG. S3. **Cryogenic wiring.** Control and readout signals are carried to and from the Sycamore chip with a set of cables, filters, attenuators, and amplifiers.

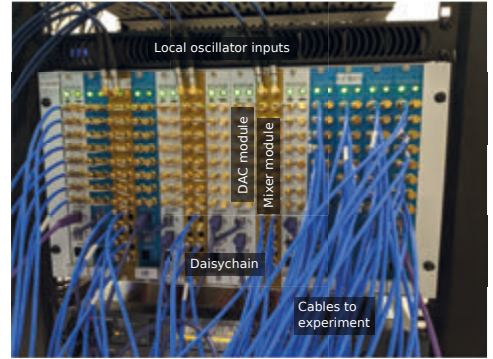


FIG. S4. **Electronics chassis.** Each chassis supports 14 DAC and/or mixer modules. Local oscillators are connected at the top of each mixer module. A set of daisychain cables connects from each ADC module to the next. Control signals exit the chassis through coaxial cables.

quantum noise limit [17, 18].

Inside the cryostat the signal is amplified by an impedance matched lumped element Josephson parametric amplifier (IMPA) [19] on the mixing chamber stage followed by a Low Noise Factory cryogenic HEMT amplifier at 3 K. At room temperature the signal is further amplified before it is mixed down with an IQ mixer producing a pair of intermediate frequency (IF) signals $I(t)$ and $Q(t)$. The IF signals are amplified by a pair of variable gain amplifiers to fine-tune their level, and then digitized by a pair of custom 1 GS/s, 8-bit analog to digital converters (ADC). The digitized samples I_n and Q_n are processed in an FPGA which combines them into a complex phasor

$$z_n = I_n + iQ_n = E_n \exp(i(\omega n dt + \phi))$$

where dt is the sample spacing, ω is the IF frequency, ϕ is the phase that depends on the qubit state, and E_n is the envelope of the reflected readout signal. The envelope is measured experimentally once and then used by the FPGA in subsequent experiments as the optimal demodulation window w_n to extract the phase of the reflected readout signal [20, 21]. The FPGA multiplies z_n by $w_n \exp(-i\omega n dt)$, and then sums over time to produce a final complex value $\exp(i\phi)$

$$\sum_{n=0}^{N-1} z_n w_n \exp(-i\omega n dt) \propto \exp(i\phi)$$

In the absence of noise, the final complex value would always be one of two possible values corresponding to the qubit states $|0\rangle$ and $|1\rangle$. However, the noise leads to Gaussian distributions centered at those two points. The size of the clouds is determined mostly by the noise of the

IMPA and cryogenic HEMT amplifier, while the separation between the clouds' centers is determined by the resonator probe power and duration. The signal to noise ratio of the measurement is determined by the clouds' separation and width [21, 22].

The 54 qubits are divided into nine frequency multiplexed readout groups of six qubits each. Within a group, each qubit is coupled to its own readout resonator, but all six resonators are coupled to a shared bandpass Purcell filter [21, 23, 24]. All qubits in a group can be read-out simultaneously by frequency-domain multiplexing [1, 25] in which the total probe signal is a superposition of probe signals at each of the readout resonators' frequencies. The phase shifts of these superposed signals are independently recovered in the FPGA by demodulating the complex IQ phasor with each intermediate frequency. In other words, we know what frequencies are in the superposed readout signal and we compute the Fourier coefficients at those frequencies to find the phase of each reflected frequency component.

IV. XEB THEORY

We use cross entropy benchmarking (XEB) [5, 26] to calibrate general single- and two-qubit gates, and also to estimate the fidelity of random quantum circuits with a large number of qubits. XEB is based on the observation that the measurement probabilities of a random quantum state have a similar pattern to laser "speckles", with some bitstrings more probable than others [27, 28]. The same holds for the output state of random quantum circuits. As errors destroy the speckle pattern, this is enough to estimate the rate of errors and fidelity in an experiment. Crucially, XEB does not require the reconstruction of experimental output probabilities, which would need an exponential number of measurements for increasing number of qubits. Rather, we use numerical simulations to calculate the likelihood of a set of bitstrings obtained in an experiment according to the ideal expected probabilities. Below we describe the theory behind this technique in more detail.

A. XEB of a small number of qubits

We first consider the use of XEB to obtain the error rate for single- and two-qubit gates. As explained above, for a two-qubit XEB estimation we use sequences of cycles, each cycle consisting of two sufficiently random single-qubit gates followed by the same two-qubit gate.

The density operator of the system after application of a random circuit U with m cycles can be written as a sum of two parts

$$\rho_U = \varepsilon_m |\psi_U\rangle \langle \psi_U| + (1 - \varepsilon_m) \chi_U, \quad D = 2^n. \quad (2)$$

Here $|\psi_U\rangle = U |\psi_0\rangle$ is the ideal output state and χ_U is an operator with unit trace that along with ε_m describes

the effect of errors. For a depolarizing channel model $\chi_U = I/D$ and ε_m has the meaning of the depolarization fidelity after m cycles. Nevertheless, in the case of small number of qubits, the part of the operator χ_U has nonzero matrix elements between the states with no error and the states with the error. However, if we undo the evolution of each random circuit and average over an ensemble of circuits such cross-terms are averaged out and we expect

$$\overline{U^\dagger \chi_U U} = \frac{I}{D}. \quad (3)$$

Here and below we use the horizontal bar on the top to denote averaging over the ensemble of random circuits. Because of this property it is possible to establish the connection between the quantity ε_m and the depolarization fidelity after m cycles.

From Eqs. (2) and (3) we get

$$\overline{U^\dagger \rho_U U} = \overline{\varepsilon_m} |\psi_0\rangle \langle \psi_0| + (1 - \overline{\varepsilon_m}) \frac{I}{D}. \quad (4)$$

This is a depolarizing channel. From this and the exponential decay of fidelity we get

$$\overline{\varepsilon_m} = p_c^m, \quad (5)$$

connecting $\overline{\varepsilon_m}$ to the depolarization fidelity p_c per cycle.

The noise model (2) is very general in the context of random circuits. To provide some insight about the origin of this model we consider a specific case with pure systematic error in the two-qubit gate. In this case the resulting pure state after the application of the random circuit U with the error can be expanded into the direction of the ideal state vector and the orthogonal direction

$$\tilde{U} |\psi_0\rangle = \xi_m |\psi_U\rangle + \sqrt{1 - |\xi_m|^2} |\varphi_{\tilde{U}}\rangle, \quad (6)$$

where

$$\langle \psi_U | \varphi_{\tilde{U}} \rangle = 0, \quad \langle \varphi_{\tilde{U}} | \varphi_{\tilde{U}} \rangle = 1. \quad (7)$$

For the ensemble of random circuits U the error vector is distributed completely randomly in the plane orthogonal to the ideal vector $U |\psi_0\rangle$ (see Fig. S5). This condition of orthogonality is the only constraint on the vector $|\varphi_{\tilde{U}}\rangle$ that involves $|\psi_U\rangle$. Therefore we expect

$$\overline{U^\dagger |\varphi_{\tilde{U}}\rangle \langle \varphi_{\tilde{U}}| U} = \frac{1}{D-1} (I - |\psi_0\rangle \langle \psi_0|). \quad (8)$$

Also

$$\overline{U^\dagger \left(\xi_m \sqrt{1 - |\xi_m|^2} |\psi_U\rangle \langle \varphi_{\tilde{U}}| + h.c. \right) U} = 0. \quad (9)$$

This gives the connection between the error vector $|\varphi_{\tilde{U}}\rangle$ and the operator χ_U

$$(1 - \varepsilon_m) \chi_U - \frac{1 - \varepsilon_m}{D} |\psi_U\rangle \langle \psi_U| = (1 - |\xi_m|^2) |\varphi_{\tilde{U}}\rangle \langle \varphi_{\tilde{U}}| + \left(\xi_m \sqrt{1 - |\xi_m|^2} |\psi_U\rangle \langle \varphi_{\tilde{U}}| + h.c. \right). \quad (10)$$

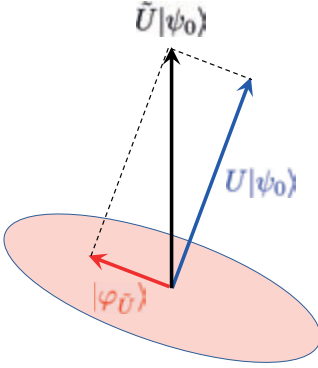


FIG. S5. **Cartoon: decomposition of the quantum state into the vector aligned with the ideal quantum state and its orthogonal complement**

The resulting equation

$$|\xi_m|^2 = \bar{\varepsilon}_m + \frac{1 - \bar{\varepsilon}_m}{D} \quad (11)$$

is to be expected, because $|\xi_m|^2$ is the average state fidelity while $\bar{\varepsilon}_m$ is the depolarization fidelity (see Sec. V). Note that Eqs. (8)–(11) lead to Eq. (4). This result can also be derived assuming that single qubit gates form a 2-design in the Hilbert space of each qubit.

We demonstrate the above findings by numerically simulating the random circuits for 2 qubits that contains single qubit gates randomly sampled from Haar measure and ISWAP-like gate

$$V(\theta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta & -i \sin \theta & 0 \\ 0 & -i \sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (12)$$

The systematic error $\Delta\theta = \theta - \pi/2$ corresponds to the deviation of the swap angle from $\pi/2$. Then assuming that the single qubit gates are error free the depolarizing channel model gives the prediction for the depolarizing fidelity per cycle

$$p_c = \frac{|\text{tr}(V(\theta)V^\dagger(\pi/2))|^2 - 1}{D^2 - 1} = \frac{1}{15}(8 \cos(\Delta\theta) + 2 \cos(2\Delta\theta) + 5). \quad (13)$$

As shown in Fig. S6 the depolarizing fidelity p_c^m for the circuit of depth m based on Eq. (13) closely matches the corresponding quantity obtained by the averaging of the squared overlap over the ensemble of random circuits (cf. (11))

$$\bar{\varepsilon}_m = \frac{D |\xi_m|^2 - 1}{D - 1}. \quad (14)$$

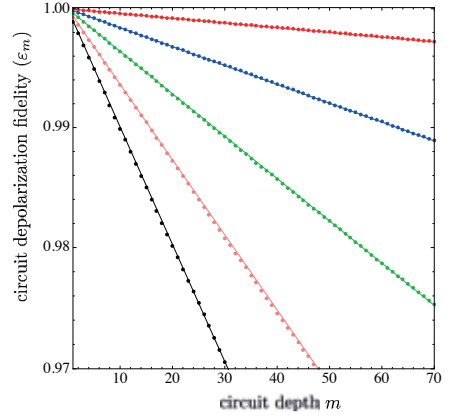


FIG. S6. **Plots of the circuit depolarizing fidelity vs the circuit depth.** Solid lines corresponds to the predictions from the depolarizing channel model (13) and points correspond to $\bar{\varepsilon}_m$ (14) obtained by the averaging of the squared overlap over the ensemble of random circuits. Different colored pots correspond to different values of the swap error $\Delta\theta = 0.01$ (red), 0.02 (blue), 0.03 (green), 0.04 (pink), 0.05 (black).

Returning to the generic case, property (3) can be extended so that for any smooth function $f(u)$ the following relation holds

$$\overline{\sum_{q \in \{0,1\}^n} f(p_s(q)) \langle q | \chi_U | q \rangle} = \sum_{q \in \{0,1\}^n} \frac{f(p_s(q))}{D} + \epsilon, \quad (15)$$

where $|q\rangle$ is a computational basis state corresponding to bitstring q , and $p_s(q) = \langle q | U \rho_0 U^\dagger | q \rangle$ is the simulated (computed) ideal probability of q . If the average is performed over a sample of random circuits of size S then the correction is $\epsilon \in O(1/\sqrt{S})$. We tested numerically for the case of $n = 2$ that relation (15) holds even for purely systematic errors in the case of a sufficiently random set of single qubit gates.

We now make the critical step of estimating the parameter p_c^m from a set of experimental realizations of random circuits with m cycles. We map each measured bitstring q with a function $f(p_s(q))$ and then average this function over the measured bitstrings. The standard XEB [5, 26] uses the natural logarithm, $f(p_s(q)) = \log(p_s(q))$. In the main text we use the linear version of XEB, for which $f(p_s(q)) = D p_s(q) - 1$. Both these functions give higher values to bitstrings with higher simulated probabilities. Another closely related choice is the Heavy Output Generation test [29], for which f is a step-function.

Under the model (2), in an experiment with ideal state preparation and measurement, we obtain the bitstring q

with probability

$$p_c^m p_s(q) + (1 - p_c^m) \langle q | \chi_U | q \rangle, \quad (16)$$

For the linear XEB, the average value of $Dp_s(q) - 1$ when sampling with probabilities given by Eq. (16) is

$$\overline{(Dp_s(q) - 1)} = p_c^m \left(D \sum_q \overline{p_s(q)^2} - 1 \right). \quad (17)$$

Similarly to Eq. (15), the horizontal bar denotes averaging over the random circuits.

The sum on the right hand side of (17) goes over all bitstrings in the computational basis, and can be obtained with numerical simulations. It can also be found analytically assuming that the random circuit ensemble approximates the Haar measure where for a given q the quantity $p_s(q)$ is distributed with the beta distribution function $(D-1)(1-p_s)^{D-2}$. In this case the right hand side in (17) equals $p_c^m(2D/(D+1) - 1)$.

The experimental average on the left hand side of (17) can be estimated with accuracy $1/\sqrt{SN_s}$ using S random circuit realizations with N_s samples each

$$\begin{aligned} \frac{1}{SN_s} \sum_{j=1}^S \sum_{i=1}^{N_s} (Dp_s^j(q_{i,j}) - 1) \\ = \overline{(Dp_s(q) - 1)} + O\left(\frac{1}{\sqrt{SN_s}}\right). \end{aligned} \quad (18)$$

This gives an estimate of p_c^m .

This estimate can be justified using Bayes rule. The log-likelihood for a set of experimental measurements $\{q_{i,j}\}$ assuming that the experimental probabilities are given by Eq. (16) is proportional to

$$\sum_{j=1}^S \sum_{i=1}^{N_s} \log(1 + p_c^m(Dp_s^j(q_{i,j}) - 1)), \quad (19)$$

where $p_s^j(q)$ is a simulated probability corresponding to the j -th circuit realization. We want to maximize the log-likelihood as a function of p_c^m . Taking the derivative with respect to p_c^m and equating to 0 we obtain

$$\sum_{j=1}^S \sum_{i=1}^{N_s} \frac{Dp_s^j(q_{i,j}) - 1}{1 + p_c^m(Dp_s^j(q_{i,j}) - 1)} = 0, \quad (20)$$

For $p_c^m \ll 1$ it is easy to solve this equation and obtain the estimate

$$p_c^m \simeq \frac{\sum_{j=1}^S \sum_{i=1}^{N_s} (Dp_s^j(q_{i,j}) - 1)}{\sum_{j=1}^S \sum_{i=1}^{N_s} (Dp_s^j(q_{i,j}) - 1)^2} \simeq \frac{\overline{(Dp_s(q) - 1)}}{D \sum_q \overline{p_s(q)^2} - 1}. \quad (21)$$

In the spirit of the XEB method, we can use other functions $f(p_s(q))$ to estimate p_c^m . One alternative is

derived from the log-likelihood of a sample $\{q_{i,j}\}$ with respect to the simulated (computed) ideal probabilities

$$\log \Pi_{j=1}^S \Pi_{i=1}^{N_s} p_s^j(q_{i,j}) = \sum_{j=1}^S \sum_{i=1}^{N_s} \log p_s^j(q_{i,j}), \quad (22)$$

which converges to the cross entropy between experimental probabilities and simulated probabilities. The experimental average of the function $f(p_s(q)) = \log p_s(q)$ under the probabilities from Eq. (16) with additional averaging over random circuits is

$$\begin{aligned} \overline{(\log p_s(q))} &\simeq p_c^m \left(\sum_q (\overline{p_s(q)} - 1/D) \log p_s(q) \right) \\ &+ \frac{1}{D} \sum_q \overline{\log p_s(q)}. \end{aligned} \quad (23)$$

As before, the sums on the right hand side can be obtained with numerical simulations and the average value on the left hand side can be estimated experimentally. This also gives an estimate of p_c^m .

Both Eq. (17) and Eq. (23) give a linear equation, from which we can obtain an estimate of the total polarization p_c^m for an experimental implementation of one quantum circuit with m cycles. We normally use multiple circuits with the same number of cycles m to estimate p_c^m , which we can do using the least squares method. Finally, we obtain an estimate of p_c from a fit of the estimates p_c^m as an exponential decay in m . This is standard in randomized benchmarking [30, 31]. One advantage of this method is that it allows us to estimate the cycle polarization p_c independently of the state preparation and measurement errors (SPAM). See also below.

B. XEB of a large number of qubits

We now consider the case of a large number of qubits $n \gg 1$. We are typically interested in estimating the fidelity F of each of a set of circuits with a given number of qubits and depth. As above, we write the output of an approximate implementation of the random quantum circuit U as

$$\rho_U = F |\psi_U\rangle \langle \psi_U| + (1 - F) \chi_U, \quad (24)$$

where $|\psi_U\rangle$ is the ideal output and $F = \langle \psi_U | \rho_U | \psi_U \rangle$ is the fidelity. We do not necessarily assume $\chi_U = I/D$, and we will ignore the small difference, of order 2^{-n} , $n \gg 1$, between the fidelity F and the depolarization fidelity p .

As for the case of small number of qubits n , we map each output bitstring q with a function $f(p_s(q))$. Given that the values $\langle q | \chi_U | q \rangle$ resulting from errors are typically uncorrelated with the chaotic ‘‘speckles’’ of $p_s(q)$, we make our main assumption

$$\sum_q \langle q | \chi_U | q \rangle f(p_s(q)) = \frac{1}{D} \sum_q f(p_s(q)) + \epsilon. \quad (25)$$

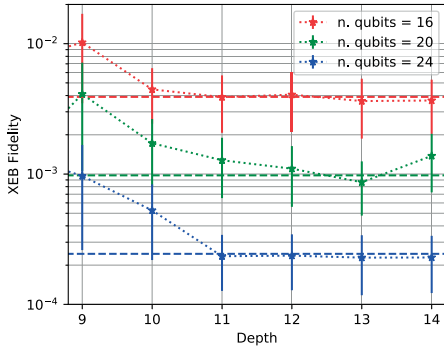


FIG. S7. Absolute value of the XEB fidelity between a random quantum circuit and the same circuit with a single Pauli error. Markers show the median over all possible positions in the circuit for both bit-flip and phase-flip errors. Error bars correspond to the first and third quartile. The dashed lines are the $1/\sqrt{D}$ theory prediction.

This equation is trivial if we assume a depolarizing model, $\chi_U = I/D$. More generally, it can be understood in the geometric context of concentration of measure [32–35] for high dimensional spaces, and from Levy’s lemma [36] we expect a typical statistical fluctuation $\epsilon \in O(1/\sqrt{D})$ with $D = 2^n$. We will only require $\epsilon \ll F$. We check Eq. (25) numerically for the output $\rho_e = |\psi_e\rangle\langle\psi_e|$ where $|\psi_e\rangle$ is the wave function obtained after a single phase-flip or bit-flip error is added somewhere in the circuit, see Fig. S7 and Ref. [26]. We have also tested this assumption numerically comparing the fidelity with the XEB estimate for a pure state $\sqrt{F}|\psi_U\rangle + \sqrt{1-F}|\psi_\perp\rangle$, see also Ref. [37] and Section X.

From Eqs. (24) and (25) we obtain Eq. (17) for linear XEB, $f(p_s(q)) = Dp_s(q) - 1$ (\mathcal{F}_{XEB} in the main text). We also obtain Eq. (23) for XEB, $f(p_s(q)) = \log p_s(q)$, with p_c^m replaced by fidelity F . As before, the sums on the right hand side can be obtained with numerical simulations and the average value on the left hand side can be estimated experimentally with accuracy $1/\sqrt{N_s}$ using N_s samples. This gives an estimate of F .

In practice, circuits of enough depth (as in the experiments reported here) exhibit the Porter-Thomas distribution for the measurement probabilities $p = \{p_s(q)\}$, that is

$$\text{Pr}(p) = D e^{-Dp}. \quad (26)$$

In this case the linear cross entropy Eq. (17) gives

$$F = \langle Dp_s(q) - 1 \rangle. \quad (27)$$

The standard deviation of the estimate of F with N_s samples from the central limit theorem is

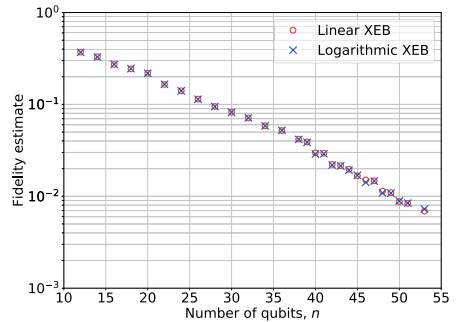


FIG. S8. Comparison of fidelity estimates obtained using linear XEB, Eq. (27) and logarithmic XEB, Eq. (28) from bitstrings observed in our quantum supremacy experiment. Standard deviation smaller than markers.

$\sqrt{(1 + 2F - F^2)/N_s}$. The cross entropy Eq. (23) gives

$$F = \langle \log Dp_s(q) \rangle + \gamma, \quad (28)$$

where γ is the Euler-Mascheroni constant ≈ 0.577 . The standard deviation of the estimate of F with N_s samples is $\sqrt{(\pi^2/6 - F^2)/N_s}$. The logarithmic XEB has a smaller standard deviation for $F > 0.32$ (it is the best estimate when $F \approx 1$), while for $F < 0.32$ the linear XEB has a smaller standard deviation (it is the best estimate for $F \ll 1$, where it relates to the maximum likelihood estimator). See Fig. S8 for comparison of the fidelity estimates produced by the linear and logarithmic XEB.

We note in passing another example for an estimator of F related to the HOG test [29] which counts the number of measured bitstrings with probabilities $p_s(q)$ greater than the median of the probabilities. The function $f(p_s(q))$ in this case returns 1 for $Dp_s(q) \geq \log(2)$, and 0 in the other case. The fidelity estimator uses the following normalization

$$F = \frac{1}{\log(2)} \langle 2n_s(q) - 1 \rangle, \quad (29)$$

where $n_s(q)$ is defined to be 1 if $Dp_s(q) \geq \log(2)$, and 0 otherwise. The standard deviation of this estimator is $\sqrt{[\log^{-2}(2) - F^2]/N_s}$, which is always larger than for the XEB. See Fig. S9 for comparison of the fidelity estimates produced by linear XEB and the HOG-based fidelity estimator. HOG test is also related to a definition of quantum volume [38].

C. Two limiting cases

Here, we consider two special cases of equation (27) and the formula (1) in the main paper. First, suppose

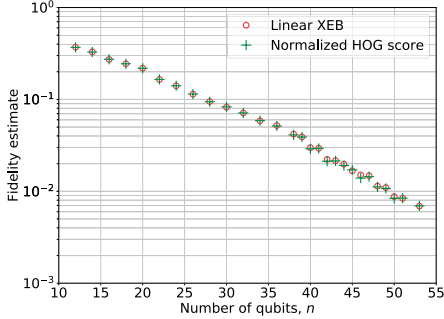


FIG. S9. Comparison of fidelity estimates obtained using linear XEB, Eq. (27) and normalized HOG score, Eq. (29) from bitstrings observed in our quantum supremacy experiment. Standard deviation smaller than markers.

bitstrings q_i are sampled from the uniform distribution. In this case $P(q_i) = 1/D$ for every i and $\mathcal{F}_{\text{XEB}} = 0$. Therefore, if the qubits are in the maximally mixed state, the estimator yields zero fidelity, as expected.

Second, suppose that bitstrings are sampled from the theoretical output distribution of a random quantum circuit. Assume that the distribution has Porter-Thomas shape. By equation (26), the fraction of bitstrings with theoretical probability in $[p, p + dp]$ is

$$\Pr(p) dp = D e^{-Dp} dp \quad (30)$$

and the total number of such bitstrings is

$$N(p) dp = D^2 e^{-Dp} dp. \quad (31)$$

Therefore, the probability that a bitstring with probability in $[p, p + dp]$ is sampled equals

$$p \cdot N(p) dp = p D^2 e^{-Dp} dp = f(p) dp \quad (32)$$

where $f(p)$ is the probability density function of the random variable defined as the ideal probability of a sampled bitstring, i.e. the random variable which is being averaged in the formula (1) of the main paper. Thus, the average probability of a sampled bitstring is

$$\begin{aligned} \langle P(q_i) \rangle &= \int_0^1 p f(p) dp = \int_0^1 p^2 D^2 e^{-Dp} dp \\ &= \frac{2}{D} \left(1 - e^{-D} \left(\frac{D^2}{2} + D + 1 \right) \right) \approx \frac{2}{D}. \end{aligned} \quad (33)$$

Substituting into equation (1) in the main paper yields $\mathcal{F}_{\text{XEB}} = 1$. The general case of a depolarizing error can be obtained from the two limiting cases by convex combination.

D. Measurement errors

We now consider how measurement errors affect the estimation of fidelity. Let us assume uncorrelated classical measurement errors, so that if the “actual” measurement result of a qubit is 0, we can get 1 with probability e_{m0} , and similarly with probability e_{m1} we get 0 for actual result 1, i.e., $p(1|0) = e_{m0}$, $p(0|0) = 1 - e_{m0}$, $p(0|1) = e_{m1}$, $p(1|1) = 1 - e_{m1}$. In this case the probability to get measurement result $q = k_1 k_2 \dots k_n$ for actual result $q' = k'_1 k'_2 \dots k'_n$ is the product of the corresponding factors. The probability of correct measurement result is then

$$\begin{aligned} p_m(q') &= (1 - e_{m0})^{n-|q'|} (1 - e_{m1})^{|q'|} \\ &\approx (1 - e_{m0})^{n/2} (1 - e_{m1})^{n/2}, \end{aligned} \quad (34)$$

where $|q'|$ is the number of 1s (Hamming distance from 00...0) in the initial bitstring q' , and in the second expression we approximated $|q'|$ with $n/2$ for large n .

Now let us make a natural assumption that if there was one or more measurement errors, $q' \rightarrow q$, then the resulting ideal probability $p_s(q)$ is uncorrelated with the actual ideal probability $p_s(q')$. Using this assumption we can write

$$F = F_U p_m \quad (35)$$

where F_U is the circuit fidelity and F is the complete (effective) fidelity. The complete fidelity F is estimated as before. The measurement fidelity p_m can be obtained independently. For instance, we can prepare a bitstring q and measure immediately to obtain the probability of a correct measurement result for q . We obtain p_m by repeating this for a set of random bitstrings. We can therefore obtain F_U from Eq. (35). As explained above, fitting the depolarization fidelity per cycle p_c for different circuit depths m is also a method to separate measurement errors.

The state preparation errors can be treated similarly, assuming that a single error leads to uncorrelated resulting distribution $p_s(q)$, so that the measurement fidelity p_m in Eq. (35) is combined with a similar factor describing the state preparation fidelity.

V. QUANTIFYING ERRORS

An important test for this experiment is predicting XEB fidelity \mathcal{F}_{XEB} based on simpler measurements of single- and two-qubit errors. Here we review how this is calculated, illustrating important principles with the example of a single qubit. The general theory is described at the end of this section.

First, we assume Pauli errors describe decoherence using a depolarizing model. This model is used, for example, to compute thresholds and logical error rates for error correction. The parameter describing decoherence

in a single qubit is the Pauli error e_P , giving a probability $e_P/3$ for applying an erroneous X, Y, or Z gate to the qubit after the gate, corresponding to a bit and/or phase flip.

Second, the depolarization model is assumed to describe the system state using simple classical probability. The probability of no error for many qubits and many operations, corresponding to no change to the system state, is then found by simply multiplying the probability of no error for each qubit gate. This is a good assumption for RB and XEB since a bit- or phase-flip error effectively decorrelates the state. The depolarization model assumes that when there is an error with probability e_d , the system state randomly splits to all qubits states, which has Hilbert space dimension $D = 2^n$. This is described by a change in density matrix $\rho \rightarrow (1 - e_d)\rho + e_d \times \mathbb{1}/D$. Note the depolarization term has a small possibility of the state resetting back to its original state. For a single qubit where $D = 2$, this can be described using a Pauli-error type model as a probability $e_d/4$ applying a I, X, Y, or Z gate. Comparing to the Pauli model, the error probability thus needs to be rescaled by $e_d = e_P/(1 - 1/D^2)$. This gives a net polarization p of the qubit state due to many Pauli errors as

$$p = \prod_i [1 - e_P(i)/(1 - 1/D^2)]. \quad (36)$$

Third, the effect of this depolarization has to be accounted for considering the measured signal. The measured signal for randomized benchmarking is given by $RB = p(1 - 1/D) + 1/D$, which can be understood in a physical argument that a complete randomization of the state has a $1/D$ chance to give the correct final state. A cross-entropy benchmarking measurement gives $\mathcal{F}_{\text{XEB}} = p$. A measurement of p , which can have offsets and prefactors in these formulas, also includes other scaling factors coming from state preparation and measurement errors. All of these scaling issues are circumvented by applying gates in a repeated number of cycles m such that $p = p_c^m$. A measurement of the signal versus m can then directly pull out the fractional polarization change per cycle, p_c , independent of these scale factors.

Fourth, from this polarization change we can then compute the Pauli error, which is the metric that should be reported since it is the fundamental error rate that is independent of D . Unfortunately, a fidelity $1 - e_P/(1 + 1/D)$ for RB is commonly reported, which has a D -dependent correction. We recommend this practice be changed, but note that removing the $1/(1 + 1/D)$ factor decreases the reported fidelity value. We also recommend reporting Pauli error, e_P instead of entanglement fidelity $(1 - e_P)$, since it is more intuitive to understand how close some quantity is to 0 than to 1. Table I summarizes the different error metrics and their relations.

This general model can also account for non-depolarizing errors such as energy decay, since quantum states in an algorithm typically average over the entire Bloch sphere (as in XEB), or for example when the al-

gorithm purposely inserts spin-echoes. Thus the average effect of energy decay effectively randomizes the state in a way compatible with Pauli errors. For a gate of length t_g with a qubit decay time T_1 , averaging over the Bloch sphere (2 poles and 4 equator positions) gives (to first order) an average error probability $e_a = t_g/3T_1$. Using Table I, this converts to a Pauli error $e_P = t_g/2T_1$.

A detailed theory of the D scaling factor is as follows. In order to arrive at a first order estimate on how error rates accumulate on random quantum circuits, the errors can be modeled via the set of Kraus operators. The density matrix of the system ρ after application of a gate is connected to the density matrix ρ_0 before the gate as follows:

$$\rho = \Lambda(\rho_0) = \sum_{k=0}^K A_k \rho_0 A_k^\dagger, \quad \sum_k A_k^\dagger A_k = \mathbb{1}. \quad (37)$$

For the closed-system quantum evolution with unitary U (no dephasing nor decay) the sum on the right hand side contains only one term with $k=0$ and $A_0 = U$. In general, Kraus operators describe the physical effects of many types of errors (control error, decoherence, etc.) that can explicitly depend on the gate. Knowing the Kraus operators allows us to calculate the total error budget as well as its individual components.

Conventionally, circuit fidelities are reported as a metric of its quality. To make a connection to physically observable quantities, the average fidelity can be expressed in terms of Kraus operators. In the absence of leakage errors and cross-talk the average fidelity equals

$$F = 1 - \frac{e_P}{1 + 1/D}, \quad e_P = 1 - \frac{1}{D^2} \sum_{k=0}^K |\text{tr}(U A_k^\dagger)|^2 \quad (38)$$

where $D = 2^n$ is the dimension of the Hilbert space and the quantity e_P plays a role of a Pauli error probability in the depolarizing channel model (see below).

For random circuits the effects of errors can be described by a depolarizing channel model, with Kraus operators of the form

$$\begin{aligned} A_{\mathbf{k}} &= \sqrt{\frac{e_P}{D^2 - 1}} P_{\mathbf{k}} U, \quad \mathbf{k} \neq \mathbf{0}, \\ A_0 &= \sqrt{1 - e_P} P_0 U, \\ P_{\mathbf{k}} &= \sigma_{k_1} \otimes \sigma_{k_2} \dots \otimes \sigma_{k_n} \end{aligned} \quad (39)$$

where $P_{\mathbf{k}}$ are strings of Pauli operators σ_{k_j} for individual qubits for $k_j = 1, 2, 3$ and also identity matrices σ_0 in the qubit subspace for $k_j = 0$. This form assumes that individual Pauli errors all happen with the same probability e_P .

To make a connection to experimental measurements of the cross-entropy we substitute (39) into (37) and obtain

$$\begin{aligned} \Lambda(\rho_0) &= (1 - e_P) U \rho_0 U^{-1} \\ &\quad + \frac{e_P}{D - 1/D} \left(\mathbb{1} - \frac{U \rho_0 U^{-1}}{D} \right). \end{aligned} \quad (40)$$

TABLE I. A “Rosetta stone” translation between error metrics. In single- and two-qubit RB or XEB experiments, we measure the per-gate (or per-cycle) depolarization decay constant p . The second column shows conversions from this rate to the various error metrics. The last two columns are representative comparisons for 0.1% Pauli error.

Error metric	Relation to depolarization decay constant p	n=1 (D=2)	n=2 (D=4)
Pauli error (e_p, r_P) ^a	$(1-p)(1-1/D^2)$	0.1%	0.1%
Average error (e_a, r)	$(1-p)(1-1/D)$	0.067%	0.08%
Depolarization error (e_d)	$1-p$	0.133%	0.107%

^a 1— process fidelity, or 1— entanglement fidelity

We compare this expression with the standard form of the depolarizing channel model

$$\Lambda(\rho_0) = pU\rho_0U^{-1} + (1-p)\frac{\mathbb{I}}{D}, \quad (41)$$

expressed in terms of the depolarization fidelity parameter p . Note the difference between the expressions. On the one hand, in (41) the second term corresponds to full depolarization in all directions. On the other hand, in (40) the second term describes full depolarization in all directions except for the direction corresponding to the ideal quantum state.

From (40), (41) one can establish the connection between the Pauli error rate and depolarizing fidelity parameter p

$$e_P = (1-p)(1-1/D^2) \quad (42)$$

We note that the explicit assumption of connecting Pauli errors to depolarization is needed for the small D case, typically for single- and two-qubit error measurements. Once we have measured the Pauli errors, then only a simple probabilistic calculation is needed to compute F_{XEB} in the large D case.

VI. METROLOGY AND CALIBRATION

A. Calibration overview

Quantum computations are physically realized through the time-evolution of quantum systems steered by analog control signals. As quantum information is stored in continuous amplitudes and phases, these control signals must be carefully chosen to achieve the desired result. Calibration is the process of performing a series of experiments on the quantum system to learn optimal control parameters.

Calibration is challenging for a number of reasons. Analog control requires careful control-pulse shaping as any deviation from the ideal will introduce error. Qubits require individual calibration as variations in the control system and qubits necessitate different control parameters to hit target fidelities. Optimal control parameters can also drift in time, requiring calibrations to be revisited to maintain performance. Additionally, the full

calibration procedure requires bootstrapping: using a series of control sequences with increasing complexity to determine circuit and control parameters to increasingly higher degrees of precision. Lastly, each qubit needs to perform a number of independent operations which are independently calibrated: single-qubit gates, two-qubit gates, and readout.

Our Sycamore processor offers a high degree of programmability: we can dynamically change the frequency of each qubit, as well as the effective qubit-qubit coupling between nearest neighbor qubits. This tunability gives us the freedom to enact many different control strategies, as well as account for non-uniformities in the processor’s parameters. However, these extra degrees of freedom are a double-edged sword. Additional control knobs always introduce a source of decoherence and control errors as well as an added burden on calibration.

Our approach is to systematize and automate our calibration procedure as much as possible, thus abstracting complexity away. This automation allows us to turn calibration into a science, where we can compare calibration procedures to determine optimal strategies for time, performance, and reliability. By employing calibration science to study full-system performance with different control strategies, we have been able to improve full-system fidelities by over an order of magnitude from initial attempts while decreasing the calibration time and improving reliability. Lastly, we design our calibration to be done almost entirely at the single- or two-qubit level, rather than at the system level, in order to be as scalable as possible.

1. Device registry

The device registry is a database of control variables and configuration information we use to control our quantum processors. The registry stores information such as operating frequencies, control biases, gate parameters such as duration, amplitude, parameterization of circuit models, etc. The goal of calibration is to experimentally determine and populate the registry with optimal control parameters. We typically store >100 parameters per qubit to achieve high fidelity across all of the various qubit operations. The large number of parameters and subtle interdependencies between them highlights the need for automated calibration.

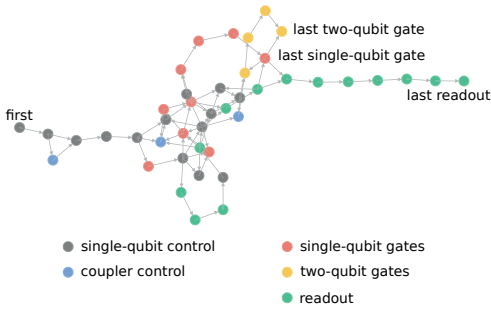


FIG. S10. **Optimus calibration graph for Sycamore.** Calibration of physical qubits is a bootstrapping procedure between different pulse sequences or “experiments” to extract control and system parameters. Initial experiments are coarse and have interplay between fundamental operations and elements such as single-qubit gates, readout, and the coupler. Final experiments involve precise metrology for each of the qubit operations: single-qubit gates, two-qubit gates, and readout.

2. Scheduling calibrations: “Optimus”

We seek a strategy for identifying and maintaining optimal control parameters for a system of physical qubits given incomplete system information. To perform these tasks, we use the “Optimus” formulation as in Ref [39], where each calibration is a node in a directed acyclic graph that updates one or more registry parameters, and the bootstrapping nature of calibration sequences is represented as directed edges between nodes. Now, calibrating a system of physical qubits becomes a well-defined graph traversal problem. The calibration graph used for the Sycamore device can be seen in Figure S10. This strategy is particularly useful for maintaining calibrations in the presence of drift, where we want to do the minimal amount of work to bring the system back in spec, and when extending the calibration procedure, as interdependencies are explicit. Typical timescales for bringup of a new Sycamore processor are approximately 36 hours upon first cooldown, and 4 hours per day thereafter for maintaining calibrations. These times are specific to current available technology, and can be significantly improved.

B. Calibration procedure

1. Device configuration

Throughout the calibration procedure, the device registry may be configured in different states in order to calibrate certain parameters. We call these different states “device configurations”, and different kinds of configu-

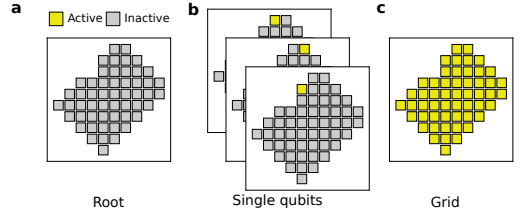


FIG. S11. **Configurations of the device over the course of calibration.** (a) In the root configuration, we start with no knowledge of the system and measure basic device parameters. (b) We create a single qubit configuration for each qubit, where all qubits except the qubit of interest are biased to near zero frequency. (c) Using knowledge learned in the single qubit configurations, we build a grid of qubits.

rations reflect our knowledge of the system at different points in the full calibration procedure. As illustrated in Figure S11, the primary difference between the different configurations is the set of “active” qubits, where active qubits are biased to an operating frequency between 5-7 GHz, and “inactive” qubits are biased near zero frequency. Following the outline above, we have three device configurations of interest:

a. Root config. The root configuration is the starting state of the system immediately after cool down and basic system verification. In this configuration, we calibrate coarse frequency vs bias curves for each readout resonator, qubit, and coupler.

b. Single qubit config. After completing root calibrations, we now know how to bias each qubit to its minimum and maximum frequencies. We create one configuration of the device registry for each qubit, where the qubit of interest is biased in a useful region (5-7 GHz) and the remaining qubits are biased to their minimum frequencies in order to isolate the qubit of interest. In each of these configurations, we fine tune the bias vs frequency curves for the qubit and its associated couplers and resonators, and also measure T_1 as a function of frequency, necessary due to background TLS defects and modes.

c. Grid config. After completing calibrations in each isolated qubit configuration, we feed the information we learned into a frequency optimization procedure. The optimizer places the biases for each qubit and coupler in a user defined grid of any desired size up to the entire chip. We then proceed to calibrate high fidelity single qubit gates, two qubit gates, and readout.

2. Root config: procedure

We begin calibration with simple frequency-domain experiments to understand how each qubit and coupler responds to its flux bias line.

- Calibrate each parametric amplifier (flux bias, pump frequency, pump power).
- For each qubit, identify its readout resonator and measure the readout signal versus qubit bias (“Resonator Spectroscopy”) [40]. Estimate the resonator and qubit frequency as a function of qubit bias.
- For each coupler, place one of its qubits near maximum frequency and the other near minimum frequency, then measure the readout signal of the first qubit as a function of coupler bias. The readout signal changes significantly as the coupler frequency passes near the qubit frequency. Identify where the coupler is near its maximum frequency, so the qubit-qubit coupling is small (a few MHz) and relatively insensitive to coupler bias.

3. Single-qubit config: procedure

After setting the biases to isolate a single qubit, we follow the procedure outlined in [41] which we will summarize here:

- Perform fixed microwave drive qubit spectroscopy while sweeping the qubit bias and detecting shifts in the resonator response, to find the bias that places the qubit at the desired resonant frequency.
- Using the avoided level crossing identified in the root config, determine the operating bias to bring the qubit on resonance with its readout resonator to perform active ground state preparation. We use a $10\ \mu\text{s}$ pulse consistent with the readout resonator ringdown time.
- Perform power Rabi oscillations to find the drive power that gives a π pulse to populate the $|1\rangle$ state.
- Optimize the readout frequency and power to maximize readout fidelity.
- Fine tune parameters (qubit resonant frequency, drive power, drive detuning [42]) for π and $\pi/2$ pulses.
- Calibrate the timing between the qubit microwave drive, qubit bias, and coupler bias.
- Perform qubit spectroscopy as a function of qubit bias to fine tune the qubit bias vs frequency curves.
- Measure T_1 vs. frequency by preparing the qubit in $|1\rangle$ then biasing the qubit to a variable frequency for a variable amount of time, and measuring the final population [43].
- Measure the response of a qubit to a detuning pulse to calibrate the frequency-control transfer function [5, 41, 44].

With the single-qubits calibrated in isolation, we have a wealth of information on circuits parameters and coherence information for each qubit. We use this information as input to a frequency placement algorithm to identify optimal operating frequencies for when the full processor is in operation.

4. Optimizing qubit operating frequencies

In our quantum processor architecture, we can independently tune each qubit’s operating frequency. Since qubit performance varies strongly with frequency, selecting good operating frequencies is necessary to achieve high fidelity gates. In arbitrary quantum algorithms, each qubit operates at three distinct types of frequencies: idle, interaction, and readout frequencies. Qubits idle and execute single-qubit gates at their respective idle frequencies. Qubit pairs execute two-qubit gates near their respective interaction frequencies. Finally, qubits are measured at their respective readout frequencies. In selecting operating frequencies, it is necessary to mitigate and make nontrivial tradeoffs between energy-relaxation, dephasing, leakage, and control imperfections. We solve and automate the frequency selection problem by abstracting it into an optimization problem.

We construct a quantum-algorithm-dependent and gate-dependent optimization objective that maps operating frequencies onto a metric correlated with system error. The error mechanisms embedded within the objective function are parasitic coupling between nearest-neighbor and next-nearest-neighbor qubits, spectrally-diffusing two-level-system (TLS) defects [43], spurious microwave modes, coupling to control lines and the readout resonator, frequency-control electronics noise, frequency-control pulse distortions, microwave-control pulse distortions, and microwave-carrier bleedthrough. Additional considerations in selecting readout frequencies are covered in Section VID. The objective is constructed from experimental data and numerics, and the individual error mechanisms are weighted by coefficients determined either heuristically or through statistical learning.

Minimizing the objective function is a complex combinatorial optimization problem. We characterize the complexity of the problem by the optimization dimension and search space. For a processor with N qubits on a square lattice with nearest-neighbor coupling, there are N idle, N readout, and $\sim 2N$ interaction frequencies to optimize. In an arbitrary quantum algorithm, all frequencies are potentially intertwined due to coupling between qubits. Therefore, the optimization dimension is $\sim 4N$. The optimization search-space is constrained by qubits’ circuit parameters and control-hardware specifications. Discretizing each qubit’s operational range to 100 frequencies results in an optimization search space of $\sim 100^{4N}$. This is much larger than the dimension of the Hilbert space of an N qubit processor, which is 2^N .

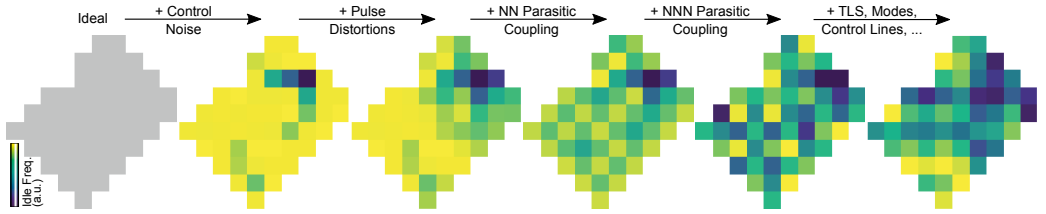


FIG. S12. **Idle frequency solutions found by our Snake optimizer with different error mechanisms enabled.** The optimizer makes increasingly complex tradeoffs as more error mechanisms are enabled. These tradeoffs manifest as a transition from a structured frequency configuration into an unstructured one. Similar tradeoffs are simultaneously made in optimizing interaction and readout frequencies. Optimized idle and interaction operating frequencies are shown in Figure S13 and optimized readout frequencies are shown in Figure S20. Color scales are chosen to maximize contrast. Grey indicates that there is no preference for any frequency.

Given the problem complexity, it is assumed that finding globally optimal operating frequencies is intractable. However, we have empirically verified that locally optimal solutions are sufficient for state-of-the-art system performance. To find local optima, we developed the “Snake” homebrew optimizer that combines quantum algorithm structure with physics intuition to exponentially reduce optimization complexity and take intelligent optimization steps. For the circuits used here, the optimizer exploits the time-interleaved structure of single-qubit gates, two-qubit gates, and readout. For our 53 qubit processor, it returns local optima in ~ 10 seconds on a desktop. Because of its favorable scaling in runtime versus number of qubits, we believe the Snake optimizer is a viable long-term solution to the frequency selection problem.

To illustrate how the Snake optimizer makes tradeoffs between error mechanisms, we plot idle frequency solutions with different error mechanisms enabled (Figure S12). Starting with an ideal processor with no error mechanisms enabled, there is no preference for any frequency configuration. Enabling frequency-control electronics noise, the optimizer pushes qubits towards their respective maximum frequencies, to minimize flux-noise susceptibility. Note that each qubit has a different maximum frequency due to fabrication variability. Enabling frequency-control pulse distortions forces a gradual transition between qubit frequencies to minimize two-qubit-gate frequency-sweep amplitudes. Enabling nearest-neighbor (NN) and next-nearest neighbor (NNN) parasitic coupling further lowers the degeneracy between qubit frequencies into a structure that resembles a multi-tiered checkerboard. Finally, enabling errors from TLS defects, spurious microwave modes, and all other known error mechanisms removes any obvious structure. A set of optimized idle and interaction frequencies is shown in Figure S13, and readout frequencies are shown in Figure S20.

5. Grid config: procedure

Calibrating a grid of qubits follows the same procedure as calibrating an isolated qubit with additional calibrations to turn off the qubit-qubit coupling.

- Achieve basic state discrimination for each qubit at its desired frequency.
- For each coupler, minimize the qubit-qubit coupling (note changing coupler biases affects qubit frequencies). For each case below, we choose the coupler bias minimizing the interaction.
 - For qubit pairs idling within 60 MHz of each other, use a resonant swapping experiment. We excite one qubit and apply flux pulses to nominally put the qubits on resonance and let the qubits interact over time [8].
 - For qubit pair idling further apart, use a conditional phase experiment. We perform two Ramsey experiments on one qubit, where the other qubit is in the ground state and the excited state, to identify the state-dependent frequency shift of the first qubit.
- Adjust the qubit biases to restore the desired qubit frequencies and proceed with qubit calibration as in the single-qubit configurations.
- Calibrate the entangling gate.
 - Estimate the qubit pulse amplitudes to reach the desired interaction frequency with their frequency versus bias calibration.
 - Fine-tune the qubit pulse amplitudes to reach resonance, compensating for pulse under-shoot.
 - Tune the coupler pulse amplitude to achieve a complete photon exchange.

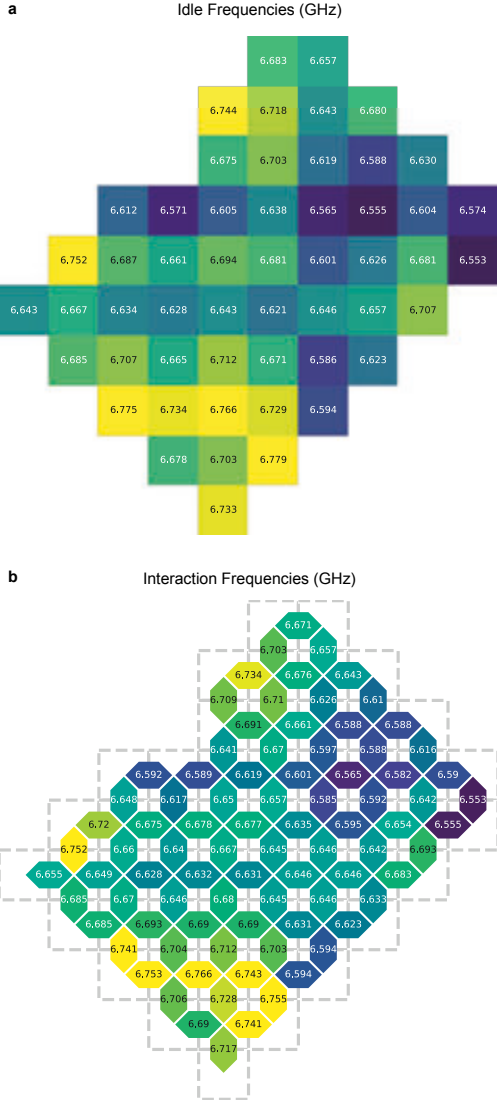


FIG. S13. **Optimized idle and interaction frequencies found by our Snake optimizer.** a, Idle frequencies, b, interaction frequencies. Readout frequencies are shown in Figure S20. These solutions are sufficient for state-of-the-art system performance. See Figure S12 to understand some of the tradeoffs that are made during optimization. Color scales are chosen to maximize contrast.

In the next two sections, we describe in more detail the fine tuning required to achieve high fidelity two qubit gates and multiqubit readout.

C. Two-qubit gate metrology

High-fidelity two-qubit gates are very hard to achieve. In an effort to make this easier, we design qubits with tunable frequencies and tunable interactions. This added control allows for immense flexibility when implementing gates. In the following subsections, we discuss a simple high-fidelity control and metrology strategy for two-qubit gates in our system.

1. The natural two-qubit gate for transmon qubits

Consider two transmon qubits at different frequencies (say 6.0 and 6.1 GHz). Here are two potential ways of generating a multi-qubit gate in this system. If the qubits are tuned into resonance, then excitations swap back-and-forth and this interaction can be modeled as a partial-iSWAP gate [45]. If the qubits are detuned by an amount close to their nonlinearity, then the 11-state undergoes an evolution that can be modeled as a controlled-phase gate (assuming the population does not leak) [46, 47]. In fact, any two-qubit control sequence that does not leak can be modeled as a partial-iSWAP followed by a controlled-phase gate.

A typical control sequence is shown Fig. S14a. Gate times of 12 ns are chosen to trade off decoherence (too slow) and leakage to higher states of the qubit (too fast). Figure S14b depicts how this operation can be decomposed as a quantum circuit. This circuit contains Z-rotations that result from the frequency excursions of the qubits, and can be expressed by the unitary:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & e^{i(\Delta_+ + \Delta_-)} \cos \theta & -ie^{i(\Delta_+ - \Delta_-)} \sin \theta & 0 \\ 0 & -ie^{i(\Delta_+ + \Delta_-)} \sin \theta & e^{i(\Delta_+ - \Delta_-)} \cos \theta & 0 \\ 0 & 0 & 0 & e^{i(2\Delta_+ - \phi)} \end{bmatrix} \quad (43)$$

These gates have an efficient mapping to interacting fermions and have been coined ‘fSim’ gates, short for fermionic simulation [48]. The long-term goal is to implement the entire space of gates (shown in Fig. S14c).

For quantum supremacy, the two-qubit gate of choice is the iSWAP gate. For example, CZ is less computationally expensive to simulate on a classical computer by a factor of two [37, 49]. A dominant error-mechanism when trying to implement an iSWAP is a small conditional-phase that is generated by an interaction of the $|11\rangle$ -state with higher states of the transmons ($|02\rangle$ and $|20\rangle$). For this reason, the fSim gate with swap-angle $\theta \approx 90^\circ$ and conditional phase $\phi \approx 30^\circ$ has become the gate of choice in our supremacy experiment. Note that small deviations from these angles are also viable quantum supremacy

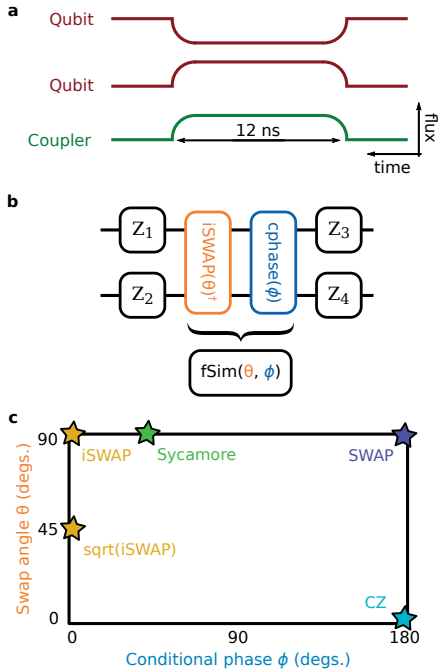


FIG. S14. **Two-qubit gate strategy.** **a**, Control waveforms for two qubits and a coupler. Each curve represents the control flux applied to the qubit's and coupler's SQUID loops as a function of time. **b**, Generic circuit representation for an arbitrary two-qubit gate using flux pulses. This family of gates have been named "fSim" gates, short for fermionic-simulation gates. Our definition of the fSim gate uses θ with the sign opposite to the common convention for the iSWAP gate. **c**, Control landscape for fSim gates as a function of the swap angle and conditional phase, up to single qubit rotations. The coordinates of common entangling gates are marked along with the Sycamore gate fSim($\theta = 90^\circ, \phi = 30^\circ$).

gates. These gates result from the natural evolution of two qubits making them easy to calibrate, high intrinsic fidelity gates for quantum supremacy.

2. Using cross entropy to learn a unitary model

We have recently introduced cross-entropy as a fidelity metric for quantum supremacy experiments. Cross-entropy benchmarking (XEB) was introduced as an analog to randomized benchmarking (RB) that can be used with any number of qubits and is independent of state-preparation and measurement errors [5, 26].

A distinct advantage of XEB is that the resulting data can be analyzed to find an optimal representation of a

unitary; this process is outlined in Fig. S15. The gate sequence for a two-qubit XEB experiment is shown in Fig. S15a. The sequence alternates between single-qubit gates on both qubits and a two-qubit gate between them. At the end of the sequence, both qubits are measured and the probabilities of bitstrings (00, 01, 10, 11) are estimated. This procedure is repeated for ~ 10 -20 instances of randomly selected single-qubit gates. The measured probabilities can then be compared to the ideal probabilities using the expression for fidelity Eq. (3) in Ref. [5].

The data from a two-qubit XEB experiment is shown in Fig. S15b (green dots). By performing additional sequences with tomography rotations prior to measurement, we can infer the decay of purity with increasing circuit depth (blue dots). For two qubits, the decay of fidelity tells us the total error of our gates while the purity decay tells us the contribution from decoherence—the difference is control error. Based on the data in green and blue, it appears that the total error is about half control and half decoherence.

So far, we have established a generic unitary model (Fig. S14b), a training dataset (Fig. S15a), and a cost-function (Fig. S15b). These three ingredients form the foundation for using optimization techniques to improve fidelity. Using a simple Nelder-Mead optimization protocol, we can maximize the XEB fidelity by varying the parameters of the unitary model. The fidelity decay curve for the optimal unitary model are shown in Fig. S15b (orange dots). The optimized results are nearly coherence limited.

The optimal control-model parameters for all pairs are shown as integrated histograms in Fig. S16a,b. Panel (a) shows the histograms for partial-iSWAP angles (~ 90 degrees) and conditional phases (~ 30 degrees). Panel (b) shows histograms for the various flavors of Z-rotations. While conceptually there are four possible Z-rotations (see Fig. S14b), only three of these rotations are needed to uniquely define the operation. These three rotations can be thought of as the detuning of the qubits before the iSWAP, the detuning after the iSWAP, and an overall frequency shift of both qubits which commutes with the iSWAP.

3. Comparison with randomized benchmarking

In Fig. S17 we show that two-qubit gate fidelity extracted using XEB agrees well with the fidelity as measured with RB, an important sanity check in validating XEB as a gate metrology tool. In two-qubit XEB, we extract the error per cycle which consists of a single-qubit gate on each qubit and a two-qubit gate between them. In Fig. S17a we show the individual RB decay curves for single-qubit gates. In panel b, we show the RB decay curve for benchmarking a CZ gate. Adding up the three errors from RB, we would expect an XEB cycle error of 0.57%. In panel c, we show the measured XEB decay curve which indicates a cycle error of 0.59%—nearly

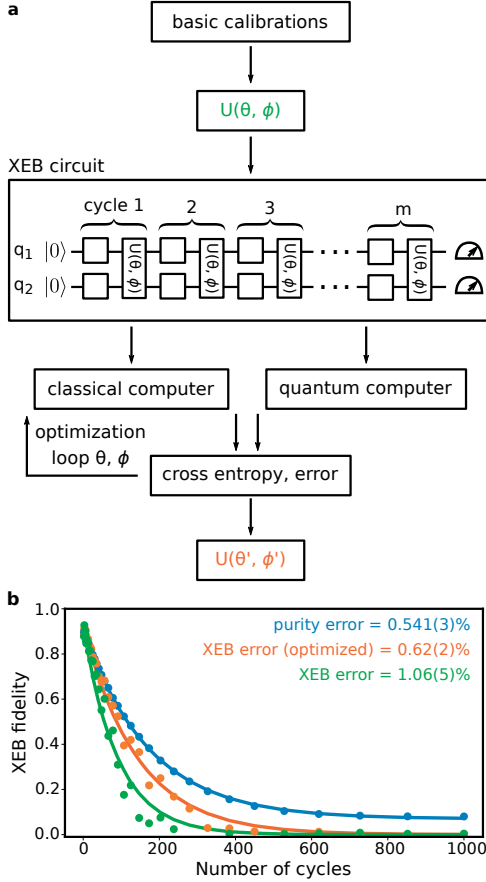


FIG. S15. **Using XEB to learn a unitary model.** **a**, Process flow diagram for using XEB to learn a unitary model. After running basic calibrations, we have an approximate model for our two-qubit gate. Using this gate, we construct a random circuit that is fed into both the quantum computer and a classical computer. The results of both outputs can be compared using cross-entropy. Optimizing over the parameters in the two qubit model provide a high-fidelity representation of the two-qubit unitary. **b**, Data from a two-qubit XEB experiment. The two-qubit purity (blue) was measured tomographically and provides the coherence-limit of the operations. The decay of the XEB fidelity is shown in green and orange. In orange, the parameters of a generic unitary model were optimized to determine a higher-fidelity representation of the unitary. All errors are quoted as Pauli errors.

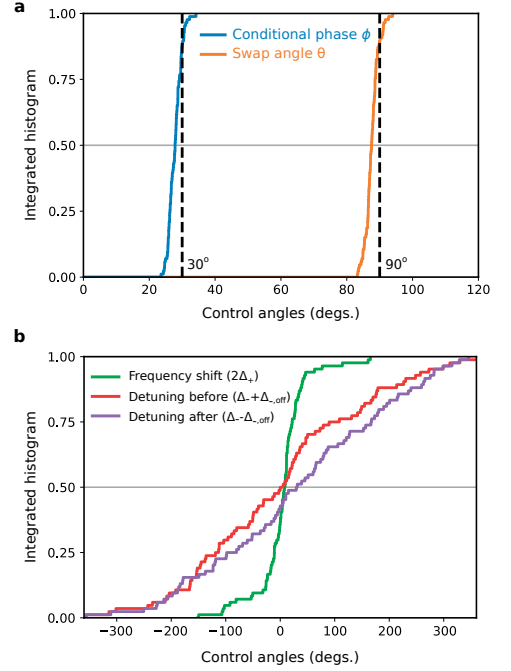


FIG. S16. **Parameters of the control model.** A generic model for two-qubit gates using flux-control has five free parameters. Using XEB we can measure these parameters with high fidelity. **a**, Integrated histogram (cumulative distribution) of the control parameters that determine the interaction between the qubits. **b**, An integrated histogram of the remaining three parameters that represent different flavors of single-qubit Z-rotations. While the first two parameters (panel a) define the entangling gate, the final three parameters (panel b) are simply measured and then kept track of during an algorithm. Intuitively, these three angles correspond to a detuning before the swap, a detuning after the swap, and an overall frequency shift which commutes through the swap; these correspond to $\Delta_- + \Delta_{\text{off}}$, $\Delta_- - \Delta_{\text{off}}$, and $2\Delta_+$ respectively in Eq. (43). Note that θ and ϕ angles are 360 degrees periodic and Z-rotation angles are 720 degree periodic.

identical to the value predicted by RB.

For single-qubit gate benchmarking on the Sycamore device used in this work (see Table II), we find that π pulse fidelities are somewhat worse than $\pi/2$ pulse fidelities, which we attribute to reflections from the imperfect microwave environment. Because the XEB gateset we have used consists only of $\pi/2$ pulses, we find that the single-qubit gate errors extracted from conventional RB, which contains π pulses, are somewhat higher than those extracted from single-qubit XEB. Using only $\pi/2$ pulses instead of π pulses in single-qubit RB brings the

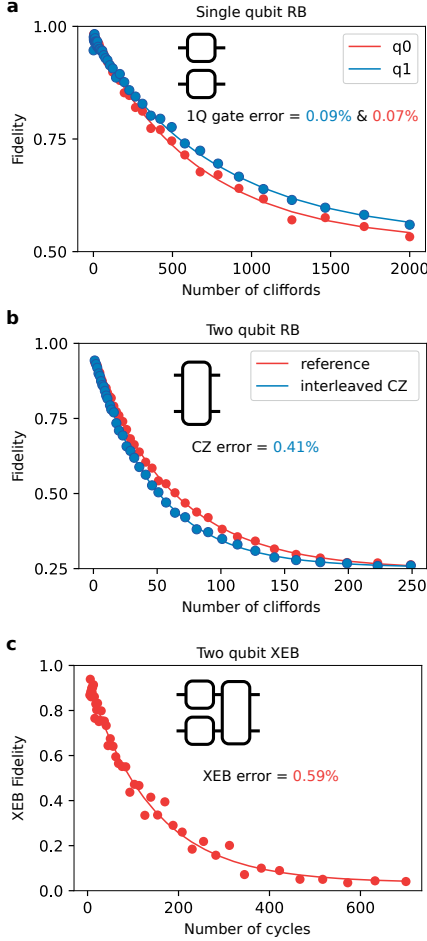


FIG. S17. **Sanity check: XEB agrees with RB.** **a**, Single-qubit randomized benchmarking (RB) data taken separately on two qubits. **b**, Two-qubit randomized benchmarking data for a CZ on the same pair of qubits. **c**, Two-qubit cross-entropy benchmarking (XEB) on the same pair of qubits. The measured XEB error (0.59% / cycle) agrees well with the prediction from single- and two-qubit RB (0.57%). All errors are quoted as Pauli errors.

extracted error close to that measured via XEB.

4. Speckle purity benchmarking (SPB)

It is experimentally useful to be able to extract state purity from XEB experiments in order to error-budget the contribution of decoherence. Conventionally, purity estimation can be done with state tomography, where the full density matrix ρ is reconstructed and used to quantify the state purity. This involves expanding a single sequence into a collection of sequences each appended with single-qubit gates. Unfortunately, full tomographic reconstruction scales exponentially in the number of qubits, both for the number of sequences needed as well as the number of measurements needed per sequence. Here, we introduce an exponentially more efficient method to extract the state purity without additional sequences.

We use a re-scaled purity definition such that a fully-decohered state has a purity of 0, and a pure state has a purity of 1. We define

$$\text{Purity} = \frac{D}{D-1} \left(\text{Tr}(\rho^2) - \frac{1}{D} \right), \quad (44)$$

which is consistent with what is defined in Ref. [50]. This can be understood as the squared length of the generalized Bloch vector in D dimensions (for a qubit, $D = 2$, this definition gives $\langle X \rangle^2 + \langle Y \rangle^2 + \langle Z \rangle^2$).

“Speckle” Purity Benchmarking (SPB) is the method of measuring the state purity from raw XEB data. Assuming the depolarizing-channel model with polarization parameter p , we can model the quantum state as

$$\rho = p |\psi\rangle \langle \psi| + (1-p) \frac{\mathbb{1}}{D}. \quad (45)$$

Here, p is the probability of a pure state $|\psi\rangle$ (which in this case is not necessarily known to us), while $1-p$ is the probability of being in the fully-decohered state ($\mathbb{1}$ is the identity operator). For the state (45), from the definition (44) it is easy to find the relation

$$\text{Purity} = p^2. \quad (46)$$

We will now work out how to obtain p^2 from a distribution of measured probabilities P_m of various bitstrings for a sequence, collected over many XEB sequences (Figs. S18a and S18b).

First, we note that for $p = 0$ the probabilities of all bitstrings are $1/D$, and the distribution is the δ -function located at $1/D$ (the integrated histogram is then the step-function – see Fig. S18b). In contrast, if $p = 1$, then the measured probabilities P_m follow the D -dimensional Porter-Thomas distribution [26]

$$\mathcal{P}_{\text{PT}}(P_m) = (D-1)(1-P_m)^{D-2}, \quad (47)$$

which has the same average $1/D$ and variance

$$\text{Var}_{\text{PT}}(P_m) = \frac{D-1}{D^2(D+1)}. \quad (48)$$

For the fully-decohered state all bitstrings have the same probability $1/D$, so in this case the variance of the distribution of probabilities is zero. For the state (45) with an arbitrary p , the histogram of probabilities P_m will be described by the distribution (47) shrunk towards the average $1/D$ by the factor p . Consequently, the variance of the experimental probabilities will be p^2 times the Porter-Thomas variance (48).

Thus, we can find p^2 by dividing the variance of experimentally measured probabilities P_m by the Porter-Thomas variance (48). Finally, using the relation (46) for the depolarization model (45), we can relate the variance of the experimental probabilities P_m to the average state purity

$$\text{Purity} = \text{Var}(P_m) \frac{D^2(D+1)}{D-1}. \quad (49)$$

With these convenient relations, we can directly compare the XEB fidelity $\mathcal{F}_{\text{XEB}} = p$ to $\sqrt{\text{Purity}}$ from SPB on the same scale, and check their dependence $p = p_c^m$ on the number of cycles m . Without systematic control errors, the XEB and SPB results should coincide. Experimentally, we always have control errors which lead us to incorrectly predict $|\psi\rangle$, so control errors give XEB a higher error than SPB. Thus, with a single XEB dataset we can extract the XEB error per-cycle, and the purity loss per-cycle with SPB. By subtracting these, we are left with the control error per-cycle. Thus, with a single experiment we can error budget total error into control error and decoherence error.

These relationships can be seen experimentally in Figure S18. Amazingly, computing the speckle purity can be done with no knowledge of the specific gate sequence performed; as long as the experiment introduces sufficient randomization of the Hilbert Space, Porter-Thomas statistics apply. Practically, SPB allows us to measure the state purity from raw XEB data with exponentially fewer number of pulse sequences as compared to full state tomography. This favorable scaling allows one to extend purity measurements to larger numbers of qubits. It is important to note that an exponential number of measurements are still required to fully characterize the probability distribution for a given sequence, as in tomography, so purity measurements of the full processor are impractical.

5. “Per-layer” parallel XEB

To execute quantum circuits efficiently, it is helpful to run as many gates as possible in parallel. We wish to benchmark our entangling gates operating simultaneously. Resulting fidelities and optimized unitaries may differ from the isolated case, where we benchmark each pair individually, due to imperfections such as control crosstalk and stray qubit-qubit interactions. In the quantum supremacy algorithm, we partition the set of two-qubit gates into four layers, each of which can be ex-

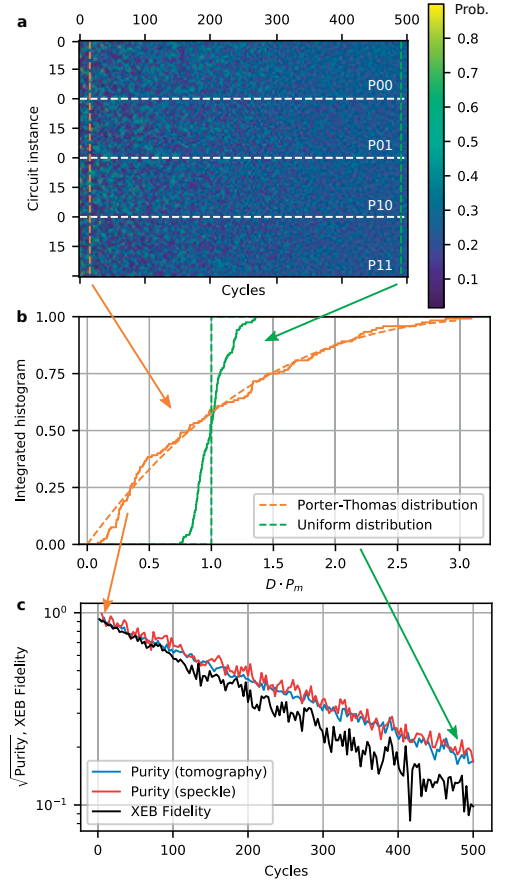


FIG. S18. “Speckle” purity extracted from XEB. **a**, Measured probabilities from XEB for a two-qubit system and 30 random circuits. Raw probabilities show a speckle pattern at low cycles (orange dashed) over circuit instance and probabilities ($|00\rangle, |01\rangle, |10\rangle, |11\rangle$). The speckle contrast decreases with cycles and thus decoherence (green dashed). **b**, Integrated histogram (cumulative distribution) of probabilities. The x-axis is scaled by the dimension $D = 2^2$, so the uniform distribution is a step function at 1.0. At low cycles, the distribution is well-described by Porter-Thomas, and at high cycles, the distribution approaches the uniform distribution. **c**, We can directly relate the variance of the distribution to the average state purity. We fit an exponential to the square root of Purity. We compare this purity-derived number per-cycle=0.00276 to a similar number per-cycle=0.00282 derived from tomographic measure of purity, and see good agreement. The error of XEB, which also includes control errors, is slightly higher at error per-cycle=0.00349.

ecuted in parallel. We then cycle through these layers interleaved with randomly chosen single-qubit gates (see Fig. 3a). However, it is intractable to directly use full-system XEB to benchmark our entangling gates for two reasons: we would simultaneously optimize over the unitary model parameters of every entangling gate, and the classical simulation would be exponentially expensive in system size.

We solve this problem with “per-layer” parallel XEB (see Ref. [51] for a related technique in the context of RB). Instead of alternating among the four layers of entanglers, where each qubit becomes entangled with each of its neighbors, we perform four separate experiments, one for each layer. The experiment sequences are illustrated in Fig. S19a. For each layer, we construct parallel sequences where the layer is repeated with interleaved single-qubit gates; nominally, each qubit only interacts with one other. Following each parallel XEB sequence, we measure all the qubits and extract the equivalent XEB data for each pair. Every two-qubit gate can be characterized in these four experiments, regardless of system size. The optimization and classical simulation are also efficient, as each pair can be analyzed individually.

We present experimental results of “per-layer” parallel XEB in Fig. S19b-c. In Fig. S19b, we compare the performance in the isolated and simultaneous (parallel) experiments. In both cases, the optimized XEB error is close to purity-limited. Simultaneous operation modestly increases the error, by roughly 0.003. This increase is primarily from purity error, which would arise from unintended interactions with other qubits, where coherent errors at the system scale manifest as incoherent errors when we focus on individual pairs. The unitaries we obtain in the simultaneous case differ slightly from the isolated case, which would arise from control crosstalk and unintended interactions. To quantify how these differences affect the gate error, we recalculate the error with the unitaries from the isolated optimization and the data from the simultaneous experiment, which increases the error. We also plot the distributions of the differences in unitary model parameters in Fig. S19c. The dominant change is in Δ_+ , a single-qubit phase.

D. Grid readout calibration

1. Choosing qubit frequencies for readout

The algorithm described in Section VIB4 generally chooses qubit idling frequencies which are far detuned from the resonator to optimize for dephasing. However, these idling frequencies are not optimal for performing readout. To address this problem, we dynamically bias each qubit to a different frequency during the readout phase of the experiment. The qubit frequencies during readout are shown in Fig. S20 (compare to Fig. S13).

To choose the qubit frequencies for readout, we first measure readout fidelity as a function of qubit frequency

and resonator drive frequency at a fixed resonator drive power, in each of the isolated single qubit configurations. This scan captures errors due to both non-optimal detuning between the qubit and resonator, as well as regions with low T_1 values due to TLSs. We then use the data for each qubit and a few constraints to optimize the placement of the qubit frequencies during readout, using the same optimization technique that was described in Section VIB4. We describe two of the important constraints and related error reduction techniques below.

First, because the coupling between qubits relies on a dispersive interaction with the coupler, the coupling would no longer be off when the qubits were detuned by a significant amount from their idling positions. Thus, we impose a constraint that qubits should not be placed near resonance during readout. Nevertheless, we found that for some pairs of qubits, we had to dynamically bias the coupler during readout to avoid any swapping transitions between the qubits during readout. This readout coupler bias is found by sweeping the coupler bias and maximizing the two-qubit readout fidelity.

Second, the pattern of the bare resonator frequencies on the chip as shown in Fig. S20 led to an unexpected problem. Pairs of readout resonators which were coupled to neighboring qubits and were also within a few MHz in frequency space were found to have non-negligible coupling. This coupling was strong enough to mediate swapping of photons from one resonator to the other. The pairs of qubits with similar resonator frequencies were all located in a diagonal chain bisecting the qubit grid, as shown by the red outline in Fig. S20. To mitigate this problem, we arrange the qubit frequencies for these qubits so that the resonator eigenfrequencies are as far apart as possible. The resulting spectral separation is not quite enough to eliminate all deleterious effects, so in addition, we use correlated discrimination on the eight of the qubits in this chain. In other words, we use the results of all eight detector values to determine which one of $2^8 = 256$ states the eight qubits were in. All other qubits in the grid are discriminated as isolated qubits.

2. Single qubit calibration

After placing the qubit frequencies for readout, we calibrate and fine tune the readout parameters for each qubit. For each qubit, we use a $1\ \mu\text{s}$ drive pulse and a $1\ \mu\text{s}$ demodulation window. We summarize the procedure for choosing the remaining parameters as follows:

- Choose the resonator drive frequency to maximize the separation between measurements performed with the qubit in either $|0\rangle$ and $|1\rangle$ [15].
- Choose the resonator drive power to hit a target separation between $|0\rangle$ and $|1\rangle$, so that the error due to this separation is below a 0.3% threshold. We do not choose the readout power to maximize

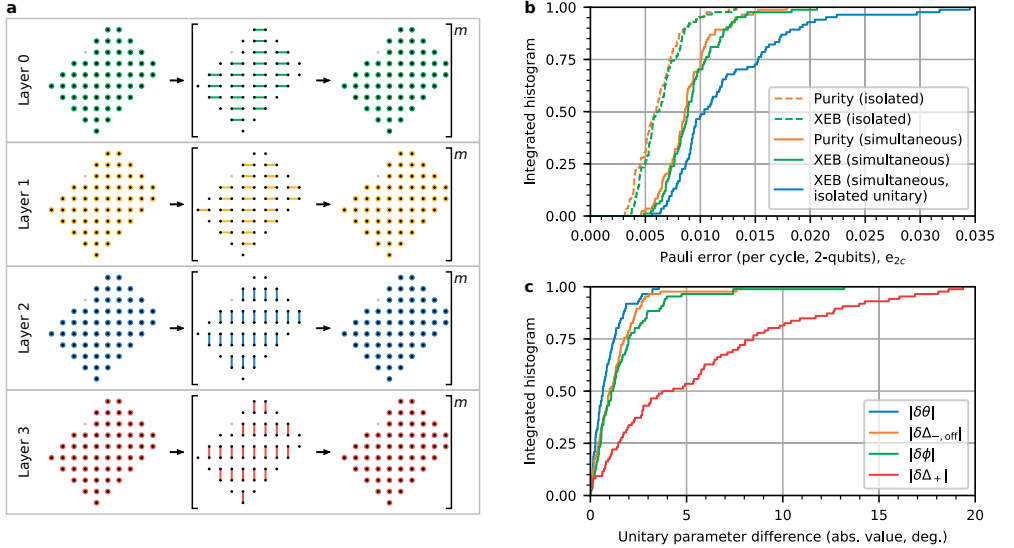


FIG. S19. **Parallel XEB.** **a**, Schematics of four device-wide sequences, one for each entangler layer. Black points are active qubits, colored circles are single-qubit gates, and colored lines are two-qubit gates m times. Compare to Fig. 3a, main text, where the layers are interleaved. **b**, Integrated histograms of Pauli error e_{2c} (see Fig. 2a, main text). These include isolated results, where each entangler is measured in its own experiment, and simultaneous (parallel) results. Purity is “speckle” purity. **c**, Difference, δ , in unitary model parameters (Eq. 43) between the unitaries obtained in the isolated and simultaneous experiments. $\delta\Delta_{-}$ is not plotted because it has a negligible effect on the unitary when $\theta \approx 90$ degrees.

the separation as doing so would saturate our amplifiers, and cause unwanted transitions of the qubit state [16, 21, 52, 53].

- Find the optimal demodulation weight function by measuring the average detector voltage as a function of time during the course of the readout pulse [15, 20].
- Finally, choose the discrimination line between the measurement results for $|0\rangle$ and $|1\rangle$, except as noted in the previous section where we need to apply correlated discrimination.

After completing these calibrations, we check each qubit’s readout fidelity by preparing either $|0\rangle$ or $|1\rangle$ and reading the qubit out. We define the identification error to be the probability that the qubit was not measured in the state we intended to prepare. We achieve 0.97% median identification error for the $|0\rangle$ state, and 4.5% for $|1\rangle$, when each qubit is measured in isolation. The full distribution is shown in dashed lines in Fig. S21a. We conjecture that the error in $|0\rangle$ is due to thermal excitation during preparation or measurement, and that the error in $|1\rangle$ is due to energy relaxation during readout.

3. Characterizing multi-qubit readout

To assess the fidelity of multi-qubit readout, we prepare and measure 150 random classical bitstring states with 53 qubits, with 3000 trials per state. We find that 13.6% of all trials successfully identified the prepared state. We can decompose this overall fidelity in two ways. First, we plot in solid lines in Fig. S21 the errors for each qubit during simultaneous readout, averaged over the 150 random bitstrings. We find that the median errors increase from 0.97% for $|0\rangle$ and 4.5% for $|1\rangle$ in isolation, to 1.8% and 5.1% for simultaneous readout. We do not yet understand the root causes of this increase in error. In addition, we show in Fig. S21 the distribution of errors among the multiqubit results. We see that the most likely error is one lost excitation in the measured state.

E. Summary of system parameters

Table II reports aggregate values for qubit and pair parameters in our processor. A complete table of single-qubit parameter values by qubit is available in supporting online materials, Ref. [54], and illustrated in Figs. S22

TABLE II. Aggregate system parameters

Parameter	Median	Mean	Stdev.	Units	Figure
Qubit maximum frequency	6.924	6.933	0.114	GHz	S22
Qubit idle frequency	6.661	6.660	0.057	GHz	S13
Qubit frequency at readout	5.750	5.766	0.360	GHz	S20
Readout drive frequency	4.618	4.588	0.076	GHz	S20
Qubit anharmonicity	-208.0	-208.0	4.7	MHz	S22
Resonator linewidth $\kappa/2\pi$	0.64	0.69	0.23	MHz	S22
Qubit-resonator coupling $g/2\pi$	72.3	72.1	2.8	MHz	S22
T_1 at Idle Frequency	15.54	16.04	4.00	μs	S22
Readout error $ 0\rangle$ isolated / simultaneous	0.97 / 1.8	1.2 / 2.3	0.8 / 2.1	%	S21
Readout error $ 1\rangle$ isolated / simultaneous	4.5 / 5.1	5.0 / 5.5	1.8 / 2.2	%	S21
1Q RB ^a e_1	0.19	0.22	0.10	%	S23
1Q RB ^a e_1 ($\pi/2$ gateset)	0.15	0.16	0.06	%	S23
1Q RB ^a tomographic e_1 purity	0.14	0.15	0.04	%	S23
1Q XEB e_1 isolated / simultaneous	0.13 / 0.14	0.15 / 0.16	0.05 / 0.05	%	3a (main) S23
1Q XEB e_1 purity isolated / simultaneous	0.11 / 0.11	0.11 / 0.12	0.03 / 0.03	%	S23
2Q XEB e_2 isolated / simultaneous	0.30 / 0.60	0.36 / 0.62	0.17 / 0.24	%	3a (main)
2Q XEB e_{2c} isolated / simultaneous	0.64 / 0.89	0.65 / 0.93	0.20 / 0.26	%	3a (main)
2Q XEB e_{2c} purity isolated / simultaneous	0.59 / 0.86	0.62 / 0.89	0.20 / 0.24	%	S19
Measurement e_m isolated / simultaneous	2.83 / 3.50	3.05 / 3.77	1.09 / 1.61	%	3a (main)

^a RB data taken at a later date

through [S24](#). Single-qubit metrics represent a sample size of 53. Two-qubit metrics represent 86 pairs.

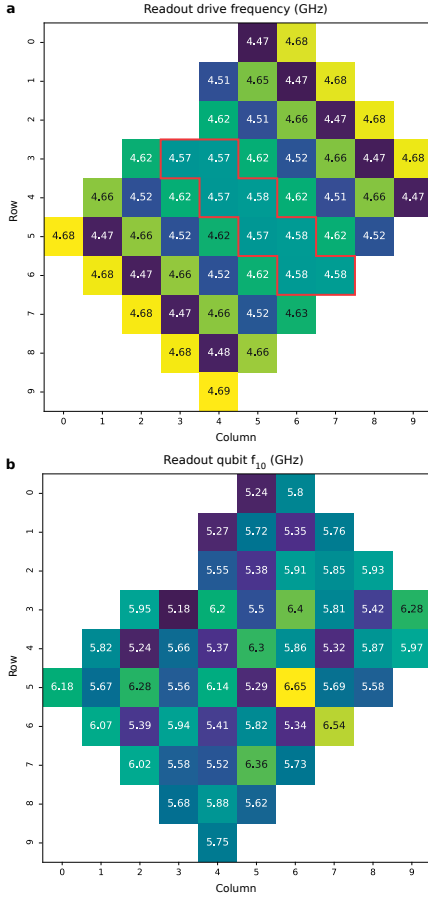


FIG. S20. **a**, Drive frequencies for the readout resonators for each qubit. The red outline shows the area where we had to perform correlated discrimination because of unwanted cross-couplings between the resonators. **b**, Qubit frequencies during readout, found using a frequency optimization procedure.

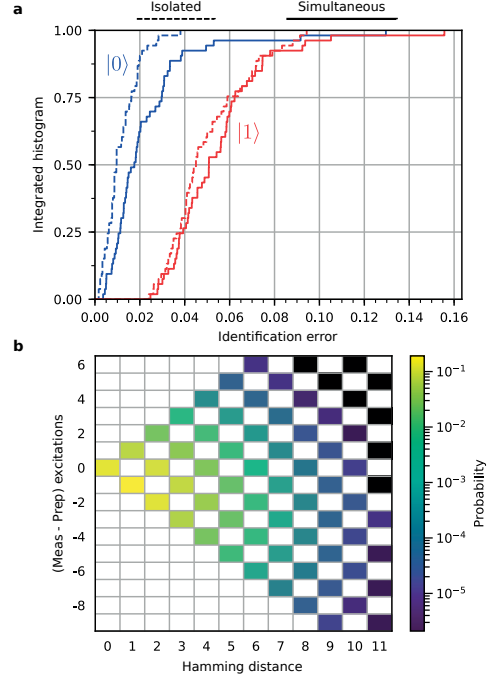


FIG. S21. **Readout errors.** **a**, Histogram of readout errors for each qubit when prepared in $|0\rangle$ or $|1\rangle$, and readout in isolation or simultaneously. **b**, Distribution of errors in multi-qubit readout. The x-axis Hamming distance is the number of bits that are different between measured and prepared states, while the y-axis is the difference in the number of 1's in the states. For example, if we prepare $|011\rangle$ and measure $|101\rangle$, the Hamming distance is 2 and the difference in the number of excitations is 0.

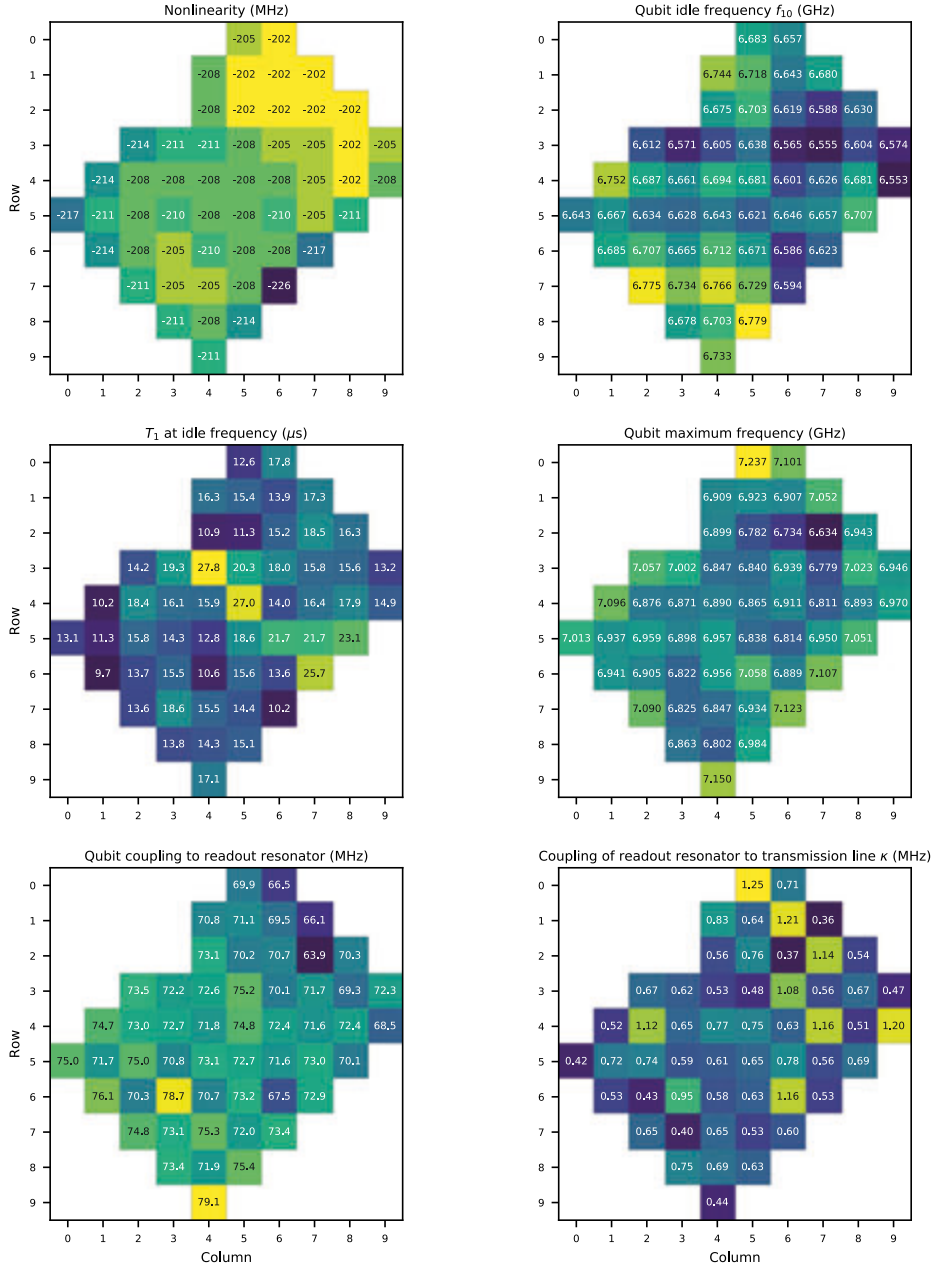


FIG. S22. Typical distribution of single-qubit parameters over the Sycamore processor.

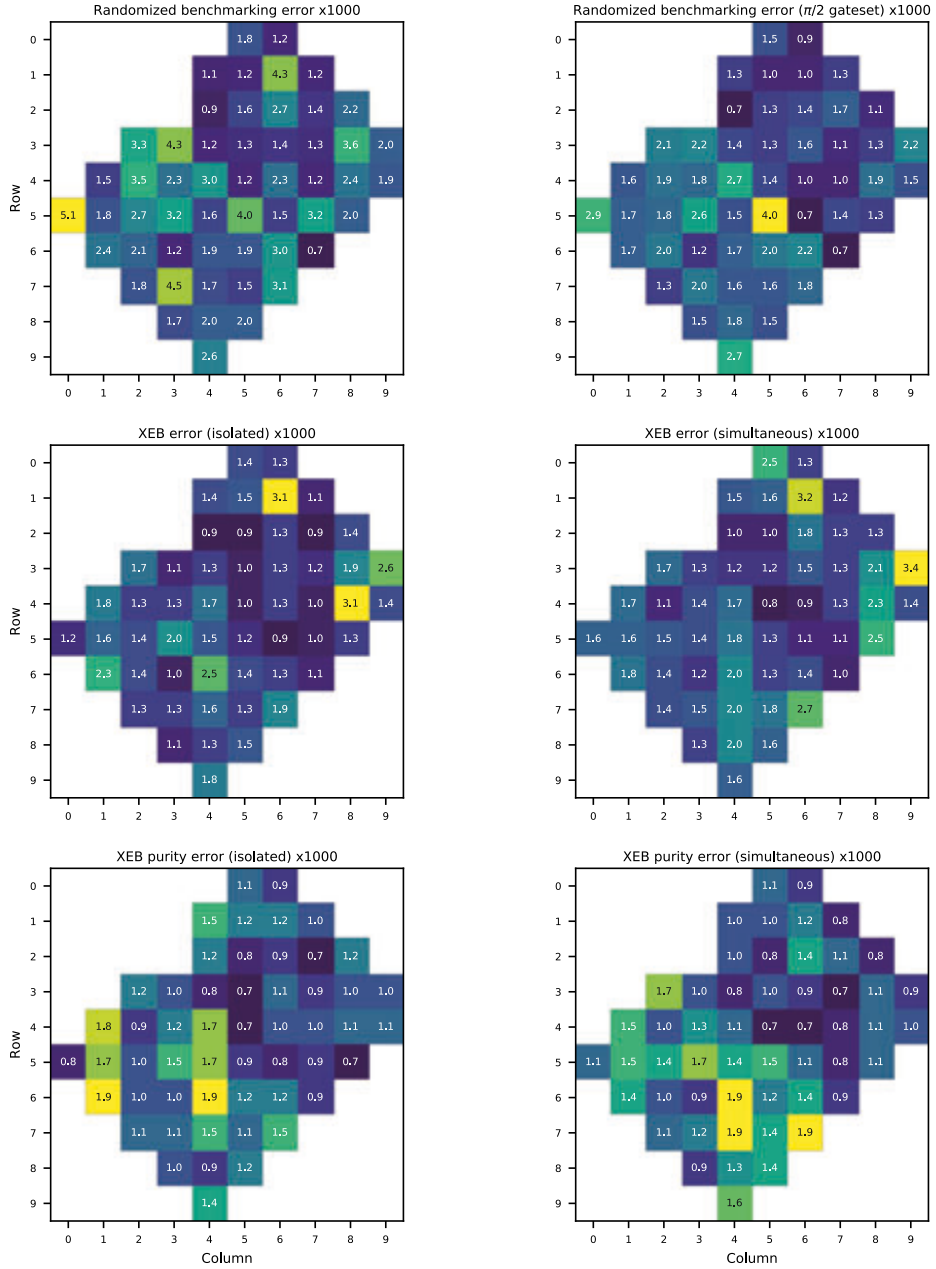


FIG. S23. Typical distribution of single-qubit gate benchmarking errors over the Sycamore processor, for both isolated and simultaneous operation.

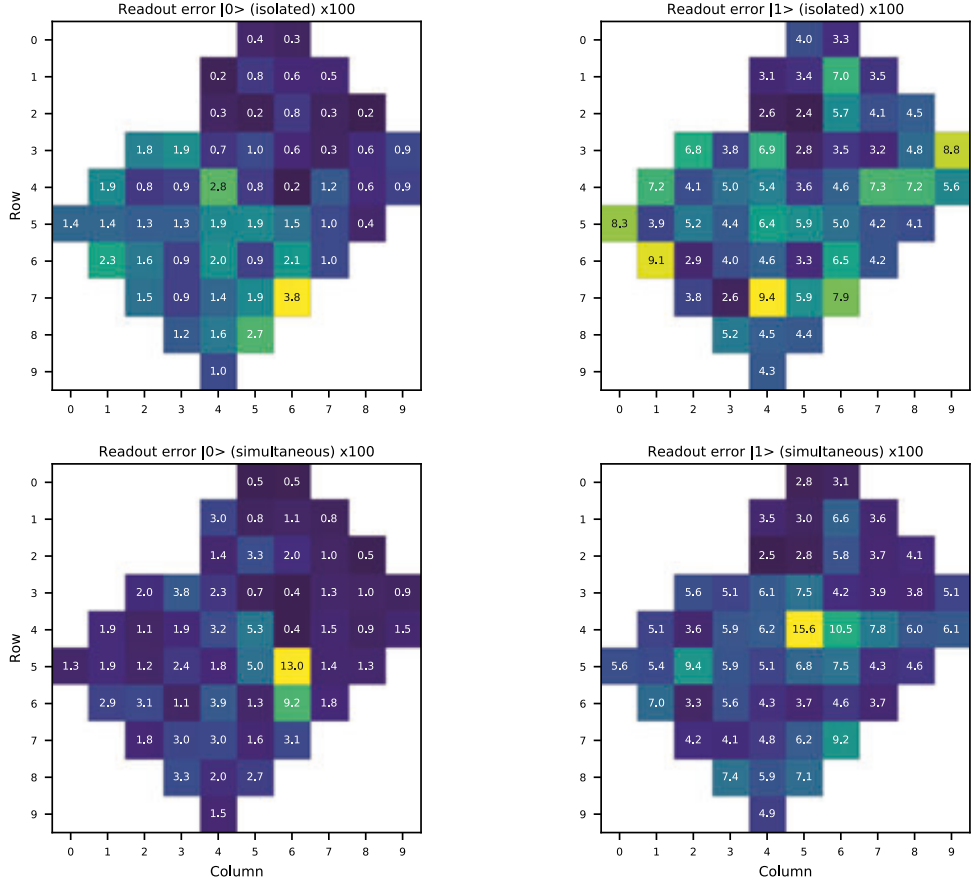


FIG. S24. Typical distribution of readout errors over the Sycamore processor, for both isolated and simultaneous operation.

VII. QUANTUM CIRCUITS

A. Background

We sample the output of random quantum circuits (RQCs) with two use cases in mind: performing a computational task beyond the reach of state-of-the-art supercomputers (quantum supremacy); and estimating the experimental fidelity (performance evaluation).

In order for the RQCs to cover both use cases, we define a circuit family with a varying number of qubits n and cycles m . Our quantum supremacy demonstration uses RQCs with a large number of qubits $n = 53$ and high depth $m = 20$. Large number of qubits hinders wave function (Schrödinger) simulation and high depth impedes tensor network (Feynman) simulation (see Sec. XB). We find that the most competitive classical simulator for our hardest RQCs is the Schrödinger-Feynman algorithm (SFA, see Sec. XA) which copes well with high depth circuits on many qubits.

SFA takes as input an n -qubit quantum circuit and a *cut* which divides $n = n_1 + n_2$ qubits into two contiguous partitions with n_1 and n_2 qubits. The algorithm computes the output state as the sum over simulation paths formed as the product of the terms of the Schmidt decomposition of all cross-partition gates. By the distributive law there are r^g such simulation paths for a circuit with g cross-partition gates of Schmidt rank r . Consequently, the algorithm achieves runtime proportional to $(2^{n_1} + 2^{n_2})r^g$. Circuit cuts with n_1 , n_2 and g that make the simulation task tractable are called *promising cuts*. The most promising cut for our largest RQCs runs parallel to the shorter axis of the device starting in the vicinity of the broken qubit. The sum over the simulation paths can be interpreted as tensor contraction. In this view, the r^g factor can be thought of as the *bond dimension* associated with the circuit partitioning, i.e. the cardinality of the index set ranged over in the contraction corresponding to all cross-partition gates. SFA is described in more detail in [37] and section X.

B. Overview and technical requirements

The two use cases for our RQCs give rise to a tension in technical requirements at the heart of quantum supremacy. On the one hand, supremacy RQC sampling should by definition be prohibitively hard to simulate classically. On the other hand, performance evaluation entails classical simulation of the RQCs. To resolve the conflict, we note that the fidelity of a RQC experiment depends primarily on the number and quality of the gates. By contrast, the simulation cost is highly sensitive to minor perturbations in the circuit. Consequently, experiment fidelity for RQCs that cannot be simulated directly may be approximated from the experiment fidelity of similar RQCs obtained as the result of transformations that reduce simulation cost without significantly

affecting experiment fidelity (see Section VIIG).

Performance evaluation using XEB provides another design consideration. The procedure requires knowledge of the cross-entropy of the theoretical output distribution of the circuit. An analytical expression for this quantity has been derived in [26] for circuits whose measurement probabilities approach the Porter-Thomas distribution. We find that our RQCs satisfy this assumption when the circuit depth is larger than 12, see Fig. S35a. Note that high circuit depth also increases the cost of classical simulation.

C. Circuit structure

A RQC with n qubits generally utilizes qubits 1 through n in the qubit order shown in Fig. S27 with small deviations from this default qubit ordering in some circuits. The qubit order has been chosen to ensure that for most RQCs with fewer than 51 qubits, there is a partitioning of the qubits into two similarly sized blocks connected by only five couplers. The next larger RQC, with 51 qubits, has seven couplers along the most promising circuit cut. Since the cost of SFA grows exponentially in the number of gates across the partitions our circuit geometry leads to a steep increase in the simulation cost of 51-qubit RQCs relative to the circuits with fewer qubits. This creates a sizeable gap in the computational hardness between most of our evaluation circuits and the quantum supremacy circuits ($n = 53$).

In the time dimension, each RQC is a series of m full cycles and one half cycle followed by measurement of all qubits. Every full cycle consists of two steps. In the first step, a single-qubit gate is applied to every qubit. In the second step, two-qubit gates are applied to pairs of qubits. Different qubit pairs are allowed to interact in different cycles. Specifically, in the supremacy RQCs we loop through the direct neighbors of every qubit over the eight-cycle sequence ABCDCDAB and in the evaluation RQCs we use the four-cycle sequence EFGH where A, B, ..., H are coupler activation patterns shown in Fig. S25. The sequence is repeated in subsequent cycles. The cost of SFA simulation is highly sensitive to the specific sequence employed in a circuit, see VIIG2. Border qubits have fewer than four neighbors and no gate is applied to them in some cycles. The half cycle preceding the measurement consists of the single-qubit gates only. The overall structure of our RQCs is shown in Fig. 3 of the main paper.

D. Randomness

Single-qubit gates in every cycle are chosen randomly using a pseudo-random number generator (PRNG). The generator is initialized with a seed s which is the third parameter for our family of RQCs. The single-qubit gate applied to a particular qubit in a given cycle depends only

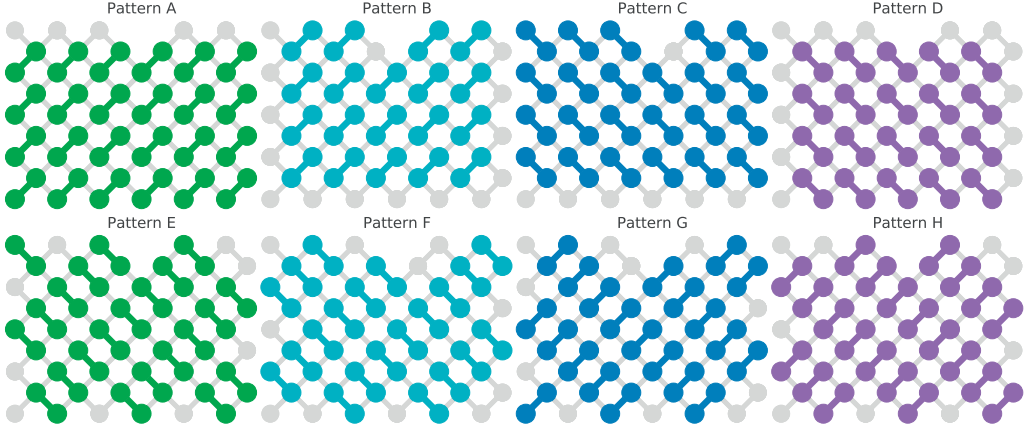


FIG. S25. **Coupler activation patterns.** Coupler activation pattern determines which qubits are allowed to interact simultaneously in a cycle. Quantum supremacy RQCs utilize the staggered patterns shown in the top row in the sequence ABCDCDAB, repeated in subsequent cycles. Performance evaluation RQCs employ the patterns shown in the bottom row in the sequence EFGH, likewise repeated in subsequent cycles. The former sequence makes SFA simulation harder by facilitating prompt transfer of entanglement created at promising circuit cuts into the bulk of each circuit partition.

on s . Consequently, two RQCs with the same s apply the same single-qubit gate to a given qubit in a given cycle as long as the qubit and the cycle belong in both RQCs as determined by their size n and depth m parameters.

Conversely, the choice of single-qubit gates is the sole property of our RQCs that depends on s . In particular, the same two-qubit gate is applied to a given qubit pair in a given cycle by all RQCs that contain the pair and the cycle.

E. Quantum gates

In our experiment, we configure three single-qubit gates. Each one is a $\pi/2$ -rotation around an axis lying on the equator of the Bloch sphere. Up to global phase, the gates are

$$X^{1/2} \equiv R_X(\pi/2) = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -i \\ -i & 1 \end{bmatrix}, \quad (50)$$

$$Y^{1/2} \equiv R_Y(\pi/2) = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, \quad (51)$$

$$W^{1/2} \equiv R_{X+Y}(\pi/2) = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -\sqrt{i} \\ \sqrt{-i} & 1 \end{bmatrix} \quad (52)$$

where $W = (X + Y)/\sqrt{2}$ and $\sqrt{\pm i}$ denotes the principal value of the square root. The first two belong to the single-qubit Clifford group, while $W^{1/2}$ is a non-Clifford gate. Single-qubit gates in the first cycle are chosen independently and uniformly at random from the set of the

three gates above. In subsequent cycles, each single-qubit gate is chosen independently and uniformly at random from among the gates above except the gate applied to the qubit in the preceding cycle. This prevents simplifications of some simulation paths in SFA. Consequently, there are $3^{n/2} 2^{nm}$ possible random choices for a RQC with n qubits and m cycles.

Two-qubit gates in our RQCs are not randomized, but are determined by qubit pair and cycle number. The gates preserve the number of ground and excited states of the qubits which gives their matrices block diagonal structure with 1×1 , 2×2 and 1×1 blocks. Therefore, up to global phase they belong to $U(1) \oplus U(2) \oplus U(1)/U(1)$ and thus can be described by five real parameters (see Fig. S16, and Eq. 43). Each gate in this family can be decomposed into four Z-rotations described by three free parameters and the two-parameter fermionic simulation gate

$$\text{fSim}(\theta, \phi) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\theta) & -i \sin(\theta) & 0 \\ 0 & -i \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 0 & e^{-i\phi} \end{bmatrix} \quad (53)$$

which is the product of a fractional iSWAP and controlled phase gate (see Fig. S14b).

In our experiment, we tune up the two-qubit gates close to $\theta \approx \pi/2$ and $\phi \approx \pi/6$ radians and then infer more accurate values of all five parameters for each qubit pair using XEB. Consequently, all five parameters of the two-qubit gate depend on the qubit pair. While inferred unitaries are suitable for RQC sampling, future applications of the Sycamore processor, for example, in

quantum chemistry, will require precise targeting of the entangling parameters [48, 55]. The three parameters which control the Z-rotations implicit in the two-qubit gates can be canceled out with active Z-rotations turning an arbitrary five-parameter gate into pure fSim(θ, ϕ). In our RQCs, we have decided not to apply such correction gates. This choice affords us greater number of interactions within the available circuit depth budget and introduces additional implicit non-Clifford single-qubit gates into the RQCs.

The Z-rotations have two origins. First, they capture the phase shifts due to qubit frequency excursions during the two-qubit gate. Second, they account for phase changes due to different idle frequencies of the interacting qubits. The latter introduces dependency of the three parameters defining the Z-rotations on the time at which the gate is applied. By contrast, for a given qubit pair θ and ϕ do not depend on the cycle.

The fSim($\pi/2, \pi/6$) gate is the product of a non-Clifford controlled phase gate and an iSWAP which is a two-qubit Clifford gate.

F. Programmability and universality

Programmability of Sycamore rests on our ability to tune up a variety of gate sets including sets that are universal for quantum computation. For example, the set of gates employed in our quantum supremacy demonstration is universal, as we show in this section.

The proof consists of two parts. First, we show that the CZ gate can be obtained as a composition of two fSim gates and single-qubit rotations. Second, we outline how the well-known proof that the H and T gates are universal for SU(2) [56] can be adapted for $X^{1/2}$ and $W^{1/2}$. The conclusion follows from the fact that the gate set consisting of the CZ gate and SU(2) is universal [57].

1. Decomposition of CZ into fSim gates

Here, we show how to decompose a controlled-phase gate into two fSim gates and several single-qubit gates. The fSim gate is native to our hardware and can be decomposed into

$$\text{fSim}(\theta, \phi) = e^{-i\theta(X \otimes X + Y \otimes Y)/2} e^{-i\phi(I - Z) \otimes (I - Z)/4}, \quad (54)$$

where the iSWAP angle $\theta \simeq \pi/2$ and the controlled-phase angle $\phi \simeq \pi/6$. The controlled-phase part can be further decomposed into

$$\begin{aligned} & e^{-i\phi(I - Z) \otimes (I - Z)/4} \\ &= e^{-i\phi/4} e^{i\phi(Z \otimes I + I \otimes Z)/4} e^{-i\phi Z \otimes Z/4}. \end{aligned} \quad (55)$$

To simplify notations, we introduce the two-qubit gate

$$\begin{aligned} \Upsilon(\theta, \phi) &= e^{-i\theta(X \otimes X + Y \otimes Y)/2} e^{-i\phi Z \otimes Z/4} \\ &= e^{i\phi/4} e^{-i\phi(Z \otimes I + I \otimes Z)/4} \text{fSim}(\theta, \phi), \end{aligned} \quad (56)$$

which is equivalent to the fSim gate up to single-qubit Z rotations. The sign of θ in $\Upsilon(\theta, \phi)$ can be changed by the single-qubit transformation,

$$Z_1 \Upsilon(\theta, \phi) Z_1 = \Upsilon(-\theta, \phi), \quad (57)$$

where $Z_1 = Z \otimes I$ ($Z_2 = I \otimes Z$ works equally well).

Multiplying two Υ gates with opposite values of θ on both sides the operator $X_1 = X \otimes I$, we have

$$\begin{aligned} \Upsilon(-\theta, \phi) X_1 \Upsilon(\theta, \phi) &= e^{i\theta Y \otimes Y/2} X_1 e^{-i\theta Y \otimes Y/2} \\ &= \cos \theta X_1 + \sin \theta Z \otimes Y. \end{aligned} \quad (58)$$

With the identity (58), we have

$$\begin{aligned} \Upsilon(-\theta, \phi) e^{i\alpha X_1} \Upsilon(\theta, \phi) &= \cos \alpha \left(\cos \frac{\phi}{2} I \otimes I - i \sin \frac{\phi}{2} Z \otimes Z \right) + i \sin \alpha \left(\cos \theta X \otimes I + \sin \theta Z \otimes Y \right) \\ &= \left(\cos \alpha \cos \frac{\phi}{2} I + i \sin \alpha \cos \theta X \right) \otimes I - i Z \otimes \left(\cos \alpha \sin \frac{\phi}{2} Z - \sin \alpha \sin \theta Y \right), \end{aligned} \quad (59)$$

where $0 \leq \alpha \leq \pi/2$ is to be determined. We introduce the Schmidt operators

$$\Gamma_1(\alpha) = \cos \alpha \cos(\phi/2) I + i \sin \alpha \cos \theta X, \quad (60)$$

$$\Gamma_2(\alpha) = \cos \alpha \sin(\phi/2) Z - \sin \alpha \sin \theta Y, \quad (61)$$

and the unitary (59) takes the simple form

$$\Upsilon(-\theta, \phi) e^{i\alpha X_1} \Upsilon(\theta, \phi) = \Gamma_1 \otimes I - i Z \otimes \Gamma_2. \quad (62)$$

The Schmidt rank of this unitary is two. Therefore, it is equivalent to a controlled-phase gate (also with Schmidt

rank two) up to some single-qubit unitaries. The two non-zero Schmidt coefficients of the unitary (59) are equal to the operator norms of $\Gamma_{1,2}$.

The target controlled-phase gate that we want to decompose into the fSim gate is

$$\text{diag}(1, 1, 1, e^{-i\delta}) = e^{-i\delta(I - Z) \otimes (I - Z)/4}, \quad (63)$$

where $0 \leq \delta \leq 2\pi$. It has two non-zero Schmidt coefficients $\cos(\delta/4)$ and $\sin(\delta/4)$. For example, we set the operator norm of Γ_2 to be equal to the second Schmidt

coefficient of the target unitary

$$|\Gamma_2(\alpha)| = \sqrt{(\cos \alpha \sin(\phi/2))^2 + (\sin \alpha \sin \theta)^2} = \sin(\delta/4), \quad (64)$$

and the parameter α can be determined

$$\sin \alpha = \sqrt{\frac{\sin(\delta/4)^2 - \sin(\phi/2)^2}{\sin(\theta)^2 - \sin(\phi/2)^2}}. \quad (65)$$

This equation has a solution if and only if one of the following two conditions is satisfied

$$|\sin \theta| \leq \sin(\delta/4) \leq |\sin(\phi/2)|, \quad (66)$$

$$|\sin(\phi/2)| \leq \sin(\delta/4) \leq |\sin \theta|. \quad (67)$$

A large set of controlled-phase gates can be implemented with the typical values of θ and ϕ of the fSim gate, except for those that are very close to the identity.

To fix the local basis of the first qubit in Eq. (59), we introduce two X rotations of the same angle

$$e^{-i\xi X/2} \Gamma_1(\alpha) e^{-i\xi X/2} = \cos(\delta/4) I, \quad (68)$$

$$e^{-i\xi X/2} Z e^{-i\xi X/2} = Z, \quad (69)$$

where the angle ξ is

$$\xi = \arctan\left(\frac{\tan \alpha \cos \theta}{\cos(\phi/2)}\right) + \frac{\pi}{2} \left(1 - \text{sgn}(\cos(\phi/2))\right). \quad (70)$$

To fix the local basis of the second qubit in Eq. (59), we introduce two X rotations of opposite angles

$$e^{i\eta X/2} \Gamma_2(\alpha) e^{-i\eta X/2} = \sin(\delta/4) Z, \quad (71)$$

where the angle η is

$$\eta = \arctan\left(\frac{\tan \alpha \sin \theta}{\sin(\phi/2)}\right) + \frac{\pi}{2} \left(1 - \text{sgn}(\sin(\phi/2))\right). \quad (72)$$

Applying these local X rotations before and after the gate sequence in Eq. (59), we have

$$e^{-i(\xi X_1 - \eta X_2)/2} \Upsilon(-\theta, \phi) e^{i\alpha X_1} \Upsilon(\theta, \phi) e^{-i(\xi X_1 + \eta X_2)/2} = \cos(\delta/4) I \otimes I - i \sin(\delta/4) Z \otimes Z, \quad (73)$$

which is the desired controlled-phase gate up to some single-qubit Z rotations.

The target controlled-phase gate equals to the CZ gate for $\delta = \pi$. We numerically checked that the decomposition (73) yields the CZ gate for all 86 fSim gates (with different values of θ and ϕ) in our device.

2. Universality for SU(2)

Here, we show how the argument for the well-known result that the H and T gates are universal for SU(2) [56] can be adapted for the $X^{1/2}$ and $W^{1/2}$ gates. At the core of the argument lies the observation that $T \equiv R_Z(\pi/4)$ followed by $HTH \equiv R_X(\pi/4)$ is a single-qubit rotation by angle α which is an irrational multiple of π . Specifically, α is such that

$$\cos \frac{\alpha}{2} = \cos^2 \frac{\pi}{8} = \frac{1}{2} \left(1 + \frac{1}{\sqrt{2}}\right). \quad (74)$$

By Theorem B.1 in Appendix B of [56], α/π is irrational because the monic minimal polynomial with rational coefficients of $e^{i\alpha}$

$$x^4 + x^3 + \frac{1}{4}x^2 + x + 1 \quad (75)$$

is not cyclotomic (since not all its coefficients are integers).

Similarly, $W^{1/2} \equiv R_{X+Y}(\pi/2)$ followed by $X^{1/2} \equiv R_X(\pi/2)$ is a single-qubit rotation by angle β such that

$$\cos \frac{\beta}{2} = \cos^2 \frac{\pi}{4} - \frac{1}{\sqrt{2}} \sin^2 \frac{\pi}{4} = \frac{1}{2} \left(1 - \frac{1}{\sqrt{2}}\right). \quad (76)$$

The monic minimal polynomial with rational coefficients of $e^{i\beta}$ is (75), the same as that of $e^{i\alpha}$. Therefore, β is also an irrational multiple of π . The rest of the universality argument for H and T also applies in the case of $X^{1/2}$ and $W^{1/2}$.

G. Circuit variants

Since XEB entails classical simulation, it is hard or impossible to use it to estimate experimental fidelity of circuits which are hard or impossible to simulate classically. As described above, we designed our RQCs to ensure that an effective partitioning for SFA exists for circuits with fewer than 51 qubits. This gives rise to a significant gap in the cost of classical simulation between quantum supremacy circuits and most of our performance evaluation circuits. This gap facilitates performance evaluation of the Sycamore processor near the quantum supremacy frontier. In practice, however, we would like greater control over the simulation hardness, for two reasons. First, performance evaluation is still very costly for large n approaching the supremacy frontier. Second, we would like to be able to estimate the fidelity of supremacy RQCs more directly, even though classical simulation of this case is unfeasible by definition.

In order to achieve more fine-grained control over the cost of classical simulation of our RQCs, we exploit the fact that the experimental fidelity depends primarily on

Circuit variant	Gates elided	Sequence of patterns
non-simplifiable full	none	ABCD CDAB
non-simplifiable elided	some	ABCD CDAB
non-simplifiable patch	all	ABCD CDAB
simplifiable full	none	EFGH
simplifiable elided	some	EFGH
simplifiable patch	all	EFGH

TABLE III. **Circuit variants.** Six variants of RQCs employed in quantum supremacy demonstration (non-simplifiable full) and performance evaluation (remaining five variants) classified by transformations applied in order to control the cost of classical simulation. The eight coupler activation patterns A, B, ..., H are shown in Fig. S25.

the number and quality of the gates while the simulation cost is highly sensitive to the structure of the quantum circuit. Therefore, we approximate the experimental fidelity of RQCs which are hard or impossible to simulate from the fidelity of similar RQCs obtained as the result of transformations that reduce simulation cost without significantly affecting experimental fidelity.

We employ two such transformations. Each decreases simulation cost by reducing the bond dimension of promising circuit cuts. The first one removes some or all cross-partition gates. We say that the removed gates have been *elided* and term the transformation *gate elision*. The second transformation changes the sequence of coupler activation patterns shown in Fig. S25 to enable the formation of *wedges* which reduce the bond dimension by slowing the spread of entanglement generated at the circuit cut.

The two transformations complete the description of RQCs used in our experiment. Consequently, each RQC is uniquely determined by five parameters: number of qubits n , number of cycles m , PRNG seed s , number of elided gates and the sequence of coupler activation patterns.

1. Gate elision

The most straightforward way to reduce the cost of classical simulation of a RQC is to remove a number of cross-partition gates across the most promising circuit cut. In order to enable independent propagation by the SFA of the wave function of each circuit partition for the first few cycles, the gates are elided beginning with the initial cycle. Each elided gate reduces the bond dimension of the partitioning by a factor of two or four, see Section X.

We refer to RQCs with a small number of elided gates as *elided circuits*. A particularly dramatic speedup is possible when all two-qubit gates across the partitions are elided leading to two disconnected circuits running in parallel. We refer to such disconnected RQCs as *patch circuits*. Base RQCs in which no gates have been elided are referred to as *full circuits*.

If the error probability of the elided two-qubit gate is similar to the error probability of the two-qubit identity gate which it is replaced with, the circuit resulting from gate elision exhibits fidelity that is similar to the fidelity of the original circuit. This assumption holds when the two-qubit gate errors are dominated by the same decoherence processes that govern the single-qubit gate errors such as finite T_1 and T_2 . Indeed, for circuit sizes where XEB on full circuits is possible, we have observed good agreement between fidelity estimates produced for patch, elided and full circuits. For harder circuits, we have observed good agreement between fidelity estimates for patch and elided circuits. See Section VIII for detailed discussion of these results.

2. Wedge formation

The most competitive algorithm for our hardest circuits, SFA (see Sec. X A) scales proportionally to the bond dimension of the circuit partitioning which is equal to the product of Schmidt rank of all cross-partition gates (see Sec. X D). The Schmidt decomposition of most two-qubit gates in our RQCs consists of four terms (a few gates can be replaced with simpler gates with Schmidt rank of two, see Section X). Therefore most cross-partition gates contribute a factor of four to the bond dimension of the partitioning. However, when two consecutive cross-partition gates share a qubit forming a *wedge* as shown in Fig. S26, the Schmidt decomposition of the resulting three-qubit unitary also has only four terms. In other words, the second cross-partition gate does not generally produce substantial new entanglement (as quantified by the Schmidt rank) among the partitions in excess of the entanglement produced by the first gate. Consequently, every wedge reduces the bond dimension of the partitioning by a factor of four.

The eight-cycle sequence ABCD CDAB and the four constituent coupler activation patterns A, B, C and D shown in Fig. S25 have been designed to prevent formation of wedges across promising circuit cuts. In other words, the sequence ensures that entanglement created in a given cycle by cross-partition gates is transferred into the bulk of each partition in the following cycle.

On the other hand, the four-cycle sequence EFGH enables formation of wedges and thus efficient simulation of RQCs using SFA. We employ the latter sequence in most evaluation circuits and use the former eight-cycle sequence for the quantum supremacy circuits and largest evaluation circuits, see Table III.

VIII. LARGE SCALE XEB RESULTS

In Section VI, we have detailed the device calibration processes used for individual components such as qubits, couplers, and coupled pairs of qubits. We have also introduced cross-entropy benchmarking (XEB) as a method

Circuit variant	n	m	Single-qubit gates	All two-qubit gates	Cross-partition two-qubit gates
non-simplifiable full	53	20	1113	430	35
non-simplifiable elided	53	20	1113	408	13
non-simplifiable patch	53	20	1113	395	0
simplifiable full	38	14	570	210	18
simplifiable elided	38	14	570	204	12
simplifiable patch	38	14	570	192	0

TABLE IV. **Gate counts.** Number of gates in selected random quantum circuits employed for quantum supremacy demonstration and performance evaluation of the Sycamore processor.

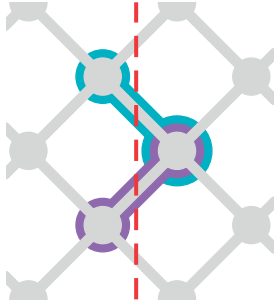


FIG. S26. **Cross-partition wedge.** Two consecutive cross-partition gates which share a qubit form a *wedge*, as illustrated here with gates highlighted in turquoise and magenta. Schmidt rank of a single two-qubit gate is at most four. Therefore, generally wedges are not efficient at increasing entanglement across partitions and can be simulated efficiently by the SFA.

that allows us to evaluate the performance of a quantum system. In this section, we describe how we use a few circuit variations to benchmark our Sycamore processor at a larger scale. In particular, we present a modular version of XEB with “patch circuits” that does not require exponential classical computation resources for estimating XEB fidelities \mathcal{F}_{XEB} of larger systems. We also describe the effect of choice of unitary model on large-scale \mathcal{F}_{XEB} , as well as how we use patch circuits to monitor the stability of the full system.

A. Limitations of full circuits

We first discuss what we refer to as “full circuits”, where for a given set of qubits, all possible two-qubit gates participate in the circuit. With full circuits, we benchmarked the system as a function of size, where as discussed below the classical resources and techniques used to compute the \mathcal{F}_{XEB} is a function of the number of qubits. The order in which each qubit was added is labeled in Fig. S27. The rationale behind this ordering is explained in Section VII. At each system size, we executed 10 randomly generated circuit instances on the

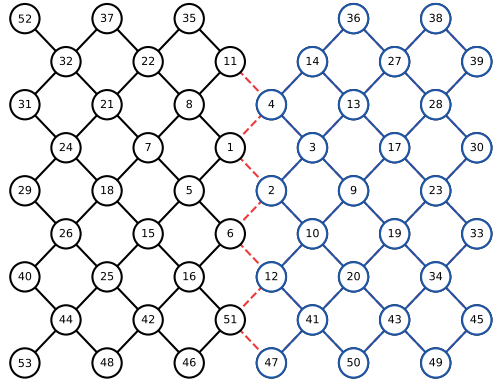


FIG. S27. **Qubit ordering for large-scale XEB experiments.** Illustration of the order in which qubits are added for large-scale experiments. The partition between left (black) and right (blue) qubits along the boundary (dashed red lines) is used in patch and elided circuits, as explained below.

processor and sampled output bitstrings 500k times for each circuit (unless otherwise specified). To minimize potential instance-to-instance fluctuations, we chose the gate sequences in a persistent, “stable” manner: using a known seed for a random number generator, for each circuit, each time a new qubit is added, we maintain the same gateset for all the “existing” qubits and new gates are only introduced to qubits and pairs associated with the added qubit (see Section VII for details).

Once a sufficient number of bitstrings are collected, \mathcal{F}_{XEB} can be calculated for each system size, following the method described in Section IV. As the system size increases, the computational complexity of XEB analysis grows exponentially, which can be qualitatively divided into three regimes. For system size from 12 to 37 qubits, XEB analysis was carried out by evolving the full quantum state (Schrödinger method) on a high-performance server (88 hyper-threads, 1.5TB memory in our case) using the “qsim” program. At 38 qubits we used a n1-ultramem-160 VM in Google’s cloud (160 hyperthreads, 3.8TB memory). Above 38 qubits, Google’s large-scale cluster computing became necessary, and in addition a

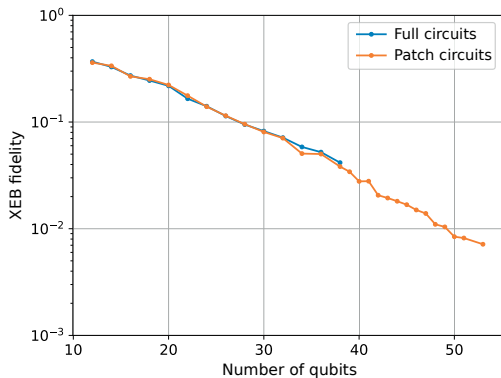


FIG. S28. **Comparison between XEB with patch circuits and full circuits.** Full vs. patch circuit benchmarking up to 38 qubits with 14 cycles, showing close agreement to within the intrinsic fluctuations of the system. We plot the results for patch circuits out to 53 qubits.

hybrid Schrödinger-Feynman approach, the “qsimh” program, was used to improve the efficiency: in this case, we break the system up into two patches, where each patch can be efficiently computed via the Schrödinger method and then connected by a Feynman path-integral approach (see Section X for more details). Finally we used a Schrödinger algorithm in the Jülich supercomputer for some circuits up to 43 qubits.

In order to reduce the computational cost, we introduce two modified circuit types in the following sections. By using slightly simplified gate sequences, these two methods can provide good approximate predictions of system performance all the way out to the “quantum supremacy” regime.

B. Patch circuits: a quick performance indicator for large systems

The simplest approach to large-scale performance estimation is referred to as “patch circuits,” which predicts the performance of the full system by multiplying together the fidelities of non-interacting subsystems, or “patches”. In this work, we use two such subsystems, where each patch is roughly half the size of the full system. The two subsystems are run simultaneously, so that effects such as gate and measurement crosstalk between patches are included, but the two patches are analyzed separately when computing the fidelity. The two patches are defined by the gates removed along their boundary, as illustrated in Fig. S27. For sufficiently large systems, these removed two-qubit gates represent a small portion of the whole circuit. As a consequence, \mathcal{F}_{XEB} of the full system can be estimated as the product of the fidelities

of the two subsystems; compared with full circuits, the main missing factor is the absence of entanglement between the two patches.

We evaluate the efficacy of using patch circuits by comparing it against full circuits with the same set of qubits. The experimental results can be seen in Fig. 4a (main text), where we show fidelities measured by these two methods for systems from 12 qubits to 53 qubits, in an interleaved fashion. We re-plot this data here in Fig. S28 as well. As expected, the fidelities obtained via patch XEB show a consistent exponential decay (up to fluctuations arising from qubit-dependent gate fidelities and a small amount of system fluctuations) as a function of system size. For every system size investigated, we found that patch and full XEB provide fidelities that are in good agreement with each other, with a typical deviation of $\sim 5\%$ of the fidelity itself (we attribute the worst-case disagreement of 10% at 34 qubits due to a temporary system fluctuation in between the two datasets, which was also seen in interleaved measurement fidelity data). Theoretically, one would expect patch circuits to result in $\sim 10\%$ higher fidelity than full circuits due to the slightly reduced gate count. We find that patch circuits perform slightly worse than expected, which we believe is due to the fact that the two-qubit gate unitaries are optimized for full operation and not patch operation. In any case, agreement between patch and full circuits shows that patch circuits can be a good estimator for full circuits, which is quite remarkable given the drastic difference in entanglement generated by the two methods. These results give us a good preview of the system performance in all three regimes discussed earlier.

The advantage of using patch circuits lies in its exponentially reduced computational cost, as it only requires calculating \mathcal{F}_{XEB} of subsystems at half the full size (or less if a larger number of smaller patches is used). This allows for quick estimates of large-scale system performance on a day-to-day basis, including for system and circuit sizes in the “quantum supremacy” regime. As a consequence, we typically use patch circuits as a quick system performance indicator, which we use for rapid turnarounds between system calibration and performance evaluation, as well as for monitoring full system stability (see Section VIII H). We also note that patch circuits can be used well beyond 50 qubits, and in fact can be extended to arbitrary numbers of qubits while keeping the analysis time at most linear in the number of qubits (or even constant if the patches can be analyzed in parallel), assuming that the patch size stays roughly constant and more non-interacting patches are added as the number of qubits grows.

C. Elided circuits: a more rigorous performance estimator for large systems

For a more rigorous prediction of full \mathcal{F}_{XEB} , we introduce a more sophisticated approach referred to as “elided

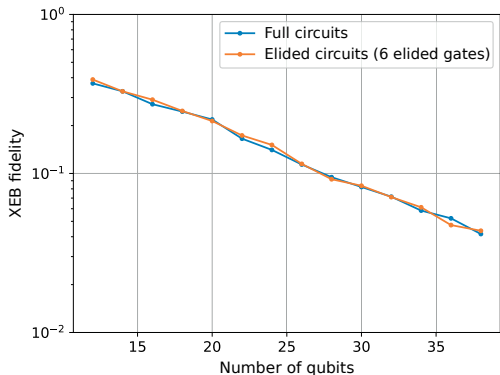


FIG. S29. **Comparison between XEB with elided circuits and full circuits.** Full vs. elided circuit benchmarking up to 38 qubits at 14 cycles, showing close agreement to within the intrinsic fluctuations of the system.

circuits”. Similar to patch circuits, we partition a given set of qubits into two subsets separated by a boundary, but elide (remove) only a fraction of the two-qubit gates along this boundary during a few early cycles of the sequence (more specifically, we elide the earliest gates in time, meaning early layers will have none of their gates along the boundary while later layers will have all of their usual gates across the boundary). Accordingly, the two subsets of qubits are no longer isolated from each other and we cannot simply compute their fidelities separately and multiply. Rather, we must still compute the evolution of the full system. Given that a sufficient number of gates are elided, we can take advantage of the “weak link” between patches with a hybrid analysis technique: we compute each patch via the Schrödinger method and then connect them with a Feynman path-integral approach (see Section X for more details on this “qsimh” program).

Compared with patch circuits, elided circuits more closely approach a description of the full system performance under a full circuit: in addition to capturing issues such as control and readout crosstalk, elided circuits allow entanglement to form between the two weakly connected subsystems. It covers essentially all the possible processes that occur in the full circuit, and therefore can be used to predict system performance at a dramatically reduced computational cost, albeit significantly costlier than patch circuits.

In order to validate the use of elided circuits as a system performance estimator, we evaluated its accuracy via a direct comparison with full circuits. In Fig. S29 we show two sets of fidelities from interleaved full and elided circuit experiments. For every system size investigated, using elided circuits yields a fidelity value that is in good agreement with the one obtained with the corre-

sponding full circuits. The average ratio of elided circuit fidelity to full circuit fidelity over all verification circuits was found to be 1.01, with a standard deviation of 5%, dominated by system fluctuations. It is this agreement that certifies elided circuits as a precise predictor for full circuits (within a systematic relative uncertainty of 5%), which we rely on to extrapolate the system performance in the regimes where full circuit analysis is too expensive to perform (i.e., Fig. 4b of the main text).

Compared with full circuits, elided circuits can result in a reduced amount of quantum entanglement in the system. The amount of reduced entanglement can be bounded from above by counting the number of iSWAP gates across the boundary: one iSWAP gate generates at most two units of bipartite entanglements (ebits). This upper bound translates directly into the exponential cost of a Schrödinger-Feynman simulation. For elided circuits with 50 qubits and 14 cycles, the full circuit has approximately 25 ebits of entanglement, while with 6 elisions the elided circuit has at most 12 ebits entanglement between the two patches. For the 53-qubit elided circuits used in the main paper, there were enough iSWAPs across the boundary that the amount of entanglement between patches for full vs. elided circuits should be close, giving us even more confidence in using elided circuits to predict the fidelity of the circuit used to claim quantum supremacy.

D. Choice of unitary model for two-qubit entangling gates

In Section VI, we discussed how the two-qubit gate unitaries can be measured by two different approaches: isolated two-qubit XEB and per-layer simultaneous two-qubit XEB. These two methods resulted in two different unitary models when deducing the best-fit unitary. Since we must specify the two-qubit gate unitary matrices in order to compute \mathcal{F}_{XEB} of the larger system, a natural question is which unitary model should be used. To address this question, we point out that full XEB on the large system occurs in repeated cycles, where during each two-qubit gate layer, all the two-qubit gates in the same orientation take place at the same time (see Fig. 3 in the main text). As a consequence, the two-qubit gate layers during simultaneous pair XEB in Fig. S19 emulate the corresponding layer when running full XEB on a large system. Accordingly, learning the unitaries in parallel operation captures any small coherent modifications induced by the simultaneous application of the other two-qubit gates, such as flux control crosstalk and dispersive shifts from stray interactions. This is evident from the fact that by re-learning the two-qubit unitary parameters, the errors extracted from simultaneous pair XEB become purity-limited (see Fig. S19). This correspondence assures us that unitary parameters extracted from simultaneous pair XEB provides a more accurate description of the full system when full XEB is performed.

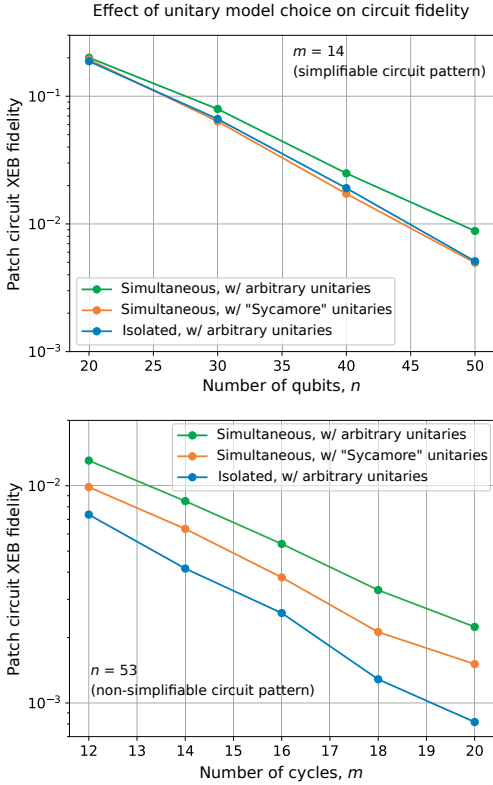


FIG. S30. **Effect of unitary model on full system fidelity.** **a.** Patch circuit fidelity versus number of qubits and choice of unitary model. **b.** Same but versus number of cycles and for the non-simplifiable supremacy circuits. Blue: patch XEB fidelities using the unitaries deduced from the best-fit fSim unitary from isolated pair XEB. Green: patch XEB fidelities using the unitaries deduced from the best-fit fSim unitary from per-layer simultaneous pair XEB. Orange: patch XEB fidelities using the unitaries deduced from the best-fit "Sycamore unitary" ($\theta = \pi/2$, $\phi = \pi/6$) from per-layer simultaneous pair XEB. As expected, the best fidelities arise from fitting to the most general unitary in parallel operation, although the fidelities are high enough to achieve quantum supremacy with the Sycamore unitary model as well.

In Fig. S30, we show patch circuit fidelities at different system sizes, where the fidelity is evaluated using three different unitary models: the best-fit unitaries from isolated pair XEB, the best-fit unitaries from simultaneous pair XEB, and the best-fit "Sycamore" unitaries from simultaneous pair XEB. The Sycamore unitaries are the unitaries obtained when keeping the swap angle fixed at $\theta = \pi/2$ and conditional phase fixed at $\phi = \pi/6$ for all

qubits, and then fitting only for two single-qubit phase terms. For the purpose of benchmarking the system fidelity for the operations we performed, we have focused on using unitaries learned from simultaneous pair XEB, which provide the most accurate description of the system. The validity of this approach is experimentally verified—for the same gate sequences, using the simultaneous pair XEB unitaries leads to the best full-system fidelity values at every system size. This is direct evidence that the unitaries learned from simultaneous pair XEB form a more accurate description of the system than those from isolated pair XEB.

On the other hand, in order to be useful for generic quantum algorithms, it will be desirable to use calibrated gatesets that are independent of the specific gate sequences used. For this purpose, it is important to check the circuit fidelity under the other two unitary models, where the two-qubit gate unitaries were calibrated in more generic settings. One can see that fidelities calculated from these two unitary models still demonstrate nearly as good performance despite the addition of small coherent control errors. They differ from the fidelities using the simultaneous pair XEB unitaries by less than a factor of 2 at 50 qubits (fidelity goes from 9×10^{-3} to 5×10^{-3} at 50 qubits). This is remarkable since it suggests going from a 2-qubit setting to 50-qubit setting, our full system calibration precision degrades only by a factor of < 2 despite the system size increasing by a factor of 25. This high precision in gate calibration gives us confidence to use our processors in NISQ algorithms.

E. Understanding system performance: error model prediction

In this section, we perform additional analysis to compare the measured fidelities to that predicted from the constituent gate and measurement errors.

The most commonly used error model in quantum computing theory is the digital error model. Analogous to the independent noise model in classical information theory, the digital error model is based on the assumption that there are no space and time correlations between errors of quantum gates [26, 58, 59]. If this assumption is valid, it should be possible to construct the fidelity of a large quantum system from the fidelities of its constituent parts: single- and two-qubit gates, and measurement. It is important to point out that the gate fidelity metric that should be used here is the entanglement fidelity, $1 - e_P$ (see Section V for more details). This is the correct quantity to describe the fidelity of quantum operations since, in contrast to other metrics such as the commonly used average fidelity, it is independent of the dimension of the Hilbert space.

In Fig. S31, we show fidelities as a function of both system size and number of cycles (circuit depth), measured with patch circuits. In each plot, we compare the measured fidelities to the predicted fidelities, which

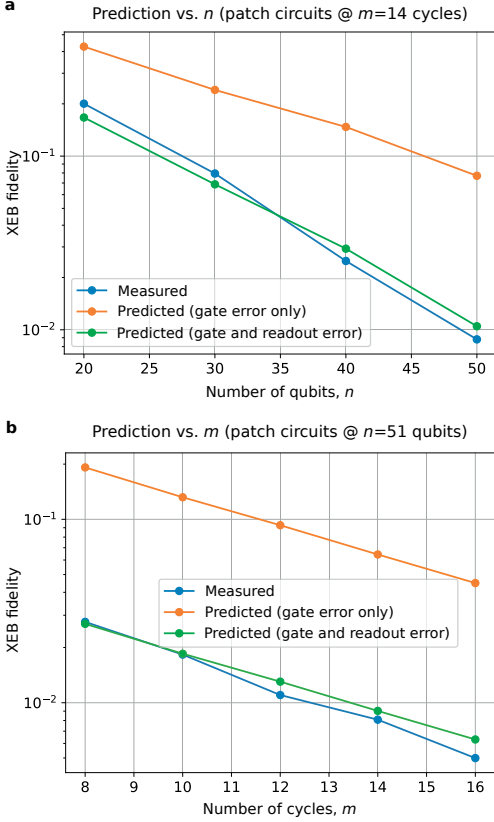


FIG. S31. **Predicted vs. measured large-scale XEB fidelity.** **a**, Data and two predictions for 14-cycle patch circuits vs. number of qubits. Predictions are based on the product of single- and two-qubit gate entanglement fidelities under simultaneous operation. Orange is the prediction based only on gate errors during parallel operation, but without taking measurement error into account. Green is the same but multiplied by the measured readout fidelities. **b**, Same as the first panel, but vs. number of cycles at a fixed number of qubits $n = 51$. Again, the prediction from simultaneous gate fidelities and measurement fidelity is a good prediction of the actual system performance.

are calculated from a simple multiplication of individual gate entanglement fidelities as measured during simultaneous operation, along with the measurement fidelities obtained during simultaneous measurement. We note that the measured readout fidelities actually also automatically include the effect of state preparation errors as well. More explicitly, if a circuit contains the set of single-qubit gates G_1 , the set of two-qubit gates G_2 , and

the set of qubits Q , then we approximate the fidelity F as

$$F = \prod_{g \in G_1} (1 - e_g) \prod_{g \in G_2} (1 - e_g) \prod_{q \in Q} (1 - e_q), \quad (77)$$

where e_g are the individual gate Pauli errors and e_q are the state preparation and measurement errors of individual qubits. It is evident that there is a good agreement between the measured and predicted fidelities, with deviations of up to only 10-20%. Given that the sequence here involves tens of qubits and ~ 1000 quantum gates, this level of agreement provides strong evidence to the validity of the digital error model.

This conclusion can be further strengthened by the close agreement between the fidelities of full circuits, patch circuits, and elided circuits. Even though these three methods differ only slightly in the gate sequence, they can result in systems with drastically different levels of computational complexity and entanglement between subsystems. The agreement between the fidelities measured by these different methods, as well as the agreement with the predicted fidelity from individual gates, gives compelling evidence confirming the assumptions made by the digital error model. Moreover, these assumptions remain valid even in the presence of quantum entanglement.

The validation of the digital error model has crucial consequences, in particular for quantum error correction. The absence of space or time correlations in quantum noise has been a commonly assumed property in quantum error correction since the very first paper on the topic [58]. Our data is evidence that such a property is achievable with existing quantum processors.

F. Distribution of bitstring probabilities

In Section IV, we motivate two different estimates for fidelity F , one based on the cross entropy, Eq. (28), and the other based on linear cross entropy, Eq. (27). In this section, we examine the probabilities of sampled bitstrings and compare them against theoretical distributions. We use bitstring samples from non-supremacy region to demonstrate the analysis methodology, then apply it to the sample in the supremacy region.

The theoretical PDF for the bitstring probability p with linear XEB is

$$P_l(x|F) = (Fx + (1 - F))e^{-x}$$

where $x \equiv Dp$ is the probability p scaled by the Hilbert space dimension D , and F is the linear cross entropy fidelity. The PDF for $\log p$ is

$$P_c(x|F) = (1 + F(e^x - 1))e^{x - e^x}$$

where $x \equiv \log(Dp)$ and F is the cross entropy fidelity.

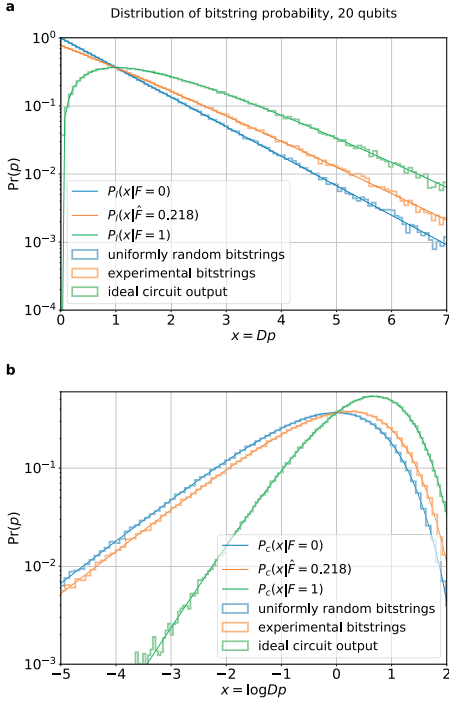


FIG. S32. **Histograms of ideal probabilities.** The ideal probability p is calculated from the final state amplitudes of a (20-qubit 14-cycle) random circuit. The blue, orange, and green histogram is the ideal probabilities of bitstrings sampled uniformly at random, from the experiment, and ideal output, respectively. **a**, The distribution of Dp and theoretical curves $P_l(x|F_i)$ normalized to histogram counts for $F_l = 0, \hat{F}_l, 1$, respectively. **b**, The distribution of $\log(Dp)$ and theoretical curves $P_c(x|F_c)$ for $F_c = 0, \hat{F}_c, 1$, respectively.

From a set of bitstrings $\{q_i\}$, the fidelity is estimated from the ideal probabilities $\{p_i = p_s(q_i)\}$ as

$$\hat{F}_l = \langle Dp \rangle - 1, \quad (78)$$

$$\hat{F}_c = \langle \log(Dp) \rangle + \gamma, \quad (79)$$

where γ is the Euler-Mascheroni constant, see Sec. IV B.

Figure S32 shows the distribution of $\{p_i\}$ from 0.5 million bitstrings obtained in an experiment with a 20-qubit 14-cycle random quantum circuit. For comparison, we produce 0.5 million bitstrings sampled uniformly at random and 0.5 million bitstrings sampled from the output distribution of the ideal circuit and show them in the same figure. The theoretical distribution curves are also shown, where the fidelity estimated from data is fed into

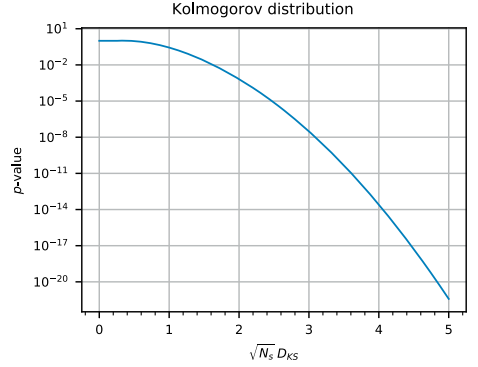


FIG. S33. **The Kolmogorov distribution function.** This function is used to compute p -value from a given D_{KS} and number of samples N_s .

the curve $P_l(x|\hat{F})$ and $P_c(x|\hat{F})$.

We see good agreements between experiment and theory. To quantify the agreements, we use the Kolmogorov-Smirnov test [60] to characterize the goodness of fit of data $\{p_i\}$ to theoretical PDFs. First we compute the Kolmogorov-Smirnov statistics D_{KS} , that is, the distance between data and theory as the supremum of point-wise distances between the empirical cumulative distribution function of data $\text{ECDF}(p)$ and the theoretical cumulative distribution function $\text{CDF}(p)$:

$$D_{KS} = \sup_i |\text{ECDF}(p_i) - \text{CDF}(p_i)|.$$

We then convert the distance D_{KS} to a p -value using the Kolmogorov distribution shown in Fig. S33. The p -value is used for rejecting the null hypothesis that the data $\{p_i\}$ is consistent with the theoretical distribution. The whole Kolmogorov-Smirnov test is done using the scipy package [61] and checked against R package ks.test [62]. Both packages produce consistent results.

We test the ideal probabilities of bitstrings observed in the experiment $\{p_i\}$ against 2 theoretical distributions, one with estimated fidelity $F = \hat{F}$ and one with fidelity $F = 0$. The Kolmogorov-Smirnov statistics D_{KS} and the p -value of every circuit are shown in figure S34. Note that the p -values for $F = 0$ are not shown because they are $\ll 10^{-20}$ due to the large $D_{KS} \approx 0.07$ with $N_s = 5 \times 10^5$ points in the sample. That is evident from reading off Fig. S33.

We reject the null hypothesis that the experimental bitstrings are consistent with the uniform random distribution with very high confidence for this (20-qubit 14-cycle) random circuit.

Now we turn our attention to the supremacy circuits.

We use random circuits with gate elisions for checking the distributions because it is exponentially expen-

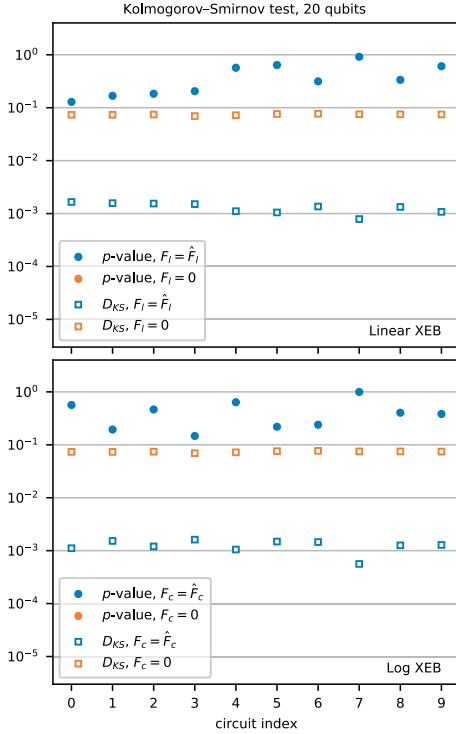


FIG. S34. The Kolmogorov-Smirnov test results for each of 10 circuits for a (20-qubit 14-cycle) random circuit. See text for the definition of D_{KS} and p -value. The upper plot is for linear XEB, and the lower one is for log XEB.

sive to calculate the ideal theoretical probability of a bitstring without gate elisions. The effect on fidelity from gate elisions is well understood, see Sec. VIII C. The gate elisions are chosen to minimize the effect while making the classical estimation feasible, see Sec. VII G 1. We sample $N_s = 3 \times 10^6$ bitstrings $\{q_i | i = 1 \dots N_s\}$ from each of 10 (53-qubit 20-cycle) random circuits, and compute the theoretical ideal probabilities of each bitstring $\{p_i | i = 1 \dots N_s\}$.

The distributions of Dp and $\log(Dp)$ from one such circuit along with the corresponding theoretical curves are shown in Fig. S35.

We again use the Kolmogorov-Smirnov test to characterize the goodness of fit of data $\{p_i\}$ to theoretical PDFs with estimated fidelity $F = \hat{F}$ and zero fidelity $F = 0$. The Kolmogorov-Smirnov statistics D_{KS} and the p -value of every circuit are shown in figure S36.

The p -value for the null hypothesis of zero fidelity is

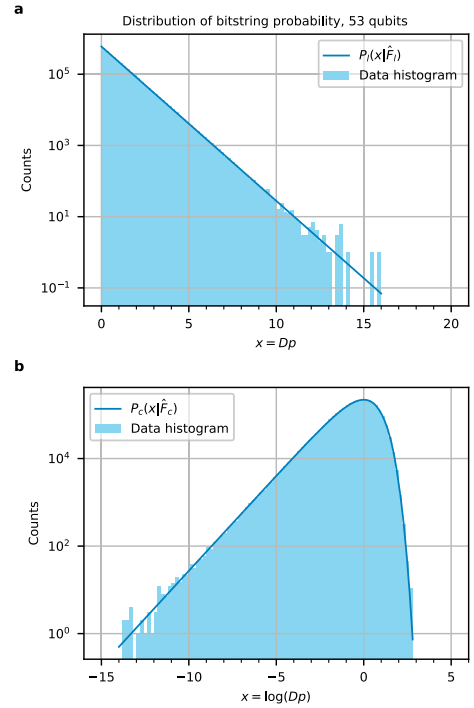


FIG. S35. Distribution of bitstring probabilities from a 53-qubit 20-cycle circuit. We calculate the theoretical probabilities of experimentally-observed bitstrings. **a**, The distribution of Dp and the theoretical curve $P_I(x|\hat{F}_I)$ normalized to histogram counts. **b**, The distribution of $\log(Dp)$ with theoretical curve $P_c(x|\hat{F}_c)$.

generally small for every circuit, with a maximum of 0.045 for circuit number 1. We say that the null hypothesis of zero fidelity is rejected better than a 95% confidence level for each circuit. On the other hand, the p -value of null hypothesis of estimated fidelity \hat{F} is generally large. The p -value is between 0.18 and 0.98 for linear XEB, and between 0.33 and 0.98 for log XEB. That indicates that the empirical cumulative distribution functions $\text{ECDF}(p_i)$ from data is quite consistent with the theoretical CDF $(p_i|\hat{F})$.

As will be seen in Fig. S38 in section VIII G below, the fidelity of individual circuits are consistent with each other within the statistical uncertainties. Therefore it makes sense to do a Kolmogorov-Smirnov test on all samples combined, containing 30 million bitstrings. The estimated fidelities from the combined sample are $\hat{F}_I = 2.24 \times 10^{-3}$ and $\hat{F}_c = 2.34 \times 10^{-3}$, respectively. The

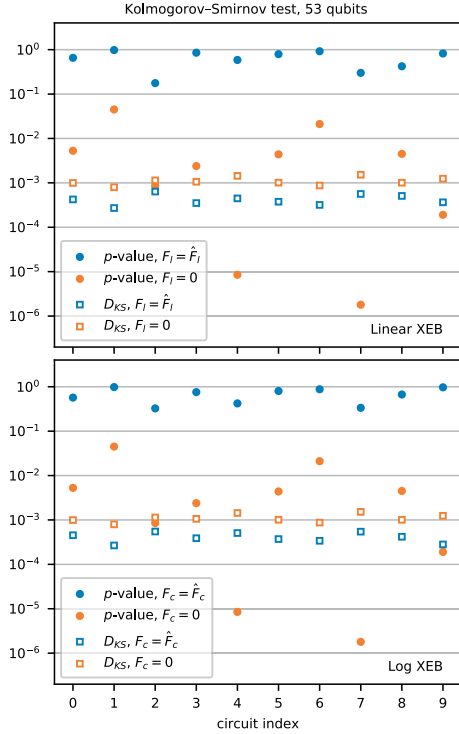


FIG. S36. The Kolmogorov-Smirnov test results for random circuits with 53 qubits. The upper plot is for linear XEB, and the lower one is for log XEB.

	D_{KS}		p -value	
	$F = \hat{F}$	$F = 0$	$F = \hat{F}$	$F = 0$
Linear XEB	1.3×10^{-4}	9.6×10^{-4}	0.66	$< 2.2 \times 10^{-16}$
Log XEB	9.5×10^{-5}	9.6×10^{-4}	0.95	$< 2.2 \times 10^{-16}$

TABLE V. The Kolmogorov-Smirnov test results on combined samples.

D_{KS} and p -values are listed in table V. The p -value for the null hypothesis of $F = 0$ is very small: p -value = 3×10^{-24} from scipy, and p -value $< 2.2 \times 10^{-16}$ from R. We note the more conservative value in the table. The null hypothesis of $F = 0$ is rejected with much higher confidence levels than individual circuits.

G. Statistical uncertainties of XEB measurements

In this section we check the statistical uncertainties of estimated linear XEB and XEB fidelities against theoretical predictions.

The statistical uncertainties of \hat{F}_l and \hat{F}_c are estimated from data using the standard error-on-mean formula as

$$\begin{aligned}\hat{\sigma}_{F_l} &= D\sqrt{\text{Var}(p)/N_s}, \\ \hat{\sigma}_{F_c} &= \sqrt{\text{Var}(\log p)/N_s},\end{aligned}$$

where $\text{Var}(x)$ is the variance estimator of sample $\{x_i\}$. Because the distribution of p and $\log p$ have finite variances both experimentally and theoretically, we can use the bootstrap procedure [63] to verify the estimate of statistical uncertainties.

The fidelity distribution from 4000 bootstrap samples are shown in Fig. S37. The distribution of \hat{F}_l and \hat{F}_c are each fit to a Gaussian distribution function using maximum likelihood.

The Kolmogorov-Smirnov test on the Gaussian fit produces p -values of 0.99 and 0.41 for \hat{F}_l and \hat{F}_c bootstrap distributions, respectively. It indicates that the central limit theorem is at work and the distributions are consistent with Gaussian distributions.

The estimated statistical uncertainty, the standard deviation of the bootstrap distribution, and the σ parameter of the Gaussian fit are compared against each other to verify that the statistical uncertainty estimate is minimally biased. For the example circuit used in the figures, the three parameters are 5.78, 5.78, 5.78 ($\times 10^{-3}$) for $\hat{\sigma}_{F_l}$, respectively. The same parameters for $\hat{\sigma}_{F_c}$ are 7.40, 7.46, 7.46 ($\times 10^{-3}$). The relative differences are less than 1%, consistent with the expected agreement of parameters for 4000 bootstrap samples.

We repeat the bootstrap procedure on all ten 53-qubit 20-cycle circuits with 2500 bootstrap resamples. The statistical uncertainty estimates are all within 3.1% of the bootstrap standard deviation.

The combined linear cross entropy fidelity and statistical uncertainty of 10 random circuits is calculated using inverse-variance weighting to be $\hat{F}_l = (2.24 \pm 0.18) \times 10^{-3}$. The theoretical prediction of the statistical uncertainty, $\sqrt{(1 + 2F - F^2)/N_s}$, is 1.8×10^{-4} , which agrees with the experimental estimate. As a comparison, the combined cross entropy fidelity is $\hat{F}_c = (2.34 \pm 0.23) \times 10^{-3}$. The theoretical prediction of statistical uncertainty, $\sqrt{(\pi^2/6 - F^2)/N_s}$, is 2.3×10^{-4} , which agrees with the experimental estimate as well. Thus, the cross entropy fidelity and linear cross entropy fidelity estimators produce consistent results. Furthermore, the statistical uncertainty of the linear cross entropy estimator is smaller, as expected from its theoretical formula.

In Fig. S38, we also show the linear XEB fidelities and 5σ statistical uncertainties of all 10 elided circuit instances for each circuit depth from Fig. 4b of the main text. Variations between the fidelities of different circuit instances are consistent with the expected statistical noise due to the finite number of samples. In the last

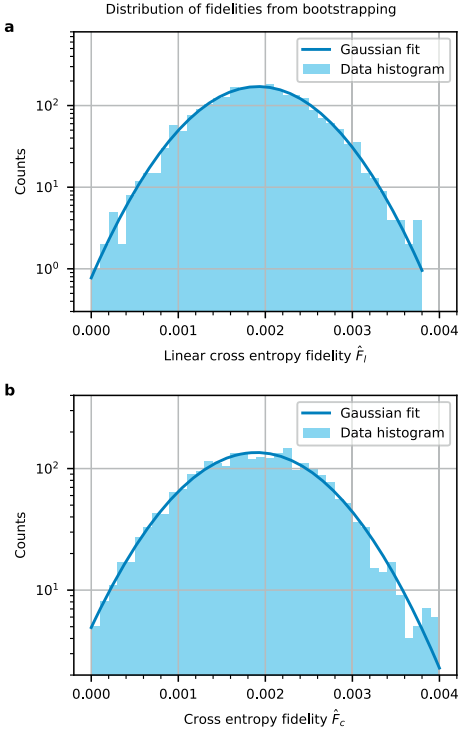


FIG. S37. Distribution of fidelity from 4000 bootstrap samples. **a**, The distribution of bootstrap \hat{F}_l . The theoretical curve is a Gaussian fit normalized to histogram counts. **b**, The distribution of bootstrap \hat{F}_c , with Gaussian fit.

panel, we also show the smaller statistical uncertainties of the fidelity averaged over the 10 circuit instances for each depth.

H. System stability and systematic uncertainties

In addition to statistical errors, XEB fidelity is also subject to systematic drift as the system performance may fluctuate and/or degrade over time. To quantify these mechanisms, we perform a patch circuits time stability measurement on 53 qubits using a circuit of 16 cycles and 1 million bitstrings for 17.4 hours after calibration. In between these measurements, we measured the fidelity of other 53-qubit circuits with 16 to 20 cycles. The analyzed results are shown in Fig. S39. The statistical uncertainties of the fidelities are estimated to be 1.29×10^{-4} , as indicated by the error bars.

We repeated the stability measurements twice, with different circuits and on different days. Fig. S39 shows the one that exhibits greater degradation as a conservative estimate of the effect. The measurement indicates a degradation of fidelity within the range of time. A linear fit with $F = p_0 + p_1 t$ results in estimated parameters $\hat{p}_0 = (5.51 \pm 0.055) \times 10^{-3}$, $\hat{p}_1 = (-6.87 \pm 0.64) \times 10^{-5}$, and a correlation coefficient of \hat{p}_0 and \hat{p}_1 , ρ , to be -0.76. The χ^2 per degree of freedom is 26.3/11.

The p -value for the χ^2 for 11 degrees of freedom is 0.0058, indicating that it is not a very good fit. Because the correctness of the estimates of statistical uncertainties has been verified in Section VIII G, this is attributed to systematic fluctuation in addition to degradation. It is supported by the larger variance of fidelity than the 1σ band in Fig. S39.

The 1σ band depends on the statistical uncertainties of fidelities and the variance of time on the x -axis, but is independent of the variance of fidelity. To take the variance of fidelity into account, we use the variance of the residuals of the linear fit as an estimator of the variance of fidelity. The standard deviation of residuals is estimated to be 1.84×10^{-4} , which is added to σ_{p_0} in quadrature to be the total σ_{p_0} . The estimate is total $\sigma_{p_0} = 1.92 \times 10^{-4}$, 3.5 times larger than the statistical-only σ_{p_0} of 5.5×10^{-5} .

The uncertainty on a fidelity measured at time t can be estimated by the standard error propagation, assuming that t is uncorrelated with either p_0 or p_1 .

$$\sigma_F = [\sigma_{p_0}^2 + 2t\sigma_{p_0}\sigma_{p_1}\rho + \sigma_{p_1}^2 t^2]^{1/2} \quad (80)$$

The value of σ_F as well as the ratio σ_F/F in the range of measured fidelities monotonically decreases. We take $\max(\sigma_F/F)$ as the estimate of relative systematic uncertainty for fidelities measured in the same run. The value is found to be 4.4% and is used in subsequent analysis.

The physical origin of the observed system fluctuations can be attributed to many possible channels: $1/f$ flux noise, qubit T_1 fluctuations, control signal drift, etc. We speculate that the dominant mechanism is the moderate interaction between a small number of TLS's and a few qubits at their idling and/or readout biases. In Fig. S40a, we show the result of measuring per-layer simultaneous pair XEB at a fixed depth of 14 cycles repeatedly over time. The quantity plotted is the ratio of the worst pair fidelity to best fidelity observed over the course of 30 minutes. This type of repetitive measurement allows us to pinpoint which pairs dominate the fluctuations in full system fidelity. Note that because we used fidelity at a fixed cycle depth rather than the one extracted from the exponential decay, these numbers contain the effect of fluctuating measurement fidelity as well.

As shown in Fig. S40a, the depth-14 fidelity of most pairs fluctuates downward by only $\sim 1\%$ at depth 14, which translates to either a $\sim 1\%$ fluctuation in measurement fidelity for a pair, or a $\sim 0.08\%$ fluctuation in the two-qubit gate fidelity for a pair. Before finding the unstable TLS defect in Fig. S40b, a single qubit dominated the fluctuations in full system fidelity seen

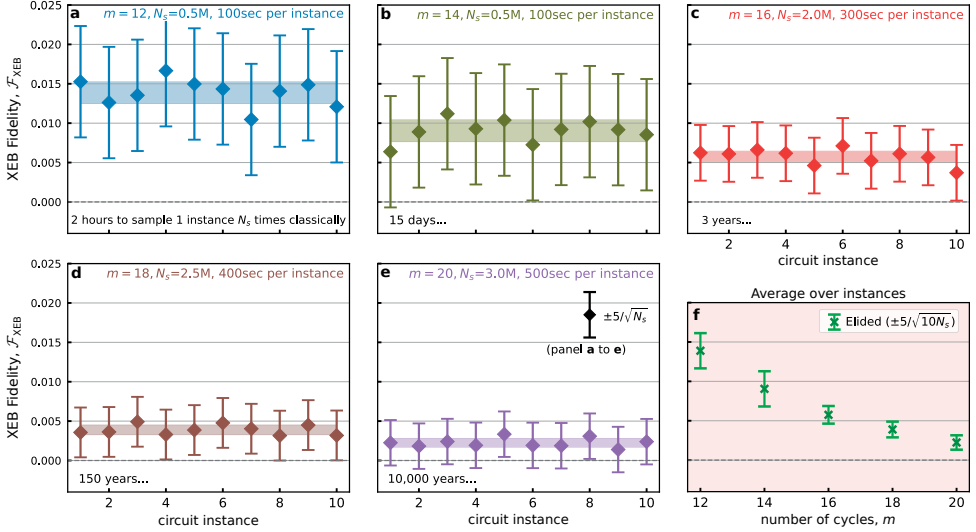


FIG. S38. Per-instance elided circuit fidelities and statistical uncertainties. XEB fidelities of all 10 elided circuit instances for each circuit depth from Fig. 4b of the main text. **a** to **e**, Here, each panel corresponds to a single circuit depth m . In these panels, $\pm 5\sigma$ statistical error bars, where $\sigma = 1/\sqrt{N_s}$, are shown for each of the individual circuit instance fidelities. Also shown is a band corresponding to $\pm\sigma$ for a single instance, but about the mean fidelity of the 10 instances, showing that the variations between circuits can be explained by statistical fluctuations from the finite number of samples. **f**, Fidelity averaged over all 10 circuits along with $\pm 5\sigma$ error bars are shown (the same quantity is plotted in Fig. 4b of the main text but on a log scale), where in this case $\sigma = 1/\sqrt{10N_s}$. Here, for all circuit depths, the mean fidelity is more than 5σ above 0.001.

in Fig. S40c. After we moved this problematic qubit far from the fluctuating TLS, the fluctuations in fidelity during the actual quantum supremacy experiment (Fig. S39) were dominated by a handful of pairs containing qubits in the “degenerate” readout region (described in section VI). For these qubits, due to constraints from readout crosstalk we had little freedom in what readout detunings we could choose, and so the best we could do was to put some qubits near defects or transmon-resonator transition modes during readout. We speculate that this is where the remaining dominant fluctuations originate.

I. The fidelity result and the null hypothesis on quantum supremacy

We use the mean fidelity of ten 53-qubit 20-cycle circuits as the final benchmark of the system. In section VIII G we estimated the fidelity and statistical uncertainty to be $(2.24 \pm 0.18) \times 10^{-3}$ using the linear cross entropy. In section VIII H we estimated the relative systematic uncertainty due to drift to be 4.4%. Combining these 2 estimations we arrive at the final fidelity as $(2.24 \pm 0.10(\text{syst.}) \pm 0.18(\text{stat.})) \times 10^{-3}$.

As we show in section X, a noisy sampling of a ran-

dom quantum circuit at fidelity $F = 10^{-3}$ requires 5000 years with a classical computer with CPU power equivalent to 1 million cores, and it scales linearly with fidelity F . It takes a quantum computer less than an hour to complete the same noisy sampling. Therefore we form the null hypothesis that the fidelity of the quantum computer is $F \leq 10^{-3}$, and the alternative hypothesis that $F > 10^{-3}$. If the alternative hypothesis is true, we can say that a classical computer can not perform the same noisy sampling task as the quantum computer.

The total uncertainty on fidelity is estimated with addition in quadrature of systematic uncertainty and statistical uncertainty. The mean fidelity of 10 random circuits with 53 qubits and 20 cycles is $(2.24 \pm 0.21) \times 10^{-3}$. The null hypothesis is therefore rejected with a significance of 6σ .

While our analysis of the uncertainty in \mathcal{F}_{XEB} was computed from both statistical and systematic errors, some care should be taken in the consideration of systematic errors as they pertain to the claim of quantum supremacy. Systematic errors should be included if we wish to use the XEB fidelity value, for example comparing fidelities of patch, elided and full circuits. However for quantum supremacy, a false claim would arise if \mathcal{F}_{XEB} was zero, but we obtained a non-zero value because of a

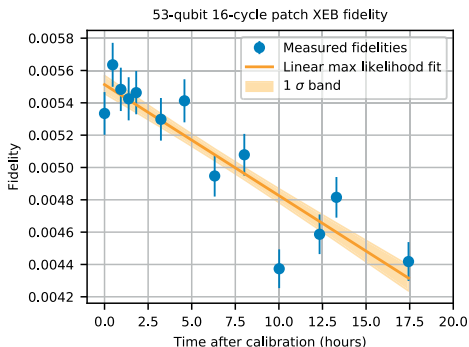


FIG. S39. Stability of repeated 53-qubit 16-cycle patch circuit benchmarking over 17.4 hours, without any system recalibration. Statistical error bars from the finite bitstring sample number are included. The intrinsic system fluctuations are likely dominated by a small number of TLSs moderately coupled to a few qubits at their idling and/or readout biases.

fluctuation. Systematic fluctuations produce a change in magnitude of XEB, as seen in the data in this section, which is thus a multiplicative-type error that does not change the XEB fidelity value when it is zero. A false positive is only produced by an additive-type statistical fluctuations and thus it is the only mechanism that should be considered when computing the uncertainty. Therefore, the 6σ significance of our claim should be considered as conservative.

Some skeptics have warned that a quantum computer may not be possible [64], for example due to the fragility of quantum information at large qubit number and exponentially large Hilbert space. The demonstration here of quantum behavior at 10^{16} Hilbert space is strong confirmation that nothing unusual or unexpected happens to our current understanding of quantum mechanics at this scale.

IX. SENSITIVITY OF XEB TO ERRORS

An important requirement for a procedure used to evaluate quantum processors, such as XEB, is sensitivity to errors. Qubit amplitudes are complex variables and therefore quantum errors are inherently continuous. Nevertheless, they can be given a discrete description, for example in the form of a finite set of Pauli operators. The digital error model is used for instance in quantum error correction where errors are discretized by syndrome extraction. In this section we examine the impact of both discrete and continuous errors on the fidelity estimate obtained from the XEB algorithm.

The XEB procedure uses a set of random quantum

circuits $\mathcal{U} = \{U_1, \dots, U_S\}$ with n qubits and m cycles. Every circuit is executed N_s times on the quantum processor under test. Each execution of the circuit U_j applies the quantum operation Λ_j , which is an imperfect realization of U_j , to the input state $|0\rangle\langle 0|$. The result of the experiment is a set \mathcal{B} of SN_s bitstrings $q_{i,j}$ sampled from the distributions $p_e(q_{i,j}) = \langle q_{i,j} | \rho_j | q_{i,j} \rangle$ where $\rho_j = \Lambda_j(|0\rangle\langle 0|)$ is the output state in the experiments with circuit U_j . For each bitstring $q_{i,j}$, a simulator computes the ideal probability $p_s(q_{i,j}) = |\langle q_{i,j} | \psi_j \rangle|^2$ where $|\psi_j\rangle = U_j |0\rangle$ is the ideal output state of the circuit U_j . Finally, XEB uses Eq. (27) or (28) to compute an estimate $\mathcal{F}_{\text{XEB}}(\mathcal{B}, \mathcal{U})$ of fidelity $F(|\psi_j\rangle\langle \psi_j|, \rho_j) = \langle \psi_j | \rho_j | \psi_j \rangle$ averaged over circuits \mathcal{U} . The result quantifies how well the quantum processor is able to realize quantum circuits of size n and depth m . See section IV for more details on XEB.

The estimate $\mathcal{F}_{\text{XEB}}(\mathcal{B}, \mathcal{U})$ is a function of bitstrings \mathcal{B} obtained in experiment and of the set of quantum circuits \mathcal{U} used to compute ideal probabilities. This enables a test of the sensitivity of the method to errors by replacing the error-free reference circuits $\mathcal{U} = \{U_1, \dots, U_S\}$ with circuits $\mathcal{U}_E = \{U_{1,E}, \dots, U_{S,E}\}$ where $U_{j,E}$ is the quantum circuit obtained from U_j by the insertion at a particular location in the circuit of a gate E representing the error. We identify errors inserted at different circuit locations that lead to the same output distribution since XEB cannot differentiate between them.

We first consider the impact of a discrete single-qubit Pauli error E placed in a random location in the circuit. In Fig. S41 we plot $\mathcal{F}_{\text{XEB}}(\mathcal{B}, \mathcal{U}_E)$ where \mathcal{B} are bitstrings observed in our experiment and \mathcal{U}_E are quantum circuits modified by the insertion of an additional X or Z gate following an existing single-qubit gate. Each fidelity estimate corresponds to a different circuit location where the error gate has been inserted. For every n , the highest fidelity values correspond to the insertion of the Z gate in the final cycle of the circuit. They have no impact on measurements and thus are equivalent to absence of error. The corresponding fidelity estimates match the estimates for the unmodified circuits.

The probability of only seeing the error E is approximately $q = ep$ where e is the probability of E arising at the particular circuit location and p is the probability that no other error occurs. The fraction q of executions realize circuit $U_{j,E} \in \mathcal{U}_E$ yielding bitstrings \mathcal{B}_E while the remaining fraction $1 - q$ yield bitstrings \mathcal{B}_* . XEB averages over circuit executions, so

$$\mathcal{F}_{\text{XEB}}(\mathcal{B}, \mathcal{U}_E) = q \mathcal{F}_{\text{XEB}}(\mathcal{B}_E, \mathcal{U}_E) + (1 - q) \mathcal{F}_{\text{XEB}}(\mathcal{B}_*, \mathcal{U}_E). \quad (81)$$

Since bitstrings \mathcal{B}_E originated in a perfect realization of \mathcal{U}_E we have $\mathcal{F}_{\text{XEB}}(\mathcal{B}_E, \mathcal{U}_E) \simeq 1$ with high probability. Also, assuming the circuits randomize the output quantum state sufficiently, we have $\mathcal{F}_{\text{XEB}}(\mathcal{B}_*, \mathcal{U}_E) \simeq 1/\sqrt{D}$, where $D = 2^n$, see Eq. (25) and Fig. S7. Therefore, for large n

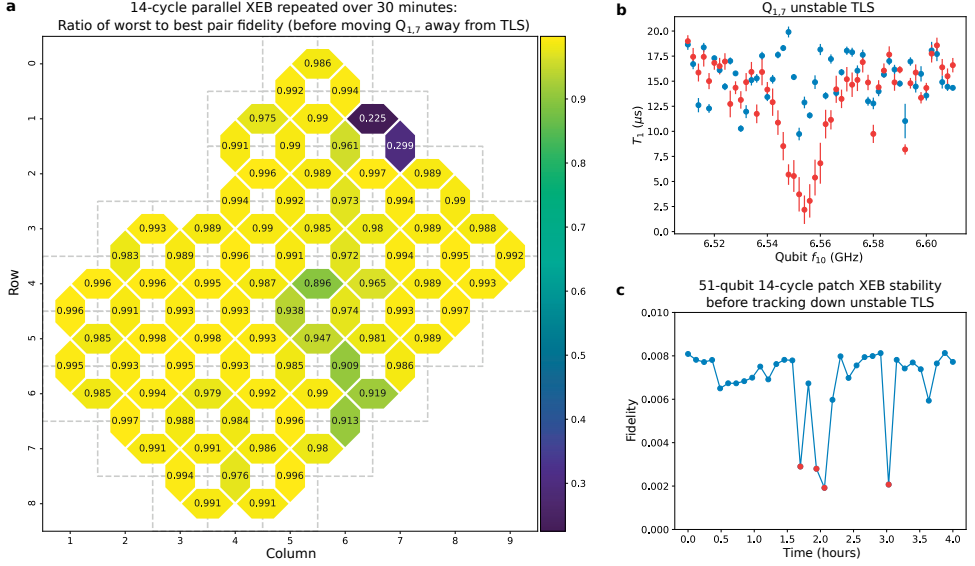


FIG. S40. **Identifying sources of fluctuations with repetitive per-layer simultaneous pair XEB.** **a**, Per-pair ratio of worst fidelity to best fidelity measured via per-layer simultaneous pair XEB at a depth of 14 cycles over the course of 30 minutes. During this time, fluctuations were dominated by a single TLS. **b**, Measured qubit T_1 vs. f_{10} for $Q_{1,7}$ at two different times a few minutes apart (red vs. blue points), showing an unstable TLS that was dominating the fluctuations in full system fidelity seen in **c**. Moving $Q_{1,7}$ far from this TLS led to the stability seen in Fig. S39.

$$\mathcal{F}_{\text{XEB}}(\mathcal{B}, \mathcal{U}_E) \simeq q + \frac{1-q}{\sqrt{D}} \simeq q \quad (82)$$

with high probability.

Now, the probability p that no error other than E occurs is approximately equal to the experimental fidelity F which is approximated by $\mathcal{F}_{\text{XEB}}(\mathcal{B}, \mathcal{U})$, so

$$\mathcal{F}_{\text{XEB}}(\mathcal{B}, \mathcal{U}_E) \simeq e \mathcal{F}_{\text{XEB}}(\mathcal{B}, \mathcal{U}) \quad (83)$$

which means that XEB result obtained using circuits modified to include E is approximately proportional to the XEB result obtained using the error-free reference circuits. Moreover, the ratio of the two XEB results is approximately equal to the probability of E .

The data in Fig. S41 agrees with the approximate proportionality in Eq. (83) and allows us to estimate the median probability of a Pauli error. Based on the drop in XEB fidelity estimate by a factor of almost 100 due to the insertion of one single-qubit Pauli error into the circuit, the probability is on the order of 1%. While more work on the gate failure model needs to be done to correctly relate Sycamore gate error rates to the probability of specific Pauli errors, we already see that e has the same

order of magnitude as our per cycle and per qubit error given by $e_{2c}/2 \simeq 0.5\%$, see Table II. A possible resolution of the factor of two discrepancy may lie in the fact that more than one gate failure can manifest itself as a particular Pauli error E in a particular circuit location.

Lastly, we consider the impact of continuous errors on XEB result. Fig. S42 shows the fidelity estimate obtained from XEB using bitstrings observed in our experiment and quantum circuits modified to include a single rotation $R_Z(\theta)$. The middle point of the plot is equal to the fidelity estimate obtained for one of the discrete errors in Fig. S41 whereas the leftmost and rightmost points correspond to the fidelity estimate obtained from XEB using the error-free reference circuit.

The analysis above illustrates how questions about the behavior and performance of quantum processors can be formulated in terms of modifications to the reference quantum circuits and how XEB can help investigate these questions. While XEB has proven itself a powerful tool for calibration and performance evaluation (see sections VI and VIII), more work is required to assess its efficacy as a diagnostic tool for quantum processors.

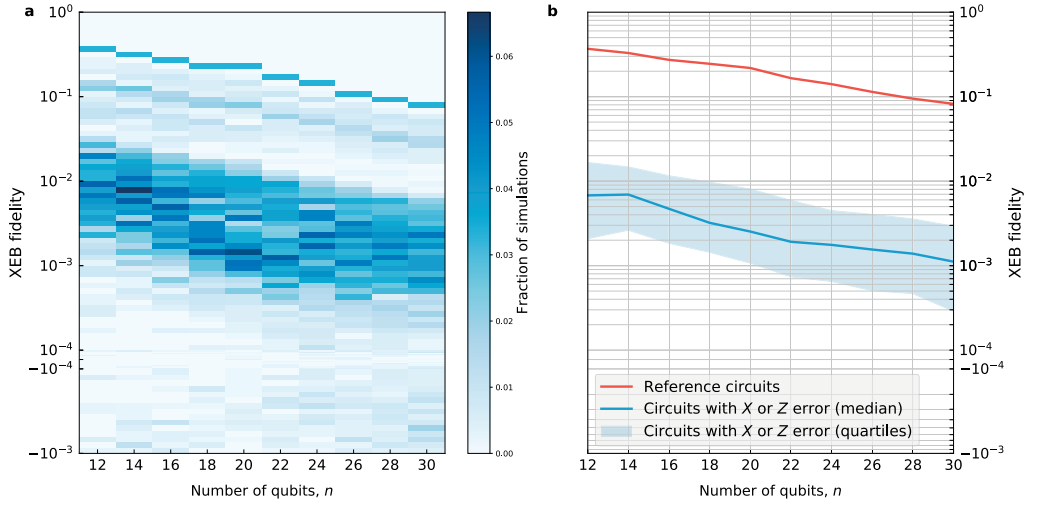


FIG. S41. **Impact of one single-qubit Pauli error on fidelity estimate from XEB.** **a**, Distributions of fidelity estimates from XEB using measured bitstrings and quantum circuits with one bit-flip or one phase-flip error. For each n , shades of blue represent the normalized histogram of the estimates obtained for the error gate placed at different circuit locations. The highest fidelity estimates correspond to phase-flip errors immediately preceding measurement and are equal to the fidelity estimates from XEB using error-free circuits. **b**, Quartiles of the distributions shown in **a** (blue) compared to the fidelity estimates from XEB using measured bitstrings and unmodified quantum circuits (red). Both plots use linear scale between 10^{-4} and 10^{-3} and logarithmic scale everywhere else.

X. CLASSICAL SIMULATIONS

A. Local Schrödinger and Schrödinger-Feynman simulators

We have developed two quantum circuit simulators: `qsim` and `qsimh`. The first simulator, `qsim`, is a Schrödinger full state vector simulator. It computes all 2^n amplitudes, where n is the number of qubits. Essentially, the simulator performs matrix-vector multiplications repeatedly. One matrix-vector multiplication corresponds to applying one gate. For a 2-qubit gate acting on qubits q_1 and q_2 ($q_1 < q_2$), it can be depicted schematically by the following pseudocode.

```
#iterate over all values of qubits q > q2
for (int i = 0; i < 2^n; i += 2 * 2^q2) {
    #iterate values for q1 < q < q2
    for (int j = 0; j < 2^q2; j += 2 * 2^q1) {
        #iterate values for q < q1
        for (int k = 0; k < 2^q1; k += 1) {
            #apply gate for fixed values
            #for all q not in [q1,q2]
            int l = i + j + k;

            float v0[4]; #gate input
```

```
float v1[4]; #gate output
```

```
#copy input
v0[0] = v[1];
v0[1] = v[1 + 2^q1];
v0[2] = v[1 + 2^q2];
v0[3] = v[1 + 2^q1 + 2^q2];

#apply gate
for (r = 0; r < 4; r += 1) {
    v1[r] = 0;
    for (s = 0; s < 4; s += 1) {
        v1[r] += U[r][s] * v0[s];
    }
}

#copy output
v[1] = v1[0];
v[1 + 2^q1] = v1[1];
v[1 + 2^q2] = v1[2];
v[1 + 2^q1 + 2^q2] = v1[3];
}
```

Here U is a 4×4 gate matrix and v is the full state vector. To make the simulator faster, we use gate fusion

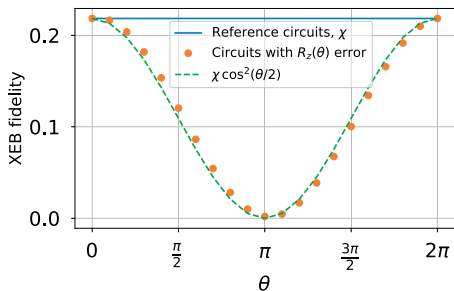


FIG. S42. **Impact of the $R_z(\theta)$ error on XEB.** Fidelity estimates computed by XEB from measured bitstrings and circuits with $n = 20$ qubits and $m = 14$ cycles modified to include $R_z(\theta)$ error applied in 10th cycle to one of the qubits as a function of θ (orange dots). Also shown is XEB fidelity computed using the same bitstrings and unmodified circuits (blue solid line) and a simple model which predicts the effect of the error (green dashed line).

qubits, n	run time in seconds
32	111
34	473
36	1954
38	8213

TABLE VI. Circuit simulation run times using qsim on a single Google cloud node (n1-ultramem-160).

[65], single precision arithmetic, AVX/FMA instructions for vectorization, and OpenMP for multi-threading. We are able to simulate 38-qubit circuits on a single Google cloud node that has 3844 GB memory and four CPUs with 20 cores each (n1-ultramem-160). The run times for different circuit sizes at depth 14 are listed in Table VI.

The second simulator, qsimh, is a hybrid Schrödinger-Feynman algorithm (SFA) simulator [37]. We cut the lattice into two parts and use the Schmidt decomposition for the 2-qubit gates on the cut. If the Schmidt rank of each gate is r and the number of gates on the cut is g then there are r^g paths, corresponding to all the possible choices of Schmidt terms for each 2-qubit gate across the cut. To obtain fidelity equal to unity, we need to simulate all the r^g paths and sum the results. The total run time is proportional to $(2^{n_1} + 2^{n_2})r^g$, where n_1 and n_2 are the qubit numbers in the first and second parts. Each part is simulated by qsim using the Schrödinger algorithm. Path simulations are independent of each other and can be trivially parallelized to run on supercomputers or in data centers. Note that one can run simulations with fidelity $F < 1$ just by summing over a fraction F of all the paths (see Ref. [37] and Sec. XD). In order to speed up the computation, we save a copy of the state after the first p

2-qubit gates across the cut, so the remaining r^{g-p} paths can be computed without re-starting the simulation from the beginning. We call the specific choice of Schmidt terms for the first p gates in the cut a *prefix*.

B. Feynman simulator

qFlex was introduced in Ref. [49] and later adapted to GPU architectures in Ref. [66] to allow efficient computation on Summit, currently the world's Top-1 supercomputer. Given a random quantum circuit, qFlex computes output bitstring amplitudes by adding all the Feynman path contributions via tensor network (TN) contractions [67, 68], and so it follows what we call a Feynman approach (FA) to circuit sampling. TN simulators are known to outperform all other methods for circuits with low depth or a large number of qubits (*e.g.*, Ref. [66] successfully simulates 121 qubits at low depth using this technique), as well as for small sample sizes (N_s), since simulation cost scales linearly with N_s .

TN simulators compute one amplitude (or a few amplitudes; see below) per contraction of the entire network. In order to sample bitstrings for a given circuit, a set of random output bitstrings is chosen before the computation starts. Then, the amplitudes for these bitstrings are computed and either accepted or rejected using *frugal rejection sampling* [37]. This ensures that the selected subset of bitstrings is indistinguishable from bitstrings sampled from a quantum computer. The cost of the TN simulation is therefore linear in the number of output bitstrings. This makes TN methods more competitive for small sets of output bitstrings.

The optimization of qFlex considers a large number of factors to achieve the best time-to-solution on current supercomputers, an approach that often diverges from purely theoretical considerations on the complexity of TN contractions. More precisely, qFlex implements several features such as:

- **Avoidance of distributed tensor contractions:** by “cutting” the TN (slicing some indexes), the contraction of the TN is decomposed into many *paths* that can be contracted locally and independently, therefore avoiding internode communication, which is the main cause for the slowdown of distributed tensor contractions.
- **Contraction orderings for high arithmetic intensity:** TN contraction orderings are chosen so that the expensive part of the computation consists of a small number of tensor contractions with high arithmetic intensity. This lowers the time-to-solution.
- **Highly efficient tensor contractions on GPU:** the back-end TAL-SH library [69] provides fully asynchronous execution of tensor operations on

qubits	cycles	\mathcal{F}_{XEB} (%)	N_s	nodes	runtime	PFlop/s*		efficiency (%)		power (MW)	energy (MWh)
						peak	sust.	peak	sust.		
53	12	0.5	1M	4550	1.29 hours	235.2	111.7	57.4	27.3	5.73	8.21
		1.4	0.5M		1.81 hours**						11.2**
		1.4	3M		10.8 hours**						62.7**
	14	2.22×10^{-6}	1M		0.72 hours	347.5	252.3	84.8	61.6	7.25	6.11
		0.5	1M		67.7 days**						$1.18 \times 10^{4**}$
		1.0	0.5M		67.7 days**						$1.18 \times 10^{4**}$
		1.0	3M		1.11 years**						$7.07 \times 10^{4**}$

TABLE VII. **Runtimes, efficiency and energy consumption for the simulation of random circuit sampling of N_s bitstrings from Sycamore with fidelity \mathcal{F} using qFlex on Summit.** Simulations used 4550 nodes out of 4608, which represents about 99% of Summit. Single batches of 64 amplitudes were computed on each MPI task using a socket with three GPUs (two sockets per node); given that one of the 9100 MPI tasks acts as master, 9099 batches of amplitudes were computed. For the circuit with 12 cycles, 144/256 paths for these batches were computed in 1.29 hours, which leads to the sampling of about 1M bitstrings with fidelity $\mathcal{F} \approx 0.5\%$ (see Ref. [49] for details on the sampling procedure); runtimes and energy consumption for other sample sizes and fidelities are extrapolated linearly in N_s and \mathcal{F} from this run. At 14 cycles, 128/524288 paths were computed in 0.72 hours, which leads to the sampling of about 1M bitstrings with fidelity 2.22×10^{-6} . In this case, one would need to consider 288101 paths on all 9099 batches in order to sample about 1M (0.5M) bitstrings with fidelity $\mathcal{F} \approx 0.5\%$ (1.0%). By extrapolation, we estimate that such computations would take 1625 hours (68 days). For $N_s = 3M$ bitstrings and $\mathcal{F} \approx 1.0\%$, extrapolation gives us an estimated runtime of 1.1 years. Performance is higher for the simulation with 14 cycles, due to higher arithmetic intensity tensor contractions. Power consumption is also larger in this case. Job, MPI, and TAL-SH library initialization and shutdown times, as well as initial and final IO times are not considered in the runtime, but they are in the total energy consumption. *Single precision. **Extrapolated from the simulation with a fractional fidelity.

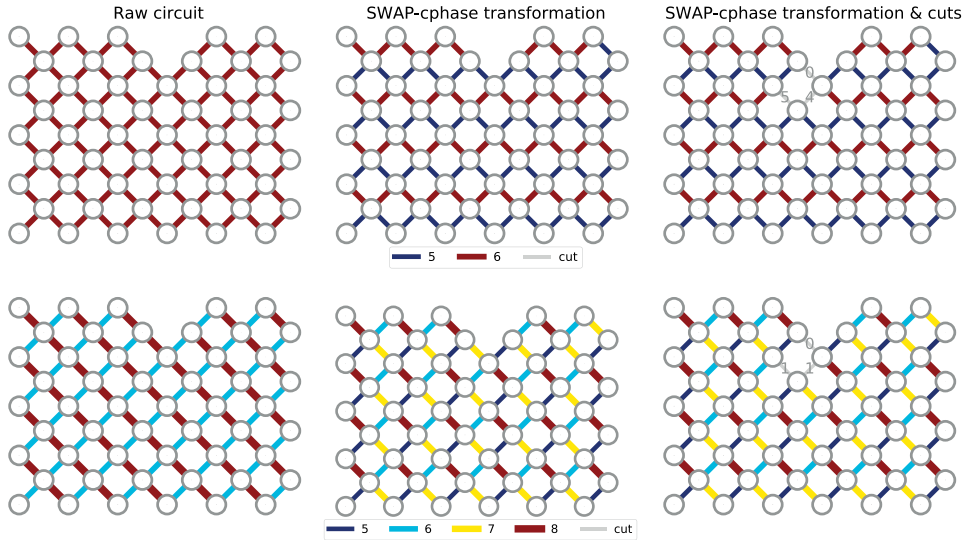


FIG. S43. **Logarithm base 2 of the bond (index) dimensions of the tensor network to contract for the simulation of sampling from Sycamore with 12 cycles (top) and 14 cycles (bottom) using qFlex.** The left plots represent the tensor network given by the circuit. The middle plots represent the tensor network obtained from a circuit where fSim gates have been transformed, when possible (see main text). The right plots represent the tensor network after the gate transformations and cuts (gray bonds) have been applied; the \log_2 of the bond dimensions of the indexes cut are written explicitly. For 12 cycles, there are $2^5 \times 2^1 \times 2^2 = 2^8 = 256$ cut instances (paths); for 14 cycles, there are $2^7 \times 2^7 \times 2^5 = 2^{19} = 524288$ cut instances.

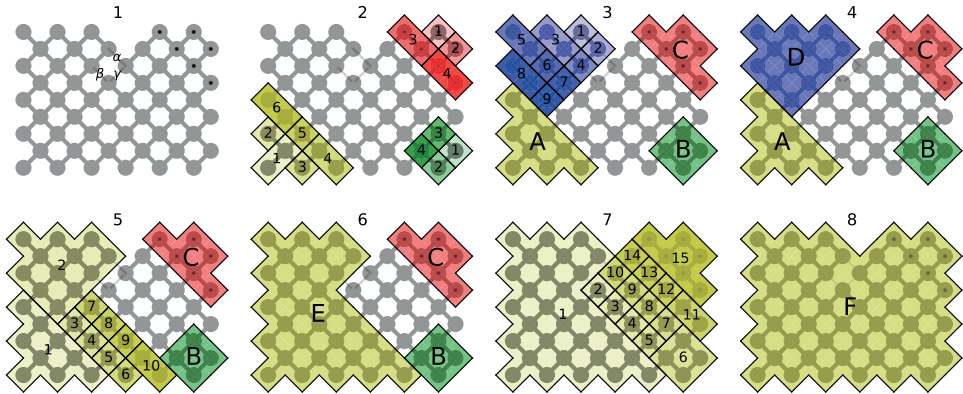


FIG. S44. TN contraction ordering for the computation of a batch of amplitudes for the simulation of Sycamore with 12 and 14 cycles. Dotted qubits are used for fast sampling; the output index is left open. Three indexes are cut, with remaining bond dimensions given in Fig. S43, and all possible cut instances are labelled by variables α , β , and γ (panel 1). Tensors A , B , and C are independent of cut instances, and so are contracted only once (panels 2 and 3) and reused several times. Given a particular instance of α and β , tensors D (panels 3 and 4) and subsequently E (panels 5 and 6) are contracted; tensor E will be reused in the inner loop. For each instance of γ (inner loop), tensor F is contracted (panels 7 and 8), which gives the contribution to the batch of amplitudes (open indexes on C and specified output bits otherwise) from a particular (α, β, γ) instance (path). The sequence of tensor contractions leading to building a tensor are enumerated, where each tensor is contracted to the one formed previously. For simplicity, the contraction of two single-qubit tensors onto a pair before being contracted with others (*e.g.*, tensor 10 in the yellow sequence of panel 5) is not shown on a separate panel; these pairs of tensors are computed first and are reused for all cut instances.

GPU and fast tensor transposition, allowing out-of-core tensor contractions for instances that exceed GPU memory. This achieves very high efficiency (see Table VII) on high arithmetic intensity contractions.

In addition, qFlex implements two techniques in order to lower the cost of the simulation:

- **Noisy simulation:** the cost of a simulation of fidelity $\mathcal{F} < 1$ ($\mathcal{F} \approx 5 \times 10^{-3}$ in practice) is lowered by a factor $1/\mathcal{F}$, *i.e.*, is linear in \mathcal{F} [37, 49].
- **Fast sampling technique:** the overhead in applying the frugal rejection sampling mentioned above is removed by this technique, giving an order of magnitude speedup [49]. This involves the computation of the amplitudes of a few correlated bitstrings (*batch*) per circuit TN contraction.

As shown in Table VII, qFlex is successful in simulating Sycamore with 12 cycles on Summit, sampling 1M bitstrings with fidelity close to 0.5% in 1.29 hours. At 14 cycles, we perform a partial simulation and extrapolate the simulation time for the sampling of 1M bitstrings with fidelity close to 0.5% using Summit, giving an estimated 68 days to complete the task. Sampling 3M bitstrings at 14 cycles with fidelity close to 1.0% (average experimentally realized fidelity) would take

an estimated 1.1 years to complete. Other estimates for different sample sizes and fidelities can be found in Table VII. At 16 cycles and beyond, however, the enormous amount of Feynman paths required so that the computation does not exceed the 512 GB of RAM of each Summit node makes the computation impractical.

The contraction of the TNs involved in the computation of amplitudes from Sycamore using qFlex is preceded by a simplification of the circuits, which allows us to decrease the bond (index) dimension of some of the indexes of the TN. This comes from the realization that $\text{fSim}(\theta = \pi/2, \phi) = -i \cdot [\text{R}_z(-\pi/2) \otimes \text{R}_z(-\pi/2)] \cdot \text{cphase}(\pi + \phi) \cdot \text{SWAP}$ (see Sections VI and VII E); note that the SWAP gate can be applied either at the beginning or at the end of the sequence. We apply this transformation to all fSim gates at the beginning (end) of the circuit that affect qubits that are not affected by any other two-qubit gate before (after) in the circuit. The SWAP is then applied to the input (output) qubits and their respective one-qubit gates trivially, and the bond dimension remaining from this gate is 2, corresponding to the cphase gate, as opposed to the bond dimension 4 of the original fSim gate. Note that in practice this identity is only approximate, since $\theta \approx \pi/2$; we find that transforming all gates described above causes a drop in fidelity to about 95%.

After the above simplification is applied, we proceed to cut (slice) some of the indexes of the TN (see Ref. [49] for details). The size of the slice of the index involved in each cut (the effective bond dimension of the index) is variable, and is chosen differently for different number of cycles on the circuit. Cutting indexes decomposes the contraction of the TN into several simpler contractions, whose results are summed after computing them independently on different nodes of the supercomputer.

Fig. S43 shows the bond dimensions of the TN corresponding to the circuits with 12 and 14 cycles simulated. We can see the decrease in bond dimension after the fSim simplification is applied, as well as the remaining bond dimension on the indexes cut for each case.

Finally, we contract the tensor network corresponding to the computation of a set of amplitudes (for fast sampling) for a particular batch of output bitstrings. The contraction ordering, which is chosen (together with the size and position of the cuts) in order to minimize the time-to-solution of the computation (which involves a careful consideration of the memory resources used and the efficiency achieved on the GPUs) is shown in Fig. S44. The computation can be summarized in the following pseudo-code, where α , β , and γ are variables that denote the different instances of the cuts:

```
# Qubits on C are used for fast sampling.
# size_of_batch amps. per circuit contraction.
size_of_batch = 2**num_qubits(C)

# Placeholder for all amplitudes in the batch.
batch_of_amplitudes = zeros(size_of_batch)

# Start contracting...
contract(A) # Panel 2
contract(B) # Panel 2
contract(C) # Panel 2

# alpha labels instances of 1st cut
for each alpha {

    # beta labels instances of 2nd cut
    for each beta {
        contract(D) # Panels 3 & 4
        contract(E) # Panels 5 & 6

        # gamma labels instances of 3rd cut
        for each gamma {
            contract(F) # Panels 7 & 8

            # Add contribution from this
            # path (alpha, beta, gamma).
            batch_of_amplitudes += F
        }
    }
}
```

Dotted qubits on Fig. S44 denote the region used for fast sampling, where output indexes are left open. The

circuit TN contraction leads to the computation of 64 amplitudes of correlated bitstrings (tensor F). Note that computing only a fraction \mathcal{F} of the paths results in amplitudes with a fidelity roughly equal to \mathcal{F} . Computing a set of perfect fidelity batches of amplitudes, where the number of batches is smaller than the number of bitstrings to sample also provides a similar fidelity \mathcal{F} in the sampling task, where \mathcal{F} is equal to the ratio of the number of batches to the number of bitstrings in the sample. A hybrid approach (fraction of batches, each only with a fraction of paths), which we use in practice, also provides a similar sampling fidelity. See Refs. [37, 49] and Section X A for more details.

A new feature of qFlex, implemented for this work, is the possibility to perform out-of-core tensor contractions (of tensors that exceed GPU memory) over more than one GPU on the same node. Although the arithmetic intensity requirements to achieve high efficiency are now higher (about an arithmetic intensity of 3000 for an efficiency close to 90% over three GPUs, as opposed to 1000 for a similar efficiency using a single GPU), the fact that a large part of a node is performing a single TN contraction lets us work with larger tensors, which implies reducing the number of cuts, as well as increasing the bond dimension of each cut; this, in turn, achieves better overall time-to-solution for sampling than simulations based on TNs with smaller tensors and with a lower memory footprint during their contraction (which could perhaps show a higher GPU efficiency due to the simultaneous use of each GPU for independent TNs). It is worth noting that the TN contraction ordering presented in Fig. S44 provides us with the best time-to-solution after considering several possibilities for the simulation of sampling from Sycamore using qFlex for both 12 and 14 cycles. This is generally not the case, since different numbers of cycles generate different TNs, which generally have different contraction schemes for best simulation time-to-solution.

Sampling of random circuits on Sycamore is difficult to simulate with TN simulators at 16 cycles and beyond. Indeed, FA simulators suffer from an exponential scaling of runtime with circuit depth. For qFlex, this is manifested in the large size of the tensors involved in the circuit TN contraction (this size grows exponentially with the number of cycles of the circuit), which require a large number of cuts in order not to exceed the RAM of a computation node, and which in turn generates an impractical number of Feynman paths. For other simulators, such as the one presented in Ref. [70], the number of projected variables is expected to be so large that the computation time (which increases exponentially with the number of projected variables) on a state-of-the-art supercomputer makes the computation impractical; see Section X E for a detailed analysis. For TN-based simulators that attempt the circuit contraction distributed over several nodes (without cuts) [71], we expect the size of the largest tensor encountered during the

TN contraction (which grows exponentially with depth) to exceed the RAM available on any current supercomputer. Not having enough memory for a simulation is the problem that led to developing FA simulators in the first place, for circuits of close to 50 qubits and beyond, for which the Schrödinger simulator (see Section XC) requires more memory to store the wave function than available. FA simulators give best performance as compared to other methods in situations with a large number of qubits and low depth. For circuits where both the number of qubits and the number of cycles are considered large enough to make the computation expensive, and contribute equally in doing so (formally, each linear dimension of the qubit grid is comparable to the time dimension), like the supremacy circuits considered in this work, we expect SFA of Section XA to be the leading approach for sampling from a random circuit, given a large enough sample size ($\sim 1\text{M}$ in this work); note the linear dependence of the runtime of FA with sample size, which is absent for SFA.

C. Supercomputer Schrödinger simulator

We also performed supercomputer Schrödinger simulations in the Jülich Supercomputing Centre. For a comprehensive description of the universal quantum computer simulators JUQCS-E and JUQCS-A, see Refs. [72] and [73].

For a given quantum circuit U designed to generate a random state, JUQCS-E [73] executes U and computes (in double precision floating point) the probability distribution $p_U(j)$ for each output or bitstring $j \in \{0, \dots, D-1\}$, where $D = 2^n$, n denoting the number of qubits. JUQCS-E can also compute (in double precision floating point) the corresponding distribution function $P_U(k) = \sum_{j=0}^k p_U(j)$ and sample bitstrings from it. We denote by \mathcal{U} the set of m states generated by executing the circuit U . A new feature of JUQCS-E, not documented in Ref. 73, allows the user to specify a set \mathcal{Q} of M bitstrings for which JUQCS-E calculates $p_U(j)$ for all $j \in \mathcal{Q}$ and saves them in a file.

Similarly, for the same circuit U , JUQCS-A [73] computes (with adaptive two-byte encoding) the probability distribution $p_A(j)$ for each bitstring $j \in \{0, \dots, D-1\}$. Although numerical experiments with Shor's algorithm for up to 48 qubits indicate that the results produced by JUQCS-A are sufficiently accurate, there is, in general, no guarantee that $p_A(j) \approx p_U(j)$. In this sense, JUQCS-A can be viewed as an approximate simulator of a quantum computing device.

In principle, sampling states with probabilities $p_A(j)$ requires the knowledge of the distribution function $P_A(k) = \sum_{j=0}^k p_A(j)$. If D is large, and $p_A(j) \approx O(1/D)$, as in the case of random states, computing $P_A(k)$ requires the sum over j to be performed with sufficiently high precision. For instance, if $D = 2^{39}$, $p_A(j) \approx O(10^{-12})$ and even with double precision arithmetic (≈ 16 dig-

its), adding $D = 2^{39}$ small numbers requires some care. Note that in practice, each MPI process only calculates a partial sum, which helps to reduce the loss of significant digits. JUQCS-A can compute $P_A(k)$ in double precision and sample bitstrings from it. We denote by \mathcal{A} the set of M bitstrings generated by JUQCS-A after executing the circuit U . Activating this feature requires additional memory, effectively reducing the maximum number of qubits that can be simulated by three. This reduction of the maximum number of qubits might be avoided as follows. In the case at hand, we know that all $p_A(j) \approx O(1/D)$. Then, since $p_A(j)$ is known, one might as well sample the states from a uniform distribution, list the weight $w_A(j) = N p_A(j)$ for each generated state j and use these weights to compute averages. We do not pursue this possibility here because for the present purpose, it is essential to be able to compute $p_U(j)$ and therefore, the maximum number of qubits that can be studied is limited by the amount of memory that JUQCS-E, not JUQCS-A, needs to perform the simulation.

For an XEB comparison, the quantities of interest are

$$\alpha_{U,U} \equiv \log D + \gamma + \sum_{j=0}^{D-1} p_U(j) \log p_U(j), \quad (84)$$

$$\alpha_{A,U} \equiv \log D + \gamma + \sum_{j=0}^{D-1} p_A(j) \log p_U(j), \quad (85)$$

$$\alpha_{A,A} \equiv \log D + \gamma + \sum_{j=0}^{D-1} p_A(j) \log p_A(j), \quad (86)$$

$$\alpha_{\mathcal{X},U} \equiv \log D + \gamma + \frac{1}{M} \sum_{j \in \mathcal{X}} \log p_U(j), \quad (87)$$

where \mathcal{X} is one of the four sets \mathcal{U} , \mathcal{A} , \mathcal{M} (a collection of bitstrings generated by the experiment), or \mathcal{C} (obtained by generating bitstrings distributed uniformly). If M is sufficiently large ($M = 500000$ in the case at hand), we may expect that $\alpha_{U,U} \approx \alpha_{U,U}$ and $\alpha_{A,U} \approx \alpha_{A,U}$.

In addition to the cross entropies Eqs. (84)–(87), we also compute the linear cross entropies

$$\hat{\alpha}_{U,U} \equiv \sum_{j=0}^{D-1} p_U(j) (D p_U(j) - 1), \quad (88)$$

$$\hat{\alpha}_{A,U} \equiv \sum_{j=0}^{D-1} p_A(j) (D p_U(j) - 1), \quad (89)$$

$$\hat{\alpha}_{A,A} \equiv \sum_{j=0}^{D-1} p_A(j) (D p_A(j) - 1), \quad (90)$$

$$\hat{\alpha}_{\mathcal{X},U} \equiv \frac{1}{M} \sum_{j \in \mathcal{X}} (D p_U(j) - 1). \quad (91)$$

Table VIII presents simulation results for the α 's defined by Eqs. (84)–(87) and for the $\hat{\alpha}$'s defined by Eqs. (88)–(91), obtained by running JUQCS-E and JUQCS-A on the supercomputers at the Jülich Supercomputer Centre. For testing quantum supremacy using

these machines, the maximum number of qubits that a universal quantum computer simulator can handle is 43 (45 on the Sunway TaihuLight at Wuxi China [73]).

The fact that in all cases, $\alpha_{U,U} \approx \alpha_{A,A} \approx 1$ supports the hypothesis that the circuit U , executed by either JUQCS-E or JUQCS-A, produces a Porter-Thomas distribution. The fact that in all cases, $\alpha_{U,U} \approx 1$ supports the theoretical result that replacing the sum over all states by the sum over $M = 500000$ states yields an accurate estimate of the former (see Section IV). Although $\alpha_{A,A} \approx 1$ in all cases, using the sample \mathcal{A} generated by JUQCS-A to compute $\alpha_{A,U}$ shows an increasing deviation from one, the deviation becoming larger as the number of qubits increases. In combination with the observation that $\alpha_{A,A} \approx 1$, this suggests that JUQCS-A produces a random state, albeit not the same state as JUQCS-E. Taking into account that JUQCS-A stores the coefficients of each of the basis states as two single-byte numbers and not as two double precision floating point numbers (as JUQCS-E does), this is hardly a surprise.

From Table VIII it is clear that the simulation results for $\alpha_{\mathcal{X},U}$ and $\hat{\alpha}_{\mathcal{X},U}$ where $\mathcal{X} = \mathcal{A}, \mathcal{M}, \mathcal{C}$ are consistent. The full XEB fidelity estimates $\alpha_{\mathcal{M},U}$ and $\hat{\alpha}_{\mathcal{M},U}$, that is the values computed with the bitstrings produced by the experiment, are close to the fidelity estimates of the probabilistic model, patch XEB, and elided XEB, as seen in Fig. 4(a) of the main text.

For reference, in Tables IX and X we present some technical information about the supercomputer systems used to perform the simulations reported in this appendix and give some indication of the computer resources used.

D. Simulation of random circuit sampling with a target fidelity

A classical simulator can leverage the fact that experimental sampling from random circuits occurs at low fidelity \mathcal{F}_{XEB} by considering only a small fraction of the Feynman paths (see Secs. X A and X B) involved in the simulation [37], which provides speedups of at least a factor of $1/\mathcal{F}_{\text{XEB}}$. This is done by Schmidt decomposing a few two-qubit gates in the circuit and counting only a fraction of their contributing terms (*paths*). A key assumption here is that the different paths result in orthogonal output states, as was studied in Ref. [37] and later in Ref. [49]. In what follows, we argue that, provided the generation of paths through decomposing gates, the Schmidt decomposition is indeed the optimal approach to achieving the largest speedup, *i.e.*, that the fidelity kept by considering only a fraction of paths is largest when keeping the paths with the largest Schmidt coefficient. This is different from proving the optimality of the Schmidt decomposition of a single gate, since here we refer to the fidelity of the entire output state, and decomposed gates are embedded in a much larger circuit. In addition, we show that, for the two-qubit gates used in this work, the speedup is very close to linear in \mathcal{F}_{XEB}

(and not much larger), since their Schmidt spectrum is close to flat. We close this section by relating the present discussion to Section VII G 2, where the formation of simplifiable gate patterns in some two-qubit gate tilings of the circuit is introduced.

In summary, this section provides a method to simulate approximate sampling with a classical computational cost proportional to \mathcal{F}_{XEB} . Sec. XI argues, based on complexity theory, that this scaling is optimal. We note that Refs [76–78] propose an alternative method to approximately sample the output distribution at low fidelity. In essence, this method relies on the observation that, for some noise models, the high weight Fourier components of the noisy output distribution decay exponentially to 0. Then this method proposes to estimate low weight Fourier components with an additive error which is polynomial in the computational cost. Nevertheless, Ref. [79] shows that all Fourier components of the output distribution of random circuits are exponentially small, and therefore they can not be estimated in polynomial time with this method. The conclusion is then that the noisy output distribution can be approximated by sampling bitstrings uniformly at random, the distribution for which all Fourier components are 0. This is consistent with Ref. [26] and Secs. IV and VIII E, but it will produce a sample with $\mathcal{F}_{\text{XEB}} = 0$, while the output of the experimental samples at 53 qubits and $m = 20$ still has $\mathcal{F}_{\text{XEB}} \geq 0.1\%$

1. Optimality of the Schmidt decomposition for gates embedded in a random circuit

Consider a two-qubit gate V_{ab} acting on qubits a and b . We would like to replace it by a tensor product operator $M_a \otimes N_b$. The final state of the ideal circuit is

$$|\psi\rangle := U_2 V_{ab} U_1 |0^n\rangle \quad (92)$$

where $U_1(U_2)$ is a unitary composed by all the gates applied before (after) V_{ab} . The final normalized state of the circuit with the replacement by $M_a \otimes N_b$ is

$$|\phi_{M,N}\rangle := U_2(M_a \otimes N_b)U_1|0^n\rangle / \|U_2(M_a \otimes N_b)U_1|0^n\rangle\|. \quad (93)$$

We would like to find M, N which maximize the fidelity of the two states, given by

$$\langle\psi|\phi_{M,N}\rangle = \langle 0^n|U_1^\dagger V_{ab}^\dagger|\beta\rangle / \sqrt{\langle\beta|\beta\rangle}, \quad (94)$$

where

$$|\beta\rangle \equiv (M_a \otimes N_b)U_1|0^n\rangle \quad (95)$$

As the overlap is invariant if we multiply $(M_a \otimes N_b)$ by a constant, we fix the normalization $\text{tr}[(M_a \otimes N_b)^\dagger(M_a \otimes N_b)] = 1$.

We now make the assumption that the circuit is random (or sufficiently scrambling) and that the V_{ab} is a gate

TABLE VIII. Simulation results for various α 's as defined by Eqs. (84)–(87), obtained by JUQCS-E and JUQCS-A. The results for the $\hat{\alpha}$'s defined by Eqs. (88)–(91) are given in parenthesis. The set of bitstrings \mathcal{M} has been obtained from experiments. In the first column, the number in parenthesis is the circuit identification number. Horizontal lines indicate that data is not available (and would require additional simulation runs to obtain it).

qubits	$\alpha_{U,U}$	$\alpha_{A,A}$	$\alpha_{U,U}$	$\alpha_{A,U} (\hat{\alpha}_{A,U})$	$\alpha_{M,U} (\hat{\alpha}_{M,U})$	$\alpha_{C,U} (\hat{\alpha}_{C,U})$
30	1.0000	1.0000	0.9997	0.8824 (0.8826)	0.0708 (0.0711)	+0.0026 (+0.0017)
39(0)	1.0000	1.0000	0.9992	0.4746 (0.4762)	0.0281 (0.0261)	−0.0003 (−0.0011)
39(1)	1.0000	1.0000	1.0002	—; (—)	0.0350 (0.0362)	—; (—)
39(2)	1.0000	1.0000	0.9996	—; (—)	0.0351 (0.0332)	—; (—)
39(3)	1.0000	1.0000	0.9999	—; (—)	0.0375 (0.0355)	—; (—)
42(0)	1.0000	1.0001	0.9998	0.4264 (0.4268)	0.0287 (0.0258)	−0.0024 (−0.0001)
42(1)	1.0000	1.0000	1.0027	—; (—)	0.0254 (0.0273)	—; (—)
43(0)	1.0000	1.0001	1.0013	0.3807 (0.3784)	0.0182 (0.0177)	−0.0010 (−0.0003)
43(1)	1.0000	1.0000	—	—; (—)	0.0217 (0.0204)	—; (—)

TABLE IX. Specification of the computer systems at the Jülich Supercomputing Centre used to perform all simulations reported in this appendix. The row “maximum # qubits” gives the maximum number of qubits n that JUQCS-E (JUQCS-A) can simulate on a specific computer.

Supercomputer	JURECA-CLUSTER [74]	JURECA-BOOSTER [74]	JUWELS [75]
CPU	Intel Xeon E5-2680 v3 Haswell	Intel Xeon Phi 7250-F Knights Landing	Dual Intel Xeon Platinum 8168
Peak performance	1.8 PFlop/s	5 PFlop/s	10.4 PFlops/s
Clock frequency	2.5 GHz	1.4 GHz	2.7 GHz
Memory/node	128 GB	96 GB + 16 GB (MCDRAM)	96 GB
# cores/node	2 × 12	64	2 × 24
# threads/core used	1	1	3
maximum # nodes used	256	512	2048
maximum # MPI processes used	4096	32768	32768
maximum # qubits	40 (43)	41 (44)	43 (46)

placed sufficiently in the middle of the computation that the reduced density matrix of qubits a and b of $U_1|0^n\rangle$ shows maximal mixing between the two. In more detail, let

$$\varepsilon := \left\| \text{tr}_{\setminus(a,b)}(U_1|0^n\rangle\langle 0^n|U_1^\dagger) - \frac{I}{4} \right\|_2, \quad (96)$$

with $\|X\|_2 := \text{tr}(X^\dagger X)^{1/2}$ the Hilbert-Schmidt norm and $\text{tr}_{\setminus(a,b)}$ the partial trace of all qubits except a and b .

Using Eq. (96) and Eq. (94), we find

$$\begin{aligned} \langle \psi | \phi_{M,N} \rangle &= \text{tr}(\text{tr}_{\setminus(a,b)}(U_1|0^n\rangle\langle 0^n|U_1^\dagger)V_{ab}^\dagger(M_a \otimes N_b)) \quad (97) \\ &= \frac{1}{4} \text{tr}[V_{ab}^\dagger(M_a \otimes N_b)] \pm \|(M_a \otimes N_b)\|_2 \|V_{ab}\|_2 \varepsilon. \end{aligned}$$

As $\|(M_a \otimes N_b)\|_2 = 1$ and $\|V_{ab}\|_2 = 2$, we find

$$\langle \psi | \phi_{M,N} \rangle = \frac{1}{4} \text{tr}[V_{ab}^\dagger(M_a \otimes N_b)] \pm 2\varepsilon. \quad (98)$$

Refs. [80, 81] proved that for a random circuit U_1 of depth D in one dimension, $\varepsilon \leq (4/5)^D$. In two dimensions we expect ε to go to zero even faster with depth, so

we can ignore the second term of Eq. (98) for sufficiently large depth.

We now want to find M_a , N_b which are optimal for

$$\max_{M_a, N_b: \|M_a\|_2 = \|N_b\|_2} \text{tr}[V_{ab}^\dagger(M_a \otimes N_b)]. \quad (99)$$

At this point, we have reduced the problem to finding the optimal decomposition of the gate as a standalone operator.

Consider the operator Schmidt decomposition of V_{ab} :

$$V_{ab} = \sum_i \lambda_i R_{a,i} \otimes S_{b,i}, \quad (100)$$

where $R_{a,i}$ ($S_{b,i}$) are orthonormal set of operators in the Hilbert-Schmidt inner product, i.e. $\text{tr}(R_{a,i}^\dagger R_{a,j}) = \text{tr}(S_{a,i}^\dagger S_{a,j}) = \delta_{ij}$. The Schmidt singular values $\lambda_1 \geq \lambda_2 \geq \dots$ are in decreasing order. Then it follows that the solution of Eq. (99) is λ_1 , with optimal solution $M_a = R_{a,1}$ and $N_b = S_{b,1}$. Indeed we can write Eq. (99) as

$$\max_{|x\rangle, |y\rangle} \langle x | \bar{V} | y \rangle \quad (101)$$

TABLE X. Representative elapsed times and number of MPI processes used to perform simulations with JUQCS-E and JUQCS-A on the supercomputer indicated. Note that the elapsed times may fluctuate significantly depending on the load of the machine/network.

qubits	gates	JUQCS-E			JUQCS-A		
		Supercomputer	MPI processes	Elapsed time	Supercomputer	MPI processes	Elapsed time
30	614	BOOSTER	128	0:02:28	CLUSTER	128	0:05:23
39	802	CLUSTER	4096	0:42:51	CLUSTER	4096	1:38:42
42	864	JUWELS	16384	0:51:16	JUWELS	8192	2:15:48
43	886	JUWELS	32768	1:01:53	JUWELS	32768	1:32:19

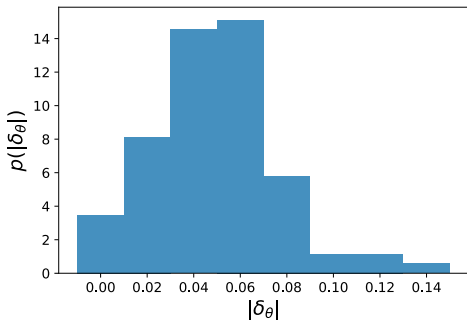


FIG. S45. **Probability distribution of the deviations $|\delta_\theta|$ from $\theta \approx \pi/2$ for fSim gates.** The magnitude of δ_θ is directly related to the runtime speedup low fidelity classical sampling can take from exploiting the existence of paths with large Schmidt coefficients. In practice, $|\delta_\theta| \approx 0.05$ radians on average, which imposes a bound of less than an order of magnitude on this potential speedup for the circuits, gates, and simulation techniques considered in this work.

where the maximum is over all unit vectors $|x\rangle, |y\rangle$ in $(\mathbb{C}^2)^{\otimes 2}$ and \bar{V} is the matrix

$$\bar{V} := \sum_i \lambda_i (R_{a,i} \otimes I) |\Phi\rangle \langle \Phi| (S_{b,i}^\dagger \otimes I) \quad (102)$$

with $|\Phi\rangle = \sum_i |i\rangle \otimes |i\rangle$. This can be verified using the fact that any unit vector $|x\rangle$ in $(\mathbb{C}^d)^{\otimes 2}$ can be written as $|x\rangle = (L \otimes I) |\Phi\rangle$ for a matrix L acting on (\mathbb{C}^d) s.t. $\|L\|_2 = 1$. The result follows by noting that λ_i are the singular values of \bar{V} .

The argument above easily generalizes to the problem of finding the optimal operator of Schmidt rank k for replacing the unitary gate. In that case the optimal choice is $\sum_{i=1}^k \lambda_i R_{a,i} \otimes S_{b,i}$.

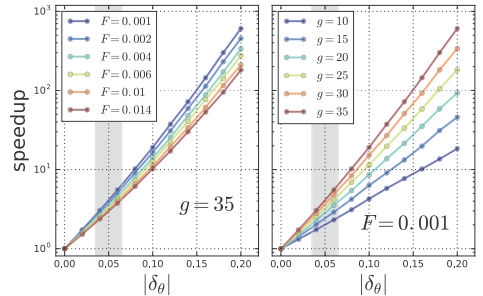


FIG. S46. **Classical speedup given by the imbalance in the Schmidt coefficients of the gates decomposed.** The speedup is computed by comparison with the case where $\theta = \pi/2$ exactly. The classical simulation has a target fidelity \mathcal{F} , and g fSim gates are decomposed. For simplicity, we assume $\theta = \pi/2 + \delta_\theta$ is the same for all gates, as well as $\phi = \pi/6$. *Left:* speedup at different target fidelities for fixed $g = 35$. Note that the speedup decreases with \mathcal{F} ; this is due to the fact that at very low fidelity, considering a few paths with very high weight might be enough to achieve the target fidelity, while for larger values of \mathcal{F} , paths with a smaller weight have to be considered, and so a larger number of them is needed per fractional fidelity increase. *Right:* speedup for fixed fidelity $\mathcal{F} = 0.001$ for different values of g . As expected, the speedup is greater as g increases, since the weight of the highest contributing paths increases exponentially with g . The largest speedup is achieved at large g and small \mathcal{F} . For $g = 35$ and $\mathcal{F} \gtrsim 0.001$, we find speedups well below an order of magnitude, given that $|\delta_\theta| \approx 0.05$ radians in practice (shaded area); this case is representative of our simulation of Sycamore with $m = 20$ (see Section X A) targeting the fidelity measured experimentally.

2. Classical speedup for imbalanced gates

We now want to analyze the Schmidt spectrum of the two-qubit gates used in this work. The fSim(θ, ϕ) gate is introduced in Section VII E. This gate, which is presented in matrix form in Eq. (53), has the following Schmidt

singular values:

$$\lambda_1 = \sqrt{1 + 2 \cdot |\cos(\phi/2) \cos \theta| + \cos^2 \theta} \quad (103)$$

$$\lambda_2 = \lambda_3 = \sin \theta \quad (104)$$

$$\lambda_4 = \sqrt{1 - 2 \cdot |\cos(\phi/2) \cos \theta| + \cos^2 \theta}, \quad (105)$$

where normalization is chosen so that $\sum_i \lambda_i^2 = 4$. In practice, we have $\theta \approx \pi/2$ and $\phi \approx \pi/6$, and so we obtain $\lambda_i \approx 1$, $\forall i \in \{1, 2, 3, 4\}$, which gives a flat spectrum.

In the case that $\theta = \pi/2 \pm \delta_\theta$, the spectrum becomes imbalanced, as expected. When considering the decomposition of a number g of $\text{fSim}(\pi/2 \pm \delta_\theta, \phi \approx \pi/6)$ gates, the set of weights of all paths is equal to the outer product of all sets of Schmidt coefficients (one per gate). Achieving a fidelity $\mathcal{F}_{\text{XEB}} > 0$ implies (in the optimal case) including the largest contributing paths, and so the advantage one can get from this is upper bounded by the magnitude of the largest weight, which is equal to $\prod_{\alpha=1}^g \lambda_{\alpha, \max}^2$, where α labels the gates decomposed and $\lambda_{\alpha, \max}$ is the largest Schmidt coefficient for gate α . In practice, $|\delta_\theta|$ has values of around 0.05 radians (see Fig. S45). The geometric mean of λ_{\max} is about 1.047, which gives an upper bound of 1.047^{2g} to the speedup discussed here. For the largest value of g considered in this work, *i.e.*, the decomposition of $g = 35$ gates using the SFA simulator (Section X A) on a circuit of $m = 20$ cycles, we obtain a value of $1.047^{2 \times 35} = 25.4$. Note that the speedup obtained in practice (as compared to run-times over circuits with perfectly flat gate Schmidt decompositions) for fidelities of the order of 0.1% and larger is expected to be far smaller than this value, given that one has to consider a large number of paths, from which only an exponentially small number will have a weight close to 25.4.

We can get a better estimate for the speedup achieved in practice, beyond the upper bound of about a factor of 25 that decomposing $g = 35$ gates with typical parameters would give. For simplicity, let us assume that all g gates have the same values of θ and ϕ . Then the weight of each path arising from this decomposition can be written as $W_i = W_{(a,b,c)} = \lambda_1^{2a} \lambda_2^{2b} \lambda_3^{2c}$, where $a + b + c = g$, and that the number of paths for each choice of (a, b, c) is equal to $\#(a, b, c) = \sum_{k=0}^b \text{multinomial}(a, b-k, k, c) = 2^b \times \text{multinomial}(a, b, c)$. After sorting all 4^g weights (and paths) by decreasing value, given a target fidelity, \mathcal{F} , one now has to consider the first S paths (*i.e.*, those with the largest weight), up to the point where the sum of their weights $\sum_{i=1}^S \frac{W_i}{4^g}$ matches the target fidelity. The normalization factor 4^g guarantees that if one were to consider all paths, the fidelity would be unity, as expected. Compared to the case where we consider a number $\mathcal{F} \times 4^g$ of paths, as for a flat Schmidt spectrum, this provides a speedup equal to $\frac{S}{\mathcal{F} \times 4^g}$. We show the speedup achieved this way in Fig. S46. For the case where we would achieve the largest speedup in the simulations considered in this work, namely the simulation of Sycamore at $m = 20$ cycles and a fidelity $\mathcal{F} \approx 0.2\%$ with $g = 35$ gates decomposed (see Section X F), we estimate that the speedup

obtained this way would be well below an order of magnitude, since $|\delta_\theta|$ typically takes values of about 0.05 radians.

3. Verifiable and supremacy circuits

So far we have considered the decomposition of gates one by one, *i.e.*, where the total number of paths is equal to the product of the Schmidt rank of all gates decomposed. However, by fusing gates together in a larger unitary, one can provide some speedup to the classical simulation of the sampling task.

The rationale here comes from the realization that a unitary that involves a number of qubits q cannot have a rank larger than $4^{\min(q_l, q_r)}$ when Schmidt decomposed over two subsets of qubits of size q_l and q_r , with $q_l + q_r = q$. Therefore one might reduce exponentially the number of paths by fusing gates such that the resulting unitary reaches on either side (l or r) a number of qubits that is smaller than the product of the ranks of the fused gates to be decomposed. This is at the heart of the formation of *wedges* of Section VII G 2. These wedges denote particular sequences of consecutive two-qubit gates that only act upon three qubits. Fusing these two-qubit gates together generates 4 paths, as opposed to a naive count of 4^2 paths if one decomposes each gate separately. Each wedge identified across a circuit cut provides a speedup by a factor of 4.

In this work, we define two classes of circuits: *verifiable* and *supremacy* circuits. Verifiable circuits present a large number of wedges across the partition used with the SFA simulator (Section X A) and are therefore classically simulatable in a reasonable amount of time. These circuits were used to perform full XEB over the entire device up to depth $m = 14$ (see Fig. 4a of the main article and Sections VII and VIII), which involves perfect fidelity computations. On the other hand, supremacy circuits are designed so that the presence of wedges and similar sequences is mitigated, therefore avoiding the possibility of exploiting this classical speedup.

It is natural to apply the ideas presented here beyond *wedges*. It is also easy to look for similar structures in the circuits algorithmically. This way, we find that for the supremacy circuits there is a small number of such sequences. On the sequence of cycles DCD (see Fig. S25), three two-qubit gates are applied on qubits 16, 47, and 51 (see Fig. S27 for numbering). These three gates can be fused in one. Then, if the two gates between qubits 47 and 51 are decomposed (as is done with the SFA simulations of Section X A used in Fig. 4 of the main article), this technique provides a speedup of a factor of 4. The sequence of layouts DCD appears twice for circuits of $m = 20$, which provides a total speedup of $4^2 = 16$ in the simulation of the supremacy circuits. This particular decomposition is currently not implemented, and the estimated timings of Section X A and Fig. 4 of the main article do not take it into account.

Beyond this, one has to go to groups of several cycles of the circuit (more than two) in order to identify regions where the fusion of several gates provides any advantage of this kind. In our circuits, the resulting unitaries act upon a large number of qubits, which makes explicitly building the unitary impractical.

E. Treewidth upper bounds and variable elimination algorithms

We explained in Section XB that the Feynman method to compute individual amplitudes of the output of a quantum circuit can be implemented as a tensor network when quantum gates are interpreted as tensors. All indexes of the tensor network have dimension two because indexes correspond to qubits. Similarly, Ref. [68] showed that a quantum circuit can be mapped directly to an undirected graphical model. In the undirected graphical model, vertices or variables correspond to tensor indexes, and cliques correspond to tensors. Individual amplitudes can be computed using a variable elimination algorithm on the undirected graphical model, which is similar to a tensor contraction on a tensor network. The variable elimination algorithm depends on the ordering in which variables are eliminated or contracted. If we define the *contraction width* of an ordering to be the rank of the largest tensor formed along the contraction, the *treewidth* of the undirected graph is equal to the minimum contraction width over all orderings. Therefore, the complexity of a tensor network contraction grows in the optimal case exponentially with the treewidth, and the treewidth can be used to study the complexity of Feynman methods for simulating quantum circuits [67]. Ref. [68] showed that for diagonal gates the undirected graphical model is simpler, potentially lowering its treewidth, and hence improving the complexity. This simplification is not achievable in the tensor network view without including hyperedges, *i.e.*, edges attached to more than two tensors. Ref. [68] also introduced the use of QuickBB to find a heuristic contraction ordering [82]. If allowed to run for long enough, QuickBB finds the optimal ordering, together with the treewidth of the graph. However, note that obtaining the treewidth of a graph is an NP-hard problem, and so in practice a suboptimal solution is considered for the simulations described here.

Once the width of a contraction is large enough, the largest tensor it generates is beyond the memory resources available. This constraint was overcome in Ref. [70] by *projecting* a subset of p variables or vertices in the undirected graphical model into each possible bistring of 0 and 1 values. This generates 2^p similar subgraphs, each of which can be contracted with lower complexity and independently from each other, making the computation embarrassingly parallelizable. Choosing the subset of variables that, after projection, opti-

mally decreases the treewidth of the resulting subgraph is also NP-hard. However, Ref. [70] developed a heuristic approach that works well in practice. The algorithm proceeds as follows:

1. Run QuickBB for S seconds on the initial graph. This gives a heuristic contraction ordering, as well as an upper bound for the treewidth.
2. For each variable, estimate the cost of contracting the subgraph after projection. The estimate is done with the ordering inherited from the previous step.
3. Choose to project the variable which results in the minimum contraction cost.
4. Repeat steps 2 and 3 until the cost is within reasonable resources.
5. Once all variables have been chosen and projected, run QuickBB for S seconds on the resulting subgraph to try to improve the contraction ordering inherited from step 1 and lower the contraction cost.

In the top panel of Fig. S47 we show the contraction width as a function of the number of variables that are projected for the supremacy circuits used in this paper. In order to decrease the contraction width to 28 or below (a tensor with 28 binary indexes consumes 2 GB of memory using single precision complex numbers), we need to project between 8 and 63 variables, depending on the depth of the circuits. In addition, we report the result of the projection procedure on the Bristlecone circuits considered in Refs. [49, 83] and available at <https://github.com/sboixo/GRCS> for depths $(1+32+1)$ and $(1+40+1)$, since these cases were benchmarked in Ref. [83]. We obtain a contraction width equal to 28 after 10 projections for Bristlecone at depth $(1+32+1)$, and width 26 after 22 projections for Bristlecone at depth $(1+40+1)$, consistent with the results in Ref. [83]. Even though Ref. [70] uses $S = 60$, we run QuickBB for 1800 seconds (30 minutes) every time, in order to decrease the contraction width of the Bristlecone simulations to values that match the memory requirements reported in Ref. [83]. Note that Ref. [83] neither reports the value of S used nor the contraction widths found; however, with $S = 1800$ we are able to match the scaling of time complexity reported, as is explained below.

To estimate the runtime of the computation of a single amplitude using this algorithm on the circuits presented in this work, we use the following scaling formula:

$$T_{\text{VE}} = C_{\text{VE}}^{-1} \cdot 2^p \cdot (\text{cost after } p \text{ projections}) / n_{\text{cores}}, \quad (106)$$

where VE refers to the variable elimination algorithm with projections described in this section, C_{VE} is a constant factor, p is the number of variables projected, and n_{cores} is the number of cores used in the computation.

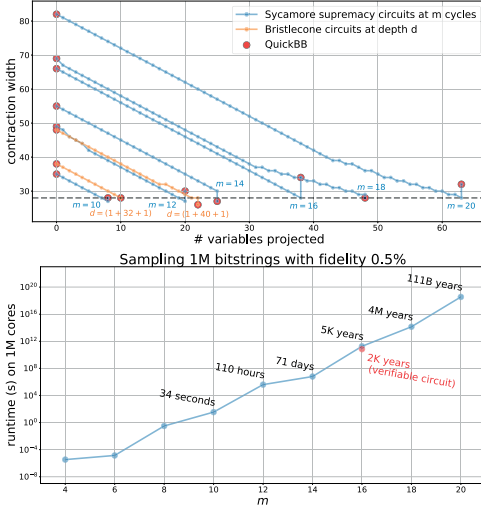


FIG. S47. Contraction widths and estimated runtimes for classical sampling using the variable elimination algorithm with projected variables of Ref. [70] for Sycamore supremacy circuits. *Top*: contraction width as a function of the number of variables projected using the algorithm of Ref. [70]. We project enough variables in order to decrease the width to 28 or lower. Note that often the second QuickBB run does not decrease the treewidth (and might even increase it), in which case the resulting contraction ordering it is ignored. *Bottom*: estimated runtimes for the classical sampling of 1M bitstrings from the supremacy circuits with fidelity 0.5% using the contraction ordering found by QuickBB at the end of the projection procedure shown in the top panel. The red data point shows the estimated runtime for a verifiable circuit; note that the heuristic algorithm analyzed here provides some speedup in this case. Our time estimates assume the use of fast sampling, although it is so far unclear whether this technique can be adapted to the algorithm described here. Failure to do so would result in a slowdown of about an order of magnitude.

The cost of the full contraction of each subgraph is estimated as the sum of 2^{rank} , where the rank refers to the number of variables involved in each individual contraction along the full contraction of the subgraph. We obtain the value of C_{VE} from the runtimes reported in Ref. [70], which shows that a single amplitude of Bristlecone at depth $(1+32+1)$ takes 0.43 seconds to compute on 127,512 CPU cores with 10 projected variables, and at depth $(1+40+1)$ it takes 580.7 seconds with 22 projected variables using the same number of cores. We use the benchmark at depth $(1+32+1)$ because it provides the largest value for C_{VE} (lowest time estimates), which is equal to 52.7 MHz; the benchmark at depth $(1+40+1)$ gives $C_{VE} = 51.6$ MHz. In order to sample 1M bitstrings

from a random circuit with fidelity 0.5%, we need to compute 5000 amplitudes.

We present our estimates for Sycamore supremacy circuits in the bottom panel of Fig. S47. Note that depth $(1+40+1)$ in Refs. [70, 83] is equivalent to $m=20$ cycles here because of the denser layout of two-qubit gates. Furthermore, computation times reported previously are for circuit variations less complex than for Sycamore, arising from changes in complexity such as CZ vs. fSim gates and differing patterns; with this change of gates, depth $(1+40+1)$ in Refs. [70, 83] is actually equivalent to $m=10$ cycles here. Finally, note that we present optimistic estimates, since we are assuming that the fast sampling technique discussed in Section X B is applicable here. To the best of our knowledge, it is not known how to apply this technique for the heuristic variable elimination algorithm discussed here; in the absence of an implementation of this technique, in order to successfully apply rejection sampling we would instead need to compute a few independent amplitudes per sampled bitstring, which would increase the estimated times by about an order of magnitude (see Section X B and Refs. [37, 84] for more details). According to our estimates, sampling from supremacy circuits at $m = 16$ and beyond is out of reach for this algorithm. Interestingly, we find some speedup for the simulation of verifiable circuits, as is shown in Fig. S47 for $m = 16$ (red data point).

Finally, note that the undirected graphical model derived from the supremacy circuits can take advantage of the structure of the Sycamore gates (fSim plus single-qubit R_z rotations). Due to the fact that $\text{fSim}(\theta \approx \pi/2, \phi) \approx -i \cdot [R_z(-\pi/2) \otimes R_z(-\pi/2)] \cdot \text{cphase}(\pi + \phi) \cdot \text{SWAP}$, the Sycamore gate corresponds to a subgraph of only two variables, which explicitly represents the diagonal cphase and the logical SWAP. This simplification, used in our estimates, results in an undirected graphical model that is simpler than that one generated by arbitrary two-qubit gates. See Fig. S48 for an example.

F. Computational cost estimation for the sampling task

We find that the most efficient simulator for our hardest circuits is the SFA simulator (see Sec. X A). In order to estimate the computational cost associated with simulating a 53 qubit circuit with 20 cycles, where no gates are elided on the cut, we use a Google cloud cluster composed of 1000 machines with 2 vCPUs and 7.5 GB of RAM each (n1-standard-2). We use n1-standard-2 because this is the smallest non-custom machine with sufficient RAM for simulating the two halves of the circuit. In 20 cycles, the circuit contains 35 gates across the cut. All cross gates have a Schmidt rank of 4 except for the last four gates which can be simplified to cphase with a Schmidt rank of 2. To obtain a perfect fidelity simulation

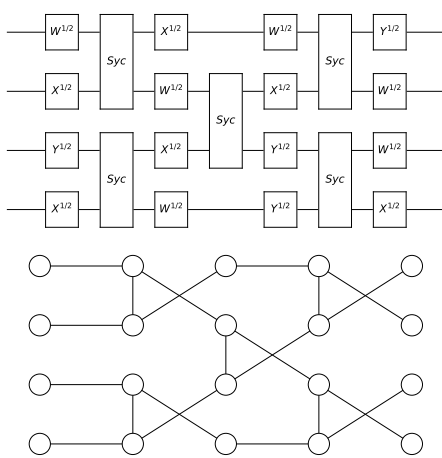


FIG. S48. Circuit with Sycamore gates (top) and its corresponding undirected graphical model (bottom). Each non-diagonal single-qubit gate introduces a new vertex or variable. Note that, even though two-qubit gates are generally represented by a clique with four vertices or variables, Sycamore gates can be simplified as a cphase followed by a SWAP. The cphase is represented as an edge between two existing variables. The SWAP, however, provides more complexity to the graph as it swaps the corresponding variables.

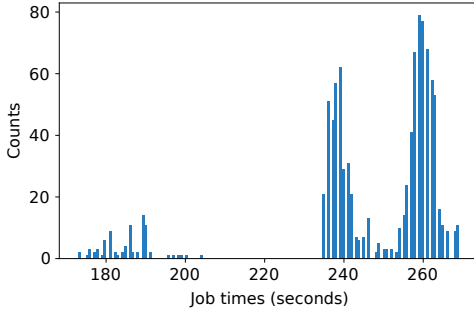


FIG. S49. Qsimh execution time for a 53 qubit circuit with 20 cycles for the first 1000 prefix values. The average job time (t_{prefix}) is calculated to be 246 seconds.

we would need to simulate all $4^{31} \times 2^4$ paths. We configure qsimh according to Ref. [37] to have a prefix of 30 cross gates, thus requiring 4^{30} separate qsimh runs. The first 1000 paths of the required 4^{30} were used for timing purposes. In Figure S49 we plot the distribution of simulation times with qsimh consuming two hyperthreads. The average job time is 246 seconds resulting in a calculated 1.6×10^{14} core hours for a simulation of the circuit

qubits, n	cycles, m	total #paths	fidelity	run time
53	12	$4^{17}2^4$	1.4%	2 hours
53	14	$4^{21}2^4$	0.9%	2 weeks
53	16	$4^{25}2^3$	0.6%	4 years
53	18	$4^{28}2^3$	0.4%	175 years
53	20	$4^{31}2^4$	0.2%	10000 years

TABLE XI. Approximate qsimh run times using one million CPU cores extrapolated from the average simulation run time for 1000 simulation paths on one CPU core.

with 0.002 fidelity [85]. Extrapolated run times for other circuits with 53 qubits are shown in Table XI. To calculate a total cost for the largest circuit we multiply the Google Cloud preemptible n1-standard-2 price in zone us-central-1 of \$0.02 per hour, 246 seconds average run time, 0.002 target fidelity, and 4^{30} qsimh runs. This results in an estimated cost of 3.1 trillion USD. For perfect fidelity simulations (necessary for XEB), an extrapolation to a fidelity value of 100% gives a good estimate of the run time. We believe these estimates are a lower bound on costs and simulation time due to the fact that these calculations are likely to compete with each other if they are run on the same nodes.

As a final remark, note that a hypothetical implementation of the decomposition discussed at the end of Section XD3 could decrease the computation time presented here by a factor of 16.

G. Understanding the scaling with width and depth of the computational cost of verification

1. Runtime scaling formulas

Here we study the scaling of the runtime of the classical computation of exact amplitudes from the output wave function of a circuit with m cycles and n qubits on Sycamore, assuming a supercomputer with 1M cores. This computation is needed in order to perform XEB on the circuits run. We consider two algorithms: a distributed Schrödinger algorithm (SA) [72, 73] (see Section XC) and a hybrid Schrödinger-Feynman algorithm (SFA) [37] that splits the circuit in two patches and time evolves each of them for all Feynman paths connecting both patches (see Section XA). The latter is embarrassingly parallelizable. Note that these scaling formulas provide rough estimates presented with the intent of building intuition on the scaling of runtimes with the width and depth of the circuits, and that the finite size effects of the circuits can give discrepancies of an order of magnitude or more for the circuit sizes considered in this work.

For SA, the runtime is directly proportional to the size of the wave function on n qubits. This is equal to 2^n . In addition, the runtime is proportional to the number of gates applied, which scales linearly with n and m . For

this reason, we propose the scaling:

$$T_{\text{SA}} = C_{\text{SA}}^{-1} \cdot mn \cdot 2^n, \quad (107)$$

where the constant C_{SA} is fit to runtimes observed experimentally when running on a supercomputer, and scaled to 1M cores.

For SFA the runtime is proportional to the number of paths connecting both patches, as well as to the time taken to simulate each pair of patches. When using the *supremacy* two-qubit gate layouts (ABCDAB...), each fSim gate bridging between the two patches (cross-gates) generates a factor of 4 in the number of paths. The number of cross-gates scales with \sqrt{n} (we assume a two-dimensional grid) and with m . The time taken to simulate each patch is proportional to $2^{n/2}$, where $n/2$ estimates the number of qubits per patch, and the exponential dependence comes from a linear scaling of the runtime with the size of the wave function over that patch. The runtime therefore scales as:

$$T_{\text{SFA, supremacy}} = C_{\text{SFA}}^{-1} \cdot 2 \cdot 2^{\frac{n}{2}} \cdot 4^{B \cdot m \sqrt{n}}, \quad (108)$$

where the extra factor of two accounts for the fact that, for every path, two patches have to be simulated. The constant C_{SFA} , with units of frequency, is the effective frequency with which 1M cores simulate paths and is fit from experimentally observed runtime. The constant B accounts for the average number of cross-gates observed per cycle, which depends on the two-dimensional grid considered and on the two-qubit gate layouts used. For Sycamore, with the supremacy layouts, we find 35 cross-gates for $n = 53$ and $m = 20$, which gives $B = 0.24 \approx 1/4$.

For SFA, using the *verifiable* two-qubit gate layouts (EFGHEFGH...), the main difference with the supremacy circuits case is the fact that most of the cross-gates can be fused in pairs, forming three-qubit gates we refer to as *wedges* (see Sec. VIIG 2 and XD 3). Each cross-wedge generates only 4 paths, as opposed to the 4^2 paths the two independent fSim gates would have generated. Since every 4 cycles provide 7 cross-gates, and from those 7 gates, 6 are converted into 3 wedges, we count only 4^4 paths, as opposed to a naive count of 4^7 for those 4 cycles. In turn, the exponent in the last factor of Eq. 108 is corrected by the fraction $\frac{4}{7}$. This results in:

$$T_{\text{SFA, verifiable}} = C_{\text{SFA}}^{-1} \cdot 2 \cdot 2^{\frac{n}{2}} \cdot 4^{\frac{4}{7} B \cdot m \sqrt{n}}. \quad (109)$$

2. Assumptions and corrections

There are several assumptions considered in Section XG 1 and other details that can either (1) contribute to a somewhat large discrepancy between the runtimes predicted by the scaling formulas and the actual runtimes potentially measured experimentally, or (2) be ignored with no significant impact on the

accuracy of the predictions. Here we discuss the ones we consider most relevant.

Concerning SA, the algorithm is benchmarked in practice on up to 100K cores. Since this is a distributed algorithm, the scaling with number of cores is not ideal and therefore the constant C_{SA} can only be estimated roughly. We assume perfect scaling in our estimates for runtime on 1M cores, *i.e.*, the runtime on 1M cores is the one on 100K cores divided by 10; this is of course an optimistic estimate, and runtimes should be expected to be larger.

For memory requirement estimates, we assume a 2 byte encoding of complex numbers. Beyond about 49 qubits there is not enough RAM on any existing supercomputer to store the wave function. In those cases, runtimes are given for the unrealistic, hypothetical case that one *can* store the wave function.

SFA is embarrassingly parallelizable, and so it does not suffer from non-ideal scaling. However, there are other factors to take into account. First, we have written no explicit dependence of the time to simulate patches of the circuit with m ; the number of cycles m only plays a role when counting the number of paths to be considered. SFA stores several copies of the state of a patch after its evolution at different depths, iterating over paths over several nested loops. For this reason, most of the time is spent iterating over the inner-most loop, which accounts for the last few gates of the circuit and is similar in cost for all depths. This implies that the amortized time per path is considered approximately equal for all depths and the direct m dependence was correctly ignored.

A factor contributing to the discrepancy between the predicted runtimes of the scaling formulas of Section XG 1 and those expected in practice is due to finite size effects. While these scaling formulas consider the average number of cross-gates encountered per cycle, different cycles have layouts that contribute a few more (or less) gates than others. Since the runtime dependency is exponential in the number of gates, this might cause discrepancies of around an order of magnitude. Furthermore, for verifiable circuits, wedges form over groups of two cycles; this coarse graining exacerbates finite size effects. For the sake of simplicity in the scaling formulas, we do not perform any corrections to include these factors. However, in order to mitigate the propagation of finite size effect errors, we consider different constants $C_{\text{SFA, supremacy}}$ and $C_{\text{SFA, verifiable}}$, that we fit independently.

Finally, we refer to runtimes of our simulations on a hypothetical supercomputer with 1M cores. While this is a realistic size for a Top-5 supercomputer currently, a core-hour can vary significantly between different CPU types. Again, we only intend to provide rough estimates in order to build intuition on the dependence of runtimes with circuit width and depth.

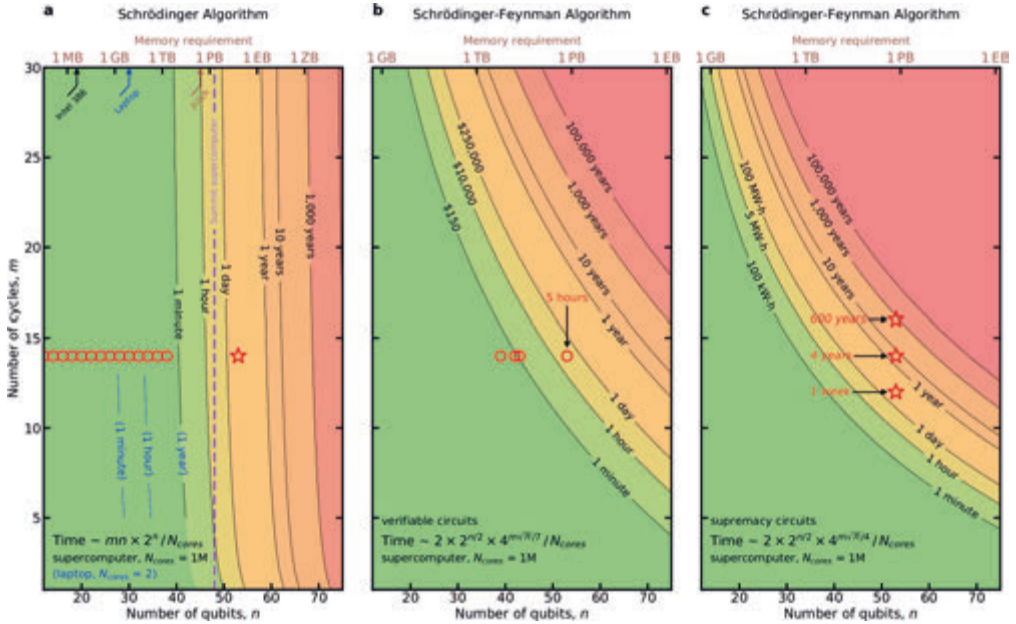


FIG. S50. **Scaling of the computational cost of XEB using SA and SFA.** **a**, For a Schrödinger algorithm, the limitation is RAM size, shown as vertical dashed line for the Summit supercomputer. Circles indicate full circuits with $n = 12$ to 43 qubits that are benchmarked in Fig. 4a of the main paper. 53 qubits would exceed the RAM of any current supercomputer, and is shown as a star. **b**, For the hybrid Schrödinger-Feynman algorithm, which is more memory efficient, the computation time scales exponentially in depth. XEB on full verifiable circuits was done at depth $m = 14$ (circle). **c**, XEB on full supremacy circuits is out of reach within reasonable time resources for $m = 12, 14, 16$ (stars), and beyond. XEB on patch and elided supremacy circuits was done at $m = 14, 16, 18$, and 20.

3. Fitting constants

In the case of SA, we fit the constant C_{SA} with a runtime of 0.1 hours for the simulation with $n = 43$ and $m = 14$. This runtime is obtained by assuming ideal scaling when extrapolating a runtime of 1 hour on nearly 100K nodes (2^{15} MPI processes, 3 cores per process), as reported in Sec. X C. This gives a value of

$$C_{SA} = 0.015 \times 10^6 \text{ GHz}. \quad (110)$$

For SFA, we consider $B = 1/4$ for simplicity. In order to fit C_{SFA} , we consider a runtime of 5 hours and 4 years for the case with $n = 53$ and $m = 14$ for verifiable and supremacy circuits, respectively (see Fig. 4 of the main text). This gives:

$$\begin{aligned} C_{SFA, \text{ verifiable}} &= 0.0062 \times 10^6 \text{ GHz} \\ C_{SFA, \text{ supremacy}} &= 3.3 \times 10^6 \text{ GHz}. \end{aligned} \quad (111)$$

As discussed above, these fits provide times estimated for a supercomputer with 1M cores. Contour plots showing the dependency of runtime with n and m are presented in Fig. S50.

4. Memory usage scaling

Let us conclude with a discussion of the memory footprint of both algorithms. For these estimates, we assume a 2-byte encoding of complex numbers, as opposed to 8 bytes (single precision) or 16 bytes (double precision). This results in a lower bound for the memory usage of these two algorithms. These estimates need an extra factor of 4 (8) when using single (double) precision. SA stores the wave function of the state on all qubits. For this reason, it needs $2^n \times 2 = 2^{n+1}$ bytes. SFA simulates the wave function of both halves of the system ($n/2$ qubits) per path, one at a time. This requires $2^{n/2} \cdot 2$ bytes per path. In practice, the use of checkpoints implies the need to store more than one wave function per path; for simplicity, and in the same optimistic spirit of other as-

sumptions, we ignore this fact. If 1M cores are used and each path is simulated using a single core, the total memory footprint is estimated to be $10^6 \times 2^{\frac{3}{2}+1}$ bytes. State-of-the-art supercomputers have less than 3 PB of memory.

H. Energy advantage for quantum computing

With the end of Dennard scaling for CMOS circuits, gains in computing energy efficiency have slowed significantly [86]. As a result, today's high performance computing centers are usually constrained by available energy supplies rather than hardware costs. For example, the Summit supercomputer at Oak Ridge National Laboratory has a total power capacity of 14 MW available to achieve a design specification of 200 Pflop/s double-precision performance. We took detailed energy measurements with qFlex running on Summit. The energy consumption grows exponentially with the circuit depth, as illustrated in Table VII.

For a superconducting quantum computer, the two primary sources of energy consumption are:

1. **A dilution refrigerator:** our refrigerator has a direct power consumption of ~ 10 kW, dominated by the mechanical compressor driving the 3 K cooling stage. The power required to provide chilled water cooling for the compressor and pumps associated with the refrigerator can be an additional 10 kW or more.
2. **Supporting electronics:** these include microwave electronics, ADCs, DACs, clocks, classical computers, and oscilloscopes that are directly associated with a quantum processor in the refrigerator. The average power consumption of supporting electronics was nearly 3 kW for the experiments in this paper.

We estimate the total average power consumption of our apparatus under worst-case conditions for chilled water production to be 26 kW. This power does not change appreciably between idle and running states of the quantum processor, and it is also independent of the circuit depth. This means that the energy consumed during the 200 s required to acquire 1M samples in our experiment is $\sim 5 \times 10^6$ J (~ 1 kWh). As compared to the qFlex classical simulation on Summit, we require roughly 7 orders of magnitude less energy to perform the same computation (see Table VII). Furthermore, the data acquisition time is currently dominated by control hardware communications, leading to a quantum processor duty cycle as low as 2%. This means there is significant potential to increase our energy efficiency further.

XI. COMPLEXITY-THEORETIC FOUNDATION OF THE EXPERIMENT

The notion of quantum supremacy was originally introduced by John Preskill [87]. He conceived of it as “the day when well controlled quantum systems can perform tasks surpassing what can be done in the classical world”. For the purpose of an experimental demonstration we would like to refine the definition.

Demonstrating quantum supremacy requires:

1. A well defined computational task, i.e. a mathematical specification of a computational problem with a well defined solution.

Comment: This requirement, standard in computer science, excludes tasks such as “simulate a glass of water”. However, it would include finding the ground state energy of an H_2O molecule to a given precision governed by a specific Hamiltonian. Note that a mathematical specification of a computational problem calls for highly accurate control resulting in measurable system fidelity.

2. Programmable computational device

Comment: Many physics experiments estimate the values of observables to a precision which can not be obtained numerically. But those do not involve a freely programmable computational device and the computational task is often not well defined as required above. Ideally, we would even restrict ourselves to devices that are computationally universal. However, this would exclude proposals to demonstrate quantum supremacy with BosonSampling [88] or IQP circuits [89].

3. A scaling runtime difference between the quantum and classical computational processes that can be made large enough as a function of problem size so that it becomes impractical for a supercomputer to solve the task using any known classical algorithm.

Comment: What is impractical for classical computers today may become tractable in ten years. So the quantum supremacy frontier will be moving towards larger and larger problems. But if a task is chosen such that the scaling for the quantum processors is polynomial while for the classical computer it is exponential then this shift will be small. Establishing an exponential separation requires substantial efforts designing and benchmarking classical algorithms [26, 49, 65–68, 70, 72, 73, 83], and support from complexity theory arguments [26, 29, 90]. Sampling the output of random quantum circuits is likely to exhibit this scaling separation as a function of the number of qubits for large enough depth. In this context, we note that quantum analog simulations that estimate an observable in the thermodynamic limit typically do not define a problem size parameter.

The requirements above are satisfied by proposals of quantum supremacy emerging from computer science, such as BosonSampling [88], IQP circuits [89], and random circuit sampling [5, 26, 29, 90, 91]. They are also implicit in the “Extended Church-Turing Thesis”: any “reasonable” model of computation can be efficiently simulated, as a function of problem size, by a Turing machine.

We note that formal complexity proofs are asymptotic, and therefore assume an arbitrarily large number of qubits. This is only possible with a fault tolerant quantum computer and therefore near term practical demonstrations of quantum supremacy must rely on a careful comparison with highly optimized classical algorithms on state-of-the-art supercomputers.

So far we have argued for quantum supremacy by comparing the running time of the quantum experiment with the time required for the same task using the best known classical algorithms, running on the most powerful supercomputers currently available. The fastest known algorithm for exact sampling (or for computing transition probabilities) runs in time exponential in the treewidth of the quantum circuit [67, 68]; for a depth D circuit on a rectangular lattice of sizes l_x and l_y , the treewidth is given by $\min(\min(l_x, l_y)D, l_x l_y)$. For approximate simulation in which one only requires a given global fidelity F , the classical cost is reduced linearly in F [37]. As classical algorithms and compute power can be improved in the future, the classical cost benchmark is a moving target.

A complementary approach to back up supremacy claims consists of giving complexity-theoretic arguments for the classical hardness of the problem solved (in our case sampling from the output distribution of a random circuit of a given number of qubits, depth and output fidelity). Previous work gave hardness results for sampling *exactly* from the output distribution of different classes of circuits [26, 88, 92–94]. Most relevant to us are Refs. [90, 91, 95], which proved that it is classically intractable (unless the polynomial hierarchy collapses to its third level, which is considered extremely unlikely [96]) to sample from the exact probability distribution of outcomes of measurements in random circuits. We note the distribution of circuits considered in [90, 91, 95] is different from ours.

An important clarification is that such results are asymptotic, i.e. they show that, unless the polynomial hierarchy collapses, there are no polynomial-time classical algorithms for sampling from output measurements of certain quantum circuits. But they cannot be used directly to give concrete lower bounds for quantum computations of a fixed number of qubits and depth. Refs. [97–99] tackled this question using tools from fine-grained complexity, giving several finite size bounds.

There are also results arguing for the hardness of *approximate* sampling (see e.g. [26, 88, 89, 93]), where the task is only to sample from a distribution which is close to the ideal one. As the quantum experiment will never

be perfect, this is an important consideration. However those results are weaker than the ones for exact sampling, as the hardness assumptions required have been much less studied (and in fact were introduced with the exact purpose of arguing for quantum supremacy). Another drawback is that the results only apply to the situation where the samples come from a distribution very close to the ideal one (i.e. with high fidelity with the ideal one). This is not the regime in which our experiment operates.

With these challenges in mind, we consider an alternative hardness argument in this section, which will allow us to lower bound the classical simulation cost of noisy quantum circuits by the cost of the ideal one. On one hand, our argument will be more restrictive than previous results in that we will assume a particular noise model for the quantum computer (one, however, which models well the experiment). On the other hand, it will be stronger in two ways: (1) it will apply even to the setting in which the output fidelity of the experimental state with the ideal one can be very small, but still the product of total fidelity with exact computational cost is large; and (2) it will be based on more mainstream complexity assumptions in contrast to the tailor-made conjectures required in e.g. [88, 89, 93] to handle the case of small adversarial noise.

A. Error model

Our error model is the following. We assume that the quantum computer samples from the following output distribution:

$$r_{U,F}(x) := F |x\rangle\langle x| U |0\rangle\langle 0| + (1 - F)/2^n, \quad (112)$$

with U the circuit implemented. In words, we assume global depolarizing noise. Ref. [26] argues that Eq. (112) is a good approximation for the output state of random circuits (see Sec. IV and Section III of [26]); this form has also been verified experimentally on a small number of qubits. In the experiment, F is in the range $10^{-2} - 10^{-3}$.

We note that while we assume a global white noise model in this section, we do not assume it in the rest of the paper, neither for validating the cross entropy test nor in the comparison with state-of-the-art classical algorithms (and indeed the algorithm considered in Section X samples from an approximate distribution different from the one in Eq. (112)).

B. Definition of computational problem

Before stating our result, let us define precisely the computational problem we consider. We start with the ideal version of the problem with no noise:

Circuit Sampling: The input is a description of a n qubit quantum circuit U , described by a sequence of one- and two-qubit gates. The task of the problem is to sample

from the probability distribution of outcomes $p_U(x) := |\langle x|U|0\rangle|^2$.

Circuit sampling is an example of a *sampling problem* [100]. A classical algorithm for circuit sampling can be thought of, without loss of generality, as a function A mapping $m \in \text{poly}(n)$ bits $r = (r_1, \dots, r_m)$ to n bits such that

$$\frac{1}{2^m} |\{(r_1, \dots, r_m) \text{ s.t. } A(r_1, \dots, r_m) = x\}| = \tilde{p}_U(x), \quad (113)$$

with $\tilde{p}(x)$ an approximation of $p_U(x)$ to $l \in \text{poly}(n)$ bits of precision. So when r is chosen uniformly at random, the output of A are samples from p (up to rounding errors which can be made super-exponentially small).

Assuming the polynomial hierarchy does not collapse, it is known that Circuit Sampling cannot be solved classically efficiently in n , meaning any algorithm A satisfying Eq. (113) must have superpolynomial circuit complexity, for several classes of circuits (such as short depth circuits [94], IQP [92] and Boson Sampling [88]). We might also be interested in the average case of circuit sampling (for a restricted class of circuits).

Random Circuit Sampling: The input is a set of quantum circuits \mathcal{U} on n qubits. The task is to sample from $p_U(x) := |\langle x|U|0\rangle|^2$ for most circuits $U \in \mathcal{U}$.

Ref. [90] proved that an efficient (in terms of n) classical algorithm for this task for random circuits would also collapse the polynomial hierarchy. As every realistic quantum experiment will be somewhat noisy, it is relevant to consider a variant of this task allowing for small deviations from ideal. One possible formulation is the following:

ε -Approximate Random Circuit Sampling: The input is a set of quantum circuits \mathcal{U} on n qubits. The task is to sample for most circuits $U \in \mathcal{U}$, from any distribution q_U s.t. $d_{\text{VD}}(q_U, p_U(x)) \leq \varepsilon$, where d_{VD} is the variational-distance between the distributions p, q [101] and $p_U(x) := |\langle x|U|0\rangle|^2$.

Refs. [26, 88, 89] put forward new complexity-theoretic assumptions about the $\#P$ -hardness of certain problems and proved they imply that several restricted classes of circuits are hard to approximately sample for ε sufficiently close to zero. However, we cannot use these results here as the ε we achieve is far from zero. We will resort to the following different variant of approximate circuit sampling.

Unbiased-Noise F -Approximate Random Circuit Sampling: The input is a set of quantum circuits \mathcal{U} on n qubits. The task is to sample from the distribution $r_{U,F}$ given by Eq. (112), for most circuits $U \in \mathcal{U}$.

We note that there are alternatives for defining the computational problem for which supremacy is achieved without having to use sampling problems. These have the advantage that it is possible to verify, for each problem

instance, that the task was achieved (whereas while it is in principle possible to verify that one is sampling from the correct distribution by estimating the frequencies of outcomes, this is unfeasible in practice for high entropy distributions with $> 2^{50}$ outcomes as the one we consider here).

One such problem (considered on Refs. [26, 29]) is the following:

b -Heavy Output Generation: Given as input a number $b > 1$ and a random circuit U on n qubits (drawn at random from a set of circuits \mathcal{U}), generate output strings x_1, \dots, x_k s.t.

$$\frac{1}{k} \sum_{j=1}^k |\langle x_j|U|0\rangle|^2 \geq \frac{b}{2^n} \quad (114)$$

Ref. [29] argues for the hardness of this task for every $b > 1$, although here again one has to resort to rather bold complexity-theoretic conjectures. Cross entropy benchmarking allows us to estimate b for a reasonable value of k (though the classical time needed to compute $|\langle x_j|U|0\rangle|^2$ still grows very fast), see Sec. IV. In terms of known algorithms, the complexity of solving Heavy Output Generation is equivalent to the complexity of sampling k samples from a noisy distribution corresponding to the same b value.

The experiment we report in this paper can be interpreted as showing quantum supremacy in solving the b -Heavy Output Generation with $b = 1 + F$ and F the fidelity of the output quantum state.

C. Computational hardness of unbiased-noise sampling

To state our result, we use the complexity class Arthur-Merlin, which is a variant of the class NP and is denoted by $\text{AM}[T]$. It is defined as the class of problems for which there is an Arthur-Merlin one-round protocol of the following form: given an instance of a problem in $\text{AM}[T]$ (which Arthur would like to decide if it is a YES or NO instance), Arthur first sends random bits to Merlin. Merlin (which is computationally unbounded) then sends back a proof to Arthur. Finally Arthur uses the proof and decides in time T if he accepts. In the YES case, Arthur accepts with probability larger than $2/3$. In the NO case, he accepts with probability no larger than $1/3$.

Theorem 1 *Assume there is a classical algorithm running in time T and using m bits of randomness that samples from the distribution $r_{U,F}(x)$ given by Eq. (112), for a given quantum circuit U on n qubits and $F \geq 0$. Then for every integer L , there is an $\text{AM}[LT + 2Lm]$ protocol for deciding, given $\lambda > 0$, whether*

$$|\langle 0|U|0\rangle|^2 \geq \lambda \left(1 + \frac{2}{L}\right) + \frac{2(1-F)}{FL2^n} \quad (115)$$

or

$$|\langle 0|U|0\rangle|^2 \leq \lambda \left(1 - \frac{2}{L}\right) - \frac{2(1-F)}{FL2^n} \quad (116)$$

Before giving the proof, let us discuss the significance of the result. We are interested in the theorem mostly when $L = c/F$ with c a small constant (say 10). Noting that for a random circuit, with high probability, $|\langle 0|U|0\rangle|^2 \geq 2^{-n}/5$ [95], the theorem states that if we can sample classically in time T from the distribution given in Eq. (112), then we can calculate a good estimate for $|\langle 0|U|0\rangle|^2$ in time $10T/F$ (with the help from an all-powerful but untrustworthy Merlin). It is unlikely that Merlin can be of any help for this task for random circuits, as estimating $|\langle 0|U|0\rangle|^2$ for random circuits is a #P-hard problem [90], and it is believed #P is vastly more complex than AM (which is contained on the third level of the polynomial hierarchy [96]). Therefore we conclude that global white noise leads to no more than a linear decrease in fidelity in classical simulation time (which is in fact optimal as it is achieved by the method presented in Ref. [37]).

Ref. [102] proposed a similar, but more demanding, conjecture about the non-existence of certain AM protocols for estimating transition probabilities of random circuits. This conjecture was applied to show that the output bits of our supremacy experiment can be used to produce certifiable random bits.

We note Theorem 1 does not establish a lower bound on the classical computation cost of calculating a transition amplitude with additive error $\delta/2^n$, for small constant $\delta > 0$. What it does is to show that the sampling problem with unbiased noise is as hard as this task, up to a linear reduction in F in complexity.

Concerning the hardness of computing $|\langle 0|U|0\rangle|^2$ it is known that this problem is #P hard for random circuits to additive error $2^{-\text{poly}(n)}$ [90]. This implies that there is no subexponential-time algorithms for this task (unless #P collapses to P). For finite size bounds, which are more relevant to our experiment, the result of Ref. [97] is the most relevant. It shows that under the Strong Exponential Time Hypothesis (SETH) [103], there are quantum circuits on n qubits which require $2^{(1-o(1))n}$ time for estimating $|\langle 0|U|0\rangle|^2$ to additive error $2^{-(n+1)}$ [104]. Together with Theorem 1, we find there is a quantum circuit U on n qubits for which the distribution $r_{U,F}$ (given by Eq. (112)) cannot be sampled in time $F2^{(1-o(1))n}$, unless SETH is false.

It is an open question to show a similar lower bound to the one proved in Ref. [97] for estimating the transition probability of random circuits. Even more relevant for this work, it would be interesting to study if one can show a lower bound of the form $2^{(1-o(1))\text{treewidth}}$ for a random quantum circuit, under a suitable complexity-theoretic assumption, as the depth of the construction in [97] is relatively high.

D. Proof of Theorem 1

The proof will follow along similar lines to previous work [88, 89, 93]. We will use approximate counting (which can be done in AM) to show that a sampling algorithm for $r_{U,F}$ running in time T implies an AM protocol to compute $r_{U,F}(0)(1 \pm 1/L)$, with classical verification of order LT . Since the noise is unbiased, i.e. $r_{U,F}(0) = F|\langle 0|U|0\rangle|^2 + (1-F)/2^n$, we can subtract it and find an AM protocol for estimating $|\langle 0|U|0\rangle|^2$ as stated in the theorem.

In more detail, suppose there is a classical algorithm for sampling from $r_{U,F}$ given by a function A mapping $m \in \text{poly}(n)$ bits $r = (r_1, \dots, r_m)$ to n bits such that

$$\frac{1}{2^m} |\{(r_1, \dots, r_m) \text{ s.t. } A(r_1, \dots, r_m) = x\}| = r_{U,F}(x). \quad (117)$$

Let $a(r_1, \dots, r_m)$ be a function which is 1 if $A(r_1, \dots, r_m) = 0^n$ and zero otherwise.

We start with the following lemma, showing the existence of A implies an $AM[LT + 2Lm]$ protocol for estimating $r_{U,F}(0)$:

Lemma 1 *Assume there is an algorithm A given by Eq. (117). Then for every θ and L there is an $AM[LT + 2Lm]$ protocol which determines if (i) $r_{U,F}(0) \geq \theta(1 + 2/L)$ (YES instance) or (ii) $r_{U,F}(0) \leq \theta(1 - 2/L)$ (NO instance).*

Proof: The protocol is the following:

1. For every $t \in [Lm]$, Arthur chooses a function at random $h_t \in H_{Lm,t}$ from a family $H_{Lm,t}$ of 2-universal linear hash functions from $\{0,1\}^{Lm}$ to $\{0,1\}^t$ [96]. Then he communicates his choice of (h_1, \dots, h_{Lm}) to Merlin.
2. Merlin sends an Lm -bitstring w to Arthur and an integer $s \in [Lm]$.
3. Arthur verifies that $h_s(w) = 0$ and

$$a(w_{1,1}, \dots, w_{1,m}) \wedge \dots \wedge a(w_{L,1}, \dots, w_{L,m}) = 0.$$

He rejects if any of the three equations is not satisfied. Then he checks if $\theta \leq 2^{-m} 20^{1/L} 2^{s/L} (1 + 2/L)^{-1}$, accepting if it is the case and rejecting otherwise.

The cost to compute $a(w_{1,1}, \dots, w_{1,m})$ is T , and the cost to compute $h_s(w)$ is less than $2Lm$, so the total verification time of the AM protocol is $LT + 2Lm$.

Let us analyze the completeness and soundness of the protocol.

Completeness: Suppose we have a YES instance, $r_{U,F}(0) \geq \theta(1 + 2/L)$. Let us show that Merlin can send w and s which makes Arthur accept with high probability.

Let M be the number of solutions of $a(r_1, \dots, r_m) = 0$ (i.e. $M = 2^m r_{U,F}(0)$). Then $a(r_{1,1}, \dots, r_{1,m}) \wedge \dots \wedge a(r_{L,1}, \dots, r_{L,m})$ has M^L solutions, M for each copy of the function a . As part of the proof Merlin sends s satisfying $20 \geq M^L/2^s \geq 10$ (such a value always exists as s can be an arbitrary integer less than or equal to Lm).

Let us apply Lemma 2 (stated below) with $q = Lm$, $t = s$, $\delta = 1/2$, and S the set of solutions, so $|S| = M^L$. Then indeed $|S|/2^s > 10 > 1/\delta^3$. Therefore, with high probability, the number of solutions of

$$a(x_{1,1}, \dots, x_{1,m}) \wedge \dots \wedge a(x_{L,1}, \dots, x_{L,m}) \wedge h_s(x) \quad (118)$$

is in the interval $[(1/2)M^L/2^s, 2M^L/2^s]$. Since $(1/2)M^L/2^s \geq 1$, there is a string w s.t. $a(w_{1,1}, \dots, w_{1,m}) \wedge \dots \wedge a(w_{L,1}, \dots, w_{L,m}) \wedge h_s(w) = 0$, which Merlin also sends to Arthur as part of the proof.

Since $M = 2^m r_{U,F}(0) \geq 2^m \theta(1 + 2/L)$ and $M^L/2^s \leq 20$,

$$20 \geq \frac{M^L}{2^s} \geq \frac{2^{Lm}}{2^s} \theta^L \left(1 + \frac{2}{L}\right)^L, \quad (119)$$

so indeed $\theta \leq 2^{-m} 20^{1/L} 2^{s/L} (1 + 2/L)^{-1}$ and Arthur will accept with high probability.

Soundness: Suppose we have a NO instance, $r_{U,F}(0) \leq \theta(1 - 2/L)$. Let us show that no matter which witnesses w , s Merlin sends, Arthur will only accept with a small probability. Merlin must send s such that

$$\theta^L \leq (20)^{2-Lm} 2^s (1 + 2/L)^{-L}, \quad (120)$$

otherwise Arthur rejects. By Lemma 2 (stated below), the number of solutions of

$$a(x_{1,1}, \dots, x_{1,m}) \wedge \dots \wedge a(x_{L,1}, \dots, x_{L,m}) \wedge h_s(x) \quad (121)$$

will be in the interval $[(1/2)M^L/2^s, 2M^L/2^s]$, with $M = 2^m r_{U,F}(0) \leq 2^m \theta(1 - 2/L)$. Since

$$\begin{aligned} 2M^L/2^s &\leq 2(2^{-s}) 2^{Lm} \theta^L (1 - 2/L)^L \\ &\leq 40(1 - 2/L)^L (1 + 2/L)^{-L} \leq 40e^{-4} < 1, \end{aligned} \quad (122)$$

there is no solution to Eq. (121) and thus there is no w which will make Arthur accept. This finishes the proof of Lemma 1.

Reduction to AM protocol for $|\langle 0|U|0\rangle|^2$: Finally let us show how to use Lemma 1 to build the AM protocol stated in Theorem 1. Since $r_{U,F}(0) = F|\langle 0|U|0\rangle|^2 + (1 - F)/2^n$, on one hand:

$$|\langle 0|U|0\rangle|^2 \geq \lambda \left(1 + \frac{2}{L}\right) + \frac{2(1 - F)}{FL2^n} \quad (123)$$

implies that

$$r_{U,F}(0) \geq (F\lambda + (1 - F)/2^n) \left(1 + \frac{2}{L}\right). \quad (124)$$

On the other hand:

$$|\langle 0|U|0\rangle|^2 \leq \lambda \left(1 - \frac{2}{L}\right) - \frac{2(1 - F)}{FL2^n} \quad (125)$$

implies that

$$r_{U,F}(0) \leq (F\lambda + (1 - F)/2^n) \left(1 - \frac{2}{L}\right). \quad (126)$$

Setting $\theta = F\lambda + (1 - F)/2^n$ we see that the AM protocol from before can also be used to decide if Eq. (123) or Eq. (125) hold true. This ends the proof of the theorem.

Lemma 2 [96] *For $t \leq q$, let $H_{q,t}$ be a family of pairwise-independent linear hash functions mapping $\{0, 1\}^q$ to $\{0, 1\}^t$, and let $\delta > 0$. Let $S \subseteq \{0, 1\}^n$ be arbitrary with $|S| \geq \delta^{-3} 2^t$. Then with probability larger than $9/10$ over the choice of $h \in H_{n,t}$,*

$$(1 - \delta) \frac{|S|}{2^t} \leq |\{x \in S | h(x) = 0^t\}| \leq (1 + \delta) \frac{|S|}{2^t} \quad (127)$$

Moreover $h(x)$ can be evaluated in time $2n$, for every $h \in H_{n,t}$.

ACKNOWLEDGMENTS

We acknowledge Georg Goerg for consultation on statistical analyses. This research used resources of the Oak Ridge Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC05-00OR22725.

Correspondence and requests for materials should be addressed to John M. Martinis (jmartinis@google.com).

[†] Frank Arute¹, Kunal Arya¹, Ryan Babbush¹, Dave Bacon¹, Joseph C. Bardin^{1,2}, Rami Barends¹, Rupak Biswas³, Sergio Boixo¹, Fernando G.S.L. Brandao^{1,4}, David A. Buell¹, Brian Burkett¹, Yu Chen¹, Zijun Chen¹, Ben Chiaro⁵, Roberto Collins¹, William Courtney¹, Andrew Dunsworth¹, Edward Farhi¹, Brooks Foxen^{1,5}, Austin Fowler¹, Craig Gidney¹, Marissa Giustina¹, Rob Graff¹, Keith Guerin¹, Steve Habegger¹, Matthew P. Harrigan¹, Michael J. Hartmann^{1,6}, Alan Ho¹, Markus Hoffmann¹, Trent Huang¹, Travis S. Humble⁷, Sergei V. Isakov¹, Evan Jeffrey¹, Zhang Jiang¹, Dvir Kafri¹, Kostyantyn Kechedzhii¹, Julian Kelly¹, Paul V. Klimov¹, Sergey Knysh¹, Alexander Korotkov^{1,8}, Fedor Kostritsa¹, David Landhuis¹, Mike Lindmark¹, Erik Lucero¹, Dmitry Lyakh⁹, Salvatore Mandrà^{3,10}, Jarrod R. McClean¹, Matthew McEwen⁵, Anthony Megrant¹, Xiao Mi¹, Kristel Michielsen^{11,12}, Masoud Mohseni¹, Josh Mutus¹, Ofer Naaman¹, Matthew Neeley¹, Charles Neill¹, Murphy Yuezhen Niu¹, Eric Ostby¹, Andre Petukhov¹, John C. Platt¹, Chris Quintana¹, Eleanor G. Rieffel³, Pedram Roushan¹, Nicholas C. Rubin¹, Daniel Sank¹, Kevin J. Satzinger¹, Vadim Smelyanskiy¹, Kevin J. Sung^{1,13}, Matthew D. Trevithick¹, Amit Vainsencher¹, Benjamin Villalonga^{1,14}, Theodore White¹, Z. Jamie Yao¹, Ping Yeh¹, Adam Zalcman¹, Hartmut Neven¹, John M. Martinis^{1,5}

1. Google AI Quantum, Mountain View, CA, USA, 2. Department of Electrical and Computer Engineering, University of Massachusetts Amherst, Amherst, MA, USA, 3. Quantum Artificial Intelligence Lab. (QuAIL), NASA Ames Research Center, Moffett Field, USA, 4. Institute for Quantum Information and Matter, Caltech, Pasadena, CA, USA, 5. Department of Physics, University of California, Santa Barbara, CA, USA, 6. Friedrich-Alexander University Erlangen-Nürnberg (FAU), Department of Physics, Erlangen, Germany, 7. Quantum Computing Institute, Oak Ridge National Laboratory, Oak Ridge, TN, USA, 8. Department of Electrical and Computer Engineering, University of California, Riverside, CA, USA, 9. Scientific Computing, Oak Ridge Leadership Computing, Oak Ridge National Laboratory, Oak Ridge, TN, USA 10. Stinger Ghaffarian Technologies Inc., Greenbelt, MD, USA, 11. Institute for Advanced Simulation, Jülich Supercomputing Centre, Forschungszentrum Jülich, Jülich, Germany, 12. RWTH Aachen University, Aachen, Germany, 13. Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI, USA, 14. Department of Physics, University of Illinois at Urbana-Champaign, Urbana, IL, USA
-
- [1] Barends, R. *et al.* Superconducting quantum circuits at the surface code threshold for fault tolerance. *Nature* **508**, 500 (2014).
 - [2] Neill, C. *A path towards quantum supremacy with superconducting qubits*. Ph.D. thesis, University of California, Santa Barbara (2017).
 - [3] Yan, F. *et al.* Tunable coupling scheme for implementing high-fidelity two-qubit gates. *Phys. Rev. Applied* **10**, 054062 (2018).
 - [4] Chen, Y. *et al.* Qubit architecture with high coherence and fast tunable coupling. *Phys. Rev. Lett.* **113**, 220502 (2014).
 - [5] Neill, C. *et al.* A blueprint for demonstrating quantum supremacy with superconducting qubits. *Science* **360**, 195–199 (2018).
 - [6] Khezri, M., Dressel, J. & Korotkov, A. N. Qubit measurement error from coupling with a detuned neighbor in circuit QED. *Phys. Rev. A* **92**, 052306 (2015).
 - [7] Tucci, R. R. An introduction to Cartan's KAK decomposition for QC programmers. Preprint at <https://arxiv.org/abs/quant-ph/0507171> (2005).
 - [8] Dunsworth, A. *High fidelity entangling gates in superconducting qubits*. Ph.D. thesis, University of California, Santa Barbara (2018).
 - [9] Dunsworth, A. *et al.* A method for building low loss multi-layer wiring for superconducting microwave devices. *Appl. Phys. Lett.* **112**, 063502 (2018).
 - [10] Rosenberg, D. *et al.* 3D integrated superconducting qubits. *npj Quantum Inf.* **3**, 42 (2017).
 - [11] Foxen, B. *et al.* Qubit compatible superconducting interconnects. *Quantum Sci. Tech.* **3**, 014005 (2017).
 - [12] Foxen, B. *et al.* High speed flux sampling for tunable superconducting qubits with an embedded cryogenic transducer. *Supercond. Sci. Technol.* **32**, 015012 (2018).
 - [13] Blais, A., Huang, R.-S., Wallraff, A., Girvin, S. M. & Schoelkopf, R. J. Cavity quantum electrodynamics for superconducting electrical circuits: An architecture for quantum computation. *Phys. Rev. A* **69**, 062320 (2004).
 - [14] Gambetta, J. *et al.* Qubit-photon interactions in a cavity: Measurement-induced dephasing and number splitting. *Phys. Rev. A* **74**, 042318 (2006).
 - [15] Bultink, C. C. *et al.* General method for extracting the quantum efficiency of dispersive qubit readout in circuit QED. *Appl. Phys. Lett.* **112**, 092601 (2018).
 - [16] Sank, D. *et al.* Measurement-induced state transitions in a superconducting qubit: Beyond the rotating wave approximation. *Phys. Rev. Lett.* **117**, 190503 (2016).
 - [17] Clerk, A. A., Devoret, M. H., Girvin, S. M., Marquardt, F. & Schoelkopf, R. J. Introduction to quantum noise, measurement, and amplification. *Rev. Mod. Phys.* **82**, 1155–1208 (2010).
 - [18] Caves, C. M. Quantum limits on noise in linear amplifiers. *Phys. Rev. D* **26**, 1817–1839 (1982).
 - [19] Mutus, J. Y. *et al.* Strong environmental coupling in a Josephson parametric amplifier. *Appl. Phys. Lett.* **104**, 263513 (2014).
 - [20] Ryan, C. A. *et al.* Tomography via correlation of noisy measurement records. *Phys. Rev. A* **91**, 022118 (2015).
 - [21] Jeffrey, E. *et al.* Fast accurate state measurement with superconducting qubits. *Phys. Rev. Lett.* **112**, 190504 (2014).
 - [22] Sank, D. What is the connection between analog signal to noise ratio and signal to noise ratio in the IQ plane in a quadrature demodulation system? Signal Processing Stack Exchange. URL: <https://dsp.stackexchange.com/questions/24372> (2015).
 - [23] Reed, M. D. *et al.* Fast reset and suppressing spontaneous emission of a superconducting qubit. *Appl. Phys. Lett.* **96**, 203110 (2010).
 - [24] Sete, E. A., Martinis, J. M. & Korotkov, A. N. Quantum theory of a bandpass purcell filter for qubit readout. *Phys. Rev. A* **92**, 012325 (2015).
 - [25] Chen, Y. *et al.* Multiplexed dispersive readout of superconducting phase qubits. *Appl. Phys. Lett.* **101**, 182601 (2012).
 - [26] Boixo, S. *et al.* Characterizing quantum supremacy in near-term devices. *Nat. Phys.* **14**, 595 (2018).
 - [27] Wootters, W. K. Random quantum states. *Found. Phys.* **20**, 1365–1378 (1990).
 - [28] Emerson, J., Livine, E. & Lloyd, S. Convergence conditions for random quantum circuits. *Phys. Rev. A* **72**, 060302 (2005).
 - [29] Aaronson, S. & Chen, L. Complexity-theoretic foundations of quantum supremacy experiments. In *32nd Computational Complexity Conference (CCC 2017)* (2017).
 - [30] Magesan, E., Gambetta, J. M. & Emerson, J. Characterizing quantum gates via randomized benchmarking. *Phys. Rev. A* **85** (2012).
 - [31] Magesan, E., Gambetta, J. M. & Emerson, J. Robust randomized benchmarking of quantum processes. *Phys. Rev. Lett.* **106** (2011).
 - [32] Popescu, S., Short, A. J. & Winter, A. Entanglement and the foundations of statistical mechanics. *Nat. Phys.* **2**, 754 (2006).

- [33] Bremner, M. J., Mora, C. & Winter, A. Are random pure states useful for quantum computation? *Phys. Rev. Lett.* **102**, 190502 (2009).
- [34] Gross, D., Flammaria, S. T. & Eisert, J. Most quantum states are too entangled to be useful as computational resources. *Phys. Rev. Lett.* **102**, 190501 (2009).
- [35] McClean, J. R., Boixo, S., Smelyanskiy, V. N., Babbush, R. & Neven, H. Barren plateaus in quantum neural network training landscapes. *Nat. Comm.* **9**, 4812 (2018).
- [36] Ledoux, M. *The concentration of measure phenomenon*. 89 (American Mathematical Society, 2005).
- [37] Markov, I. L., Fatima, A., Isakov, S. V. & Boixo, S. Quantum supremacy is both closer and farther than it appears. Preprint at <https://arxiv.org/pdf/1807.10749> (2018).
- [38] Cross, A. W., Bishop, L. S., Sheldon, S., Nation, P. D. & Gambetta, J. M. Validating quantum computers using randomized model circuits. *Phys. Rev. A* **100**, 032328 (2019).
- [39] Kelly, J., O'Malley, P., Neeley, M., Neven, H. & Martinis, J. M. Physical qubit calibration on a directed acyclic graph. Preprint at <https://arxiv.org/abs/1803.03226> (2019).
- [40] Wallraff, A. *et al.* Strong coupling of a single photon to a superconducting qubit using circuit quantum electrodynamics. *Nature* **431**, 162 (2004).
- [41] Chen, Z. *Metrology of quantum control and measurement in superconducting qubits*. Ph.D. thesis, University of California, Santa Barbara (2018).
- [42] Chen, Z. *et al.* Measuring and suppressing quantum state leakage in a superconducting qubit. *Phys. Rev. Lett.* **116**, 020501 (2016).
- [43] Klimov, P. V. *et al.* Fluctuations of energy-relaxation times in superconducting qubits. *Phys. Rev. Lett.* **121**, 090502 (2018).
- [44] Kelly, J. *et al.* Optimal quantum control using randomized benchmarking. *Phys. Rev. Lett.* **112**, 240504 (2014).
- [45] Bialczak, R. C. *et al.* Quantum process tomography of a universal entangling gate implemented with Josephson phase qubits. *Nat. Phys.* **6**, 409 (2010).
- [46] Martinis, J. M. & Geller, M. R. Fast adiabatic qubit gates using only σ_z control. *Phys. Rev. A* **90**, 022307 (2014).
- [47] DiCarlo, L. *et al.* Demonstration of two-qubit algorithms with a superconducting quantum processor. *Nature* **460**, 240 (2009).
- [48] Kivlichan, I. D. *et al.* Quantum simulation of electronic structure with linear depth and connectivity. *Phys. Rev. Lett.* **120**, 110501 (2018).
- [49] Villalonga, B. *et al.* A flexible high-performance simulator for the verification and benchmarking of quantum circuits implemented on real hardware. Preprint at <https://arxiv.org/pdf/1811.09599v2> (2018).
- [50] Wallman, J., Granade, C., Harper, R. & Flammaria, S. T. Estimating the coherence of noise. *New J. Phys.* **17**, 113020 (2015).
- [51] Erhard, A. *et al.* Characterizing large-scale quantum computers via cycle benchmarking. Preprint at <https://arxiv.org/pdf/1902.08543> (2019).
- [52] Johnson, J. E. *et al.* Heralded state preparation in a superconducting qubit. *Phys. Rev. Lett.* **109**, 050506 (2012).
- [53] Sank, D. T. *Fast, accurate state measurement in superconducting qubits*. Ph.D. thesis, University of California, Santa Barbara (2014).
- [54] Supporting online materials.
- [55] Babbush, R. *et al.* Low-depth quantum simulation of materials. *Phys. Rev. X* **8**, 011044 (2018).
- [56] Boykin, P., Mor, T., Pulver, M., Roychowdhury, V. & Vatan, F. On universal and fault-tolerant quantum computing. *Proc. 40th Annual Symposium on Foundations of Computer Science, IEEE Computer Society Press* (1999).
- [57] DiVincenzo, D. P. Two-bit gates are universal for quantum computation. *Phys. Rev. A* **51**, 1015–1022 (1995).
- [58] Shor, P. W. Scheme for reducing decoherence in quantum computer memory. *Phys. Rev. A* **52**, R2493(R) (1995).
- [59] Knill, E. *et al.* Randomized benchmarking of quantum gates. *Phys. Rev. A* **77**, 012307 (2008).
- [60] Lehmann, E. L. & Romano, J. P. *Testing Statistical Hypotheses* (Springer-Verlag New York, 2005).
- [61] Jones, E., Oliphant, T., Peterson, P. *et al.* SciPy: Open source scientific tools for Python (2001). URL: <http://www.scipy.org/> (2016).
- [62] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria (2017). URL <https://www.R-project.org/>.
- [63] Efron, B. Bootstrap methods: Another look at the jack-knife. *Ann. Stat.* **7**, 1–26 (1979).
- [64] Kalai, G. The argument against quantum computers. Preprint at <https://arxiv.org/pdf/1908.02499> (2019).
- [65] Smelyanskiy, M., Sawaya, N. P. & Aspuru-Guzik, A. qHIPSTER: the quantum high performance software testing environment. Preprint at <https://arxiv.org/pdf/1601.07195> (2016).
- [66] Villalonga, B. *et al.* Establishing the quantum supremacy frontier with a 281 P/flop/s simulation. Preprint at <https://arxiv.org/pdf/1905.00444> (2019).
- [67] Markov, I. L. & Shi, Y. Simulating quantum computation by contracting tensor networks. *SIAM J. Comput.* **38**, 963–981 (2008).
- [68] Boixo, S., Isakov, S. V., Smelyanskiy, V. N. & Neven, H. Simulation of low-depth quantum circuits as complex undirected graphical models. Preprint at <https://arxiv.org/pdf/1712.05384> (2017).
- [69] Lyakh, D. Tensor algebra library routines for shared memory systems. URL: <https://github.com/DmitryLyakh/TAL.SH> (2019).
- [70] Chen, J. *et al.* Classical simulation of intermediate-size quantum circuits. Preprint at <https://arxiv.org/pdf/1805.01450> (2018).
- [71] Guo, C. *et al.* General-purpose quantum circuit simulator with projected entangled-pair states and the quantum supremacy frontier. Preprint at <https://arxiv.org/pdf/1905.08394> (2019).
- [72] De Raedt, K. *et al.* Massively parallel quantum computer simulator. *Comput. Phys. Commun.* **176**, 121–136 (2007).
- [73] De Raedt, H. *et al.* Massively parallel quantum computer simulator, eleven years later. *Comput. Phys. Commun.* **237**, 47–61 (2019).
- [74] Krause, D. & Thörnig, P. Jureca: Modular supercomputer at Jülich supercomputing centre. *Journal of large-scale research facilities JLSRF* **4**, 132 (2018).
- [75] Krause, D. Juwels: Modular tier-0/1 supercomputer at

- the Jülich supercomputing centre. *Journal of large-scale research facilities JLSRF* **5**, 135 (2019).
- [76] Kalai, G. & Kindler, G. Gaussian noise sensitivity and bosonsampling. *arXiv:1409.3093* (2014).
- [77] Bremner, M. J., Montanaro, A. & Shepherd, D. J. Achieving quantum supremacy with sparse and noisy commuting quantum computations. *Quantum* **1**, 8 (2017).
- [78] Yung, M.-H. & Gao, X. Can chaotic quantum circuits maintain quantum supremacy under noise? *arXiv:1706.08913* (2017).
- [79] Boixo, S., Smelyanskiy, V. N. & Neven, H. Fourier analysis of sampling from noisy chaotic quantum circuits. *arXiv:1708.01875* (2017).
- [80] Nahum, A., Vijay, S. & Haah, J. Operator spreading in random unitary circuits. *Phys. Rev. X* **8**, 021014 (2018).
- [81] Von Keyserlingk, C., Rakovszky, T., Pollmann, F. & Sondhi, S. L. Operator hydrodynamics, OTOCs, and entanglement growth in systems without conservation laws. *Phys. Rev. X* **8**, 021013 (2018).
- [82] Gogate, V. & Dechter, R. A complete anytime algorithm for treewidth. In *Proceedings of the 20th Conference on Uncertainty in Artificial Intelligence*, 201–208 (2004).
- [83] Zhang, F. *et al.* Alibaba cloud quantum development kit: Large-scale classical simulation of quantum circuits. Preprint at <https://arxiv.org/pdf/1907.11217> (2019).
- [84] Chen, M.-C. *et al.* Quantum teleportation-inspired algorithm for sampling large random quantum circuits. Preprint at <https://arxiv.org/pdf/1901.05003> (2019).
- [85] We assume 2 vCPUs per core.
- [86] Koomey, J. & Naffziger, S. Energy efficiency of computing: What's next? *Electronic Design*. URL <https://www.electronicdesign.com/microprocessors/energy-efficiency-computing-what-s-next> (2016).
- [87] Preskill, J. Quantum computing and the entanglement frontier. *Rapporteur talk at the 25th Solvay Conference on Physics, Brussels* (2012).
- [88] Aaronson, S. & Arkhipov, A. The computational complexity of linear optics. In *Proceedings of the Forty-third Annual ACM Symposium on Theory of Computing*, 333–342 (2011).
- [89] Bremner, M. J., Montanaro, A. & Shepherd, D. J. Average-case complexity versus approximate simulation of commuting quantum computations. *Phys. Rev. Lett.* **117**, 080501 (2016).
- [90] Bouldan, A., Fefferman, B., Nirkhe, C. & Vazirani, U. On the complexity and verification of quantum random circuit sampling. *Nat. Phys.* **15**, 159 (2019).
- [91] Movassagh, R. Cayley path and quantum computational supremacy: A proof of average-case #P-hardness of random circuit sampling with quantified robustness. Preprint at <https://arxiv.org/pdf/1909.06210> (2019).
- [92] Bremner, M. J., Jozsa, R. & Shepherd, D. J. Classical simulation of commuting quantum computations implies collapse of the polynomial hierarchy. *Proc. Royal Soc. A* **467**, 459–472 (2010).
- [93] Harrow, A. W. & Montanaro, A. Quantum computational supremacy. *Nature* **549**, 203 (2017).
- [94] Terhal, B. M. & DiVincenzo, D. P. Adaptive quantum computation, constant depth quantum circuits and Arthur-Merlin games. *Quant. Inf. Comp.* **4**, 134–145 (2004).
- [95] Harrow, A. W. & Mehraban, S. Approximate unitary t -designs by short random quantum circuits using nearest-neighbor and long-range gates. Preprint at <https://arxiv.org/pdf/1809.06957> (2018).
- [96] Arora, S. & Barak, B. *Computational complexity: a modern approach* (Cambridge University Press, 2009).
- [97] Huang, C., Newman, M. & Szegedy, M. Explicit lower bounds on strong quantum simulation. Preprint at <https://arxiv.org/pdf/1804.10368> (2018).
- [98] Dalzell, A. M., Harrow, A. W., Koh, D. E. & La Placa, R. L. How many qubits are needed for quantum computational supremacy? Preprint at <https://arxiv.org/pdf/1805.05224.pdf> (2018).
- [99] Morimae, T. & Tamaki, S. Fine-grained quantum supremacy of the one-clean-qubit model. Preprint at <https://arxiv.org/pdf/1901.01637> (2019).
- [100] Aaronson, S. The equivalence of sampling and searching. *Theory of Computing Systems* **55**, 281–298 (2014).
- [101] The variational distance is defined as $d_{\text{VD}}(p, q) := \sum_i |p_i - q_i|$.
- [102] Aaronson, S. Certifiable randomness from supremacy. *Manuscript in preparation*. (2019).
- [103] Calabro, C., Impagliazzo, R. & Paturi, R. The complexity of satisfiability of small depth circuits. In *International Workshop on Parameterized and Exact Computation*, 75–85 (Springer, 2009).
- [104] In terms of depth, the current construction gives a circuit U on n qubits of depth $n^{3k/2+5/2}$ for which it takes time $2^{(1-2/k)n}$ to estimate the transition probability to additive error $2^{-(n+1)}$, assuming a form of SETH stating that it takes time no less than $2^{(1-2/k)n}$ to solve k -SAT.

A 4 Quantum Optics and Quantum Information

Jian-Wei Pan

CAS Center for Excellence in Quantum Information and
Quantum Physics

University of Science and Technology of China

Contents

1. Lecture 1: Introduction to Quantum Physics and Quantum Information

Part 1: Quantum Foundation

Part 2: Quantum Communication

Part 3: Quantum Computation and Quantum Metrology

Part 4: Quantum Repeaters

2. Lecture 2: Scalable Quantum Information Processing with Photons and Atoms

Part 1: Elemental Optical Manipulations and Demonstrations of Quantum Communication

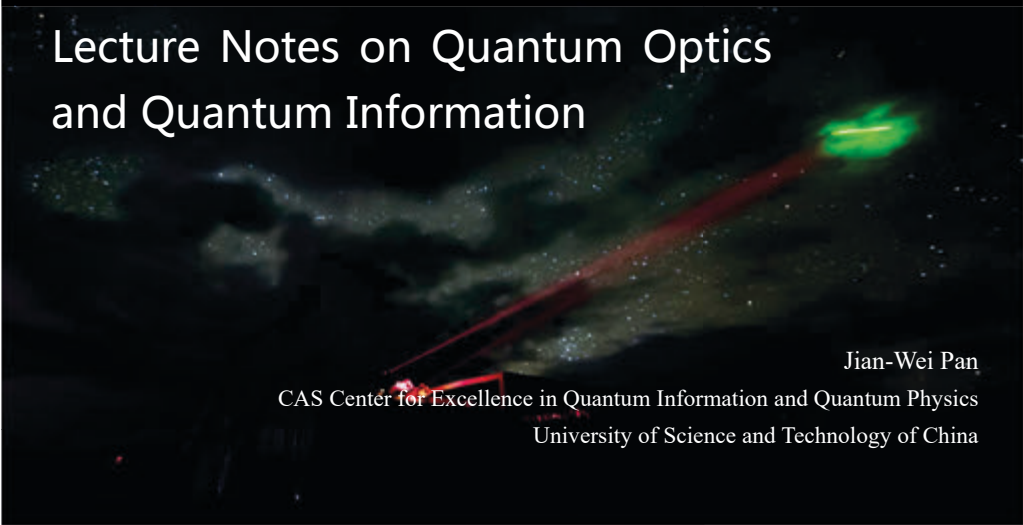
Part 2: Multi-photon Interferometry

Part 3: Demonstrations of Quantum Repeaters

Part 4: Towards Long-distance Quantum Communication

Part 5: Future Prospects

Lecture Notes on Quantum Optics and Quantum Information



Jian-Wei Pan
CAS Center for Excellence in Quantum Information and Quantum Physics
University of Science and Technology of China

Lecture 1: Introduction to Quantum Physics and Quantum Information

Part 1: Quantum Foundations

Old Chinese Legends



Wu Cheng-En (1501-1582) and his "Journey to the West"

Many interesting imaginations:

➤ “天上一日，地上一年”

One day in heaven, one year in earth

➤ Two famous immortals:

顺风耳、千里眼



Wind-following Ear Thousand-miles Eye

Old Chinese Legends



➤ The most famous hero Monkey King,
he has the abilities of:



• 分身术 Body-split



• 筋斗云 Cloud-somersault

Legend to Modern Physics &

➤ Wind-following ear, Thousand-miles eye

Electrodynamics!



Maxwell



Maxwell Equations (1864)



Bell invented telephone (1876)



Hertz



The first radio experiment (1888)

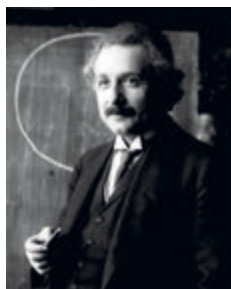


Baird invented television (1926)

Legend to Modern Physics &

➤ One day in heaven, one year in earth

Relativity



Einstein

3. Zur Elektrodynamik bewegter Körper; von A. Einstein.

Daß die Elektrodynamik Maxwells — wie dieselbe gegenwärtig aufgefaßt zu werden pflegt — in ihrer Anwendung auf bewegte Körper zu Asymmetrien führt, welche den Phänomenen nicht anzuhafien scheinen, ist bekannt. Man denke z. B. an

"On the Electrodynamics of Moving Bodies" (1905)



Twin paradox

Would it be possible to manifest the Monkey King's body-split and cloud-somersault?

Quantum mechanics!



Max Planck

Albert Einstein

Niels Bohr

Erwin Schrödinger

Werner Heisenberg

Paul Dirac

Quantum Superposition



Classical world:

or



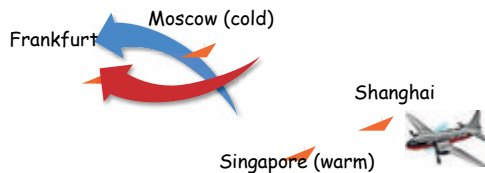
Quantum world:

+

Body-split in quantum world!

Quantum Superposition

A "quantum flight": from Stockholm to Beijing, two possible routes



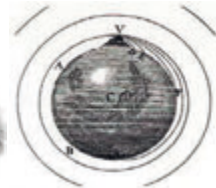
When arrived

- If I fell asleep during flight (do not know which route I take)
I will feel "both cold and warm" ➔ It seems like I took both routes in one flight?
- If I was awake during flight and checked which route I take
I will feel either cold or warm ➔ It confirms I can only take one of the routes!

In quantum world, the state of a quantum object can be affected by measurement!

When Classical Physics Meets Life Philosophy

Newton's law precisely predicts every single movement for all objects in our daily life

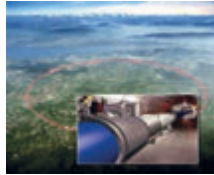


- A manifest of the beauty and power of physics!
- However, does it imply determinism?
- Does it mean everything (e.g. lectures today) is already determined from Big-bang?
- Efforts meaningless?
- Fortunately, quantum mechanics tells that your act (measurement) can affect the world!

Quantum Mechanics and Information Technology



Nuclear weapon



CERN



Atomic clock



ENICA



World Wide Web



GPS

Challenges in Information

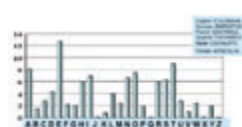
History: every advance in classical cryptography has been defeated by advances in crackina!



Ancient Greek scytale, 700 BC



Caesar cipher, 50 BC



Cracked via variations in the frequency of the occurrence of letters, by Al-Kindi (800-873)



Enigma machine broken by Alan Turing



RSA 512: cracked in 1999
 RSA 768: cracked in 2009
 RSA 1024: ?
 SHA-1: cracked in 2017 by Google

Challenges in Information

All the classical encryption methods that depend on computational complexity, can be cracked in principle!

".....human ingenuity cannot concoct a cipher which human ingenuity cannot resolve"

— *A few words on secret writing*, Edgar Allan Poe (1841)

Challenges in Computational Capacity

In 1943: "I think there is a world market for maybe five computers"

--Thomas Watson, Chairman of IBM



Colossus, weight: 1 ton, power: 8.5kw, 5 kOPS (operations per second)

In 2010's: Almost everyone owns computing power larger than the total computing power used in Apollo Program!



Apple A12, power: <5W, 5 trillion OPS



The world's data volume roughly grow 40% per year

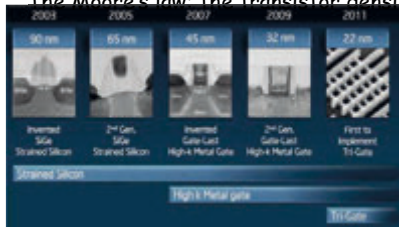
Challenges in the Computational Capacity

Classical computational bottleneck

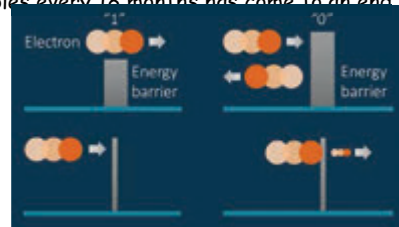
The world's total computing power is insufficient to search a target in 2^{80-90} database within a year

A technological limit

The Moore's law: the transistor density doubles every 18 months has come to an end

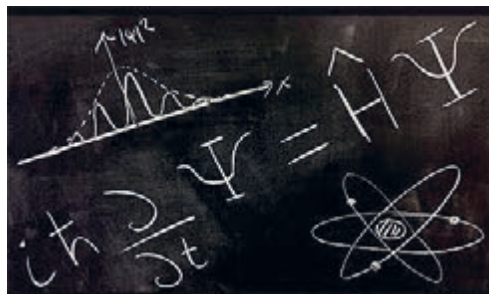


2017, 14 nm → 2022, 4 nm →
0.2 nm (atomic scale) → ???



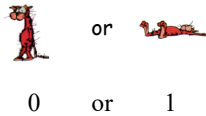
Tunneling induced leakage →
The "0/1" logic in the transistors will fail

Quantum physics, after one century's development, comes to the rescue for the problems confronted in the classical information technologies



Quantum Superposition and Quantum Bits

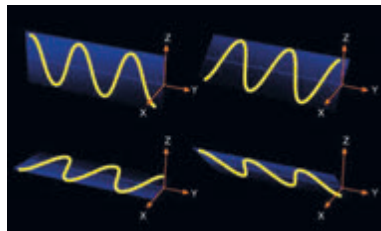
Classical Physics: "bit"



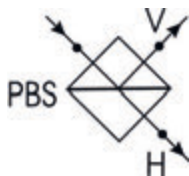
Quantum Physics: "qubit"

$$| \text{standing} \rangle + | \text{lying} \rangle$$

A qubit encoded in the polarization of a photon



Measurement Uncertainty and Non-cloning



Non-cloning theorem:
An unknown quantum state
can not be copied precisely!

Input	0, 1	?	Require:	Linear superposition principle:
		Output	00 → 00	
	0	?	10 → 11	(0+1)0 → 00+11 ≠ (0+1)(0+1)

Wootters & Zurek, Nature 299, 802 (1982)

Single-Qubit Operations

Column vector represent of
two-dimensional quantum states

Pauli matrix

Two eigenstates:

Unitary rotation:

Pauli matrix

Two eigenstates:

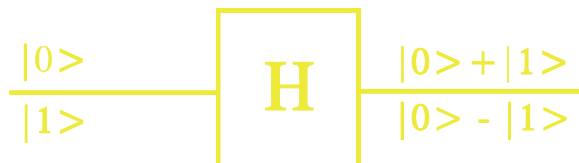
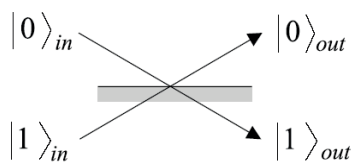
Unitary rotation:

Pauli matrix

Two eigenstates:

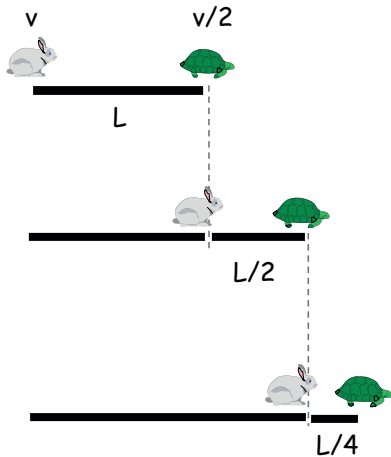
Unitary rotation:

Single-Qubit Operations



The Hadamard gate

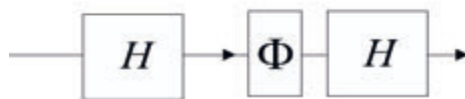
Zeno Paradox



Origin of Zeno effect

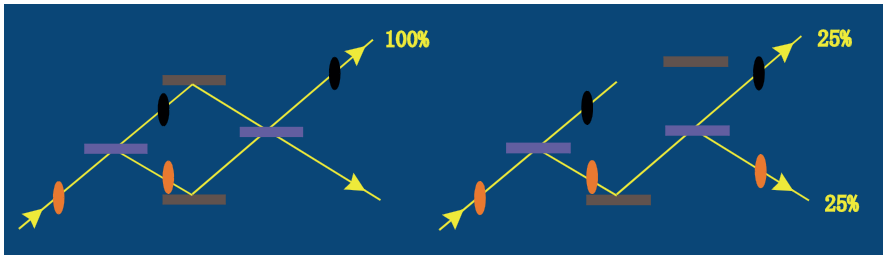
Can the rabbit overtake the turtle?

Quantum Zeno Effect



$$H\Phi H|0\rangle = \frac{1}{2}((e^{i\varphi} + 1)|0\rangle + (e^{i\varphi} - 1)|1\rangle)$$

Quantum Zeno Effect



Interaction-free measurement !

Quantum Zeno Effect



up



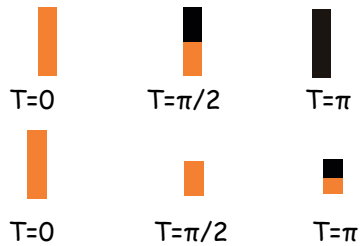
down

Considering neutron spin evolving in magnetic field, the probability to find it still in spin up state after time T is

where ω is the Larmor frequency

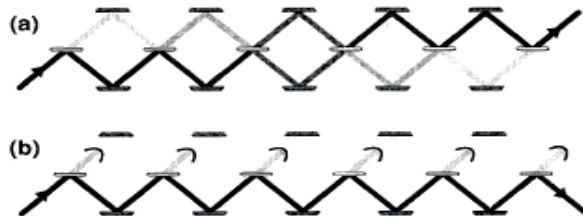
Quantum Zeno Effect

$$G(\text{cake is good}) = G_0$$



If we cut the bad part of the cake at time $T=\pi/2$, then at $T=\pi$ we have $G=1/4 \times G_0$

Experiment



In the limit of large N :

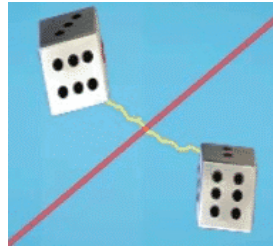
Kwiat *et al.*, PRL 74, 4763 (1995)

Quantum Entanglement

Quantum entanglement:

$$| \text{standing} \rangle | \text{standing} \rangle + | \text{lying} \rangle | \text{lying} \rangle$$

Bell states - maximally entangled states:



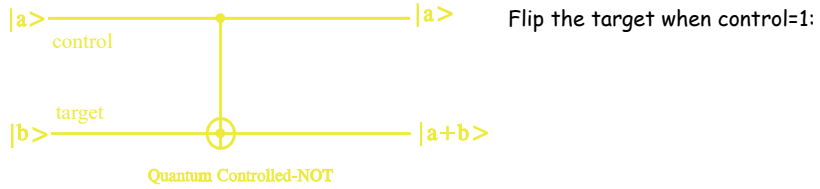
Spooky action at a distance

—Albert Einstein

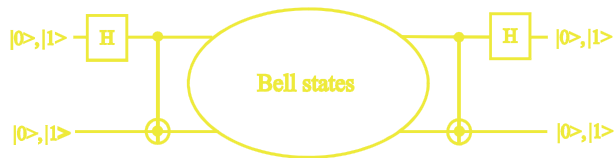
Quantum Entanglement

GHZ states: three-photon maximally entangled states

Manipulation of Entanglement

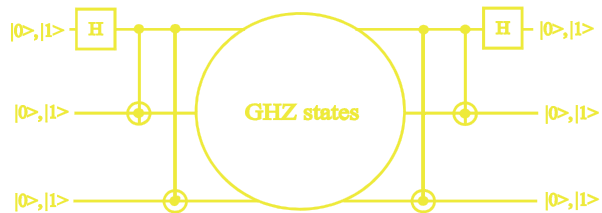


(a)



Manipulation of Entanglement

(b)



Spooky Action at a Distance?



Entangled pair

Einstein believed that :

- ☞ The outcome of a measurement on any physical system is determined prior to and independent of the measurement
 - ☞ the outcome cannot depend on any actions in space-like separated regions
- A seemingly reasonable assumptions of "local realism"

Quantum mechanics predicts that:

- ☞ Initially, the individual states of two particles are not identified
- ☞ The measurement outcome on particle A will not only determine its state, but also the state of particle B immediately!

Spooky Action at a Distance?



MAY 15, 1935

PHYSICAL REVIEW

VOLUME 47

Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?

A. EINSTEIN, B. PODOLSKY AND N. ROSEN, *Institute for Advanced Study, Princeton, New Jersey*

(Received March 25, 1935)

Bell's Inequality: Testing the Battle

Experimental testable inequality

Bell, Physics 1, 195 (1964)

Clauser *et al.*, PRL 23, 880
(1969)



- Einstein's local realism:
- Quantum mechanics:

e. g.,

Bell's Inequality: Testing the Battle

A simplified case: Sakurai's Bell Inequality



Singlet state: anti-correlation of measurement results of two sides

	Alice			Bob			
measurement directions	A	b	c	a	b	c	probability
pre-determined outcomes	0	0	0	1	1	1	P_1
	0	0	1	1	1	0	P_2
	0	1	0	1	0	1	P_3
	0	1	1	1	0	0	P_4
	1	0	0	0	1	1	P_5
	1	0	1	0	1	0	P_6
	1	1	0	0	0	1	P_7
	1	1	1	0	0	0	P_8

Pick three arbitrary directions **a**, **b**, and **c**:

$$P(a, b) = P_3 + P_4$$

$$P(a, c) = P_2 + P_4$$

$$P(c, b) = P_3 + P_7$$

And

Bell's Inequality: Testing the Battle

Local realism requires:

Quantum-mechanical prediction:

For example , the inequality would require

An unsatisfactory feature

In the derivation of BI such a local realistic and thus classical picture can explain perfect correlations and is only in conflict with *statistical prediction of quantum mechanics*

Conflict with Local Realism

Consider a three-photon GHZ state written in basis

Linear polarization basis

Circular polarization basis

$$\sigma_x : |H'\rangle = \frac{1}{\sqrt{2}}(|H\rangle + |V\rangle),$$

$$\sigma_y : |R\rangle = \frac{1}{\sqrt{2}}(|H\rangle + i|V\rangle),$$

$$|V'\rangle = \frac{1}{\sqrt{2}}(|H\rangle - |V\rangle).$$

$$|L\rangle = \frac{1}{\sqrt{2}}(|H\rangle - i|V\rangle).$$

Conflict with Local Realism

$$\sigma_{1y}\sigma_{2y}\sigma_{3x} : |\psi_{123}\rangle = \frac{1}{2}(R_1L_2H'_3 + L_1R_2H'_3 + R_1R_2V'_3 + L_1L_2V'_3)$$

$$\sigma_{1y}\sigma_{2x}\sigma_{3y} : |\psi_{123}\rangle = \frac{1}{2}(R_1H'_2L_3 + L_1H'_2R_3 + R_1V'_2R_3 + L_1V'_2L_3)$$

$$\sigma_{1x}\sigma_{2y}\sigma_{3y} : |\psi_{123}\rangle = \frac{1}{2}(H'_1R_2L_3 + H'_1L_2R_3 + V'_1R_2R_3 + V'_1L_2L_3)$$

Therefore state is the eigenstate of operators
with value -1

Conflict with Local Realism

- **EPR reality criterion:** the individual value of any local operator is predetermined
- There exists an element of local reality S_{ix} corresponding to operator

$$\sigma_{ix} (i = 1, 2, 3).$$

All six of the elements of reality
 S_{ix} and S_{iy} have to be there, each
with the values +1 and -1!

$$S_{1y}S_{2y}S_{3x} = -1,$$

$$S_{1y}S_{2x}S_{3y} = -1,$$

$$S_{1x}S_{2y}S_{3y} = -1.$$

What Outcomes Are Possible?

Consider measurement of 45° linear polarization basis

Local realism:

$$\begin{aligned} S_{1x}S_{2x}S_{3x} &= S_{1x}(S_{1y})^2 S_{2x}(S_{2y})^2 S_{3x}(S_{3y})^2 \\ &= (S_{1x}S_{2y}S_{3y})(S_{1y}S_{2x}S_{3y})(S_{1y}S_{2y}S_{3y}) \\ &= -1 \end{aligned}$$

Possible outcomes:

$$V_1'V_2'V_3', H_1'H_2'V_3', H_1'V_2'H_3', V_1'H_2'H_3'$$

What Outcomes Are Possible?

Quantum physics

$$\begin{aligned} |\psi_{123}\rangle &= \frac{1}{2}(H_1'H_2'H_3' + H_1'V_2'V_3' + V_1'H_2'V_3' + V_1'V_2'H_3') \\ &\Rightarrow S_{1x}S_{2x}S_{3x} = 1! \end{aligned}$$

Possible outcomes:

$$H_1'H_2'H_3', H_1'V_2'V_3', V_1'H_2'V_3', V_1'V_2'H_3'$$

Whenever local realism predicts a specific result definitely to occur for a measurement for one of the photons based on the results for the other two, quantum physics definitely predicts the opposite result

First Observation of Quantum Entanglement



John Wheeler

Two quanta emitted in the annihilation of a positron-electron pair, with zero relative angular momentum, are polarized at right angles to each other

Ann. New York Acad. Sci. 48, 219 (1946)



Chien-Shiung Wu

First observation of quantum entanglement

The Angular Correlation of Scattered Annihilation Radiation*

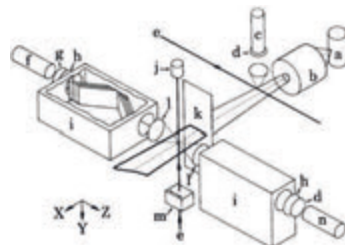
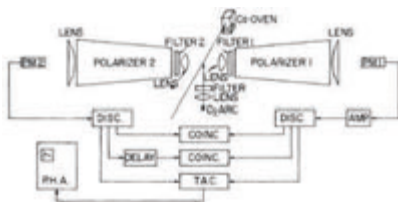
C. S. WU AND I. SHAKNOV

Pupin Physics Laboratories, Columbia University, New York, New York

November 21, 1949

Phys. Rev. 77, 136 (1950)

Experimental Test of Bell Inequality

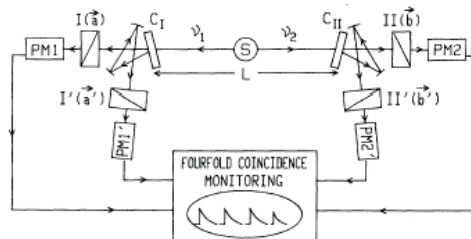


- Freedman & Clauser, PRL 28, 938 (1972)
- Fry & Thompson, PRL 37, 465 (1976)

Two measurement sites are not space-like separated

Experimental Test of Bell Inequality

Aspect *et al.*, PRL 49, 1804 (1982)



violates a generalized inequality by 5 standard deviations

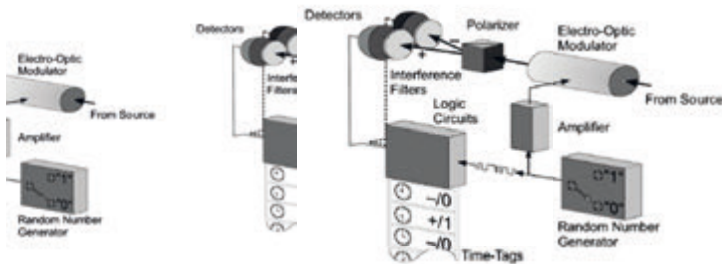
Drawbacks: 1. locality loophole
2. detection loophole

Locality Loophole

- ✗ Measurement devices may "tell" the EPR source their basis choices → the source may "select" according events to violate Bell inequality
- ✓ Solution: basis choice and emission of EPR source must be also space-like separated (i. e., fast and random switch of measurement basis)

Experimental Test of Bell Inequality

Weihs *et al.*, PRL 81, 5039 (1998)



violates CHSH inequality by 30 standard deviations

Drawback: detection loophole

Detection Loophole

❌ Detection efficiency of single photon detectors is not unity → some events cannot contribute to were not detected?

✅ Solution: high detection efficiency (>83%)

Pearle, PRD 2, 1418 (1970)

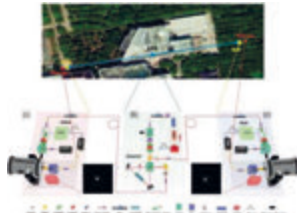
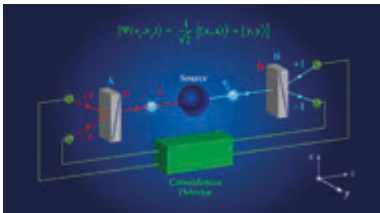
Garg & Mermin, PRD 35, 3831 (1987)

Closed detection loophole:

• Rowe *et al.*, Nature 409, 791 (2001)

Bell's Inequality: Testing the Battle

Closed both detection and locality loopholes



- Hensen *et al.*, Nature 526, 682 (2015)
- Giustina *et al.*, PRL 115, 250401 (2015)
- Shalm *et al.*, PRL 115, 250402 (2015)
- Li *et al.*, PRL 121, 080404 (2018)

Quantum mechanics is right!
But still with loopholes...

Freedom of Choice and Collapse Locality Loophole

☒ **Freedom of choice loophole:** random number generators (RNGs) could be prior correlated → the choice of measurement bases are not truly random
Brunner *et al.*, RMP 86, 419 (2014)



Schrödinger's cat

☒ **Collapse locality loophole:** measurement outcome is not defined until it is registered by a human consciousness →

Realized "events" have never been space-like separated
Kent, PRA 72, 012107 (2005)

Leggett, Compendium of Quantum Physics (Springer, 2009)

Bell-test experiment with human-observer

Solution: Bell-test experiment with human-observer!

Why need human-observer?

Though in "Westworld" : AI "thinks" she has free consciousness

- ☒ Her every actions in future have been indeed **priori determined** by the remote control station.....



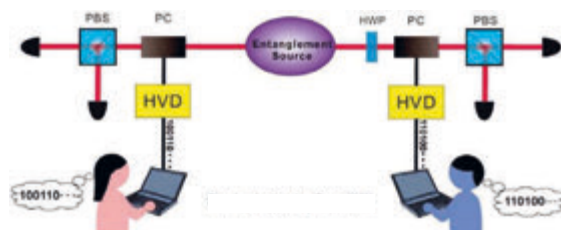
Bell's Inequality: Testing the Battle

- ☑ Basis choice by free will

- ☑ Measurement outcomes defined by consciousness

Kent, PRA 72, 012107 (2005)

Leggett, Compendium of Quantum Physics (Springer, 2009)



Requirement:

Quantum signal transit time exceeds human reaction (100ms) → entanglement distribution at a distance on the order of **one light-second** (e. g., between Earth and Moon, 1.28 ls)

Quantum Information Processing (QIP)

Test of quantum nonlocality

Coherent manipulation of quantum systems

Enabling encode and process information in quantum states, outperform classical information systems in terms of

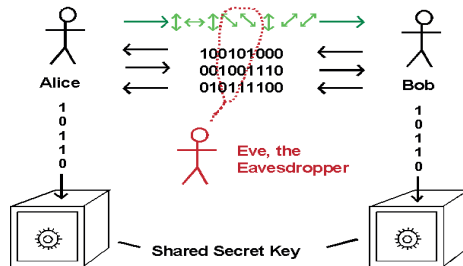
Unconditional security Computational capacities Super-resolution

Quantum communication Quantum computation
and simulation Quantum metrology

Part 2: Quantum Communication

Quantum Key Distribution (QKD)

➤ **Single-photon-based key distribution:** [Bennett & Brassard 1984 protocol]



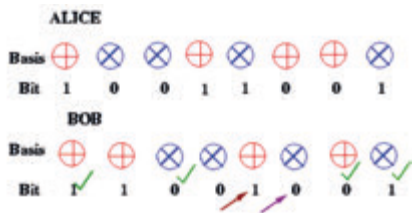
➤ **Entanglement-based key distribution:** [Ekert, PRL 67, 661 (1991)]

BB84 Protocol

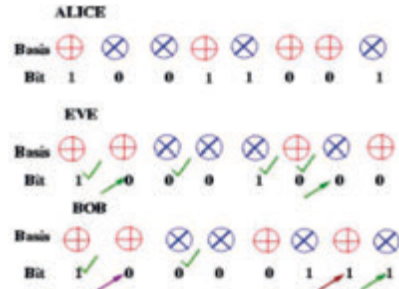
1. Alice tosses a coin several times and notes out come each time (i.e. generates a random sequence of 0s and 1s)
2. If it is head she decides to encode using a H/V basis. If it is a tail, she encodes in 45/135 basis.
3. Each bit is encoded as 0 or 1 in the chosen basis
4. Bob receives each bit and does not know the basis used to encode. He also tosses a coin and decides to decode using the basis as decided by coin toss
5. Half the time Bob's basis will be the same as Alice's in which case the qubit received will be the same provided Eve is not intercepting.
6. Alice now uses a classical channel to announce the basis that she used each time. Bob discards those where the bases are different. The remained bits are called raw key (with an efficiency of 50%).
7. Bob now announces a part of the qubits. Alice can conclude whether an eve is present

BB84 Protocol

➤ No Eve



➤ With Eve



If Eve is present, the probability that Alice and Bob can not tell is $(0.25)^N$ after they compare N raw key's value!

BB84 Security

	one-way communication	two-way communication
Upper bound	14.6%	25%
Lower bound	11.0%	18.9%

- All the error rates are brought by the eavesdropping
- When the error rate is lower than the lower bound, we can utilize some classical cryptography method to let Eve know nothing about the key
- If the error rate is higher than the upper bound, the key is insecure

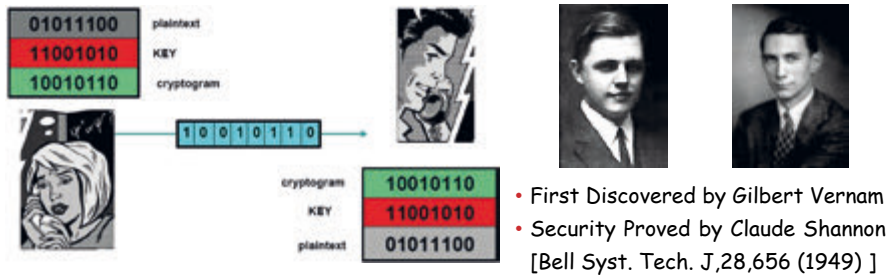
-----Gottesman and Lo, IEEE TIT-49, 457 (2003)-----

Perfect Cipher in Principle

QKD \Rightarrow Secure key

+

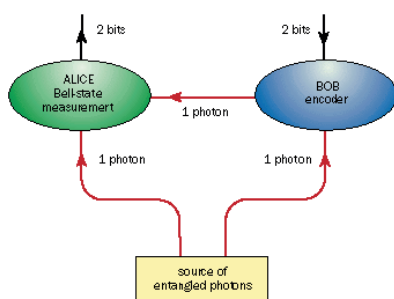
One-time pad



Unconditional security!

Dense Coding

Transmit two bits of information by sending one photon



Transformations between 4 Bell states:

Bennett & Wiesner, PRL 69, 2881 (1992)

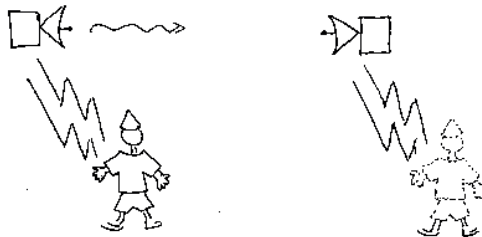
Dense Coding

1. Alice and Bob share an entangled photon pair in the state of
2. Bob chooses one of the four unitary transformation on his photon.
The information of which choice is **2 bit**.

e. g.

3. Bob sends his photon to Alice
4. Alice does a joint Bell-state measurement (BSM) on the photon from Bob and her photon.
5. With the measurement result, she can know Bob's unitary transformation and achieve the 2 bit information.

Quantum Teleportation



➤ Classical physics

Scanning and reconstructing

➤ Quantum physics

Principle of quantum measurement forbidden extracting all the information from an unknown quantum state!

Quantum Teleportation

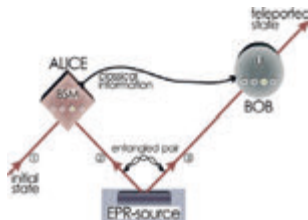
Initial state

$$|\Phi\rangle_1 = \alpha |H\rangle_1 + \beta |V\rangle_1$$

The shared entangled pair

$$|\Phi^+\rangle_{23} = \frac{1}{\sqrt{2}} (|H\rangle_2 |H\rangle_3 + |V\rangle_2 |V\rangle_3)$$

$$\begin{aligned} |\Psi\rangle_{123} &= |\Phi\rangle_1 \otimes |\Phi^+\rangle_{23} \\ &= |\Phi^+\rangle_{12} \otimes (\alpha |H\rangle_3 + \beta |V\rangle_3) + \\ &\quad |\Phi^-\rangle_{12} \otimes (\alpha |H\rangle_3 - \beta |V\rangle_3) + \\ &\quad |\Psi^+\rangle_{12} \otimes (\alpha |V\rangle_3 + \beta |H\rangle_3) + \\ &\quad |\Psi^-\rangle_{12} \otimes (\alpha |V\rangle_3 - \beta |H\rangle_3) \end{aligned}$$



BSM results on particles 1, 2	operations on particle 3

Bennett *et al.*, PRL 73, 3801 (1993)

Quantum Teleportation



Quantum version of cloud-somersault

Though nowadays we can only teleport two-particle composite system.....

Essential ingredient for distributed quantum information processing!

Part 3: Quantum Computation and Quantum Metrology

Quantum Computation

Quantum Parallelism

Bits		Qubits
0 or 1	V. S.	0 + 1
00, 01, 10 or 11		00 + 01 + 10 + 11
000, 001, 010.....		000 + 001 + 010 +

Evaluating function $f(x)$ for many different x simultaneously

Exponentially speedup!

This is what makes famous quantum algorithms, such as Shor's algorithm for factoring, or Grover's algorithm for searching

RSA Encryption and Factorizing

RSA public-key cryptosystem

- Produce a large integer N
 $m1 \times m2 = N$, (with $m1$ and $m2$ primes)
- N is made public available and is used as a key (x) to encrypt data
- $m1$ and $m2$ are the secret keys (k) enable one to decrypt the data

$$C = E_x(P)$$

$$P = D_k(C) = D_k(E_x(P))$$

X: Public Key; K: Private Key

P: Plain Text; E: Encryption; C: Ciphertext; D: Decryption

Rivest, Shamir and Adleman, MIT/LCS/TR-212, Jan. 1979

RSA Encryption and Factorizing

- To crack a code, a code breaker needs to factorize N
- The security of RSA based on the **ease** with which N can be calculated from $m1$ and $m2$, and the **difficulty** of calculating $m1$ and $m2$ from N

RSA Encryption and Factorizing

- Problem: given a number, what are its prime factors ?
e. g. a 129-digit odd number which is the product of two large primes,
11438162575788886766923577997614661201021829672124236256256184293570
693524573389783059712363958705058989075147599290026879543541
=3490529510847650949147849619903898133417764638493387843990820577
x 32769132993266709549961988190834461413177642967992942539798288533
- Best factorizing algorithm requires sources that grow exponentially in the size of the number: , with n the length of N

Shor's Algorithm

Algorithms for quantum computation: discrete logarithms and factorizing

E.g. factor a 300-digit number with

- Classical THz computer: 10^{24} steps \Rightarrow 150,000 years
- Quantum THz computer: 10^{10} steps \Rightarrow 1 second!



Peter Shor

Foundations of Computer Science, 1994 Proceedings. 35th Annual Symposium

- Code-breaking can be done in minutes, not in millennia
- Public key encryption, based on factoring, will be vulnerable!

Deutsch-Jozsa Algorithm

Deutsch's problem: two types of functions

Considering input n bits,

- Constant f : for all 2^n inputs, $f=0$ or $f=1$
- Balanced f : for 2^{n-1} inputs, $f=0$, for another 2^{n-1} inputs, $f=1$

Question: given a function f , whether is it constant or balanced ?

Classical deterministic algorithm: at most $2^{n-1}+1$ inquiries

- All outputs are the same \rightarrow constant
- At least 1 output is different from others \rightarrow balanced

Deutsch-Jozsa Algorithm

The simplest
example:
($x=0$ or 1)

Constant:

$f(0)=1$

$f(1)=1$

Balanced:

$f(0)=0$

$f(1)=1$

- Classical algorithm needs 2 inquiries
- Deutsch-Jozsa quantum algorithm:
Assume f was mapped into a quantum oracle satisfying

e. g.,

Deutsch-Jozsa Algorithm

- Prepare two qubits input state
- Perform Hadamard operation
- Perform

Key point: All possible inputs are in superposition state!

Output state:

- Measure the first qubit on $\{+/-\}$ basis: \Rightarrow constant f , \Rightarrow balanced f

Quantum algorithm only needs one inquiry

Deutsch-Jozsa Algorithm

Consider a more general case with n -bit inputs x : $x=0, 1, 2, \dots, 2^n-1$

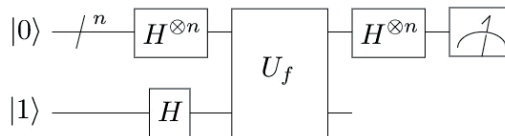
- Prepare $n+1$ qubits input state
- Perform Hadamard operation on all qubits

The binary representation of x corresponds to values of each qubits, e. g.,

- Perform

Deutsch-Jozsa Algorithm

- Measure the first n -qubit on $\{+/-\}$ basis: \Rightarrow if and only if the output is $+$, f is constant
Only needs one inquiry!



Deutsch and Jozsa, Proc. Royal Society London A 439, 553 (1992)

Deutsch problem is not a practically important problem, but Deutsch-Jozsa algorithm firstly demonstrated the superiority of quantum computation!

Grover's Search Algorithm



Lov Grover

How quickly can you find a needle in a haystack ?

The simplest example:

Which one is equal to -1 in a database?

Serial	0	1	2	3	4	5
Value	1	1	-1	1	1	1	

➤ Classically search

- Sequentially try all N possibilities
- Average search takes $N/2$ steps

➤ Quantum search

- Simultaneously try all possibilities
- Refining process reveals answer
- Average search takes $N^{1/2}$ steps

Grover's Search Algorithm

- A database is encoded with a $N \times N$ diagonal matrix R (rotate phase)

The task is to find x

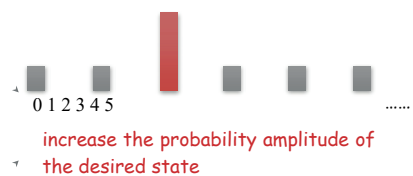
- Take a m -qubit register ($2^m = N$), and prepare the registers in an equal superposition state of all the states

- Perform rotate phase matrix R on the register

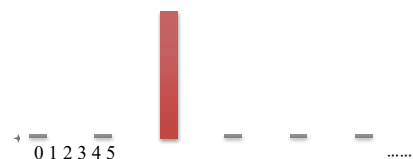


Grover's Search Algorithm

- Then perform diffusion operator D



- Iterations of operators R and D



- Measure the register to get the specific state with nearly unity probability

Grover's Search Algorithm

Formulas

- Initial state:
- Phase rotation R :

- Diffusion operator D :

Grover's Search Algorithm

Formulas

- After n iteration:

- The probability to collapse into the x
- Choose iteration steps

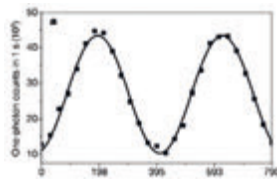
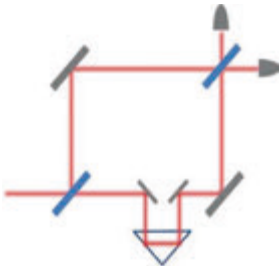
- The probability of failure:

Grover, PRL 79, 325 (1997)

Quantum Metrology

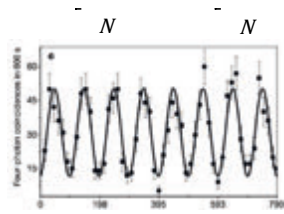
Super-resolution with multi-particle entanglement

Single particle:



Phase uncertainty with N sampling:

N -particle $NOON$ state



v. s.

Phase uncertainty with same cost of resource N :

Part 4: Quantum Repeaters

Noise Environment

Unavoidable interaction with environment and decoherence will happen

- represents the qubit state and represents the environment initial state, is the joint unitary time evolution operator
- For arbitrary qubit state:

$$(\alpha_0|0\rangle + \alpha_1|1\rangle)|E\rangle \xrightarrow{U(t)} \alpha_0|0\rangle|E_0(t)\rangle + \alpha_1|1\rangle|E_1(t)\rangle \quad \rho_q(t) = \text{Tr}_E \rho_{q+E} = \begin{bmatrix} |\alpha_0|^2 & \alpha_0 \alpha_1^* \langle E_1 | E_0 \rangle \\ \alpha_1 \alpha_0^* \langle E_0 | E_1 \rangle & |\alpha_1|^2 \end{bmatrix}$$

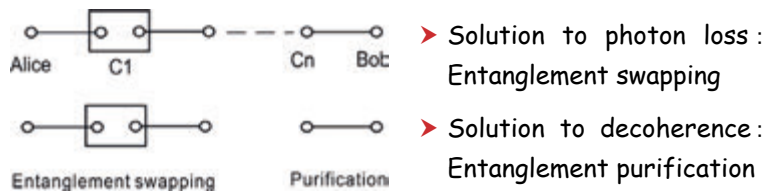
The off-diagonal element of the qubit density matrix will drop down with the rate

The maximally entangled state will be in some mixed state with a certain entanglement fidelity due to the process

Channel Loss

- ☒ Photon loss increases exponentially with channel length:
(e. g., in commercial fiber)
- ☒ For 1000 km commercial fiber, even with a perfect 10 GHz single-photon source and ideal detectors, only **0.3** photon can be transmitted on average **per century!**

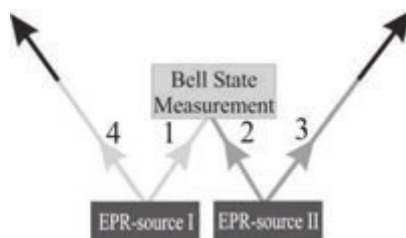
Solution: quantum repeater



Briegel *et al.*, PRL 81, 5932 (1998)

Entanglement Swapping

Entangling the remote particles which never interacted!



Zukowski *et al.*, PRL 71, 4287 (1993)

Entanglement Purification

Initially pure singlet state

Mixed state:

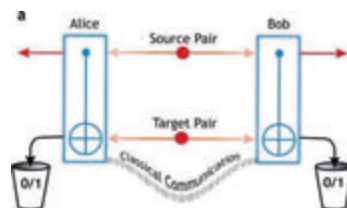
Fidelity :

Goal: to extract from a large ensemble of low-fidelity M a small sub-ensemble with sufficiently high fidelity

Scheme for Entanglement Purification

- Random bilateral Pauli rotation on each photon in the states \Rightarrow change arbitrary mixed state into Werner state:

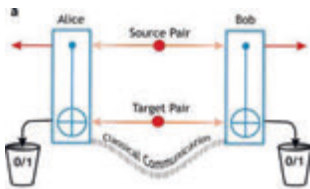
- For two same pairs of Werner states, we consider them as source pair and target pair respectively
- A unilateral is performed on each of the two pairs:



i. e., states with a large component ($F > 1/2$) of , and equal components of the other three Bell states

Scheme for Entanglement Purification

- Perform CNOT operation on source and target pairs:

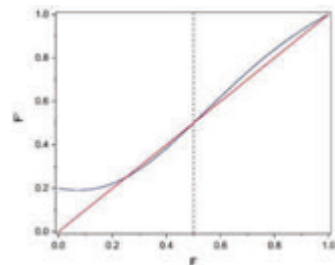
[illegible]

- Measure target pair in $\{H/V\}$ basis, keep the unmeasured source pair when measuring results are same


Scheme for Entanglement Purification

- After that, the component of \mathbf{u} of the target pair will be

- Equivalently, the fidelity is equal to
- Via several of these kind of processes, we can purify a general mixed state into a highly entangled state



Bennett *et al.*, PRL 76, 722
(1996)



Lecture 2: Scalable Quantum Information Processing with Photons and Atoms

Part 1: Elemental Optical Manipulations and Demonstrations of Quantum Communication



Optical Quantum Information Processing

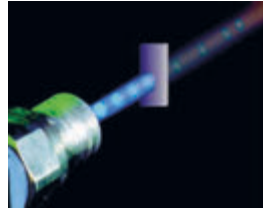
Why do we like photons?

- ☑ Flying qubit (fastest quantum information transmitter)
- ☑ Robust qubit (with weak interaction with environment)
- ☑ High-precision manipulation with off-the-shell devices

Single Photons: Generation & Detection

Generation of single photons

- Practical single-photon source is far out of reach within current technologies
- Probabilistic quasi single photon: weak coherence pulse



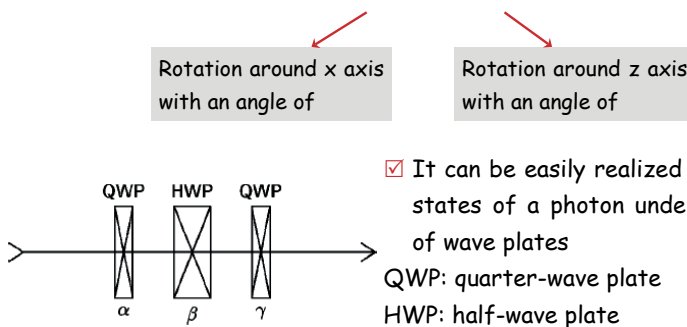
Single photon detector

- InGaAs Avalanche photo diode
- Si detector
- Superconducting nanowire detector.....

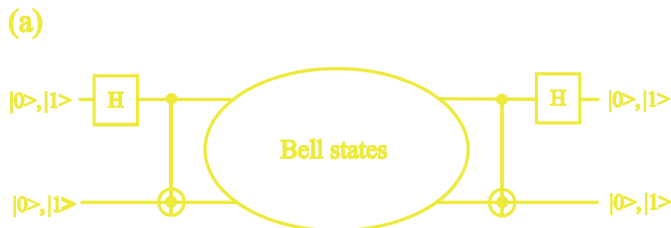


Single-qubit SU(2) Rotations

Arbitrary SU(2) rotation can be achieved by 3 elemental rotations:



Manipulation of Entanglement

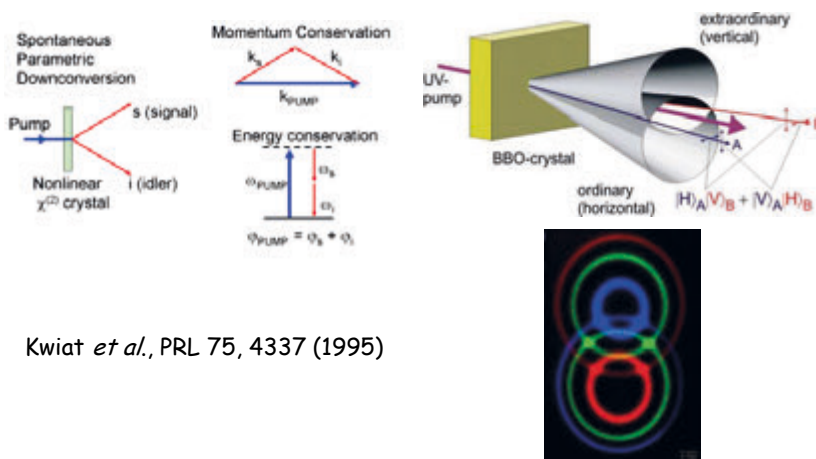


For photons, CNOT gate requires strong non-linear coupling

☒ But the coupling between photons is negligibly weak!

Probabilistic Generation of Photonic Entanglement

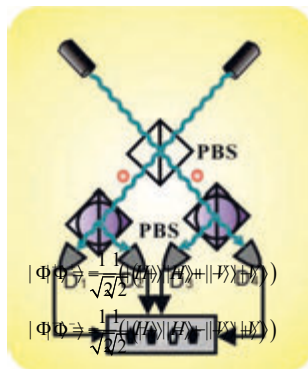
Spontaneous Parametric Down-conversion (SPDC)



Kwiat *et al.*, PRL 75, 4337 (1995)

Bell-state Measurement (BSM) with Linear Optics

Pan and Zeilinger, PRA 57, 2208 (1998)



	D1	D2
D3	$ 0\rangle 0\rangle$	$ 1\rangle 1\rangle$
D4	$ 0\rangle 1\rangle$	$ 1\rangle 0\rangle$

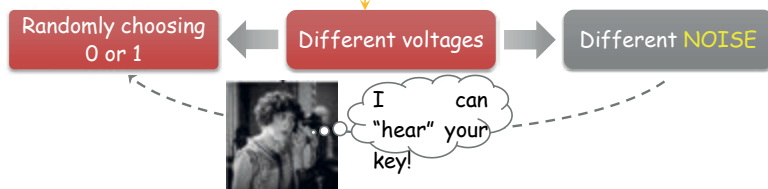
Required non-linearity of CNOT gate can be effectively induced with the help of post-selection measurements

Proof of Concept Demonstrations of QKD



First concept demonstration (32 cm)
Bennett *et al.*, J. Cryptol. 5, 3 (1992)

☒ Security loopholes due to imperfection of realistic devices



The experiment is unconditionally secure, unless you are a deaf

--Gilles Brassard

Security Loophole of QKD with Realistic Devices

Probabilistic quasi single
photon: weak coherence pulse



Eavesdrop the keys with two photon events
(Photon number splitting attack)

Brassard *et al.*, PRL 85, 1330 (2000)
Lütkenhaus, PRA 61, 052304 (2000)

Due to imperfect single-photon source:

- ❑ Not secure when distance is longer than ~10km in fiber
- ❑ Very low key rate



~10km



If so, why we need QKD?

Security of QKD with Realistic Devices

➔ Security loophole 1: imperfect single-photon source

Quasi single photon source:

Two identical photons per pulse with
probability $P/2$

Photon number splitting attack (PNS):
Eavesdrop the keys with two photon events
Brassard *et al.*, PRL 85, 1330 (2000)



- ❑ Not secure when distance is longer than ~10km
- ❑ Very low key rate

Security of QKD with Realistic Devices

⇒ Solution: Decoy-state QKD scheme: sending pulses randomly with intensity P_1 or P_2

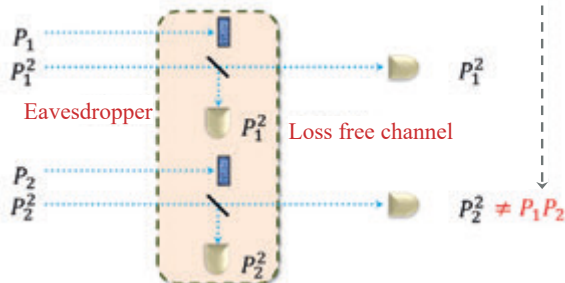
• Wang, PRL 94, 230503 (2005)

• Lo *et al.*, PRL 94, 230504 (2005)

Without
eavesdropping:



With eavesdropping:



Security of QKD with Realistic Devices

Experiments

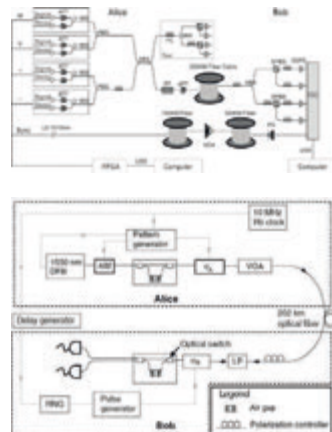
100km:

Rosenberg *et al.*, PRL 98, 010503 (2007)

Peng *et al.*, PRL 98, 010505 (2007)

200km:

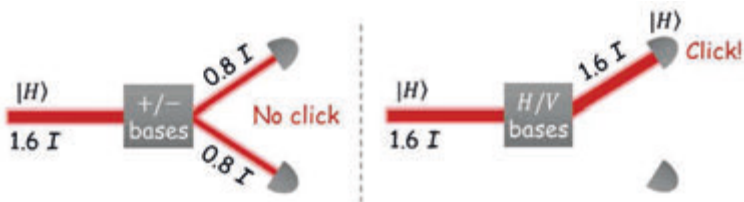
Liu *et al.*, Optics Express 18, 8587 (2010)



Security of QKD with Realistic Devices

➔ Security loophole 2: imperfect single-photon detectors

Blinking attack: can fully control detectors by specially tailored strong light [Lydersen *et al.*, Nature Photonics 4, 686 (2010)]



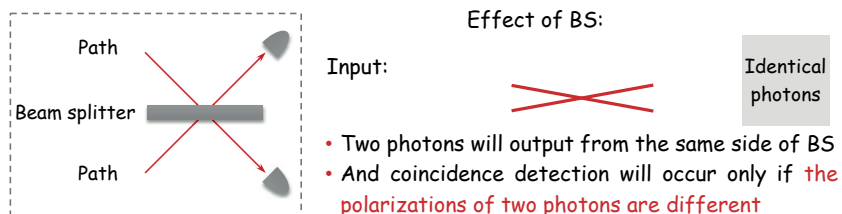
Security of QKD with Realistic Devices

➔ Solution: Measurement Device Independent QKD: Immune to any attack on detection

- Scheme: Lo *et al.*, PRL 108, 130503 (2012)

Key point: two-photon interference (HOM effect)

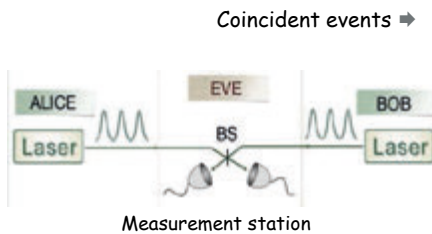
consider simultaneously input two photons with the same polarization to a BS




















Hong, Ou & Mandel, PRL 59, 2044 (1987)

Security of QKD with Realistic Devices

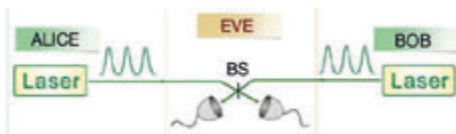
- Alice and Bob send one of four polarization states randomly to measurement station



Alice's basis	Alice's state	Bob's basis	Bob's state
+		+	
+		+	
+		+	
+		+	
×		+	 or 
×		×	 
×		×	
		×	
×		×	

- Compare their basis in public channel, keep the cases that basis choices are the same
- Share key according to anti-correlation of polarizations
- Even measurement station is fully controlled by Eve, she can only reveal the correlation information, but gains no information of the key

Security of QKD with Realistic Devices



High-precision interference between two remote independent lasers:
relative timing jitter after hundreds km fiber < 10ps

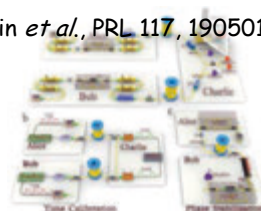
First experiment (50km):

- Liu *et al.*, PRL 111, 130502 (2013)



Extended distance:

- 200km: Tang *et al.*, PRL 113, 190501 (2014)
- 404km: Yin *et al.*, PRL 117, 190501 (2016)



Security of QKD with Realistic Devices

MDI-QKD
+
DIY light source (Do It Yourself)

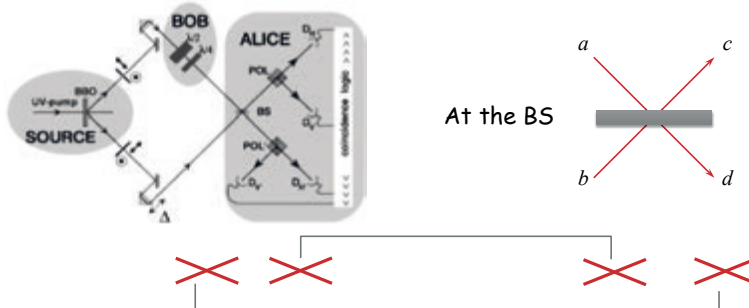


Information-theoretically secure QKD with realistic devices
can be approached properly!

In MDI-QKD.....They need only trust themselves not to have inadvertently created a side channel to Eve through incompetent design of their do-it-yourself light sources

-- Charles Bennett

Experimental Dense Coding

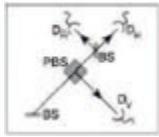


⇒ Coincidence between D_H and D_V , or $D_{H'}$ and $D_{V'}$

Similarly,

- ⇒ Coincidence between D_H and $D_{V'}$, or $D_{H'}$ and D_V (only will cause coincidence between the different outputs of the BS)
- ⇒ 2 photons in one of detectors

Experimental Dense Coding



To distinguish by coincidence, insert an additional BS before two photons arrive detectors

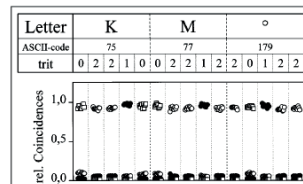
$$0 \equiv |\Phi^-\rangle \hat{=} \square$$

$$1 \equiv |\Psi^+\rangle \hat{=} \bullet$$

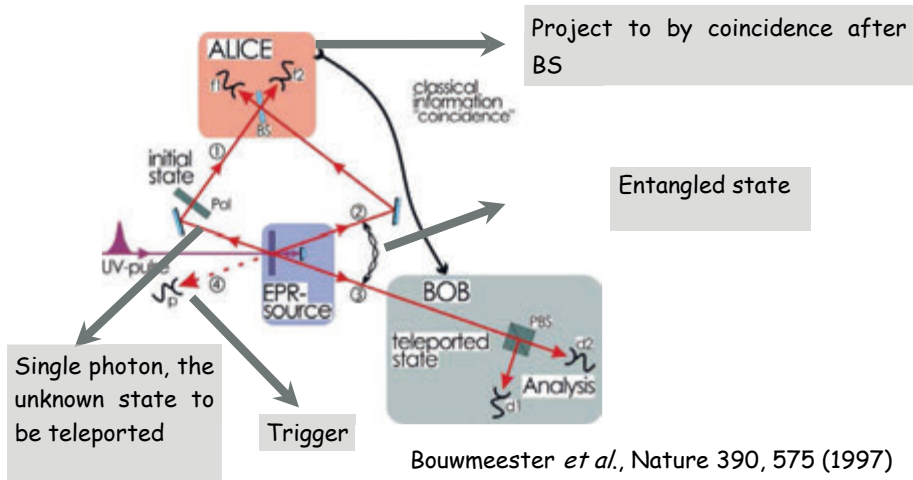
$$2 \equiv |\Psi^-\rangle \hat{=} \circ$$

So the experiment can distinguish 3 Bell states \rightarrow transmit by sending one photon

Mattle *et al.*, PRL 76, 4656 (1996)



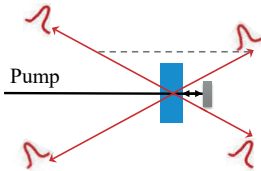
Experimental Quantum Teleportation



Experimental Quantum Teleportation

But it was not so straightforward.....

Two photons must be indistinguishable on the BS \Rightarrow be spatially and temporally overlapped on the BS perfectly

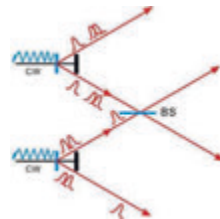


However, if the pulse duration T of pump is too long (e. g., a CW laser)

⊗ A large uncertainty of generation time of two EPR pairs ($\sim T$)

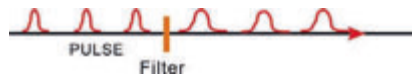
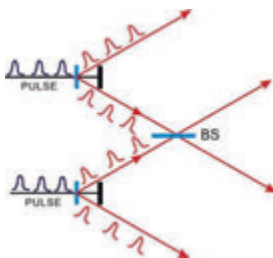
Energy conservation condition, allows some uncertainty of frequency of EPR pair

⊗ The coherent time of EPR pair (, at the order of 100fs) will be much shorter than T



Experimental Quantum Teleportation

⊗ A solution is to use short pulse laser (pump pulses duration: 200 fs)

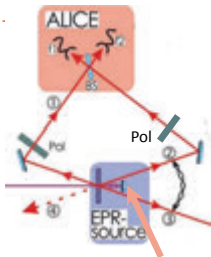


The pulse will bring some time jitter to the SPDC photon

⊗ Insert a narrow band filter can extend the coherent time (4nm results in a coherence time of 520 fs)

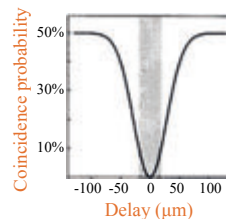
Experimental Quantum Teleportation

- ✗ The coherent time of EPR photons is **definitely shorter** than the time resolution of state-of-the-art single photon detectors in 1997 → we cannot confirm that photons were well-overlapped at the BS by detecting the arriving time
- ✓ Scan the interference fringes



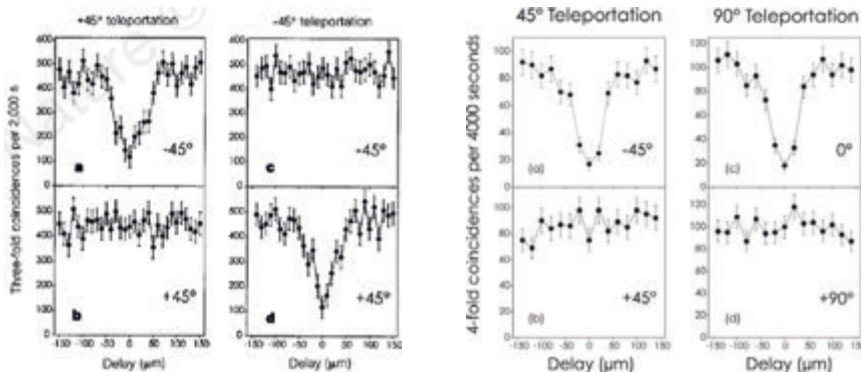
Adjusting delay between photon 1 and 2

- Insert two polarizers to make the polarizations of photon 1 and 2 the same
- Due to HOM effect, there will be no coincidence in theory when two photons are well-overlapped
- Adjust delay to find the optimal position



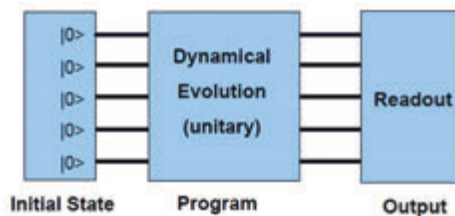
Experimental Quantum Teleportation

The experimental results



Part 2: Multi-photon Interferometry

Quantum Computaiton : Quantum Circuit Model



Single-qubit $SU(2)$ gates
+
Two-qubit controlled-NOT gates



Universal quantum computation
Lloyd, PRL 75, 346 (1995)

☒ With linear optics, nonlinearity required by CNOT gate can only be induced by post-selection.

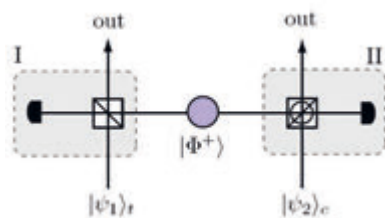
Is it sufficient for efficient quantum computation?

Quantum Computaiton : Quantum Circuit Model

Efficient quantum computation is possible with linear optics

Knill, Laflamme and Milburn (KLM), Nature 409, 46 (2001)

- ☑ Non-deterministic quantum logic operations can be performed using linear optical elements
- ☑ The success rate of the quantum logic can be arbitrarily close to one



Probabilistic photonic CNOT gate

Scheme:

- Pittman *et al.*, PRL 88, 257902 (2002)

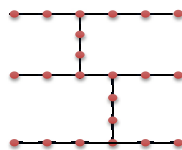
Experiments:

- O'Brien *et al.*, Nature 426, 264 (2003)
- Gasparoni *et al.*, PRL 93, 020504 (2004)
- Zhao *et al.*, PRL 94, 030501 (2005)

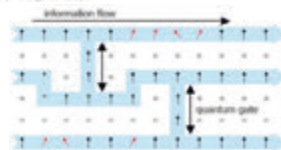
One-way Quantum Computation Model

- Universal QC resource: multi-particle entanglement

cluster state :



$$\sigma_x^{(a)} \bigotimes_{a' \in \text{neigh}(a)} \sigma_z^{(a')} |\Phi\rangle_C = \pm |\Phi\rangle_C$$



- Quantum gate is implemented by measuring particles in a certain order and in a certain basis:



Raussendorf and Briegel, PRL 86, 5188 (2001)

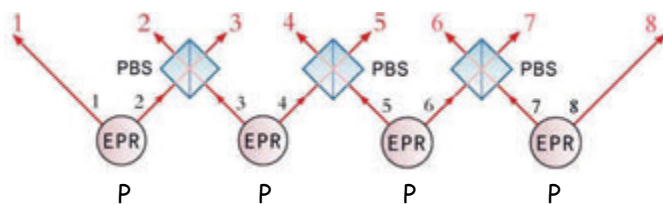
Resource Reduction for Optical Quantum Computation

- KLM scheme: $\sim 10^5$ photons, and efficiency (source, detection) threshold of more than 99%

- Browne and Rudolph, PRL 95, 010501 (2005) (cluster state)
 - Ralph *et al.*, PRL 95, 100501 (2005)
- ~10² photons

- ## Essential task: generation and manipulation of multi-photon entanglement!

Multi-photon Interferometry

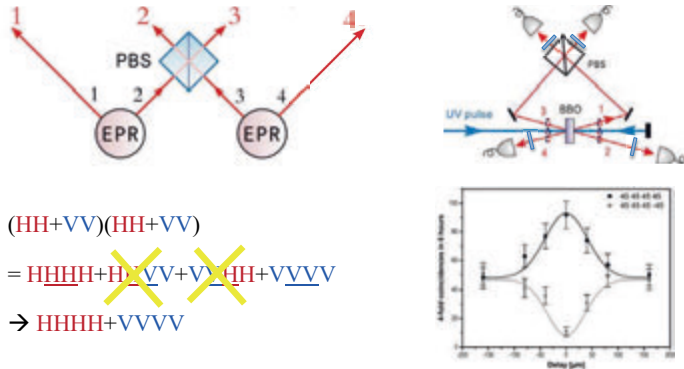


Two-photon entanglement source: $P \rightarrow$ Four-photon entanglement: $P^2/2$
 \rightarrow Six-photon entanglement: $P^3/4 \rightarrow$ Eight-photon entanglement: $P^4/8 \dots$

One must need high-brightness entanglement source!

Multi-photon Interferometry

In 2001: Brightness of entanglement source: 2500pair/s@76MHZ

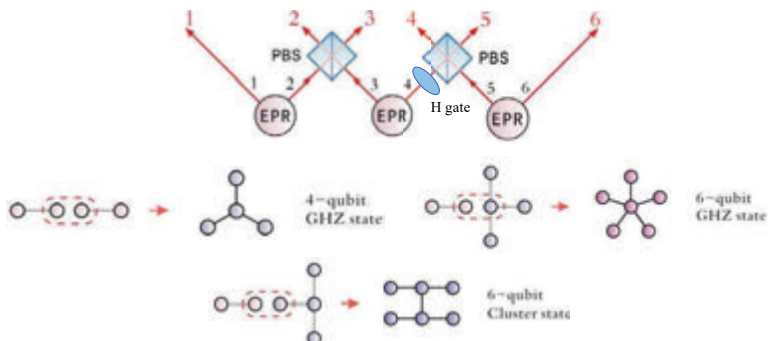


$$\begin{aligned}
 & (HH+VV)(HH+VV) \\
 & = \cancel{HHHH} + \cancel{HHVV} + \cancel{VVHH} + VVVV \\
 & \rightarrow HHHH + VVVV
 \end{aligned}$$

Pan *et al.*, PRL 86, 4435 (2001)

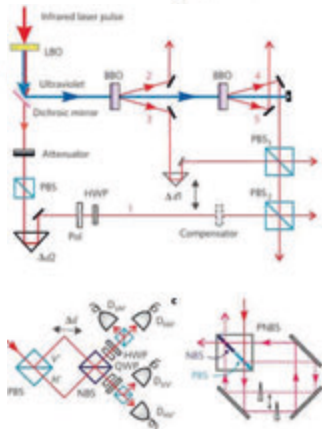
Six-photon Cluster States

In 2007: A brighter, stable laser pump source, Verdi 10W → 16W, IR ~ 2.5W
 Brightness of entanglement source: 93000pair/s@76MHZ



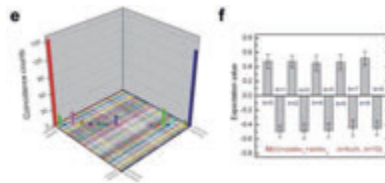
Lu *et al.*, Nature Physics 3, 91 (2007)

Hyper-entangled Schrödinger Cat States



Hyper-entangled state: Polarization and spatial modes

5-photon 10-qubit cat state



Gao *et al.*, Nature Physics 6, 331 (2010)

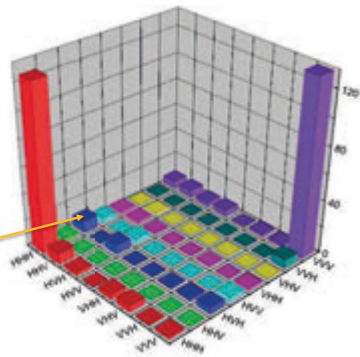
The Request for Both High Brightness & Fidelity

With higher pump

- ☑ Increase probability
- ☒ More double pair emissions →
degrades fidelity

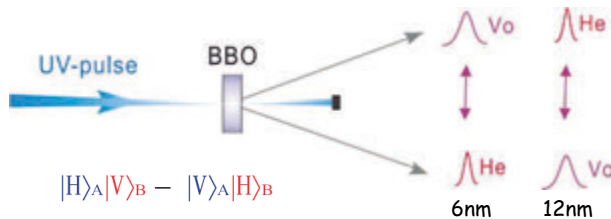


Error



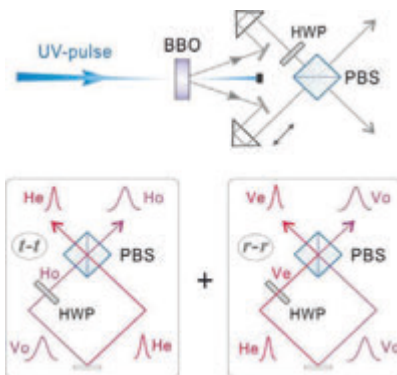
Can we have very bright source of entangled photons, meanwhile with high fidelity?

Frequency-uncorrelated Entangled Photons



- ☑ The o and e light differs in their spectral (and temporal) widths → decrease the indistinguishability thus the fidelity
- ☑ Previous experiments: narrow-band filters (~3nm) → unnecessary waste of photons

Frequency-uncorrelated Entangled Photons



- ☑ ~1 million coincidence counts per second without filter, with ~90% fidelity

Eight-Photon Entanglement

- Fidelity : 0.708
- Brightness : 9 counts per hour

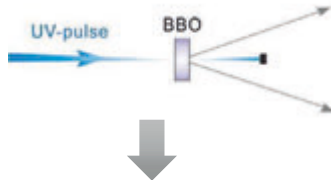


Interferometric Bell-state synthesizer:
disentangles the timing from the polarization

Yao *et al.*, Nature Photonics 6, 225 (2012)

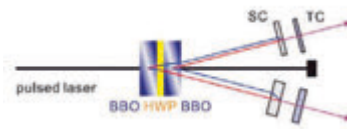
Ten-Photon Entanglement

Previous SPDC method:



❌ Only collect photons from overlaps of up and down circles

To increase count rate:

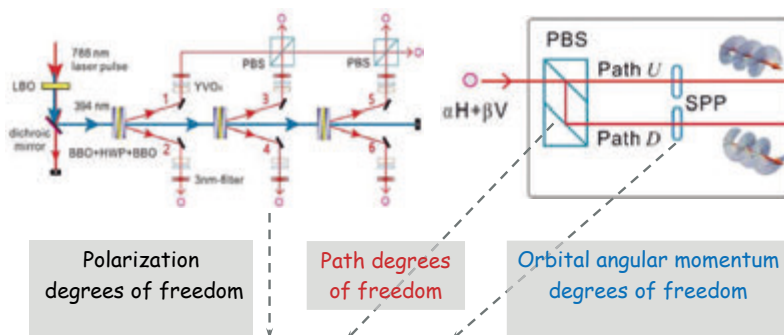


✅ Collect all photons from two separate circular beams
entangled-photon source ~ 4 times brighter than the previous result in eight-photon entanglement

Wang *et al.*, PRL 117, 210502 (2016)

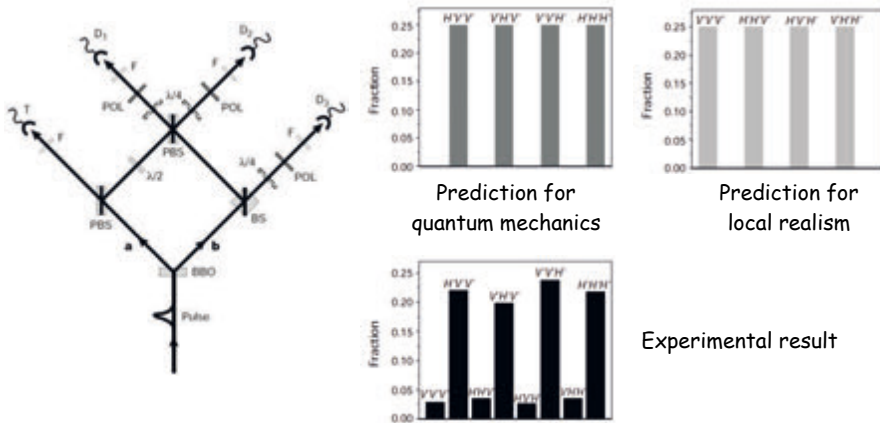
18-qubit Hyper-entanglement

6-photon 18 qubits hyper-entanglement state (largest entanglement)



Wang *et al.*, PRL 120, 260502 (2018)

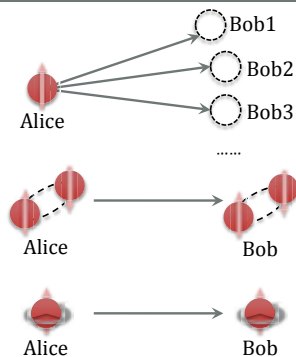
Test of Quantum Nonlocality in 3-photon GHZ entanglement



Pan *et al.*, Nature 403, 515 (2000)

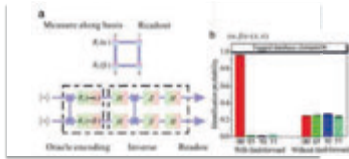
Quantum Teleportation with Multi-photon Entanglement

- **Open-destination teleportation**
Zhao *et al.*, Nature 430, 54 (2004)
- **Teleportation of a composite system**
Zhang *et al.*, Nature Physics 2, 678 (2006)
- **Teleportation of multiple degrees of freedom**
Wang *et al.*, Nature 516, 518 (2015)
- **Teleportation in high dimensions** [Luo *et al.*, PRL 123, 070505 (2019)]

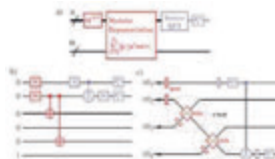


Forming an essential element in quantum computation!

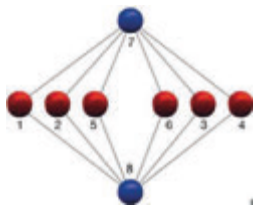
Demonstrations of quantum algorithms



Grover's searching algorithm
Chen *et al.*, PRL 99, 120503 (2007)



Shor's factoring algorithm ($15 = 3 \times 5$)
Lu *et al.*, PRL 99, 250504 (2007)



Topological quantum error correction:
Relax the qubit error threshold from 10^{-5} to 10^{-2}
Yao *et al.*, Nature 482, 489 (2012)

Demonstrations of quantum algorithms

➤ Solving linear systems of equations

➔
Key point: find out eigenvalues of $N \times N$ matrix A

Solving linear equations with 10^{24} variables with

- Taihu Light (~100PFlops, supercomputer No. 1): 100 years
- Quantum THz computer: 0.01 second!

- Classical: $\mathcal{O}(N)$ steps
- Quantum: $\mathcal{O}(\log(N))$ steps

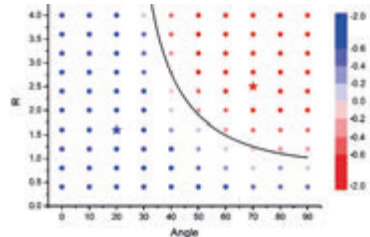
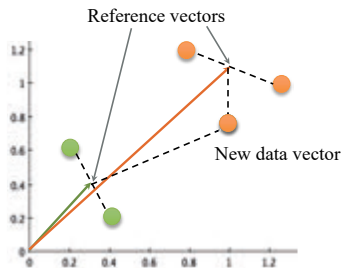


Theory: Harrow *et al.*, PRL 103, 150502 (2009)

Experiment: Cai *et al.*, PRL 110, 230501 (2013)

Demonstrations of quantum algorithms

- Quantum machine learning: recognizing the object from a landscape background
- Key point: calculate the distance of two N -dimensional vectors



Classifying 2D vectors, rate of correctness 98%

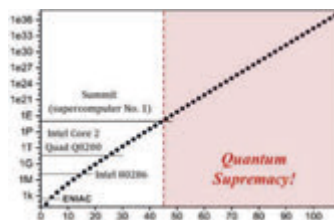
- Classical: $\mathcal{O}(N)$ steps
- Quantum: $\mathcal{O}(\log(N))$ steps

Theory: Lloyd *et al.*, arXiv:1307.0411

Experiment: Cai *et al.*, PRL 114, 110504 (2015)

Roadmap of Quantum Computing

- Milestone 1: Coherent manipulation of **~50 qubits**
 - ➡ beating classical supercomputer in specific tasks (e. g. Boson sampling and portfolio optimization, etc.)

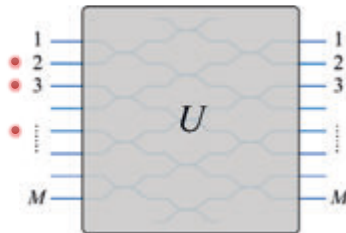


- Milestone 2: Quantum simulation using **hundreds of qubits** ➡ revealing the microscopic mechanism of condensed matter physics (e. g., high temperature superconductivity, etc.)
- Milestone 3: Universal and programmable quantum computers

Achieving quantum supremacy: Boson sampling

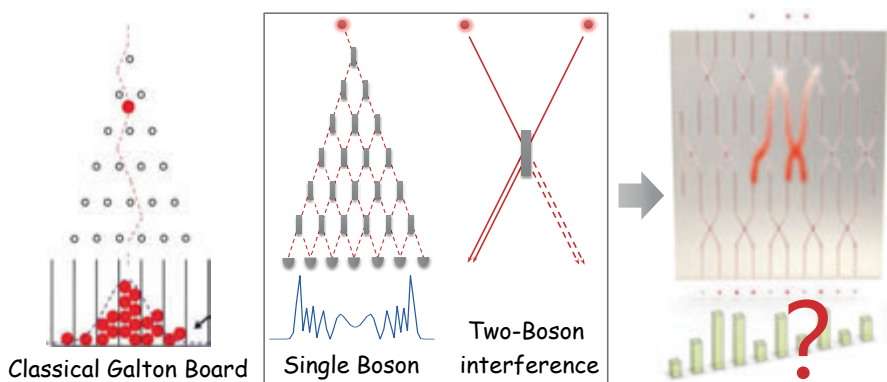
Boson sampling: a specific task in which near-term quantum computers can beat classical supercomputer!

Input identical Bosons (photons) to a multi-channel quantum circuit



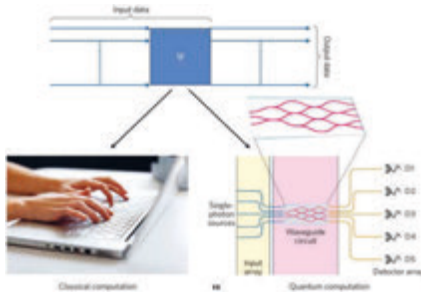
Output distribution?

Achieving quantum supremacy: Boson sampling



Aaronson and Arkhipov, *Theory of Computing*. 9, 143 (2013)

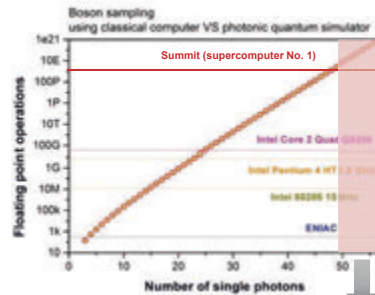
Achieving quantum supremacy: Boson sampling



Classical: $n^2 2^{n+1}$ steps
sharp P-complete problem!

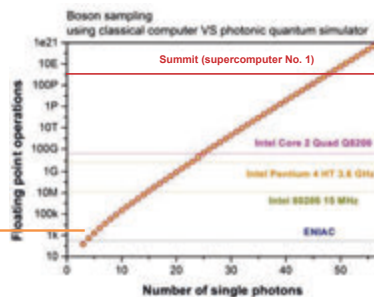
V.S.

Quantum: directly measure the
outputted photons' distribution



Quantum supremacy!

Prototype of quantum computer for Boson sampling

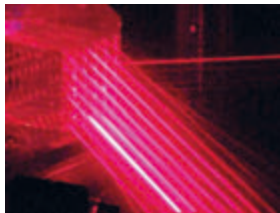
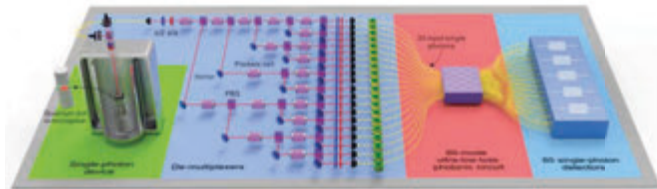


High-efficiency multiphoton boson sampling machine

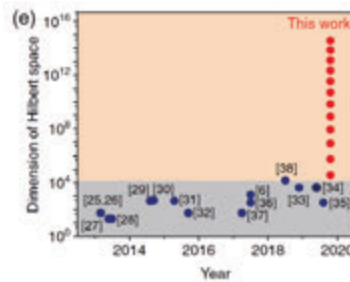
Wang *et al.*, Nature Photonics 11, 365 (2017)

Provably faster than ENIAC and TRADIC for the first time

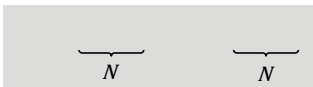
Prototype of quantum computer for Boson sampling



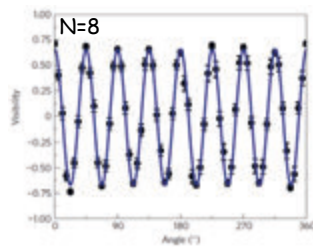
Wang *et al.*, PRL 123, 250503 (2019)



Super-resolution with Multi-photon Entanglement



- Walther *et al.*, Nature 429, 158 (2004), $N=4$
- Nagata *et al.*, Science 316, 726 (2007), $N=4$
- Resch *et al.*, PRL 98, 223601 (2007), $N=6$
- Gao *et al.*, Nature Physics 6, 331 (2010), $N=8$



- Wang *et al.*, PRL 120, 260502 (2018), $N=18$

Part 3: Demonstrations of Quantum Repeaters

Challenge towards long distance quantum communication

- Longest distance of MDI-QKD in fiber: ~400km
 - Yin *et al.*, PRL 117, 190501 (2016)
- Longest distance of quantum teleportation: ~100km
 - Yin *et al.*, Nature 488, 185 (2012), by Chinese group
 - Ma *et al.*, Nature 489, 269 (2012), by Austrian group



Challenge towards Scalable Quantum Information Processing

As mentioned in Lecture 1, we need quantum repeater to overcome

- ☒ Absorption \Rightarrow Photon loss
- ☒ Decoherence \Rightarrow Degrading entanglement quality

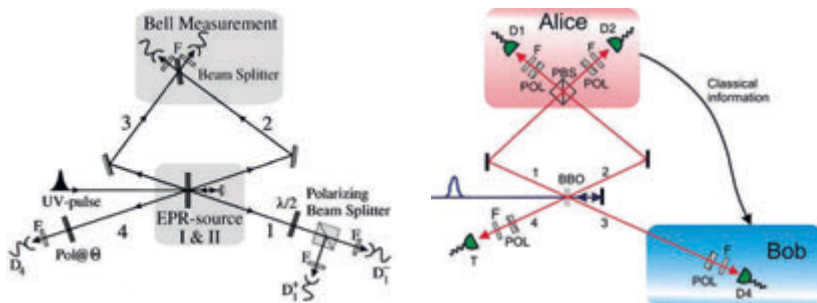
And

- ☒ Probabilistic entangled photons \Rightarrow Exponential resource cost and single photon source

Require

- Entanglement swapping with high precision
- Entanglement purification with high precision
- Quantum memory with high performance

High Precision Entanglement Swapping



First demonstration with beam splitter

Pan *et al.*, PRL 80, 3891 (1998)

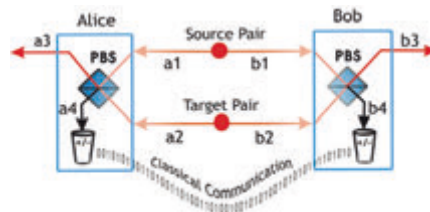
High precision fault-tolerable entanglement swapping

Pan *et al.*, Nature 421, 721 (2003)

Practical Scheme for Entanglement Purification

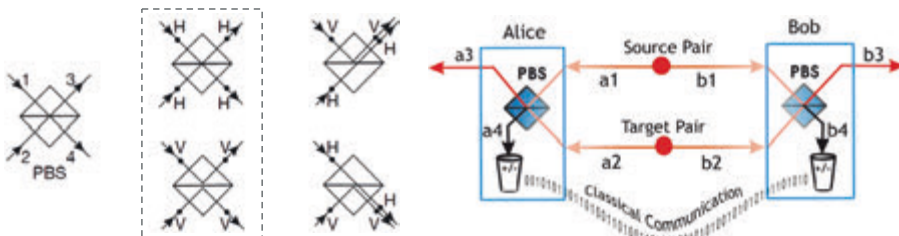
- ❑ Original entanglement purification scheme requires CNOT operation between independent photons
 - ✅ Practical scheme: non-linearity effectively induced by post-selection
- Pan *et al.*, Nature 410, 1067 (2001)

Consider a simpler case: to purify



Keep 4-fold coincidence at a3, b3, a4, b4

Practical Scheme for Entanglement Purification



4-fold coincidence after PBS

Probability				
Case				

These two cases will not result in 4-fold coincidence

Practical Scheme for Entanglement Purification

➤ For

Four-fold events

No four-fold events

↓ Probability of 50%

- After local measurements in $\{+/-\}$ base at a_4 and b_4 :

↓ Probability of $F^2/2$

Practical Scheme for Entanglement Purification

➤ For

Four-fold events

No four-fold events

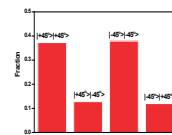
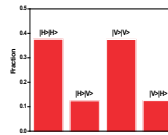
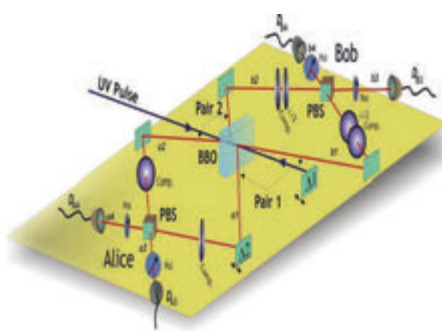
↓ Probability of 50%

- After local measurements in $\{+/-\}$ base at a_4 and b_4 :

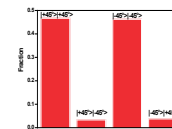
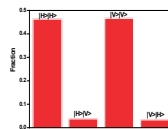
↓ Probability of $(1-F)^2/2$

- Final state:

High Precision Entanglement Purification



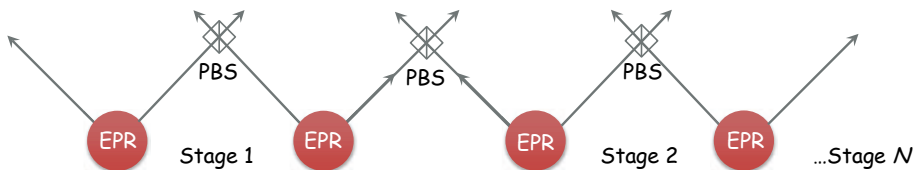
Before purification, $F=3/4$



After purification, $F=13/14$

Pan *et al.*, Nature 423, 417 (2003)

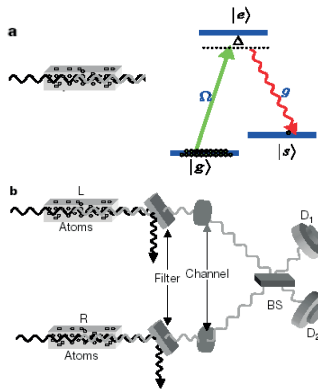
Quantum Memory



Probabilistic EPR source, Channel loss, Probabilistic entanglement purification

- ❌ Without quantum memory, the cost of resource in multi-stage experiments $\sim 1/p^n$, thus not scalable
- ✅ If we know when photon pair is created and can store them on demand, then implement entanglement purification and swapping, the total cost $\sim 1/p^2$

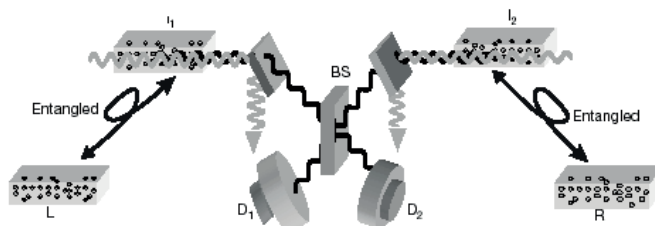
Triggered and Storable Entanglement Generation



DLCZ scheme
Duan *et al.*, Nature 414, 413 (2001)

Maximally entangled in
the number basis

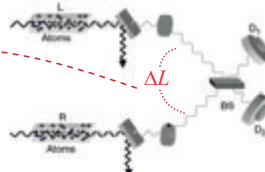
Entanglement Connection



- Apply a reverse laser pulse to transfer atomic excitation back to optical excitation
- Succeeds if D1 or D2 registers a single photon
- Fails otherwise, and repeat every step from entanglement generation

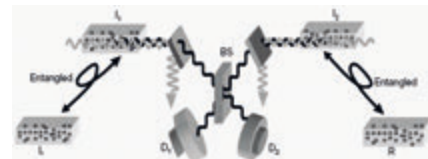
Drawbacks in DLCZ Scheme

➤ Phase stabilization



➤ Error rate grows rapidly with distance

Vacuum term becomes dominant after a few connections



➤ Short Lifetime

Achieved lifetime $\sim 30 \mu\text{s}$

Preparation time $\sim 100 \mu\text{s} \rightarrow$ lifetime needed $\sim 1 \text{ ms!}$

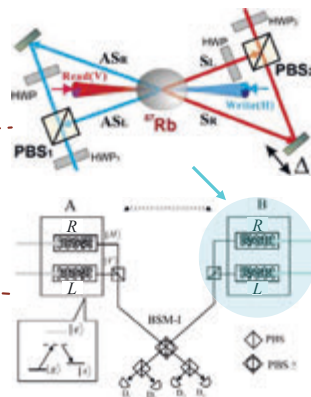
Deterministic Entanglement Generation

Solution:

➤ Phase stability:

Sub-wavelength 100nm

Sub-coherence length $\sim 1\text{m}$



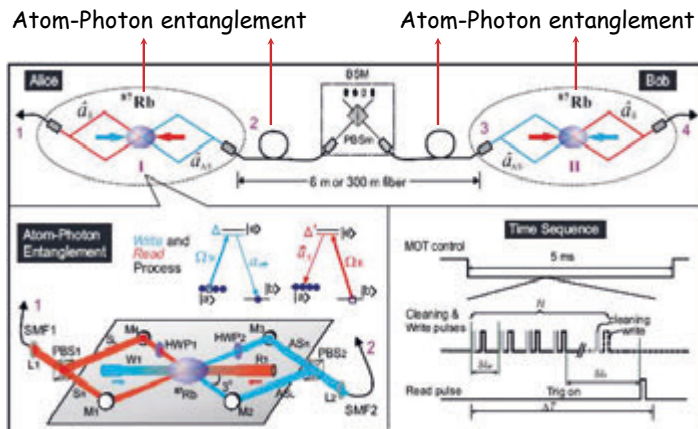
➤ Lower error rate

Vacuum term is NO more dominant

➤ Higher efficiency

Zhao *et al.*, PRL 98, 240502 (2007)

Quantum Repeater Nodes



Experiment: Yuan *et al.*, Nature 454, 1098 (2008)

Efficient and Long-lived Quantum Memory

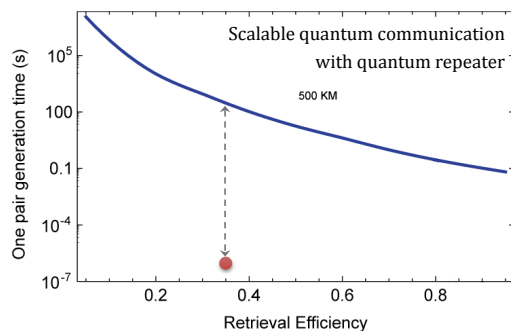
Long lifetime: storage time must be long enough to ensure every node creates an entangled pair

High retrieve efficiency: the stored quantum state must be converted into photon with sufficient high efficiency to establish remote entanglement

In 2008 experiment,

- Life time: $1\mu\text{s}$
- Retrieve efficiency: 35%

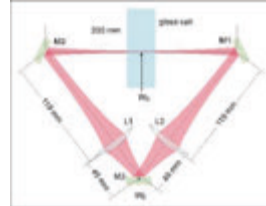
Require lifetime to be extended about 8 orders of magnitude!



Efficient and Long-lived Quantum Memory

➤ Increasing retrieval efficiency:

- ☑ Ring cavity enhancement:
increase interaction strength

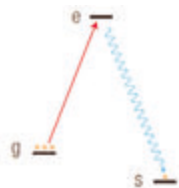


➤ To Increase life time, need to overcome:

- ☒ Inhomogeneity of magnetic field
- ☒ Loss of atoms due to gravity and atomic random motion
- ☒ Spin-wave dephasing

Efficient and Long-lived Quantum Memory

➤ Collective excitation state (spin-wave) of atomic ensemble:



Raman process:

- Absorbing a photon with momentum
- Emitting a photon with momentum

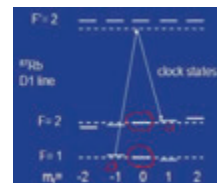


\vec{r}_j is the position of atom j

- ☒ Inhomogeneity of magnetic field:

The evolution phase given by each atom

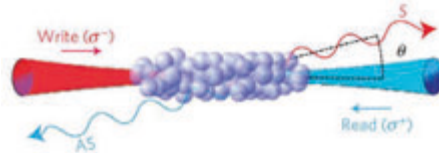
Inhomogeneous magnetic field may cause different \rightarrow uncertain additional phase



- ☑ Solution: "clock states" is not sensitive to magnetic field)

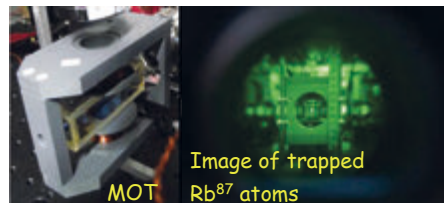
Efficient and Long-lived Quantum Memory

- ❌ Loss of atoms due to gravity and atomic random motion: atoms will diffuse or fall



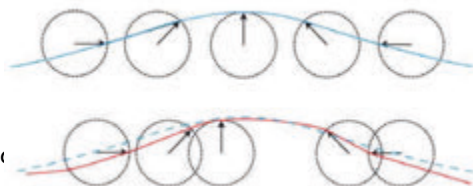
- ❑ Solution:

- Cooling atoms with optical molasses
- Write/Read in the gravitational direction



Efficient and Long-lived Quantum Memory

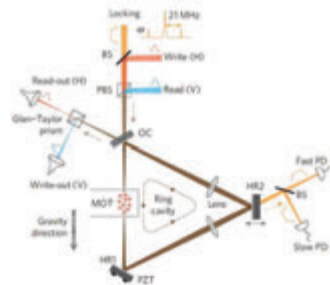
- ❌ Spin-wave dephasing
Different due to atomic random motion



- ❑ Solution: collinear recoil, smallest \Rightarrow evolution phase $\Delta \mathbf{k} \cdot \mathbf{r}$ is almost fixed to 0

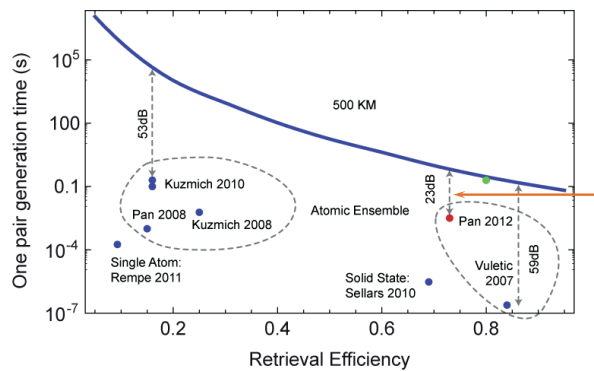
Efficient and Long-lived Quantum Memory

- ☑ Ring cavity (finesse=48)
- ☑ Clock state
- ☑ Optical molasses
- ☑ Write/Read in the gravitational direction
- ☑ Collinear configuration



Life time 3ms, retrieve efficiency 73%
 Bao *et al.*, Nature Physics 8, 517 (2012)

Efficient and Long-lived Quantum Memory

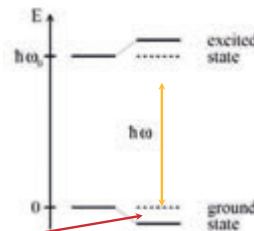


Require lifetime to be extended about 2 orders of magnitude

Efficient and Long-lived Quantum Memory

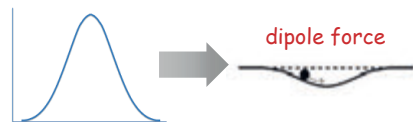
Optical trap

- Light field:
- Induced atomic dipole moment:
polarizability of atom,
(: line width, : detune)
- Interaction between and results in
energy shift of atomic levels:



(: light intensity)

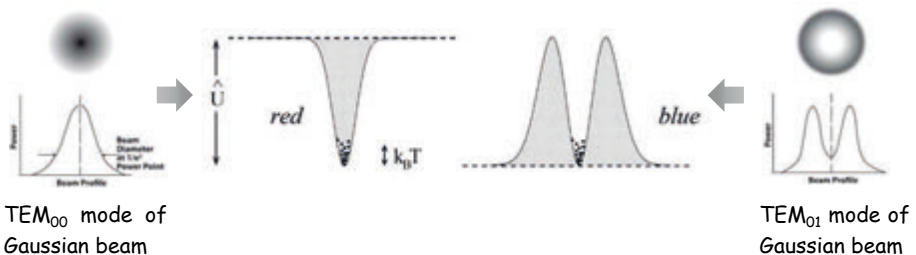
- Gradient of light intensity will provide a
dipole force



Efficient and Long-lived Quantum Memory

Red and blue detuned trap

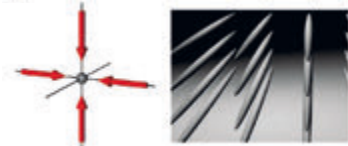
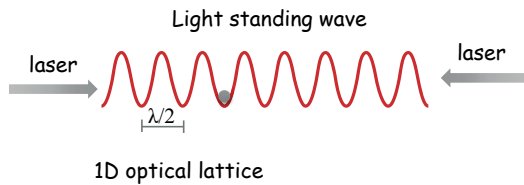
- Red detuning : a trap is formed at the position of maximal intensity
- Blue detuning : a trap is formed at the position of minimal intensity



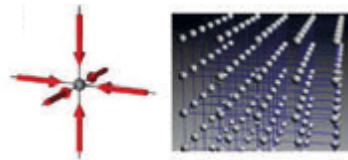
Grimm *et al.*, Adv. At. Mol. Opt. Phys. 42, 95 (2000)

Efficient and Long-lived Quantum Memory

Optical lattices



2D optical lattice



3D optical lattice

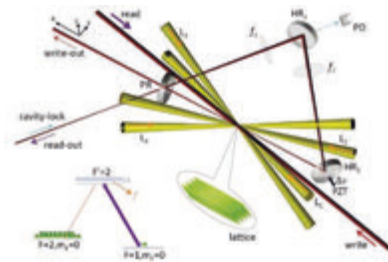
- Interference of counter-propagating laser beams
➔ a spatially periodic pattern
- "Lattice": periodic optical dipole potential
➔ atoms are cooled and collected in the locations of potential minima

Efficient and Long-lived Quantum Memory

We use:

- 3D Lattice (0~180μk, distance between adjacent wells: $dx \sim 2.8\mu\text{m}$, $dy \sim 5.9\mu\text{m}$, $dz \sim 0.54\mu\text{m}$)
- Spin-wave excitation ($\Lambda \sim 15\mu\text{m}$)

Limits atomic motion in all direction to suppresses atomic collision-induced decoherence

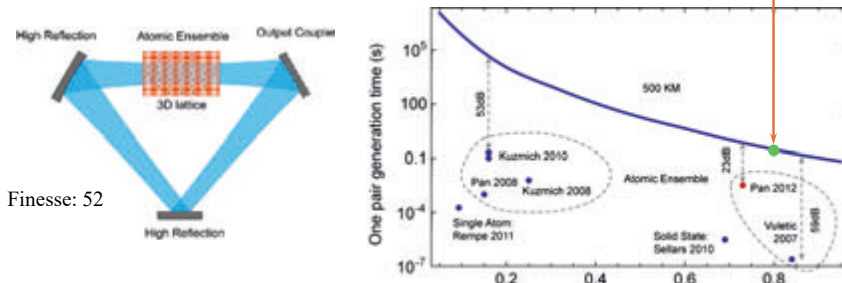


Efficient and Long-lived Quantum Memory

With ring cavity + optical lattice confinement:

Life time 220ms, retrieve efficiency 76%

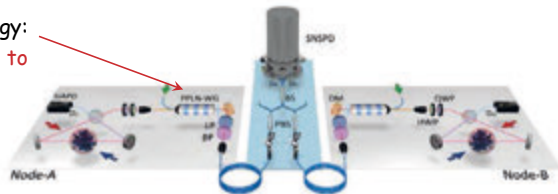
Yang *et al.*, Nature Photonics 10, 381 (2016)



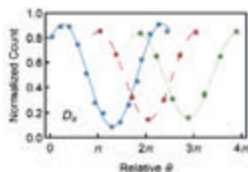
☑ Support quantum repeaters enabling quantum communication at a range of ~500km

Long-distance Entanglement of Quantum Memories

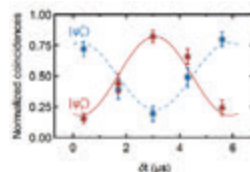
Frequency conversion technology:
shift the atomic wavelength to
telecom wavelength



50km coiled fiber with single-photon scheme



22km field-deployed fiber with two-photon scheme

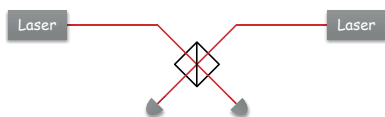


Yu *et al.*, to appear in Nature (2020)

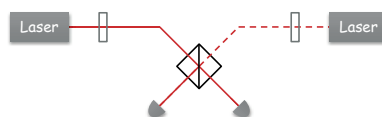
A practical quantum repeater might still need 10 more years

Part 4: Towards Long-distance Quantum Communication

Immediate Way: Twin-field MDI-QKD



V. S.



- Original MDI-QKD: efficiency of **two-photon** interference:
(channel loss of each path:)

- TF-MDI-QKD: efficiency of **single-photon** interference:
(two remote independent lasers with locked phase)

Theory of TF QKD :

- Lucamarini *et al.*, Nature 557, 400 (2018)
- Wang *et al.*, PRA 98, 062323 (2018)
- Ma *et al.*, PRX 8, 031043 (2018)

	0	0	π	π
D_0 click	0	0	π	π
D_1 click	0	π	π	0

QKD in commercial fiber can even reach 700-800km!

Immediate Way: Twin-field MDI-QKD

Major technology challenging:

Stable phase interference between two independent lasers, separated a few hundreds km

$$\delta_{ba} = \frac{2\pi}{s} (\Delta\nu L + \nu \Delta L)$$

Wavelength difference

Fiber length fluctuation

➤ Locking the frequency of two remote lasers: instability $10^{-14}/s$

Frequency and time dissemination technology [e. g., Predehl *et al.*, Science 336, 441 (2012)]

➤ Overcoming fiber length fluctuation:

Post-selecting small phase difference via phase measurement on time-multiplexing strong reference pulses

Immediate Way: Twin-field MDI-QKD

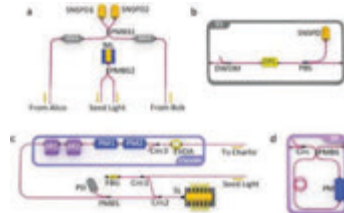
- ✓ Frequency difference of two remote lasers is less than $10^{-15}/s$
- ✓ Length perturbation of fiber is less than 100nm

Experiments (in coiled fiber):



First realization: 300km

Liu *et al.*, PRL 123, 100505 (2019)

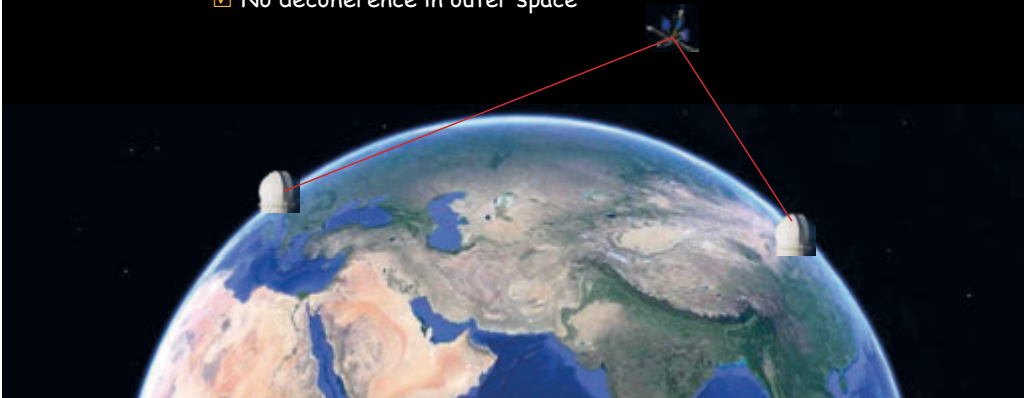


Extended distance: 502km

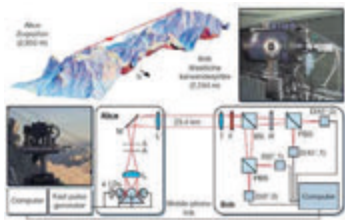
Fang *et al.*, to appear in Nature Photonics (2020)

More Efficient Way: Free-Space Quantum Communication

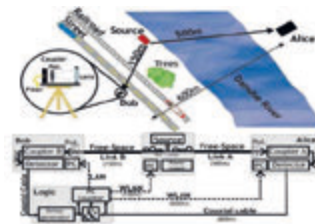
- ✓ Non-obstruction from terrestrial curve and barrier
- ✓ Effective thickness of atmosphere is only ~10km
- ✓ No decoherence in outer space



Attempt to Free-space Quantum Communication



- QKD with weak coherent pulse, 23.4km: Kurtsiefer *et al.*, Nature 419, 450 (2002)
- Security distance ~5km



- Distribution of entanglement ~600m: Aspelmeyer *et al.*, Science 301, 621 (2003)

Major question: could the quantum states of single and entangled photons still survive after passing through atmosphere?

Ground Tests for Satellite Quantum Communication

Phase 1:

Test the possibility of single photon and entangled photons passing through atmosphere



Free-space quantum entanglement distribution ~13km

Peng *et al.*, PRL 94, 150501 (2005)



Free-space quantum teleportation (16km)

Scheme: Boschi *et al.*, PRL 80, 1121(1998)

Experiment: Jin *et al.*, Nature Photonics 4, 376 (2010)

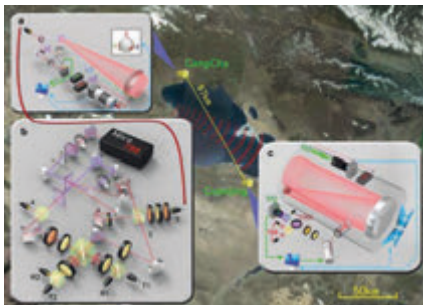
Well beyond the effective thickness of the aerosphere!

Ground Tests for Satellite Quantum Communication

Phase 2:

Test the feasibility of quantum communication via high-loss ground-to-satellite channel

Free-Space Quantum Teleportation (97km)



Channel loss:

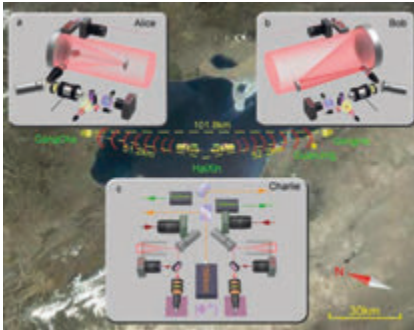
35-53dB

V. S.

Loss for an uplink of ground to satellite: 45dB

Ground Tests for Satellite Quantum Communication

Free-space quantum entanglement distribution (over 100km)



Channel loss:

66-85dB

V. S.

Loss for two-downlink
between satellite and two
ground stations: 75dB

Yin *et al.*, Nature 488, 185 (2012)

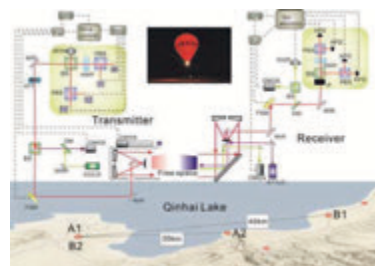
Ground Tests for Satellite Quantum Communication

Phase 3:

Direct and full-scale verifications towards ground-satellite quantum communication

- ☑ Mimicking the satellite's angular motion
- ☑ Mimicking the satellite's attitude change
- ☑ A huge loss channel (about 50 dB loss, 97 km)

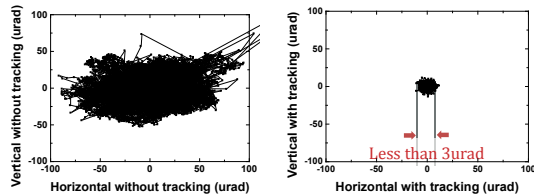
Wang *et al.*, Nature Photonics 7, 387 (2013)



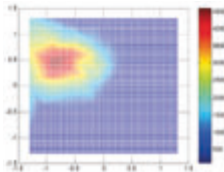
Overcoming all the demanding conditions for ground-satellite QKD

Ground Tests for Satellite Quantum Communication

☑ High precision Acquiring, Pointing and Tracking (APT)

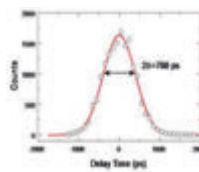


☑ Near-diffraction-limited far-field divergence angle



- Diffraction-limited divergence angles: $8\mu\text{rad}$
- Divergence angle $\sim 10\mu\text{rad}@140\text{mm}$

☑ High precision synchronization

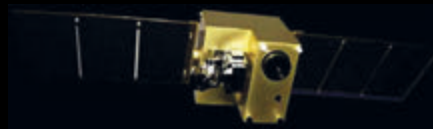


- Hundreds of kilometers
- Rapid motion
- Random movement
- Vibration

➡ Ultra-high energy resolution : detecting from the earth a single match fire lighted on the Moon

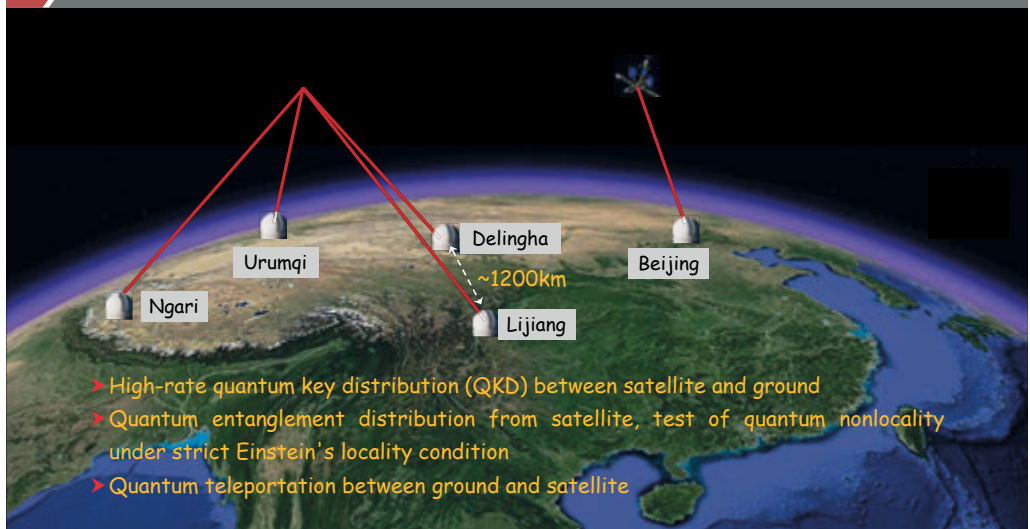
Quantum Science Satellite "Micius"

Launched on 16th Aug, 2016 in Jiuquan Satellite Launch Center

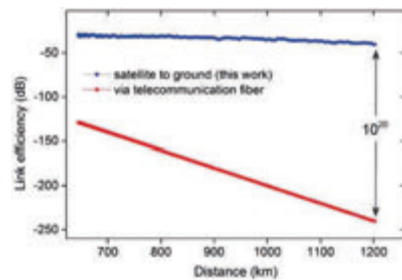


- Weight: $\sim 640\text{kg}$
- Power: 560W
- Sun-synchronous orbit, altitude 500km

Micius' Three Missions



QKD between Satellite and Ground

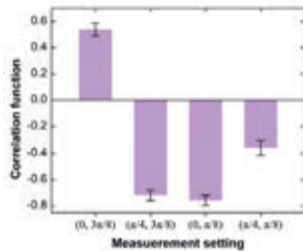


- Channel loss: <40dB
- Average final key rate:
>1kbps@650-1200km
- Recent result: ~100kbps

For telecommunication fiber channel (loss ~0.2dB/km) with same distance, the attenuation will be 20 orders higher

Liao *et al.*, Nature 549, 43 (2017)

Satellite-based Entanglement Distribution



Satellite Channel

- Loss < 70dB
- 1 pair /s

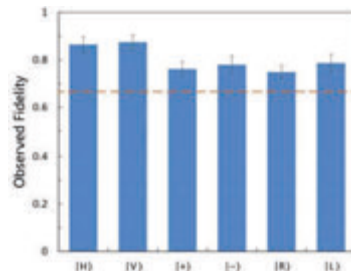
V. S.

Fiber Channel

- Loss > 190dB
- 1 pair /30000 years

- ☑ Entanglement distribution over 1200 km
 - ☑ Violate Bell's inequality at a confidence level 99.9%
- Yin *et al.*, Science 356, 1140 (2017)

Ground-to-satellite Quantum Teleportation



Satellite Channel

- Loss < 50dB
- 0.08 event/s

V. S.

Fiber Channel (1200km)

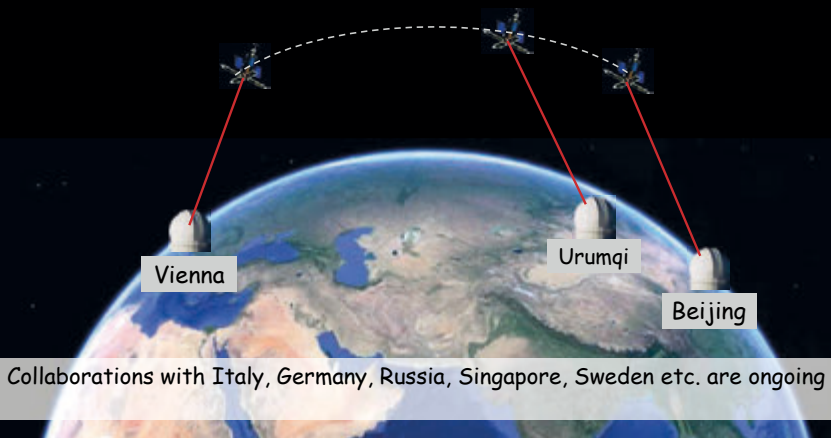
- Loss > 240dB
- 1 event /380 billion years
(20 times the Universe's lifetime)

Ground-to-satellite teleportation over 500km-1400km

Ren *et al.*, Nature 549, 70 (2017)

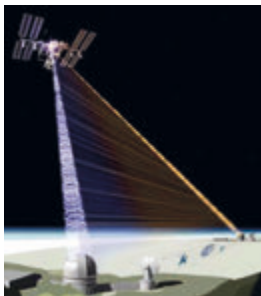
Intercontinental Quantum Key Distribution

Satellite as a trusted relay [Liao *et al.*, PRL 120, 030501 (2018)]



Entanglement-based QKD

- Longest distance of point-to-point QKD on physically separated targets on the Earth: ~100 km [PRL 98, 010504 (2007), Nature Physics 3, 481 (2007)]
- **Entanglement-based QKD**: extending the distance without compromising security



Even the satellite is controlled by your enemy, the security of QKD can still be ensured by violation of Bell's inequality!

Scheme: Ekert, PRL 67, 661 (1991) ;
Bennett *et al.*, PRL 68, 557 (1992)

Entanglement-based QKD

Experiment: over 1120 km [submitted (2019)]

- Channel Loss: 56~71dB
- Entanglement pairs received at 2/s
- Final key: 0.43 bps
- QBER: $4.51\% \pm 0.37\%$
- ➔ If load GHz entanglement source, up to 10kbits per orbit

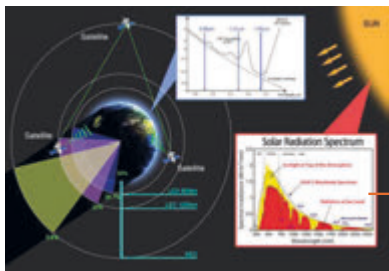


This would thus achieve the Holy Grail that all cryptographers have been dreaming of for thousands of years

--Gilles Brassard

Towards Global Quantum Communication Network

- ☒ LEO orbit, can not cover the whole earth directly
- ☒ Only working in earth's shadow
- ☑ Solution: **Quantum Constellation!** ➔ A prerequisite: working in solar radiation background



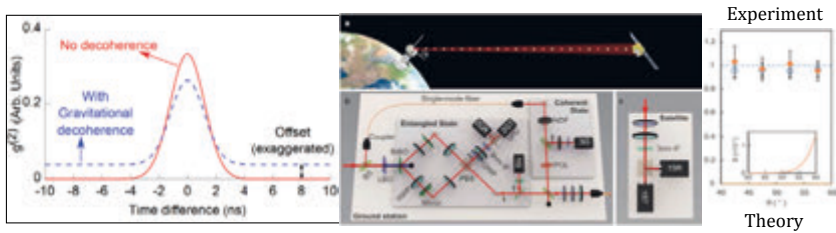
Orbit Type	Shadow Probability
Ground	50%
LEO (600km)	29.7%
GEO	0.6%

- Using C-band (1550nm) to overcome the noise from the sunlight
- 1550nm can be integrated with ground fiber networks naturally

Long-distance free-space QKD in daylight
Liao *et al.*, Nature Photonics 11, 509 (2017)

Investigation at the interface of quantum physics and gravity

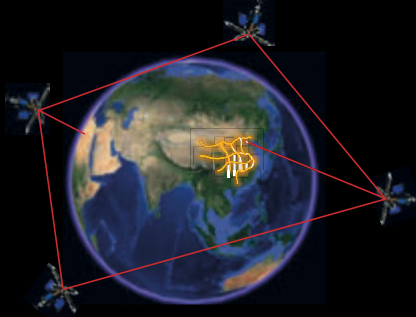
Event Formalism model: gravitationally induced decorrelation of time-energy entanglement in exotic spacetime [Scheme: Ralph *et al.*, PRA 79, 022121 (2009)]



The experiment excludes the prediction of a strong event formalism model
Xu *et al.*, Science 366, 132 (2019)

Part 5: Future Prospects

Future Prospect



Quantum Constellation + Fiber quantum communication infrastructure



"Quantum Internet"

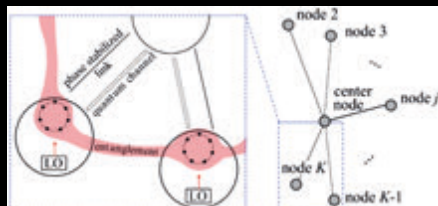
Future Prospects

Applications for the global quantum communication network

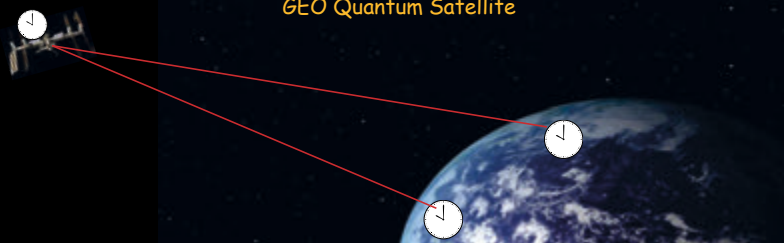
- **Enormous spatial resolution:** Combining photons from distributed telescopes worldwide by quantum teleportation in space [Gibney, Nature 535, 478 (2016)]
- **Precise atomic clocks:** Entangling atoms in remote atom clocks, short term instability is times better than classical method [Kómár *et al.*, Nature Physics 10, 582 (2014)]



Reading license plates floating in Jupiter's orbit on the Earth



Future Prospects

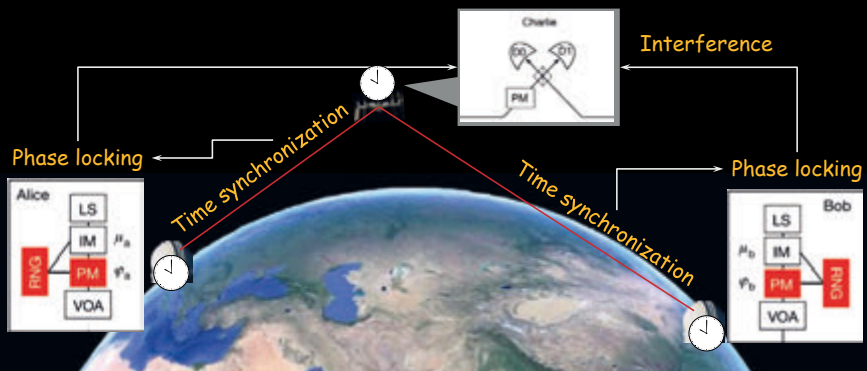


- > In GEO orbit, long mooring time, distributing much more keys per day (> 1 Gbits/year)
- > Global precise timing information sharing networks: Optical atomic clocks + Optical frequency transfer, long time instability to 10^{-19}

V. Microwave timing sharing network: long time instability below 10^{-15}
S.

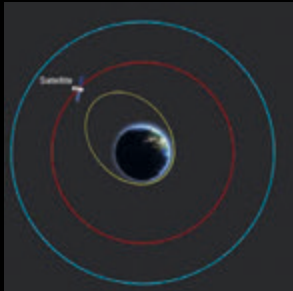
Future Prospects

- > Long distance ($\sim 1000\text{km}$ - 10000km) and precise time synchronization with global precise timing information sharing networks
- > Stabilization of phase in long distance \rightarrow Twin field MDI-QKD between ground and satellite, increasing the key rate by 4 orders of magnitude

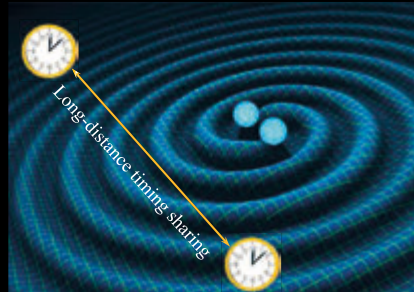


Future Prospects

Negligible magnetic and gravitational noise in GEO orbit → Fractional instability $\sim 10^{-21}$

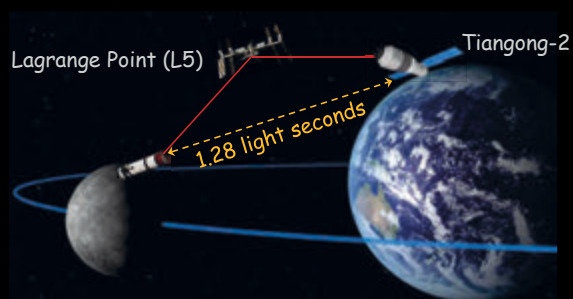


- Precisely detecting gravitational red shift at different altitude of orbits



- Detecting gravitational wave signal with lower frequency to 0.1Hz (LIGO: ~ 100 Hz)

Large-scale Bell Test with Human-observer



Entanglement distribution between Moon and Earth with China's future Moon landing project!

Pretests for Large-scale Bell Test with Human-observer

Channel loss of entanglement distribution between Earth and Moon: 100 dB

- Bell test with human supplying random measurement over simulated extremely high loss channel (103dB)
Cao *et al.*, PRL 120, 140405 (2018)

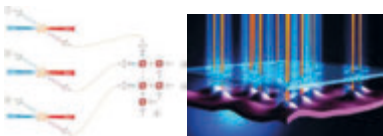
Challenging local realism with human choices

- Generating random numbers with the help of worldwide 100,000 volunteers' free will
- 12 labs run Bell tests with the random numbers

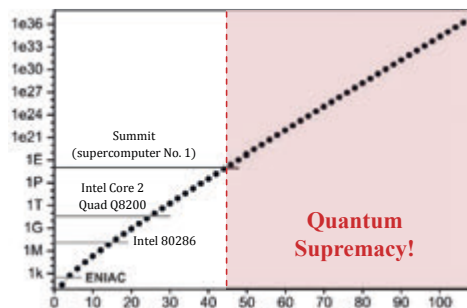
The BIG Bell Test Collaboration,
Nature 557, 212 (2018)



Scalable Optical Quantum Computation



- More entangled particles with the help of quantum memory
- Efficient quantum dot single photon emitters



- In next 1-2 years: optical quantum computer with 40-50 qubits ➡ beating classical super computer in specific tasks (e. g. Boson sampling and portfolio optimization, etc.)

A 5 Statistics for quantum estimation

D. Gross
Institute for Theoretical Physics
University of Cologne

Contents

1	Introduction	3
1.1	Statistics and physics education	3
2	Orthodox vs Bayesian approach	4
3	Statistical Models	6
3.1	Observables and the quantum design problem	7
4	Point estimation	8
4.1	Sample means	9
4.2	Maximum Likelihood Estimation	10
5	Uncertainty Quantification	12
5.1	Variance, bias, risk	12
5.2	Region estimation	16
5.3	Interpretation	17
6	Optimality criteria & “principles”	17
6.1	Unbiased estimators	18
6.2	Optimal estimators	18
7	The role of positivity	19

8	Acknowledgments	21
	References	21

1 Introduction

Characterizing and verifying components for quantum computing devices is a non-trivial task, for a number of reasons: First, by definition, quantum computers operate in a regime that is too high-dimensional to track classically. Second, quantum physics is an inherently probabilistic theory, where the underlying parameters are not directly observable, but need to be statistically estimated from the results of many complementary measurements. Third, in order to sustain a calculation of indefinite length with practical overhead, the components of a quantum computer have to be implemented in a highly accurate way. Obtaining precise statistical estimates in high-dimensional problems is known to be a very challenging task.

As a result, the field of *quantum characterization and certification* has grown over the past years from a niche to a subject area supporting a substantial community. Today, there are even start-up companies offering their *QCC* expertise on the market.

A confusingly large collection of methods have been drawn up, including *randomized benchmarking*, *gate set tomography*, *direct fidelity estimation*, *cross validation*, *compressed sensing*, and many more. That's bad news to newcomers. The *good news* is that the multitude of methods means that in many cases, the literature is not "stacked too deeply", i.e. it is not the case that one has to work through generations of papers and ideas to understand the state of the art. Maybe with the exception of randomized benchmarking, it suffices to read the defining paper and a recent follow-up to understand where we are. A very high-level overview over these methods, with an extensive list of references, can be found in the very recent review paper [1]. We will cover the scope and theory behind a selection of these approaches in the oral lecture.

In these written notes, in contrast, I will neither repeat the bird's-eyes view offered in [1], nor give an encyclopedic account of the many *QCC* schemes in existence. Instead, I want to use the opportunity to close what I feel is a gap in the available literature (and also, I am afraid, a somewhat common gap among researchers working in the field): An introduction to the *statistics* of quantum characterization.

1.1 Statistics and physics education

Physics is an empirical science. The mathematical theory of empirical inference is *statistics*. Still, statistics is not part of the standard curriculum of most physics programs. Many fields that are generally seen as less quantitative in scope – e.g. the social sciences and psychology – produce graduates that are much better versed in statistical methods than the average physicist!

Partly, physics gets away with this state of affairs, because we rely less on statistics for learning about the world than other empirical fields. There are two main reasons: *more data* and *better models*.

More data: Fully automated quantum experiments can often be repeated at rates of thousands to millions of times *per second*. In this regime, statistical fluctuations are less of a concern; the *central limit theorem* ensures that the commonly employed, simplifying assumption that relevant distributions are Gaussian is actually well-justified; and, different approaches to statistical estimation tend to agree with each other in the asymptotic setting. This justifies the quote "If your experiment needs statistics, you ought to have done a better experiment", often attributed

to Ernest Rutherford.

Better models: A less-appreciated aspect is that the role of statistical hypothesis testing for validating a theory is inverted in physics as compared to other empirical fields. More precisely, consider the standard method of *null-hypothesis significance testing* (NHST). In the social sciences or in clinical studies, the *null hypothesis* posits the *absence* of the effect – e.g. a treatment effect of a drug, or discrimination against an underrepresented group – that one actually expects to find. An experiment is taken to be conclusive if the data presents significant evidence *against* this null hypothesis of no effect. This indirect reasoning is necessary, because while one might be able to predict from general theories that a given compound should treat a medical condition, or that discrimination exists, these theories are not typically strong enough to predict the strength of the effect quantitatively. Physics, in contrast, has developed strong models that allow one to make quantitative predictions about the outcomes of experiments. (Phrased more modestly, physics restricts attention to phenomena that are sufficiently boring that such models are available.) Data analysis for the purpose of verifying a physical theory therefore typically follows a more direct logic: A test of general relativity via the bending of star light during an eclipse would be felt to be conclusive if the gathered data not only give strong evidence against a Newtonian explanation, but would also be compatible with the quantitative prediction of GR.

As we will see, this distinction has profound implications on the interpretation of the weight of statistical evidence. For example, the oft-used *p-value* statistic quantifies how unlikely the observed data would be *if the null hypothesis were true*. But of course, clinical tests are only conducted when one fully expects the null hypothesis (of no treatment effect) to be *false*. Thus the assumption under which the *p-value* has a direct interpretation is generally *not* met, making it a conceptually difficult quantity to work with.

The above remarks explain why a physicist can often be successful without speaking the language of mathematical statistics. But “often” does not mean “always”. In astrophysics, where observations cannot be repeated; in high-energy physics where all the low-hanging fruit have been picked and the remaining effects are extracted via indirect measurements in the presence of a large amount of background noise; and indeed in quantum characterization, a solid understanding of the foundations of statistical reasoning is necessary. Hence these notes. My goal is to provide a quick introduction, with an emphasis of problems that are relevant to quantum applications. This is no substitute for a textbook on mathematical statistics. These exist in abundance – as a student, I enjoyed [2, 3] (orthodox) and [4] (Bayesian).

2 Orthodox vs Bayesian approach

In math and science, we are accustomed to having “canonical”, universally accepted answers to important problems. It is therefore unexpected to learn that statistics is split into two big *schools of thought* that remain at odds with each other to a surprising degree. Both are, of course, correct as mathematical theories, but they address subtly different questions. The unsettled dispute is thus about which of the two approaches to the task of inference from noisy data gives more satisfactory results in practice.

To get an idea about the tone of the debate, check out the (otherwise excellent) Bayesian textbook [4] by the late David MacKay (freely available from www.inference.org.uk).

Search the book for the term “happily squeak” for the scathing parody the Cambridge professor uses to describe the thinking of his orthodox colleagues. I can’t point to a printed version of equally open mockery in the opposite direction, but I have certainly heard respected researchers engage in it.

Simply speaking, Bayesian statistics provides a formalism for reasoning about beliefs. Beliefs about reality combine *prior knowledge* with experimental observations. In the Bayesian ansatz, beliefs are expressed in terms of probabilities that encode how plausible any given hypothesis is, given the information available. *Bayes’ rule* (also *Bayes’ Theorem* – although mathematicians would take offense with that nomenclature) tells one how to *update* these distributions as the result of observations. The advantage of this approach is that it answers the question practitioners care about most: How likely is it that any given hypothesis about the system under study is correct? The disadvantage is that there is no canonical way of how to choose the *prior distribution* that encodes the totality of knowledge that goes into the analysis before any observations are made. Therefore, Bayesian reasoning is often criticized as *subjective*, an unattractive feature for a framework used to reason about physical reality.

In contrast, the goal of orthodox statistics is to develop inferential *methods* that are guaranteed to give sound results with high probability. For example, the typical “ $p \leq 0.05$ -rule” used in empirical studies says that the method will give a *false positive* or *type-I error* – i.e. reject the null hypothesis even though it is true – with probability at most 5%. The advantage is that no subjective prior has to be chosen. A major disadvantage is that orthodox statistics makes statements only about the *method* in general, but not about any given inference. For example, if one conducts an AIDS test with a false positive rate of 5% on a collection of pre-industrial virgins, then given a positive result, we can still be *certain* that the test subject is actually healthy. However, orthodox statistics, de-emphasizing the need to combine different sources of information, offers no systematic guidance for how to reach this conclusion. What is worse, users of statistical methods *expect* those to assign a quantitative value to the degree of certainty they should have about the tested property. Therefore, they often turn to the only quantity available, and erroneously interpret $(1 - p)$ as the probability that the result found is correct – a reading that is not justified by the theory at all!

Homework: Read newspaper articles / social science publications / physics papers and search for the above mistake. You will not have to search long.

A second problem of orthodox statistics is that the guarantees it offers are valid only if the estimator has been chosen *independently of the data*. However, a very large number of reasonable estimators has been developed, and the orthodox framework offers few generally accepted criteria for choosing among them (as we will see). Thus, in practice, the choice of which set of frequentist methods to apply to a given problem can implicitly introduce bias. A cautionary example is offered by the current controversial discussion surrounding the so-called *replication crisis* in the social sciences. There, the alleged practice of trying out various orthodox estimators until one confirms the intended result at a “statistical significance level” has been deemed sufficiently prevalent to have received a name: “*p*-hacking”.

In a nutshell, Bayesian methods force you to choose a (subjective) prior, but orthodox methods don’t answer the question practitioners want to have answered, and can easily be misused. No wonder then that the debate hasn’t been settled. The orthodox method is called *orthodox*, as it is the one that has traditionally been used in science. In recent years, as a result of an increased

awareness of common misuse of orthodox procedures and with increased availability of computational power often required for Bayesian inference, the latter has gained more prominence.

For completeness: I have no horse in this race. Either approach makes sense, *if used carefully, with a sufficient understanding of the common pitfalls*. We will primarily look at the orthodox theory here, but will occasionally mention the Bayesian point of view.

3 Statistical Models

The first step of any statistical analysis is to fix a *statistical model*. We assume that the unknown parameters of the system are given in terms an element θ of the set Θ of all possibilities. To facilitate the mathematical analysis, one typically assumes that $\Theta \subset \mathbb{R}^k$ (and, as we will see, in practice one also encounters estimators that map into a superset of Θ within \mathbb{R}^k). Inferences about θ are made via observations of random variables X_1, \dots, X_n . In the simplest case, the X_i are independently distributed – this models an independent repetition of a physical experiment, without any memory effects or drifting parameters. The distributions will, however, depend on the unknown parameter:

$$\text{Prob}_\theta[X = x] =: p(x|\theta). \quad (1)$$

The function $p(x|\theta)$ is called the *statistical model* and will be used to connect observations to the unknown parameters. By definition, for any fixed θ , the function $x \mapsto p(x|\theta)$ is a probability distribution. If, in contrast, one fixes the data parameter, one obtains the *likelihood function*

$$\mathcal{L}(\theta) : \theta \mapsto p(x|\theta)$$

which is defined on the parameter space Θ .

Example—A loaded coin: Consider a coin whose (unknown) probability to come up heads is given by $\theta \in \Theta = [0, 1]$:

$$p(x, \theta) = \begin{cases} \theta & x = H, \\ 1 - \theta & x = T. \end{cases}$$

Typical question: given an observation $(X_1, \dots, X_n) = (H, H, T, \dots, H)$, what can we say about θ ?

Example—The German tank problem: (It's a thing! Look it up!) An allied elite force gets parachuted into pre-D Day Nazi Germany. Suffering heavy losses, they manage to capture one tank, pry off the manufacturer's nameplate, and escape to friendly territory. The plate is flown to you, the head of the Royal Statistical Society. It says the captured tank had serial number 123. How many tanks $\theta = N \in \Theta = \mathbb{N}$ do the Germans have? The statistical model is

$$p(x, N) = \begin{cases} 1/N & 1 \leq x \leq N \\ 0 & x > N. \end{cases}$$

(This makes the simplifying assumption that Germans, strangely literal, label their tanks in order 1, 2, 3, ... during wartime. Well, never mind.) The example emphasizes these common characteristics of estimation problems:

1. The problem is too pressing to put off a decision;
2. Non-trivial information is available;
3. It is too costly to obtain further samples;
4. It is plain to see that we don't really know how many tanks there are, and no sophisticated mathematical method will change that fundamental limitation. In other words: Our inferences will often be subject to *a very high degree of uncertainty*!

Example—Qubit state: Consider a quantum experiment where measurements are performed on a qubit in an unknown state $\theta = \rho$. The state space Θ is thus the set of qubit density matrices, conveniently represented by the Bloch ball, i.e. the unit ball in \mathbb{R}^3 . In every run, a spin axis is chosen uniformly at random from $s \in \{1, 2, 3\}$, and the eigenprojections

$$P_{\pm}^{(s)} := \frac{1}{2}(\mathbb{I} \pm \sigma_s)$$

are measured. Thus

$$p(x = (s, \pm) | \rho) = \frac{1}{3} \text{tr } \rho P_s^{\pm}. \quad (2)$$

Given the observation $(1, +), (1, +), (2, +), (2, +), (3, +), (3, +)$, what is inferable about ρ ?

3.1 Observables and the quantum design problem

The third example underscores that, in order to treat data from quantum experiments in the framework of classical statistics, we need to choose *which observables to measure*. In classical statistics, the study of how to set up an experiment in such a way as to maximize the expected amount of information gained is called *design theory*¹. In the classical setting, design theory less prominent, as often, the choice of which data to observe is dictated by practical reasons. For quantum estimation problems, it is an unavoidable task – and many seemingly contradictory results in quantum estimation theory differ in implicit limitations imposed on the class of observables that are admitted.

When we wrote down the statistical model, Eq. (2), we assumed that the physical measurement apparatus was described by Pauli matrices. In reality, there is typically a significant degree of uncertainty about the proper quantum mechanical representation of the observables. Such *systematic* (as opposed to *statistic*) uncertainties are not directly modeled in the theory. There are a number of ways to treat this problem:

¹ A classical example of design theory are *weighing problems*, where the weight of a collection of small objects is to be determined. It turns out that the individual weights can be estimated to better precision if the objects are weighted in different groups, rather than individually. The problem of how to optimally divide up the weights led to the study of *Hadamard matrices*, i.e. matrices with orthogonal columns that have entries from ± 1 . That's where quantum computing's *Hadamard gate* got its name from. Also, *spherical* and *unitary designs*, as used e.g. in randomized benchmarking are named so because they generalize combinatorial constructions that originated in design theory of classical statistics.

To *calibrate* the detectors, one can alternate between estimating reference *states* and *observables*, until one reaches a description that is consistent with the experimental data. Indeed, note that the *Born rule* $\text{tr} \rho P$ which assigns a probability to a quantum state and a *positive operator-valued measure* (POVM) [5] element P treats ρ and P symmetrically. Thus, if one knows how to prepare a number of states ρ_i , one can set Θ to be the space of POVM elements and reach an estimate for P . This strategy is implemented e.g. in Ref. [6].

Some authors have proposed to include *both* the states and the observables into the parameter set Θ and try to fit them simultaneously from observations [7,8]. Because the Born rule is then a *quadratic* function (rather than linear, if either the state or the observable are fixed), this makes the problem considerably harder from a computational point of view. It also underscores the fact that there are *gauge degrees of freedom* in the joint description. Indeed, any change of basis

$$\rho \mapsto U \rho U^\dagger, \quad P \mapsto U P U^\dagger$$

will leave the physical predictions invariant, which implies that there will always be a large family of pairs that fit the data equally.

In experimental setups, there is often one distinguished basis in which measurements are comparatively easy to perform and to which one can, with high confidence, assign a mathematical model. For example, in quantum optics, photon counters measure in the Fock basis. In ion traps, electronic eigenstates can be identified if they scatter laser light of a distinguished frequency. Measurements in other bases are then implemented by first performing a unitary rotation, and then measuring in the distinguished basis. The uncertainty about the description of general observables thus stems from uncertainty about which time evolution is actually being implemented in the system, which will typically be noisy and whose interactions are not perfectly characterized.

The discussion emphasizes an unfortunate, but unavoidable consequence of the Born rule: States and observables are defined only in relation to each other, so the idea of estimating just one of the two is potentially ill-defined.

4 Point estimation

A function that takes observational data x_1, \dots, x_n and maps it to an

$$\hat{\theta}(x_1, \dots, x_n) \in \Theta$$

is called a *point estimator* or just *estimator*. Finding and characterizing good estimators is a major topic in orthodox statistics. Estimators are usually denoted by a “hat”, as in $\hat{\theta}$. This clashes with the convention in the quantum literature to use the hat symbol to signify an operator corresponding to classical (un-hatted) quantity. For the purpose of these notes, we adopt the statistics convention.

There are a number of optimality criteria and heuristic *principles* that can be used in order to choose estimators. However, as alluded to in the introduction, one challenge of the orthodox framework is that in general, no distinguished estimator exists. This ambiguity is particularly pronounced for high-dimensional problems (like quantum state estimation). As the issue is

less severe for the estimation of scalars, beginners, confronted with simple, one-dimensional textbook examples, might easily underestimate the conceptual difficulty!

Before presenting a more general theory, let's discuss some natural examples.

4.1 Sample means

The most natural estimator for the probability θ of the coin coming up heads is the *sample mean* or *empirical mean* – i.e. the proportion of “head events” seen in the data:

$$\hat{\theta} = \frac{1}{n} |\{i \mid x_i = H\}| = \frac{\# \text{ of heads seen}}{\text{trials}}.$$

This is a good estimator.

For the tank example, taking a “mean” over one data point seems unnatural. But we can still define an estimator in analogy to the previous example. Note that the *expected value* of the serial number given a total of N vehicles is

$$\mathbb{E}_N[X] = \frac{1}{N} \sum_{x=1}^N x = \frac{N+1}{2}. \quad (3)$$

Thus it is natural to choose \hat{N} such that the expected value matches the empirical mean (of a single observation):

$$\hat{N} = 2X - 1. \quad (4)$$

This seems reasonable.

Applying these methods to the qubit state example gives an estimator that is both very important and somewhat problematic. First, recall that the Pauli matrices for an orthogonal basis for the space of 2×2 -matrices w.r.t. the trace inner product:

$$\frac{1}{2} \text{tr} \sigma_i \sigma_j = \delta_{ij} \quad i, j = 0, \dots, 3,$$

where we have set $\sigma_0 = \mathbb{I}$, the identity matrix. It follows that we can expand

$$\rho = \frac{1}{2} \sum_{i=0}^3 c_i \sigma_i, \quad c_i = \text{tr} \rho \sigma_i.$$

In other words, the expansion coefficient c_i is just the expected value of spin measurements along the i -th axis. We can use an empirical means estimator

$$\hat{c}_i := \frac{\# (+1 \text{ outcomes for } \sigma_i) - \# (-1 \text{ outcomes for } \sigma_i)}{\# \text{ times } \sigma_i \text{ has been measured}}$$

for the c_i , and set

$$\hat{\rho} = \frac{1}{2} + \frac{1}{2} \sum_{i=1}^3 \hat{c}_i \sigma_i.$$

In quantum estimation, $\hat{\rho}$ is often called the *direct inversion estimator*, as it consists of inverting the linear map from expansion coefficients of ρ with respect to a matrix basis to observable quantities. While the empirical mean estimator works well for the simple coin example, the quantum version has a major drawback: Due to statistical fluctuations, the estimates can have negative eigenvalues (Fig. 1). In this case, they do not represent physical states.

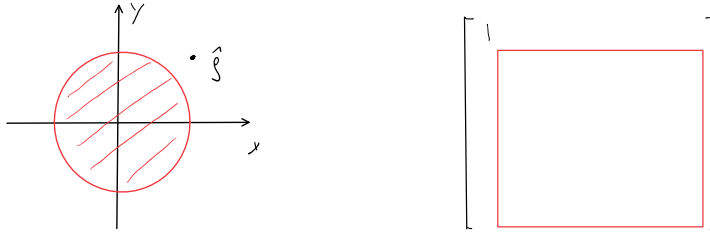


Fig. 1: Left: The *direct inversion estimator* can give non-physical estimates. Shown is a cut through the equatorial plane of the Bloch sphere. The red region is a disk of radius one, corresponding to density matrices with non-negative eigenvalues. If one estimates each coordinate individually, statistical fluctuations can cause the estimate to lie outside of that disk. Right: The effect is worse for very pure states in high dimensions – precisely the regime of interest to quantum information. Take a pure density matrix, which, in its eigenbasis, can be written as a matrix that is zero save for a 1 in the upper-leftmost entry. Any perturbation that is not positive semi-definite in the red square of the figure will cause negative eigenvalues to appear. For high dimensions, this happens with probability approaching 1. A detailed analysis can be found in Ref. [9].

This causes a number of problems. For example, quantities that are defined only for density matrices – such as the von Neumann entropy

$$S(\rho) = - \sum_i \lambda_i \log \lambda_i, \quad \{\lambda_i\} \text{ eigenvalues of } \rho$$

can therefore not be estimated as $\hat{S} = S(\hat{\rho})$ (which would be a *plug-in* or *substitution estimator*). How to handle the positivity constraints inherent to quantum formulations is a major issue in the theory, and we will come back to it in Sec. 7.

4.2 Maximum Likelihood Estimation

A general estimation strategy is to pick the model that maximizes the likelihood function given the data

$$\hat{\theta} = \operatorname{argmax}_{\theta} \mathcal{L}(\theta). \quad (5)$$

This is the widely-used *maximum likelihood estimator* (MLE). It selects the model under which the probability of seeing the data would be highest.

In everyday English, *likely* and *probable* can be used synonymously. One might thus be tempted to interpret the MLE estimate as the model choice that is “most probable” to be correct given the data. While common, *this interpretation is unjustified!* Indeed, as indicated in the introduction,

in orthodox statistics, one does not associate a probability distribution with the parameter θ . It is therefore not just a numerical, but a *categorical* mistake to interpret the likelihood as a probability, or to assign any probability at all to a choice of θ . To me, it seems deeply regrettable that the important function $\theta \mapsto p(x|\theta)$ carries a name that strongly suggests a meaning that one cannot assign to it.²

Still, the MLE estimator is extremely useful for a number of reasons. For one, the rule (5) can be directly applied to *any* statistical model. Also, under reasonable assumptions, one can use the *Cramer-Rao bound* to show MLE is optimal in the asymptotic setting (at least away from the boundary of parameter space).

Let us try to get a feeling for it by looking at examples.

Start with the loaded coin. Assume h of n trials gave heads. The likelihood is

$$\mathcal{L}(\theta) = \theta^h (1 - \theta)^{n-h}.$$

Maximizing it is equivalent to maximizing its logarithm, which tends to be better behaved. Compute:

$$\frac{\partial \log \mathcal{L}(\theta)}{\partial \theta} = \frac{h}{\theta} - \frac{n-h}{1-\theta} = \frac{h-n\theta}{\theta(1-\theta)},$$

which vanishes for $\theta = h/n$. Thus, the MLE estimator coincides with the empirical means estimator considered before. Nice! (Don't get used to it).

For the German tanks, the maximum likelihood estimate is trivially found to be

$$\hat{N} = x. \tag{6}$$

So if you find tank number 123, you conclude that there are 123 tanks. Not an answer that is likely (probable?) to boost stats funding after the war. So the claimed asymptotic optimality of the MLE estimator does not seem to extend to small sample sizes.

For quantum problems, MLE does well and is commonly used by practitioners. The main reason is that, by definition, it only returns estimates that are actually elements of Θ – e.g. density matrices. MLE thus avoids the negative eigenvalue problem of the direct inversion method. A generalization of the argument used in the coin example (detailed e.g. in Ref. [12]) shows that

² The nomenclature goes back to a paper by statistics pioneer R. Fisher that appeared 100 years ago. For a discussion of its impact and some of the ensuing debates, see e.g. Ref. [10].

One reason why I am going on and on about this topic is a bizarre encounter I had during an international conference on quantum estimation several years ago. With most of the then-active community in the audience, a highly distinguished veteran of the field was imploring the younger researchers in the audience to stop thinking about new estimators. Fisher had already conclusively settled the argument by the 1930s. He put up slides with quotes by Fisher and himself, making some of the untenable claims I quote above. The pull of the name was stronger than any doubt we tried to sow in his mind about his preferred interpretation. A postdoc from his group later mentioned in private that he, too, had long been trying to convince his boss – to no avail. What a testament to the effectiveness of scientific branding!

Given the central role of statistical methods for many actually important fields (compared to quantum information—like pharmaceutical research), and the problems caused by their misuse (as witnessed by the *replication crisis* [11]), I wonder how much damage to human understanding and wealth has been caused by seemingly innocuous naming decisions.

if the direct inversion estimator gives a physical result, then it equals the MLE estimate. If the direct inversion is outside the physical region, MLE will “project” to a point on the boundary.

It is not obvious how to implement to maximization (5) in a computationally efficient manner. The log likelihood function for quantum state estimation from independent measurements is *convex*: If P_i denotes the POVM element of the i th measurement outcome, we find

$$\log \mathcal{L}(\rho) = \sum_{i=1}^n \log \text{tr} \rho P_i,$$

which is convex, as the logarithm of a linear function is convex, as is the sum of convex functions. Convex functions defined on a convex domain (like the set of density matrices) attain their maximum at the boundary. What is more, finding the maximum of a convex function is often a computationally hard problem. In the case of quantum estimation, iterative algorithms for maximizing the log likelihood have been developed, c.f. Ref. [13]. While I am not aware of a rigorous bound on the number of steps required for convergence, these methods perform well in practice, even if one uses a highly non-linear parameterization for the state (c.f. e.g. Ref. [14]).

5 Uncertainty Quantification

As demonstrated forcefully by the tank example, estimation procedures must give information not only about a “best guess” $\hat{\theta}$ for the unknown parameter, but must also quantify and convey information about the degree of *uncertainty* that remains about its true value. Practitioners need access to an uncertainty measure in order to gauge their trust in the result. Designers of statistical methods can use uncertainty information to rank estimators: The simple idea (though often flawed in practice, see below) is to reject an estimator if there is another one that consistently provides more certain estimates.

5.1 Variance, bias, risk

The most common measure of uncertainty used in physics publication is the *mean error*. It is what the methods (like the *error propagation formula*) taught in undergraduate lab courses are aimed at estimating. In plots in physics publications, the *error bars* usually indicate this value. Here’s the idea, first for the case of a scalar variable $\Theta \subset \mathbb{R}$: For a given true value θ , the observations X_1, \dots, X_n are random, and thus so is the estimate $\hat{\theta}(X_1, \dots, X_n)$. We associate three quantities with the random variable $\hat{\theta}$:

- The expected *mean squared error* or *expected loss*

$$\mathbb{E}_{\theta}[(\hat{\theta} - \theta)^2].$$

- The *variance*

$$\text{Var}_{\theta}[\hat{\theta}] = \mathbb{E}_{\theta}[(\hat{\theta} - \mathbb{E}_{\theta}[\hat{\theta}])^2] = \mathbb{E}_{\theta}[\hat{\theta}^2] - \mathbb{E}_{\theta}[\hat{\theta}]^2. \quad (7)$$

- The *bias*

$$\mathbb{E}_\theta[\hat{\theta} - \theta]. \quad (8)$$

An estimator is *unbiased* if the bias vanishes for all true values θ . For unbiased estimators, Eq. (7) shows that variance and mean squared error are the same.

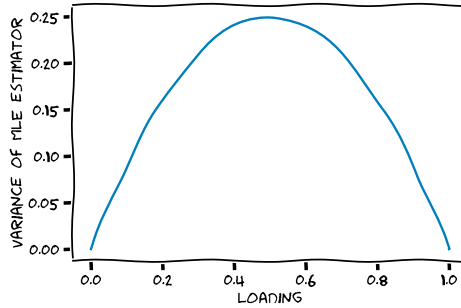


Fig. 2: The mean squared error or variance $\text{Var}_\theta[\hat{\theta}]$ of the MLE estimator for the coin over θ .

Some examples. All estimators from Sec. 4.1 are unbiased. In fact, for the tank problem, we explicitly constructed the estimator by demanding that $\mathbb{E}_N[\hat{N}] = N$. To compute the variance of the coin estimator, we write it as

$$\hat{\theta}(X_1, \dots, X_n) = \frac{1}{n} \sum_{i=1}^n \hat{\theta}(X_i)$$

and use the standard fact that for *independent* variables, the variance of the sum is the sum of the variances to find

$$\text{Var}_\theta[\hat{\theta}(X_1, \dots, X_n)] = n \text{Var}_\theta[\hat{\theta}(X)] = n(\mathbb{E}_\theta[\hat{\theta}^2] - \mathbb{E}_\theta[\hat{\theta}]^2) = \frac{p - p^2}{n} = \frac{p(p-1)}{n},$$

i.e. an inverted parabola that vanishes for $p = 0, 1$ and assumes its maximum of $1/(4n)$ at $p = 1/2$ (Fig. 2). Taking the square root, we recover the familiar $1/\sqrt{n}$ decay for the standard error in terms of the number of experiments. The sample mean estimator for the tank problem has variance

$$\text{Var}_N[2X - 1] = 4\text{Var}_N[X] = 4\mathbb{E}_N[X^2] - (N+1)^2 = \frac{4}{N} \sum_{x=1}^N x^2 - (N+1)^2 = \frac{N^2 - 1}{6},$$

so that the standard error is roughly $N/\sqrt{6}$.

The main takeaway from these calculations is that *we do not know the standard error, unless we know the true value θ !* It seems like we have encountered a cyclical dependency. To estimate the uncertainty in θ , we need θ , but we can't be sure how well we know it until we have estimated its uncertainty. To break out of this cycle, one estimates $\text{Var}[\hat{\theta}]$ directly from the data.

The simplest first guess would be the *substitution estimate*

$$\widehat{\text{Var}}_{\hat{\theta}}[\hat{\theta}] = \text{Var}_{\hat{\theta}}[\hat{\theta}]. \quad (9)$$

For large sample size, this is not altogether bad, but it tends to *underestimate* the error (Why? Hint: the curve in Fig. 2 is *concave*). For one-dimensional problems, there are often correction formulas available which lead to a (nearly) unbiased estimator for the standard error. You may recall the $1/(n-1)$ factors in the definition of the *sample variance* from your lab course, which achieve unbiasedness.

For higher-dimensional problems – like quantum estimation – the theory is much more involved. We will consider the relevant generalizations next.

5.1.1 Distance measures and loss functions for quantum estimation

The first step is to decide which distance measure to use when generalizing the loss function $(\hat{\theta} - \theta)^2$. Many distance functions for quantum states and channels are available, which have different interpretations [5]. For example, if the goal is to estimate a density matrix, the *trace distance*

$$d_{\text{tr}}(\rho, \sigma) = \frac{1}{2} \|\rho - \sigma\|_{\text{tr}} = \frac{1}{2} \text{tr} |\rho - \sigma| = \frac{1}{2} \sum_i \lambda_i(\rho - \sigma), \quad \lambda_i(A) := i\text{-th eigenvalue of } A \quad (10)$$

measures the optimal single-shot statistical distinguishability and is thus often advocated as a well-motivated distance measure. (It is the quantum version of the *total variation* distance used in classical probability theory). In practice, the 2-norm

$$d_2(\rho, \sigma) = \|\rho - \sigma\|_2 = \text{tr}(\rho - \sigma)^2 \quad (11)$$

(also *Euclidean* norm, *Frobenius norm*, or ℓ_2 -norm) is often used, mainly for the pragmatic reason that it is much easier to compute. Indeed, Eq. (11) is just a sum of squares, whereas Eq. (10) requires one to diagonalize a matrix in order to find the singular values. (In the *matrix product state* representation employed in Ref. [14] for quantum estimates, the 2-norm is essentially the only distance measure that can be computed at all).

As an aside, I personally believe that the 2-norm *is* actually a well-motivated measure to use to quantify the risk of an estimator, in particular in quantum technology applications (this might be a non-standard view). Indeed, I would argue that the trace norm as a distance measure is too pessimistic for diagnostic purposes and puts undue weight on the “tail of the spectrum”. The reason is simple: the engineering goal will usually be to prepare a fairly pure state. For the sake of the argument, consider an exactly pure target state $\rho = |\psi\rangle\langle\psi|$. Assume that the actual implementation is described by a density operator σ with a dominating eigenvector equal to $|\psi\rangle$, a second large eigenvector equal to $|\psi'\rangle$ and then a long tail of small eigenvectors. Clearly, from an engineering perspective, we would conclude that there is one dominating noise source that gives rise to $|\psi'\rangle$ and that needs to be addressed in order to rectify the deficiency of the device. The long tail represents miscellaneous small noise sources – but this type of information is unlikely to be actionable. This lack of interpretability notwithstanding, the trace norm would count the full sum of the tail eigenvalues toward the error. If there are two estimators which

both correctly identify the main error source, but one misidentifies the tail, this latter estimator would be assigned a significantly smaller power in terms of trace norm, even though it would be equally useful for the application we have in mind.

As a real-world example, consider the photonic experiment in Ref. [15]. There, the design goal was to create a six photon *symmetric Dicke state*. The reconstructed density operator has essentially three eigenvectors: The target state, plus two components that can be attributed to photon loss in the system. These findings can clearly be interpreted and do not assume any prior hypotheses about the noise processes. Further small deviations in the tail of the density operator have never been interpreted or led to improvements to the experimental setup.

5.1.2 Error estimation via resampling

How can we find the variance of the MLE estimator for quantum states? Consider the relevant case of a true state $\theta = \rho = |\psi\rangle\langle\psi|$. As argued in the caption of Fig. 1, this means that the direct inversion estimate will lie outside of the physical region with high probability, and therefore, that the maximum of the likelihood function will be attained on the boundary. In higher dimensions, the geometry of the boundary of state space close to a low-rank point is fairly complex (no longer a ball as in the qubit case). It is therefore not to be expected that analytical formulas can capture the fluctuations of the estimate.

One common solution used in many large-scale quantum estimation experiments is to resort to numerical Monte Carlo algorithms that aim to simulate the sampling noise. For example, a pioneering quantum estimation experiment was [16]. Uncertainty quantification was done via Monte Carlo resampling, which, according to one of the authors, took significantly more time to be performed than the physical experiment! Two popular methods go by the names of the *jackknife* and the *bootstrap* method, respectively, names that emphasize their general utility and the fact that few assumptions are required.

We sketch the basic ideas here. For a more thorough introduction, I recommend Chapter 5 of Ref. [2].

The bootstrap is the analogue of the substitution estimator for the variance described in (9), combined with a numerical evaluation of $\text{Var}_{\hat{\theta}}[\hat{\theta}]$. Based on the data, one constructs an estimate $\theta_0 = \hat{\theta}(X_1, \dots, X_n)$. Then, n further samples are produced on a computer, following the distribution $p(x|\theta_0)$. Let θ_1 be the result of applying the estimator $\hat{\theta}$ to these synthetic data points. Repeat the re-sampling process t times, to arrive at a collection of simulated results $\theta_1, \dots, \theta_t$. The bootstrap estimate of the uncertainty of $\hat{\theta}$ is then taken to be the empirical variance of these θ_j .

The scheme is attractive, as it is completely general, and can be implemented on a computer mainly by re-using the estimation code that has anyway been written. It comes with certain asymptotic performance guarantees [2]. However, one can argue that the bootstrap is likely to *underestimate* uncertainty, in particular for quantum estimation problems. (This is a heuristic argument – I am not aware of a rigorous bound in the literature). Indeed, consider a true state ρ that has a long tail of small eigenvalues, which are of the same order as, or smaller than the statistical uncertainty. As in Fig. 1, we would then expect that about half the eigenvalues will be negative for a direct inversion estimate, and therefore, that the maximum likelihood estimate $\hat{\rho}$ would have rank of only about half the dimension. The estimate is therefore expected to

be too pure, and thus, the re-sampling process will produce less variation than a re-run of the physical experiment. Based on the argument, the use of the bootstrap for quantum uncertainty quantification therefore has to be carefully justified, to avoid too optimistic error estimates! (For example, one could run computer simulations with known true state chosen to be close to the one that is expected experimentally, and compare the actual risk with the bootstrap variance).

The jackknife simulates additional experiments not by using a first estimate, but by restricting to randomly chosen subsets of the empirical data. Indeed, let $\hat{\theta}_i$ be the estimate produced from samples $X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n$, i.e. with the i -th data point omitted. One can then prove [2] that $(n-1)^2/n$ times the empirical variance of the $\hat{\theta}_i$ is a good asymptotic estimator of the true variance. It is plausible to me that jackknife-based uncertainty estimators do not suffer from the potential bias of the bootstrap described above. But again, I am not aware of a general analysis and would therefore recommend computer studies to validate the method for any particular use case.

5.2 Region estimation

Consider for simplicity a scalar quantity $\Theta \subset \mathbb{R}$, and fix some (unknown) true parameter θ . Assume that under θ , the estimate $\hat{\theta}$ follows a Gaussian distribution centered at θ and with standard deviation σ . Then the probability

$$\Pr_{\theta}[|\hat{\theta} - \theta| \leq x\sigma] = \text{erf}(x/\sqrt{2}) \quad (12)$$

that the estimate differs from the true value by no more than x standard deviations is given in terms of the *error function*, which quickly converges to 1 as a function of x . Thus, e.g., the “ 5σ standard” used in particle physics means that one only trusts estimators that give a wrong result only one in

$$\frac{1}{1 - \text{erf}(5/\sqrt{2})} \simeq 1.7 \times 10^6$$

times. It is this interpretation that most practitioners will have in mind when they see an error bar around a data point in the graph of a physics paper. Note that we assumed a Gaussian distribution – a property that is not always justified and rarely ever checked. (Again, in quantum estimation problems, true states close to the boundary of phase space will cause non-Gaussian distributions of estimators, for the simple fact that a Gaussian function never vanishes and is thus incompatible with compact parameter space.)

So, since we have a tendency of interpreting standard errors as confidence regions, we should really directly devise estimators that output regions instead of point estimates, and that also work in the absence of unfounded Gaussianity assumptions.

Formally, a *confidence region* $C(X)$ is a region of the parameter space Θ that is selected based on observed data $X = X_1, \dots, X_n$. It should be large enough to provide sufficient *coverage*, i.e. the true value θ must be contained in C with a given coverage probability $1 - \alpha$:

$$\text{Prob}_{\theta}[\theta \in C(X)] \geq 1 - \alpha \quad \forall \theta \in \Theta. \quad (13)$$

Simultaneously, the region should obviously be as small as possible, in order to be meaningful.

A number of approaches are available for constructing confidence regions in quantum estimation. One can use the re-sampling algorithms (bootstrap, jackknife) discussed above to create a cloud of estimates and numerically find an ellipsoid enclosing a desired fraction of them. I am not aware of a ready-to-use software implementation of this natural approach, but the open QInfer software package [17] implements a similar logic in a Bayesian context. It could either be used directly, or adapted to an orthodox paradigm. Several authors have argued for *likelihood ratio*-based methods, where the confidence region around a maximum likelihood estimate is chosen to include parameters of decreasing likelihood until the desired coverage is met [18, 19]. Analytical bounds for the radius of trace-norm and Frobenius-norm balls that form confidence regions are also available [20], which even have the property of achieving asymptotically optimal size for any given rank of the density matrix (c.f. Sec. 7).

There is a Bayesian concept – *credible regions* – that fulfills a similar role to confidence regions in orthodox theory. It is much easier to interpret, but, again, requires the specification of a prior distribution. In any case, the Bayesian approaches to quantum state estimation tend to put more of an emphasis on region estimation, so it pays to read e.g. [17, 21].

5.3 Interpretation

Confidence regions are closely related to hypothesis tests, and thus subject to the same potential sources of confusion sketched in the introduction. In particular, the statement Eq. (13) says that *for a given value θ , the probability that an experiment will lead to $\theta \notin C(X)$ is $\leq \alpha$* . This is distinct from a hypothetical statement of the form *having seen the data X , the probability that $\theta \notin C(X)$ is $\leq \alpha$* .

Let's return to the example of a particle discovery in high-energy physics at the 5σ -level. The particle either exists or it does not, and orthodox statistics does not assign a probability to it, much less $1 - 1.7 \times 10^{-7}$. But we should still have high confidence in the finding, as it was obtained by a method that, whatever the truth, will be wrong no more than one in 1.7×10^7 times.

The Bayesian framework does quantify the certainty about the correctness of hypothesis in terms of probabilities. The distinction is sometimes summarized as “orthodox methods deal with the probability of the data given the model; while Bayesians can deal with the probability of the model given the data”.

Similarly, the expected loss quantifies the squared deviation, when averaged over many hypothetical runs of an otherwise identical experiment. It is a reasonable number to use – but keep in mind that the average in its definition is *not usually performed*, so the direct *operational meaning* of the number is not very clear.

6 Optimality criteria & “principles”

Which estimator should one use? Is there a principled reason for rejecting the ML estimator (6) for the tank puzzle, which intuitively seems “too low” compared to the mean estimator (4), which feels more reasonable?

A number of optimality criteria and sundry *principles* have been developed for the purpose of comparing and selecting estimation methods. This program is reasonably successful for simple, one-dimensional models. But, as we will argue in this section, in higher-dimensional cases relevant for quantum estimation, these guidelines are often not helpful. This leads to an unsatisfactory conclusion: There is no unambiguous “optimal” or “best practice” or “general purpose” scheme. Users will have to understand the trade-offs between various methods, and will have to select one tailored to each given use case.

(However, I feel obliged to mention that in the tank case, one *can* show that the means estimator has minimal risk among all unbiased estimators, for all possible values of N . In this sense, it is objectively better than the MLE one, which we intuitively rejected. Glad this, at least, is settled!)

6.1 Unbiased estimators

There seems to be a tendency to name statistical principles in such a way that makes them hard to argue with – even if the formal definition does not adequately reflect the informal meaning of the terms. We have already encountered this problem in the case of the likelihood function. Here, we will find that the notion of *unbiasedness* is similarly problematic, in particular for quantum problems.

By Eq. (8), an estimator $\hat{\theta}$ is *unbiased* if its expected value equals the true one $\mathbb{E}_{\theta}[\hat{\theta}] = \theta$. Nobody likes to be accused of being biased – so it’s natural to reject estimators whose bias does not vanish, no?

To create some first doubt, note that, as argued in Sec. 5.3, the expectation value that appears in the definition ranges over many hypothetical repetitions of an experimental procedure. As this averaging is *not* usually performed in practice, it is therefore not clear that it has great relevance.

That’s a helpful realization, for the purpose of softening the blow that will come next. A slightly disturbing result of Ref. [22] is a simple and elegant geometric argument showing that *there are no unbiased estimators that only return quantum states!* The proof is reproduced in Fig. 3.

6.2 Optimal estimators

The various quantitative notions of uncertainty introduced in Sec. 5 suggest to look for “optimal” estimators that reduce uncertainty as much as possible.

First we have to overcome a problem of orthodox statistics: One can easily construct two estimators such that one of them performs better for some θ_1 , and the other one performs better for another parameter θ_2 . Thus, the best we can hope for is a *partial ordering*, where an estimator is considered worse if it is outperformed by another one for *every* $\theta \in \Theta$. (The situation is better in the Bayesian setting, where one can order procedures by their average performance over the prior distribution). An estimator that is dominated in this way is called *inadmissible*, as there should be no reason to ever use it. (Another term whose everyday meaning will soon turn out not to be defensible within the theory).

Disturbingly, the seemingly innocent notion of admissibility leads to highly counter-intuitive

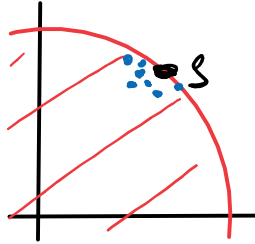


Fig. 3: There are no unbiased estimators of quantum states. Indeed, recall from basic convex geometry that an *extreme point* of a convex set S is such that it cannot be written as a non-trivial convex combination of elements of S . Pure states are extreme points of the convex set S of density matrices. So suppose $\hat{\theta}$ is an unbiased estimator. Then by definition, $\mathbb{E}_\rho[\hat{\theta}] = \rho$. But $\mathbb{E}_\rho[\hat{\theta}]$ is a convex combination of all estimates that can occur due to random fluctuations when performing a measurement on ρ (blue points). Thus we conclude that for all pure ρ , the estimate $\hat{\theta}$ must be equal to ρ with probability 1. But this is impossible by the uncertainty principle. Note for consistency that the proof fails when one tries to apply it to classical distributions. The extreme points of the classical probability simplex are precisely the deterministic configurations, for which indeed no randomness occurs.

consequences. This is demonstrated by *Stein's paradox*, which we now sketch. Suppose the aim is to estimate the mean θ of a scalar Gaussian variable $X \sim N(\theta, 1)$ with variance 1 from a single observation. It is intuitive that $\hat{\theta} = X$ is a good estimator, and indeed, it is admissible. Now assume that $n \geq 3$ independent variables $X_i \sim N(\theta_i, 1)$ are to be measured. Say the temperature in Jülich, the caffeineation level of a randomly chosen faculty member of Princeton's IAS, and the IQ of the last person to have entered Saint Peter's Basilica in Rome (all properly re-scaled to follow the indicated distribution). As these quantities have patently no connection to each other, a well-behaved theory would confirm that it's OK to estimate each of these values independently, using the rule devised for the one-dimensional case. But... this procedure is inadmissible! It is beaten by a so-called *shrinkage estimator* which, outrageously, introduces dependencies between the estimates, but outperforms the more natural method for every given set of true values $\theta \in \mathbb{R}^3$!

Lesson learned: There is no simple criterion for selecting orthodox estimators for high-dimensional problems. Case-by-case judgments will have to be made.

7 The role of positivity

What is the most important characteristic of *quantum* estimation problems, as compared to those that occur more naturally in classical applications? In my opinion, the prevalence of *positive semi-definiteness* constraints on quantum states and quantum evolution maps takes this role. Taking them into account is often algorithmically and conceptually difficult, but can also lead to significantly improvements in accuracy.

The geometry of the problem is shown in Fig. 4. To explain the idea, we consider the task of characterizing a state ρ from a set of possibly informationally *incomplete* measurements whose

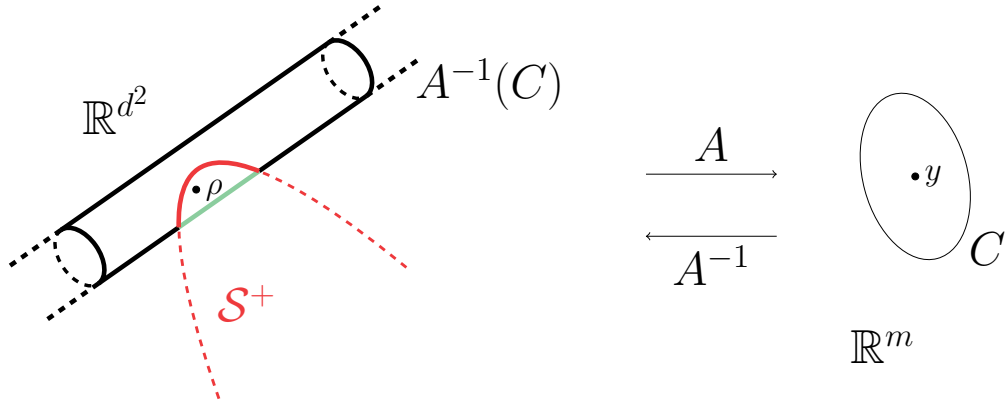


Fig. 4: Geometry of optimal error regions. Quantum states are mapped by a measurement matrix A to expectation values y . Conversely, confidence regions C for expectation values give rise to confidence regions $A^{-1}(C)$ for quantum states. While possibly unbounded, the intersection of $A^{-1}(C)$ with the set of states S^+ will often be small. Error bars correspond e.g. to the diameter of the intersection (green line).

results y follow a Gaussian distribution. If the dimension of the system is d , the state is an element of \mathbb{R}^{d^2} . The vector y of measurement outcomes belongs to \mathbb{R}^m . The “measurement map” sending states ρ to expected outcomes \bar{y} can be represented by an $m \times d^2$ -matrix A such that $\bar{y} = A\rho$. A, say, 95%-confidence region C for y can readily be found—in case of Gaussian measurement statistics, it will be an easily described ellipsoid. A first 95%-confidence region for ρ is the set of pre-images $A^{-1}(C)$ of C under the measurement map. However, in case of informational incompleteness, A is singular, and the pre-image is unbounded (a cylinder with ellipsoidal cross-section). This corresponds to infinite error bars.

At this point, we use the only prior information we have about any quantum experiment: physical states are non-negative. Thus we know that ρ must lie (with high confidence) in the intersection of the unbounded error region $A^{-1}(C)$ with the set of physical states. This intersection is precisely the error region that follows from C and physicality, and it may become small, even if $A^{-1}(C)$ is *unbounded* (Fig. 4). That such cases indeed exist follows from the theory of *compressed sensing* as applied to density matrices [23, 24]. In particular, it was demonstrated that for states of *low rank*, i.e. those with few non-zero eigenvalues, using positivity typically gives rise to unbounded improvements of uncertainty. This is all the more relevant since, as argued before, nearly-pure states are usually desired for quantum technology applications.

After the results of Ref. [23, 24] on the beneficial influence of the positivity constraint had been obtained, it became a natural question whether there is a general-purpose algorithm that can identify the smallest region in the intersection of a given uncertainty region and the set of physical states. The answer is negative: As shown in Ref. [25], finding the diameter of the intersection is NP-hard, i.e. computationally intractable. Only approximate solutions can be obtained with efficient algorithms.

Thus, in summary, the positivity of states gives rise to the unique flavor and to the difficult challenges of statistical problems in quantum estimation.

8 Acknowledgments

The author's work on quantum estimation has been supported by the ARO under contract W911NF-14-1-0098 (Quantum Characterization, Verification, and Validation), and the DFG projects GRO 4334/1,2 (SPP1798 CoSIP). Some paragraphs in these notes are based on project reports to the ARO, which were co-authored by Steve Flammia (University of Sydney) and Chris Granade (Microsoft Research).

References

- [1] J. Eisert, D. Hangleiter, N. Walk, I. Roth, D. Markham, R. Parekh, U. Chabaud, E. Kashefi, <https://arxiv.org/abs/1910.06343>
- [2] J. Shao, *Mathematical Statistics* (Springer, New York, 2003).
- [3] K. Knight, *Mathematical Statistics* (Chapman & Hall, Boca Raton, 2000).
- [4] D. MacKay, *Information Theory, Inference and Learning Algorithms* (Cambridge University Press, Cambridge, 2003).
- [5] M. Nielsen, I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
- [6] J.S. Lundeen, A. Feito, H. Coldenstrodt-Ronge, K.L. Pregnell, Ch. Silberhorn, T.C. Ralph, J. Eisert, M.B. Plenio, I.A. Walmsley, *Nature Physics* **5**, 27 (2009).
- [7] C. Stark, *Physical Review A* **89**, 052109 (2014).
- [8] D. Greenbaum, <https://arxiv.org/abs/1509.02921>
- [9] L. Knips, C. Schwemmer, N. Klein, J. Reuter, G. Toth, H. Weinfurter, <https://arxiv.org/abs/1512.06866>
- [10] S. Stigler, *Statistical Science* **22**, 598 (2007).
- [11] H. Pashler, E.-J. Wagenmakers, *Perspectives on Psychological Science* **7**, 528 (2012).
- [12] R. Blume-Kohout, *New Journal of Physics* **12**, 043034 (2010).
- [13] Z. Hradil, J. Reháček, J. Fiurasek, M. Jezcaronek, in *Quantum state estimation*, edited by M. Paris and J. Reháček, Springer, Berlin, 2004
- [14] T. Baumgratz, D. Gross, M. Cramer, M. Plenio, *Physical Review Letters* **111**, 020401 (2013).
- [15] C. Schwemmer, G. Toth, A. Niggebaum, T. Moroder, D. Gross, O. Gühne, H. Weinfurter, *Physical Review Letters* **113**, 040503 (2014).
- [16] H. Häffner, W. Hänsel, C. F. Roos, J. Benhelm, D. Chek-al kar, M. Chwalla, T. Körber, U. D. Rapol, M. Riebe, P. O. Schmidt, C. Becher, O. Gühne, W. Dür, and R. Blatt, *Nature* **438**, 643 (2005)..

- [17] C. Granade, C. Ferrie, I. Hincks, S. Casagrande, T. Alexander, J. Gross, M. Kononenko, Y. Sanders, *Quantum* **1**, 5 (2017).
- [18] R. Blume-Kohout, <https://arxiv.org/abs/1202.5270>
- [19] M. Christandl, R. Renner, *Physical Review Letters* **109**, 120403 (2012).
- [20] A. Carpentier, J. Eisert, D. Gross, R. Nickl, in *High Dimensional Probability VIII*, edited by N. Gozlan, R. Latała, K. Lounici, M. Madiman, Progress in Probability, vol 74. Birkhäuser, Basel, 2019
- [21] K.M.R. Audenaert, S. Scheel, *New Journal of Physics* **11**, 023028 (2009).
- [22] C. Schwemmer, L. Knips, D. Richart, T. Moroder, M. Kleinmann, O. Gühne, H. Weinfurter, *Physical Review Letters* **114**, 080403 (2015).
- [23] D. Gross, Y.-K. Liu, S.T. Flammia, S. Becker, J. Eisert, *Physical Review Letters* **105**, 150401 (2010).
- [24] S.T. Flammia, D. Gross, Y.-K. Liu, J. Eisert, *New Journal of Physics* **14**, 095022 (2012).
- [25] D. Suess, L. Rudnicki, D. Gross, *New Journal of Physics* **19**, 093013 (2017).

A 6 **Quantum Measurement**

David P. DiVincenzo

JARA Institute for Quantum Information

Forschungszentrum Jülich GmbH

Contents

1	Basics of Measurement	2
2	Preventing measurements from happening	8
3	A bit on real measurements: incomplete measurements	9
4	The Bohr-Rosenfeld measurement of electromagnetic field measurements in space and time.....	11
5	Creating entanglement by measurements on the vacuum	13

1 Basics of Measurement

Let us first do the usual textbook story [1] of measurement. A qubit, say, is in some state as time t evolves, $|\psi(t)\rangle = \alpha(t)|0\rangle + \beta(t)|1\rangle$, where the complex state amplitudes α and β ($|\alpha|^2 + |\beta|^2 = 1$) are evolving smoothly in time in accordance with some Schrödinger equation, in other words, according to some unitary time evolution operator $U(t)$.

When we say “measure”, the rules change completely. Then reversible, unitary evolution is suddenly, at the measurement time t_m , replaced by stochastic, irreversible, projective evolution. The state “collapses” by projection either to $|0\rangle$ (with probability $|\alpha(t_m)|^2$) or $|1\rangle$ with probability $1 - |\alpha(t_m)|^2 = |\beta(t_m)|^2$.

The arbitrariness of this prescription, the *Born rule* should bother you. What determines t_m , and why is the change instantaneous? Why is it stochastic? What if the initial state is entangled? I will answer all these questions, but the last one is easiest, and allows me to introduce some of the additional machinery of quantum theory. When we talk about entanglement, we are talking about a multi-part quantum system. How do we describe the state of such a multi-part system? Each system separately is described as a complex vector in a Hilbert space. Systems taken together are described by a vector in the *Kronecker product* (\otimes) space. This is a linear algebra concept worth knowing about, see

https://en.wikipedia.org/wiki/Outer_product#Contrast_with_the_Kronecker_product.

The vector for an unentangled state of two parts is just a Kronecker product of two vectors; an entangled state is described by a linear combination of at least two such vectors.

In this linear algebra language, operators on these vectors are square matrices. Operators on multi-part systems are also square matrices, constructed by the Kronecker product rule appropriate for matrices, this is covered in https://en.wikipedia.org/wiki/Kronecker_product.

Now we can return to Born. For the single-system, the linear algebra prescription is to apply a projection operator for each measurement outcome, namely $\Pi_0 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $\Pi_1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. My example state from above we write in this notation as $|\psi\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$. Then the probability of outcome b ($b = 0, 1$) is given by the norm squared of the resultant vector:

$$p_b = \|\Pi_b|\psi\rangle\|^2. \quad (1)$$

Now we can state generalized Born. Suppose we have a two-part state. A good example would be the maximally entangled Bell state. It can be written as a linear combination of vectors created by the Kronecker product; let us write this out once in long form:

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (2)$$

Suppose we measure only the first qubit. Generalized Born says that the projection operator on the composite space is the Kronecker product of the projection on the measured operator with

the identity operator $\mathbb{1}$, viz., $\Pi_b \otimes \mathbb{1}$. The mathematical rule for the probabilities is then the same as before, and gives the values $p_0 = p_1 = 1/2$ for the Bell state.

If the state has many parts and we only measure the first part, then the generalized Born operators are $\Pi_b \otimes \mathbb{1} \otimes \mathbb{1} \otimes \dots \otimes \mathbb{1}$. If the state has N parts (so $N - 1$ unmeasured parts), this is abbreviated $\Pi_b \otimes \mathbb{1}^{\otimes(N-1)}$. A similar abbreviated notation is used for the tensor product of identical state vectors. We apply this immediately for an important generalization of the Bell state to N parts, called the Greenberger-Horn-Zeilinger (GHZ) state:

$$|\Psi_{GHZ}\rangle = \frac{1}{\sqrt{2}} (|0\rangle^{\otimes N} + |1\rangle^{\otimes N}) \quad (3)$$

With this formalism in hand, I would like to take a first stab at the question, what really happens when a measurement is performed? I will take the point of view that what is really fundamental about quantum mechanics is unitary evolution, and that the concept of projective measurement must be nothing other than a *handy approximation*.

We will say that measurement commences when the object to be measured (a qubit, in the discussion I am about to give) interacts with another physical apparatus. This apparatus should be macroscopic. Let us provisionally model it simply as a collection of many qubits. Let us suppose that each apparatus qubit interacts with the original qubits such that their evolution is that of a CNOT, with unitary matrix in the two qubit space

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \quad (4)$$

Fig. 1 shows this evolution in a circuit diagram. The horizontal axis is time – it does take time to do a measurement! Formally the CNOTs can happen simultaneously (their matrices commute), but even so, the interaction producing any single qubit cannot be instantaneous. We will come back to this later.

Finally, suppose the apparatus qubits are prepared in a simple initial state, in fact $|0\rangle^{\otimes(N-1)}$. Then this evolution will result in a simple entangled state; if the original qubits is in an equal superposition state, then the final state is exactly the GHZ state Eq. (3).

We will say that in this situation the measurement has taken place, and that the Born rule may be applied. Why? We say that if N is large, the state of the qubit to be measured has been robustly recorded in the apparatus. This has the consequence that it is easy to macroscopically “copy” the outcome even further – as we like to say, “publish in the newspaper”. We have of course come to the Schrödinger cat paradox – that the cat standing in for our apparatus, and then the newspaper reporting its death or life, and all the people reading the newspaper, are in an ever larger GHZ state. In this state, their state of knowledge (of the readers, that is) is perfectly correlated with

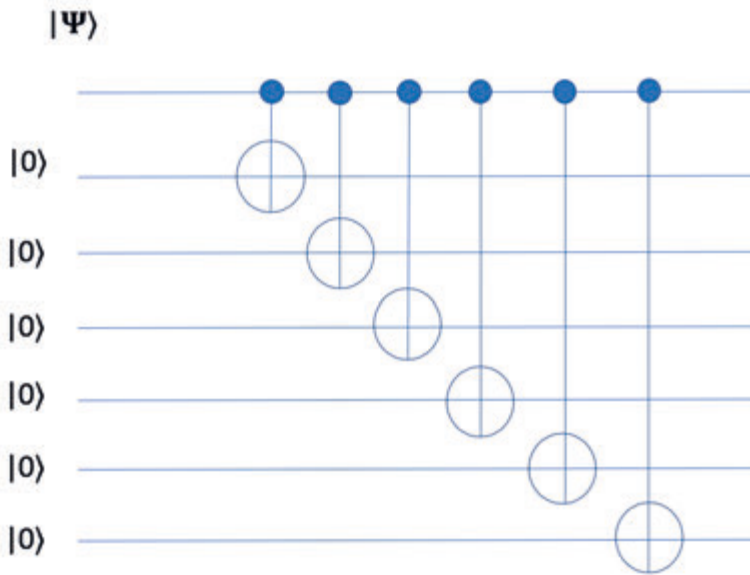


Fig. 1: Quantum circuit of measurement: repeated “copying” of the quantum state with controlled not (CNOT) operations.

the state of the original qubit. Therefore, it is a matter of computational economy to say that the outcome is “DEAD” or “ALIVE” (as announced in the headline of the newspaper), with the projective Born rule applied to the qubit. The probability factor is also the same as the norm squared of the appropriate part of the GHZ state.

Since the only unknown in the GHZ state is the duplication factor N , it is this that must evolve in time, as suggested by Fig. 1. It is interesting to inquire how it changes in time $N(t)$. Before the measurement starts, we can confidently say that $N(0) = 1$. In the figure, it is just a counter summing up how many CNOTs have happened. It is interesting to inquire what happens to $N(t)$ in the actual thought experiment as laid out by Schrödinger.

Here is what Schrödinger said [2], interspersed with my interpretation as a measurement circuit:

One can even set up quite ridiculous cases. A cat is penned up in a steel chamber, along with the following device (which must be secured against direct interference by the cat): in a Geiger counter there is a tiny bit of radioactive substance, so small,...

I will simplify this “tiny bit” of substance to just one atom in an excited state. He actually describes a more complicated situation, but I don’t think that my simplification loses anything essential.

... that perhaps in the course of the hour one of the atoms decays, but also, with equal probability, perhaps none; ...

I represent the one atom by one qubit; the 1 state is the metastable state, the 0 is the final state after decay. But this decay emits an excitation (an α particle, say) into the environment. Let's represent this by one more qubit, 0 is empty environment, 1 is environment with one excitation. We will interpret this decay as a unitary operation, which in the course of the hour would be the transformation

$$U_{decay} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & (H) & 0 \\ 0 & & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (5)$$

The last part serves to define the 2×2 Hadamard matrix H , which we will use later. We follow Schrödinger's description further:

...if it happens, the counter tube discharges...

This tells us that the evolution is not the pure unitary operation of Eq. (6), as it would be if the atom were in interstellar space: the excitation released into the environment very soon reaches the “counter”. Without knowing much about the inner workings of the counter, we can be pretty confident that it very rapidly flips a whole lot of bits (e.g., a rush of electrons through a detection circuit), which we can schematize as our CNOT network of Fig. 1. We might say that this process happens continuously, and we will deal with this by saying that the decay evolution is separated into many factors acting over a short period of time within the allotted hour:

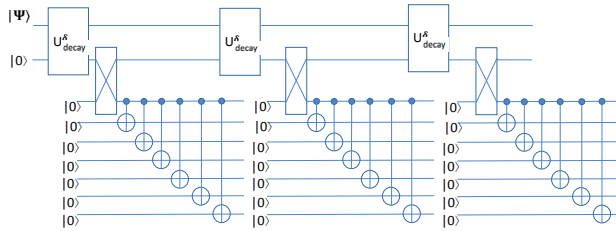


Fig. 2: A quantum-circuit depiction of the steps of the indirect measurement performed in Schrödinger's thought experiment.

$$U_{decay}^{\delta} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \delta & \sin \delta & 0 \\ 0 & -\sin \delta & \cos \delta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (6)$$

Note that if this evolution occurs repeatedly without disturbance, it will in a number of repetitions equal to $\frac{\pi}{2\delta}$ come to the U_{decay} above. But instead after every short evolution the excitation, if emitted, propagates a short distance (schematized by the SWAP in Fig. y), and then a “record is made”, that is, according to the action of the CNOT network:

$$|\Psi(\delta)\rangle = \cos \delta |10\rangle + \sin \delta |00 1_0^{\otimes n}\rangle \quad (7)$$

or after K stages of such interaction:

$$|\Psi(K\delta)\rangle = \cos^K \delta |10\rangle + \sum_{k=0}^{K-1} \sin \delta \cos^k \delta |00 1_k^{\otimes n}\rangle \quad (8)$$

The n new qubits represent the internal state of the counter; we might imagine that immediately, this is nearly macroscopic, e.g., $n \sim 10^{18}$. In fact, the idea is that we get a fresh batch in every small interval, that is, the signal comes in a particular window of time. This is the significance of the subscript k in Eq. (8) ($k = 0$ in Eq. (7)). It should be understood that when an excitation is *not* emitted into the environment, the counter qubits are not flipped, but they are still there; therefore, the state in the first term of Eqs. (7) or (8) should be thought of as a shorthand for $|1 0^{\otimes n}\rangle$.

Now Schrödinger proceeds further, and finally comes to the famous aspects of his thought experiment:

...and through a relay releases a hammer which shatters a small flask of hydrocyanic acid. ...

At this point, this is just a bit of theatrics; nothing essential changes about the state anymore, except that the record becomes “more macroscopic”, now involving everything in the enclosure, even that cat. So, perhaps now $n \sim 10^{25}$. This is making the point that once the GHZ state has reached macroscopic proportions, It is not hard to make it even more macroscopic, perhaps by announcing the decay on loudspeakers worldwide, etc. Perhaps at this point n would be 10^{50} . But nothing essential is changed about the measurement (but there can be a much longer timescale process of erasure and forgetting, see the comment on Bennett’s work on the next page).

Schödinger finishes ironically:

If one has left this entire system to itself for an hour, one would say that the cat still lives if meanwhile no atom has decayed. The first atomic decay would have poisoned it. The ψ -function of the entire system would express this by having in it the living and dead cat (pardon the expression) mixed or smeared out in equal parts.

We have not imposed the “full hour” condition, which amounts to setting $\cos^{2K} \delta = 1/2$. Using this condition, we write the final “psi-function” (with Schrödinger’s flourish):

$$|\Psi(1/2)\rangle = \frac{1}{\sqrt{2}} |1 0^{\otimes n}\rangle |\text{LIVE}\rangle + \sum_{k=0}^{K-1} \sqrt{\frac{\ln 2}{K}} 2^{-\frac{k}{2K}} |00 1_k^{\otimes n}\rangle |\text{DEAD}\rangle \quad (9)$$

We see that the theme which Schrödinger's story focuses on is the every larger size of the effective GHZ state. The N of Eq. (3) is, after the initiation of measurement, a very rapidly growing function of time. While I have offered some speculation about how big N is in Eq. (9), I would like to mention that there has been effort to give a formal definition of this N , referred to as the "cat size" [11]. The definition is given operationally: N is the maximum number of subsets into which the "cat" could be subdivided, such that an ideal (mathematical) measurement on each subset individually would yield, univocally, the correct measurement outcome. More colloquially, it is given by the volume v of the smallest voxel in an MRI scan of the cat such that each voxel makes it clear whether the cat is DEAD or ALIVE; then, $N = V_{CAT}/v$, with V_{CAT} being the total volume of the cat. It would seem that N being the total number of cells in the body of the cat would not be a bad guess.

In Eq. (9) the state has reached a condition that the probabilities of 0 and 1 for the original qubit (the radioactive atom) are equal. But we see, and this is usually taken as one of the signature features of measurement, that it is impossible to undo the measurement, meaning to turn the qubit back into a coherent superposition of 0 and 1, $1/\sqrt{2}(|0\rangle + |1\rangle)$. It is not strictly correct that this transformation is impossible. There always exists a unitary transformation U_{undo} that will rotate any state vector to any other state vector. Constructing such a transformation for the state of Eq. (9) is extremely inconvenient, as it would involve interacting in an invasive way with every degree of freedom of the cat and the rest of the apparatus.

C. Bennett [15] has pointed out that one can see that the undoing of the measurement is, at another, more profound level, impossible. To explain his insight, we imagine that, unlike in Schrödinger's original story, the enclosure is out in the open. It is further imagined that some light strikes the cat, and of the reflected light, some of it departs immediately into outer space. Thus, some of the qubits needed for the undoing operation U_{undo} have left the area at the speed of light; no material apparatus can possibly catch up with them and bring them back. With advance planning one could surround the entire earth with mirrors, but barring this, the parts necessary to undo the measurement are irretrievably gone. One can make a similar, but perhaps less compelling, story of the excitations that travel inward, that is, into the deep earth.

Bennett further observes that while the measurement cannot be undone, it is quite possible for the outcome of a completed measurement to be lost, again by the expedient of the departure of some quanta into outer space. We imagine that a measurement process has brought us to Eq. (8). Then the whole apparatus is destroyed, or maybe just falls apart due to age. (Let's leave the cat out of this.) The destruction per say does not destroy the record, to the extent that it is just a unitary transformation on the full state. But if then some quanta escape into space, then the information becomes even in principle not reconstructable on earth. Bennett surmises that this is the status of, e.g., books written in antiquity that have been lost.

I will just note here that the (relatively rare) cases where a classical record lasts for a long time is dealt with in the theory of *quantum Darwinism* [7]. Many facts about quite common real-world observations have been explained with this theory [8, 9].

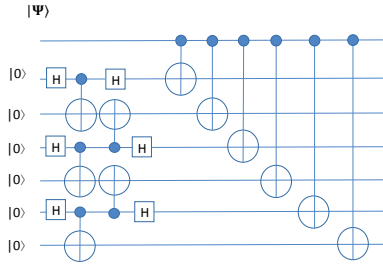


Fig. 3: A modification of the measurement circuit of Fig. 1. The specific pre-entanglement of the environment, paradoxically, prevents any production of useful copies of the measurement!

2 Preventing measurements from happening

It would appear that the idealization of Fig. 1, the repeated applications of CNOTs, is the most efficient way to build up a measurement, and we feel that something like this is going on all the time in natural processes that produce a macroscopic “cat size”. But further theoretical analysis shows something quite different could happen. Fig. 3 shows a modification of the measurement circuit. The only difference is that the initial environment, rather than itself being in the unentangled $|0\rangle^{\otimes N}$ state, is subject to a short series of interactions, viz., the depth-5 circuit shown (repeated down the line of N qubits). This makes the environment somewhat entangled, but you might imagine that it still effectively consists of a set of disconnected parts, since there is no time for a signal to propagate very far down this line of N qubits.

Nevertheless, there is an amazing fact about this operation: after the N CNOTs that we thought would create the GHZ state, the cat number is still 1, no matter how large N is! That means that an ideal measurement on even $N - 1$ of the environment qubits reveals nothing about the original qubit – it remains in an entangled state with the one remaining qubit. This situation, of arranging conditions such that it is very hard to measure a qubit, is connected with the possibility of quantum error correction codes – qubit which are hard for the environment to measure are also hard for the environment to decohere. Quantum states that have error correction power are closely related to the state here, in that they are both described by the “stabilizer theory”, about which I will say no more here.

This fact, that a straightforward “preparation” of the environment can result in a big alteration of the dynamics of measurement, is a rather new observation, only coming with the advent of quantum information theory. The multi-qubit state we have introduced is called the *cluster state*, [10] which has many interesting properties and roles in the theory of quantum computation.

While “easy” to prepare from the theoretical point of view, it does not occur naturally as far as we know, and will require the advent of reliable quantum computers to make its creation routine.

3 A bit on real measurements: incomplete measurements

I would like to come to few more technical things, of interest in current work. Let us go back to a single qubit, and to the simplest scenario outlined in Eq. (1). It is understood that in a setting of indirect measurement, the measurement operators need not be projectors. Actually, Schödinger’s cat experiment is such an indirect measurement, and the following formalism could be applied to it. In the setting of indirect measurement, outcome a can be associated with a general measurement matrix M_a . M_a is not a projector, is not Hermitian, and is not unitary. If outcome a is obtained, state $|\psi\rangle$ is transformed via a linear transformation to the unnormalized state

$$M_a|\psi\rangle. \quad (10)$$

The probability of this outcome is $||M_a|\psi\rangle|| = \langle\psi|M_a^\dagger M_a|\psi\rangle$. We alternatively consider the measurement to be the nonlinear transformation ending in the normalized state

$$\frac{M_a|\psi\rangle}{||M_a|\psi\rangle||}. \quad (11)$$

The only constraint on M_a comes from completeness, or in other words, the necessity for the sum of all probabilities to be one. This gives

$$\sum_a M_a^\dagger M_a = \mathbb{1}. \quad (12)$$

The constraint that this puts on any individual measurement operator is

$$M_a^\dagger M_a \leq \mathbb{1}, \quad (13)$$

equality only occurring in the trivial case of only one outcome, and M_a being ordinary unitary evolution.

Given this simple extension of measurement theory, I want to jump to a rather sophisticated application of these simple equations. This will involve a description of a very important measurement in quantum technology, which is the *homodyne measurement*. We will see that this is covered by the discussion just given, with the added feature that the measurement is repeated over and over. Such repetition is completely uninteresting in the case of projective measurement; a second application of the projector gives no further change compared with the first application. But the interesting, and realistic, case is one in which any one application is uninformative, so that a good measurement only results from many applications. Uninformative

will means that all (or most) of the operators M_a are close to being proportional to the identity operator,

$$M_a \approx c\mathbb{1}, \quad (14)$$

for some constant $|c| < 1$.

Homodyne measurement is usually given a rather lengthy physical explanation, which can be found in many places (e.g., Wiseman and Milburn) [12]. But I would like to jump right to a mathematical hypothesis offered by Gambetta (his Eq. (7.105) [13]), which permits us to see immediately that this physical measurement satisfies the equations given above. Gambetta's hypothesis is

$$\hat{M}_I = \left(\frac{dt}{2\pi} \right)^{1/4} \exp(-I^2 dt/4) \{1 - [i\hat{H}dt - \sqrt{\gamma}I\hat{L}dt + \frac{1}{2}\gamma\hat{L}^\dagger\hat{L}dt]\} \quad (15)$$

We have adopted two points of Gambetta's notation: we now put hats over operators on the Hilbert space, and the measurement outcome is now labeled I . The operator \hat{L} determines the measurement that is being done, for standard heterodyne measurement it is a photon destruction operator, but we will see other operators in the following. The outcome I is now understood to be real-valued, with any value between $-\infty$ to ∞ , and we see that the measurement has an infinitesimal prefactor that is a function of the time differential dt . Of course, the gaussian prefactor strongly suppresses I s far outside the range $-1/\sqrt{dt} \lesssim I \lesssim 1/\sqrt{dt}$.

Fractional powers of infinitesimals are probably unfamiliar to you. But you can confirm that these peculiar forms are necessary in order to satisfy the continuous-valued form of the completeness relation,

$$\int_{-\infty}^{\infty} \hat{M}_I^\dagger \hat{M}_I dI = \mathbb{1}, \quad (16)$$

with corrections only at order dt^2 . We will shortly see that the stochastic calculus, or the Ito calculus, is needed to describe the time evolution under repeated application of this measurement; but I will mention that one Ito rule is needed to arrive at Eq. (16), namely the replacement of the stochastic quantity $I^2 dt$ by its leading-order expectation value 1.

Let me proceed on immediately to discuss the result of repeated application of Eq. (15). We assume that in each successive interval dt of time that measurement outcome I_1, I_2, \dots, I_k is obtained. Then, as Gambetta writes (his Eq. (7.42)), the resultant wavefunction is

$$|\psi_{\mathbf{I}}(t)\rangle = \frac{\hat{M}_{I_k}(t, t-dt) \hat{M}_{I_{k-1}}(t-dt, t-2dt) \dots \hat{M}_{I_1}(t_0+dt, t_0) |\psi(t_0)\rangle}{N}. \quad (17)$$

N is a factor to keep the final state normalized. Of course, the numerator of this expression has very tiny norm (of order $dt^{k/4}$), reflecting the fact that to get any *particular* sequence of outcomes \mathbf{I} is highly unlikely. Taking one step of this normalized evolution, Gambetta shows that the evolution is given by a Schrödinger-like equation:

$$d|\psi_{\mathbf{I}}(t)\rangle = \left[-i\hat{H}dt + \sqrt{\gamma}(\hat{L} - \langle\hat{L} + \hat{L}^\dagger\rangle_t/2)(dW + \sqrt{\gamma}(\hat{L} + \hat{L}^\dagger)_t dt) - \gamma/2 \left(\hat{L}^\dagger\hat{L} + \hat{L}\langle\hat{L} + \hat{L}^\dagger\rangle_t - \frac{3}{4}\langle\hat{L} + \hat{L}^\dagger\rangle_t^2 \right) dt \right] |\psi_{\mathbf{I}}(t)\rangle. \quad (18)$$

This equation introduces a rescaled version of the Gaussian random variable I , in the form of the Wiener increment dW , another Gaussian random variable with mean 0 and variance dt . The notation $\langle \dots \rangle_t$ denotes the diagonal matrix element of the operator for state $|\psi_I(t)\rangle$. Note that this makes the evolution equation nonlinear; its nonlinearity arises entirely from the normalizing factor N in Eq. (17).

Note that the stochastic Schrödinger equation of Eq. (18) can be extended to the stochastic master equation:

$$d\hat{\rho}(t) = -i[\hat{H}, \hat{\rho}(t)]dt + \gamma \mathcal{D}[\hat{L}]\hat{\rho}(t)dt + \sqrt{\eta\gamma} \mathcal{H}[\hat{L}]\hat{\rho}(t)dW, \quad (19)$$

$$\mathcal{D}[\hat{L}]\hat{\rho} \equiv \hat{L}\hat{\rho}\hat{L}^\dagger - \frac{1}{2}(\hat{L}^\dagger\hat{L}\hat{\rho} + \hat{\rho}\hat{L}^\dagger\hat{L}), \quad (20)$$

$$\mathcal{H}[\hat{L}]\hat{\rho} \equiv \hat{L}\hat{\rho} + \hat{\rho}\hat{L}^\dagger - \text{Tr}[(\hat{L} + \hat{L}^\dagger)\hat{\rho}]\hat{\rho} = \hat{L}\hat{\rho} + \hat{\rho}\hat{L}^\dagger - \langle \hat{L} + \hat{L}^\dagger \rangle \hat{\rho}. \quad (21)$$

Eq. (19) contains the new parameter η the “measurement efficiency”. When $\eta = 1$ this equation gives the same state evolution as the Schrödinger equation Eq. (18). $\eta < 1$ gives the possibility of modeling measurements that are more noisy than necessary.

4 The Bohr-Rosenfeld measurement of electromagnetic field measurements in space and time

I will finish these notes with some topics which I consider only partially finished, that is, on which original research could still be done. The first of these involves the measurements of electric and magnetic fields in space. Suppose that you (i.e., a measuring apparatus) goes out into the vacuum, and at some time (or within some time integral) measures the strength of the electric field in the x direction. According to the Born rule as applied to the vacuum state of the quantum electromagnetic field, this measurement should typically yield some non-zero value. The mean over many measurements would indeed be zero, but there is a zero-point fluctuation, so that the observed field will have a gaussian distribution around zero.

In Ref. [4], Bohr and Rosenfeld (BR) consider several very fundamental general aspects of this measurement. They stated for the first time the observation that the measurement at a point in space-time of a field value is not defined, that is, will lead to diverging values. However, if the measurement is over a finite test volume in space and time, the statistics of the measurement values becomes well defined.

BR give an extensive discussion of the properties of the classical physical apparatus that will do the measurement. It should contain a solid body with dimensions matched to the test volume, and with a body charge and mass. The electric field couples to the charge, and the body accelerates under this force. The measurement of the displacement of the body after a specific length of time constitutes the field measurement. This displacement will be accordingly positive or negative, or we can say (anachronistically, as BS predates Schödinger’s cat by several years) that the test body is in a Schrödinger-cat superposition of forward and backward displacement.

It is clear that BR have in mind a very specific sort of classical apparatus in mind, with masses, springs, coils (for magnetic field measurements) and recording pointers, but they provide no figures of what they envision.

BR's main theoretical concern had to do with the compatibility of this measurement scenario with the commutation relations of the quantum electromagnetic field. One example of these commutators, as stated in BR, is $(r^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2)$:

$$[E_x(t_1, x_1, y_1, z_1), E_x(t_2, x_2, y_2, z_2)] = \quad (22)$$

$$i\hbar \left(\frac{-\partial^2}{\partial x_1 \partial x_2} + \frac{1}{c^2} \frac{\partial^2}{\partial t_1 \partial t_2} \right) \left[\frac{1}{r} \delta \left(t_2 - t_1 - \frac{r}{c} \right) - \frac{1}{r} \delta \left(t_1 - t_2 - \frac{r}{c} \right) \right]$$

There are similar commutators for all the field components ($E_{x,y,z}$, $B_{x,y,z}$); the electric fields commute with magnetic fields in the same coordinate direction, but not with ones in perpendicular directions.

When written in this old fashioned way, one can see that the structure of this commutator is exceedingly singular. As BR mentioned, this had led Peierls and Landau to conclude that the Heisenberg uncertainty relation would not apply, that is, that the precision of measurement of non-commuting variables would not be determined by this commutator. BR's first step, in taking on this question, was to introduce the idea, which has now become standard in field theory, that the observable should be considered an *averaged* field quantity, that is, over a volume in space and time. This “smooths off” the severe singularities of the field commutator.

In addition, this “test volume” concept was precisely matched, in BR's approach, with the size of the sensing element. They take very seriously the idea that there should be a complete specification of a classical apparatus that does the field measurements, with a test volume of a specified mass and charge density interacting with the field for a specified length of time, and a spring apparatus to determine the resulting impulse. While in other settings Bohr provided realistic technical drawings of gedanken experiments (for example, Fig. 4 is his apparatus for weighing a photon), unfortunately Bohr and Rosenfeld provided no visual specifications for the complex machines that they describe.

Without such details, it is still possible to understand qualitatively what is going on, and why the commutator above may give an accurate indication of what measurements cannot be performed together without Heisenberg uncertainty. For example, their electric field detector is basically a charged body, and its acceleration provides a measurement of the field. As Preskill [3] notes, this acceleration produces additional radiation, which travels outward along the light cone starting at space-time position 1. If space-time position 2 is on this light cone, then it will be struck by this radiation, disturbing its attempt to make a similar acceleration measurement. This is consistent with the commutator, which says that only field operators at points 1 and 2 on this light cone fail to commute.

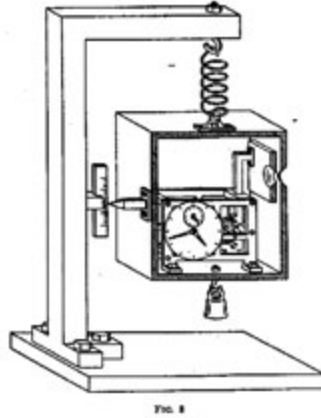


Fig. 4: An example of a very concrete concept of a thought experiment of Bohr. This is his figure for “weighing a photon”, constructed to resolve a paradox proposed by Einstein. Figure from <https://physics.stackexchange.com/questions/8295/einsteins-box-unclear-about-bohrs-retort>

5 Creating entanglement by measurements on the vacuum

I will end with some extremely paradoxical (on the face of it) observations about the quantum properties electromagnetic vacuum (discovered by [5]). They are not immediately about quantum measurement, but as with BR they use again pairs of “detectors” that interact with the field at only specific locations in space-time. The details of the detectors are very different, and the actual measurement step is considered as an explicitly separate second step. The detector is just a quantum system, situated at rest at position (x, y, z) , with two states, where an interaction with the field can be turned on and off in time:

$$\frac{\Omega}{2} \sigma_z + \epsilon(t)(\sigma^+ + \sigma^-) E_x(t, x, y, z) \quad (23)$$

The setup of “Entanglement from the Vacuum” [6] is shown in Fig. 5. In the inertial frame chosen, two probes are placed at different space positions 1 and 2 separated by distance L . The detectors are “turned on” simultaneously, meaning that $\epsilon(t)$ jumps from zero to a finite value. After a short time T , both probes are turned off, and they are henceforth decoupled from the electromagnetic field.

Resnik calculates the final state of the two probes right after the end of the interaction time T . He finds that the state has *entanglement*, meaning that the state of the two probes are correlated, in fact to a degree that is stronger than that which is possible in classical physics (this is the subject of “Bell inequality violation”, see NC, Sec. 2.6.). This means that subsequent measurement outcomes of the two probes are certainly correlated, even though they are uncorrelated before being turned on. By some quantification that I will not explain here, the amount of entanglement is at least $\exp(-(L/cT)^2)$.

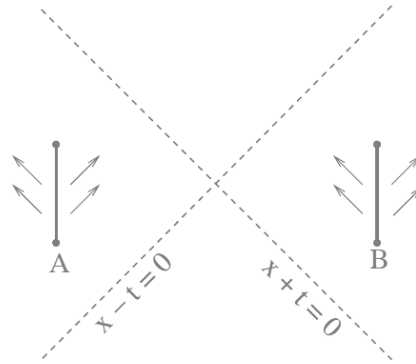


Fig. 5: Space-time diagram of Resnik's version of the Summer-Werner procedure for drawing entanglement non-causally out of the vacuum.

This seems to be in bald contradiction to a basic premise of quantum information theory, which is that LOCC (local operations and classical communication) cannot create entanglement (NC Sec. 12.5). Surely the action of turning on and off the probes is local, and the setup is arranged so that no classical signal can possibly pass from one to the other – they are outside each other's light cone.

The resolution of this paradox is that there is an unstated (or unemphasized) premise of LOCC, which is that there is no other Hilbert space involved other than that of the two qubits (or probes). But this is not the case for the quantum vacuum – it has a huge (one might say multiply infinite) Hilbert space. If one inquires where the entanglement comes from, one finds that it is pre-present in the quantum state of the vacuum. The very existence of the vacuum results in quantum correlations, even at a distance. Non-trivial correlations can be present because the quantum state of the vacuum is not an eigenstate of the electric field operator. Due to the commutator expression above, there does exist some quantum state of the electromagnetic field that is a simultaneous eigenstate of the electric field operators at the two probes. If the field were in this state, then no entanglement would result from the interaction of the probes with this field. But this is a state far from the vacuum state – it would have in fact a diverging value of the expectation value of the energy of the field, or, less precisely, it would be a state with many photons.

References

- [1] N. D. Mermin, *Quantum Computer Science: An Introduction* (Cambridge University Press) (2007).
- [2] “Die gegenwärtige Situation in der Quantenmechanik,” E. Schrödinger *Naturwissenschaften* **48**, pages 807-812, **49** 823-828, **50** 844-849 (1935) (see Sec. 5); translation by J. Trimmer, “The present situation in quantum mechanics”, available in *Quantum Theory and Measurement*, J.A. Wheeler, W.H. Zurek (eds.) Princeton University Press, 152-167 (1983).
- [3] J. Preskill lecture notes, “Measurement of Quantum Fields,” in <http://www.theory.caltech.edu/~preskill/ph205/205Chapter1-Page1-45.pdf>
- [4] N. Bohr and L. Rosenfeld, “On the Measurability of Electromagnetic Field Quantities,” in *Selected Papers of Leon Rosenfeld*, eds. Robert S. Cohen and John J. Stachel (Boston Studies in the Philosophy of Science, eds. Robert S. Cohen and Marx W. Wartofsky, Vol. XXI, D. Riedel) 357-400, (1979) [“Zur Frage der Messbarkeit der Elektromagnetischen Feldgrößen,” *Det Kgl. Danske Videnskabernes Selskab., Matematisk-fysiske Meddelelser* XII, 8 (1933)].
- [5] S. J. Summers and R. Werner, “The vacuum violates Bell’s inequalities,” *Phys. Lett. A* **110**, 257 (1985).
- [6] Benni Resnik, “Entanglement from the Vacuum”, *Found. Phys.* **33**, 167-176 (2003).
- [7] W. H. Zurek, *Quantum Darwinism* *Nat. Phys.* **5**, 181–8 (2009).
- [8] C. Jess Riedel and Wojciech H. Zurek, *Quantum Darwinism in an Everyday Environment: Huge Redundancy in Scattered Photons*, *Phys. Rev. Lett.* **105**, 020404 (2010).
- [9] C. Jess Riedel and Wojciech H. Zurek, *Redundant Information from Thermal Illumination: Quantum Darwinism in Scattered Photons*, *New J. Phys.* **13**, 073038 (2011).
- [10] H. J. Briegel; R. Raussendorf, “Persistent Entanglement in arrays of Interacting Particles”, *Physical Review Letters*. **86** (5): 910–3 (2001).
- [11] Jan Ivar Korsbakken, K. Birgitta Whaley, Jonathan Dubois, and J. Ignacio Cirac, *Measurement-based measure of the size of macroscopic quantum superpositions* *Phys. Rev. A* **75**, 042106 (2007).
- [12] H. Wiseman and G. Milburn, *Quantum Measurement and Control*, Cambridge: Cambridge University Press, doi:10.1017/CBO9780511813948 (2009).
- [13] Jay M. Gambetta, “Non-Markovian stochastic Schrödinger equations and interpretations of quantum mechanics”, PhD thesis, Griffiths University (2003).

- [14] Jay Gambetta, Alexandre Blais, M. Boissonneault, A. A. Houck, D. I. Schuster, and S. M. Girvin, Quantum trajectory approach to circuit QED: Quantum jumps and the Zeno effect *Phys. Rev. A* 77, 012112 (2008).
- [15] C. H. Bennett, Publicity, Privacy, and Permanence of Information AIP Conference Proceedings 864, 11 (2006), <https://doi.org/10.1063/1.2400875>, and private communication.

A 7 **Quantum Networks with Superconducting Circuits**

P. Kurpiers, P. Magnard, B. Royer, M. Pechal, S. Storz,
J.-C. Besse, S. Gasparinetti, J. Heinsoo, Y. Salathé, A. Akin,
A. Blais, A. Wallraff

Department of Physics

ETH Zürich, CH-8093 Zürich, Switzerland

Contents

1	Quantum communication with single photons	4
2	Deterministic Protocols	7
2.1	Deterministic Quantum State Transfer	7
2.2	Deterministic Generation of Remote Entanglement	8
3	Time-Bin Encoding Protocols	11
3.1	Time-Bin-Encoded Quantum State Transfer	12
3.2	Time-Bin-Encoded Generation of Remote Entanglement	15
4	Conclusion and Outlook	16

The development of computers and their interconnection via networks has fundamentally changed the way we process and transfer information. Similar to today's classical information processing, envisaged quantum networks can be categorized into infrastructures which distribute information over large distances via a quantum internet [1] and those which process information in local area networks (quantum LAN) [2]. Local quantum networks can be tailored for distributed quantum computation in which quantum states and entanglement are shared between multiple quantum processors [3]. As for monolithic quantum computers, these distributed quantum machines can provide exponential speed up compared to conventional computing approaches [4]. Individual quantum processors can consist of many coupled stationary, atomic or solid-state quantum systems, which store and process quantum information [5, 6]. Subroutines of the full quantum algorithm may run on separate processors [7–10], which are connected by classical and quantum channels [11]. Quantum channels used in local quantum networks are optimized for high-bandwidth quantum communication and enable error-correction across different nodes to protect quantum information against decoherence [12–16]. These requirements render direct quantum channels between two nodes an efficient way for scaling up distributed quantum computers [2, 17]. Direct quantum channels offer the possibility to transmit arbitrary quantum states directly and to generate entanglement deterministically between two nodes using single photons [2, 17, 18]. First algorithms for distributed quantum computing [19, 20] and efficient implementations of quantum gates between different nodes [21–23] have been developed. Experimentally, recent progress has been made to realize deterministic quantum communication in direct quantum channels between separate quantum systems with superconducting circuits [24–27]. Here, we present an experimental implementation of a basic quantum network with circuit quantum electrodynamics (circuit QED) systems [26, 28] using single itinerant microwave photons as information carriers [29, 30].

We first review experimental realizations of quantum communication protocols in various physical systems in the optical and microwave regime. Then we present experimental realizations of quantum communication protocols with single shaped microwave photons using a direct quantum channel. We discuss our experimental implementation based on a cavity assisted Raman transition $|f, 0\rangle \leftrightarrow |g, 1\rangle$ [31], and describe the calibration routines used to emit and absorb single photons with high probability. We discuss the direct transfer of qubit states and the generation of entanglement based on two quantum communication protocols, which use different mappings of our local quantum states to itinerant photons. First, we map quantum states, locally stored in qubits at a node of the network, to the Fock-state basis of itinerant photons and realize deterministic quantum communication between two separated quantum systems. Second, we show that photon loss and qubit decay can be detected in a protocol which maps quantum states to a time-bin superposition of a single photon.

Remote entanglement experiments have been performed in many different physical systems (Fig. 1) using optical photons with atomic ensembles [36–40, 60], trapped ions [41–45], single atoms - Bose Einstein condensates [46], vibrational states of diamonds [47], rare-earth doped crystals [48], single atoms [49, 50], nitrogen-vacancy centers [51–53], quantum dots [57, 58], or using microwave photons in superconducting circuits [24–28, 54–56]. In these experiments, schemes have been implemented which generate remote entanglement probabilistically based on joint measurements, which project the remote systems on an entangled state, based on the scattering of a single photon or two photons from two remote systems, which is/are interfered on a beam splitter and detected afterwards, and based on direct transfer of photons between the two systems. These probabilistic protocols can be implemented using

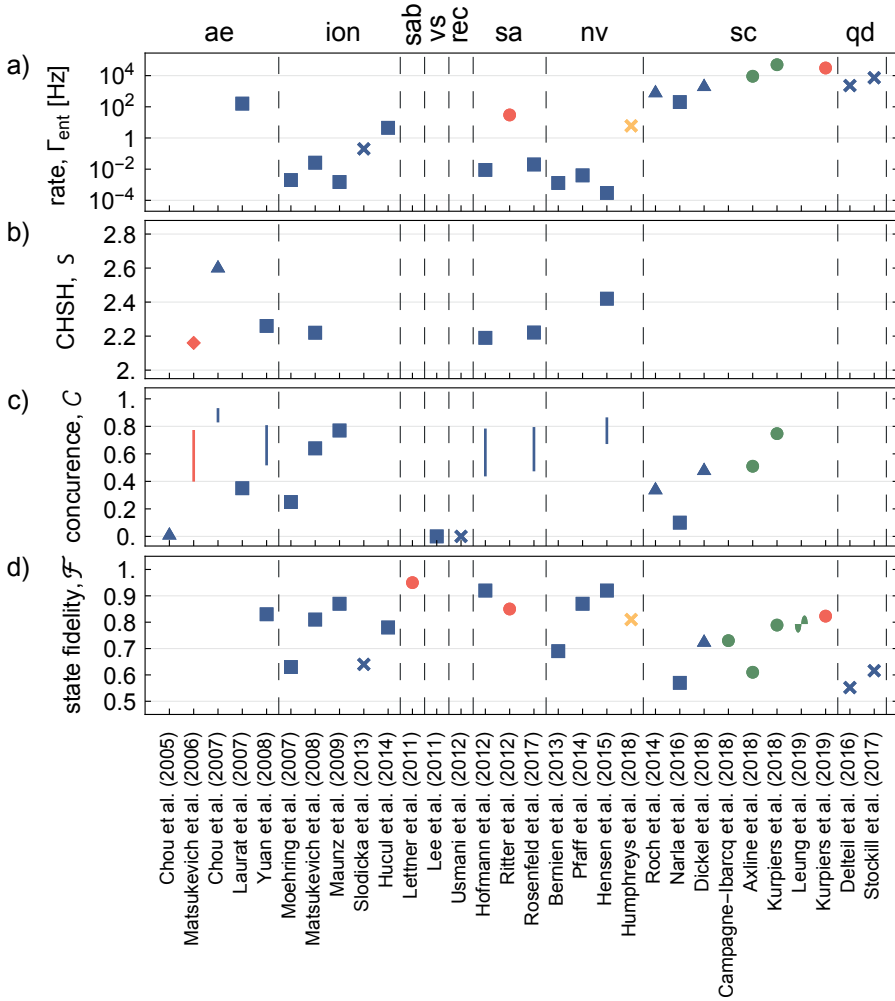


Fig. 1: Overview of performance metrics of remote entanglement experiments in different physical architectures. a) Entanglement generation rate Γ_{ent} , b) CHSH-value S [32], c) concurrence C [33,34] and d) entangled state fidelity \mathcal{F} [35]. The data is grouped by physical system: atomic ensembles (ae) [36–40], trapped ions (ion) [41–45], single atom Bose-Einstein-condensate (sab) [46], vibrational state of diamonds (vs) [47], rare-earth doped crystals (rec) [48], single atoms (sa) [49, 50], nitrogen-vacancy (nv) centers [51–53], superconducting circuits (sc) [24–28, 54–56] and quantum dots (qd) [57, 58]. The colors indicate different implementations: probabilistic unheralded (red), probabilistic heralded (blue), guaranteeing a deterministic delivery of an entangled state at a pre-specified time (yellow), and fully deterministic protocols (green). Symbols indicate different schemes to realize the remote interaction: measurement-induced (triangle), single- (cross) or two-photon (squares) interference and detection, direct transfer (diamond), direct transfer with shaped photons (circles), multi-mode characteristic of communication channel (two half ellipses). Lines in (c) are bounds [59] on C calculated from measured values of S .

heralding techniques, which allow one to re-transmit the quantum information in case an error is detected during the protocol. Heraldng protocols can also provide deterministic remote entanglement at predetermined times [61]. Fully deterministic quantum communication can be achieved by either using direct transfer of photons with a time-reversible temporal envelope [17] or by exploiting multi-mode characteristic of the communication channel [27]. An overview of figures of merit such as the entanglement generation rate Γ_{ent} , CHSH-value \mathcal{S} [32], concurrence \mathcal{C} [33, 34] and entangled state fidelity \mathcal{F} [35] of remote entanglement experiments is presented in Fig. 1.

1 Quantum communication with single photons

In our adaptation of the protocol proposed by Ignacio Cirac [17], each quantum node consists of a planar, chip-based, circuit QED architecture, in which a transmon qubit [62] is dispersively coupled to two coplanar microwave resonators [63], analogous to an atom coupled to two cavity modes (Fig. 2). One resonator is dedicated to dispersive transmon readout [64] and the second one to excitation transfer [30]. The transfer resonators of the two nodes are tuned during device fabrication and by choosing of qubit resonator detunings to have matching frequencies $\nu_{\text{T}} \approx 8.400$ GHz. All resonators are coupled to dedicated Purcell filters [65–67]. An external coaxial line of 0.9 m length, bisected with a circulator, connects the transfer circuits of both chips. With this setup, photons are routed from node A to B, and from node B to a detection line. As detailed in Ref. [30, 31], we generate a controllable light-matter interaction, by applying a coherent microwave tone to the transmon which induces an effective interaction $\tilde{g}_{\text{f0g1}}(t)$ with tunable amplitude and phase between states $|f, 0\rangle$ and $|g, 1\rangle$ (f0g1) [30, 31] with the vacuum state $|0\rangle$ and the single-photon Fock state $|1\rangle$. The two lowest energy eigenstates $|g\rangle, |e\rangle$ of the transmon form the qubit subspace and the second excited state $|f\rangle$ is used as an auxiliary level to control the matter-light interaction in our experiment. The f0g1-interaction swaps an excitation from the transmon to the transfer resonator, which then couples to a mode propagating towards node B. By controlling $\tilde{g}_{\text{f0g1}}(t)$, we shape the temporal mode of the itinerant photon to have a time-symmetric envelope. By inducing the reverse process $|g, 1\rangle \leftrightarrow |f, 0\rangle$ with the time reversed amplitude and phase profile of $\tilde{g}_{\text{f0g1}}(t)$, we absorb the itinerant photon in the transmon at node B. Ideally, this procedure returns all photonic modes to their vacuum state.

As a first step in the characterization of the excitation transfer protocol, we initialize the transmon at the receiving node B in state $|f\rangle$ with the aim to create a photon. This photon provides a reference signal to extract the transfer efficiency for photons between the sending and receiving nodes. Inducing the effective coupling $\tilde{g}_{\text{f0g1}}(t)$ with a f0g1 drive pulse truncated at time τ (R_{f0g1}^τ) we emit a symmetric photon (Fig. 3a). Here and in all following characterization measurements, we use active reset techniques to initialize both transmons in their ground states before the quantum communication protocols [68, 69] and extract the population of the transmon states using single-shot readout correcting for qubit readout errors (for details see [26]). The populations of the three lowest levels of the transmon P_g , P_e and P_f are measured immediately after truncating the emission pulse R_{f0g1}^τ at time τ (see Fig. 3b). We observe that the transmon smoothly evolves from $|f\rangle$ to $|g\rangle$ during the emission process. At the end of the protocol the emitting transmon reaches a ground state population of $P_g = 95.8\%$ limiting the photon emission efficiency.

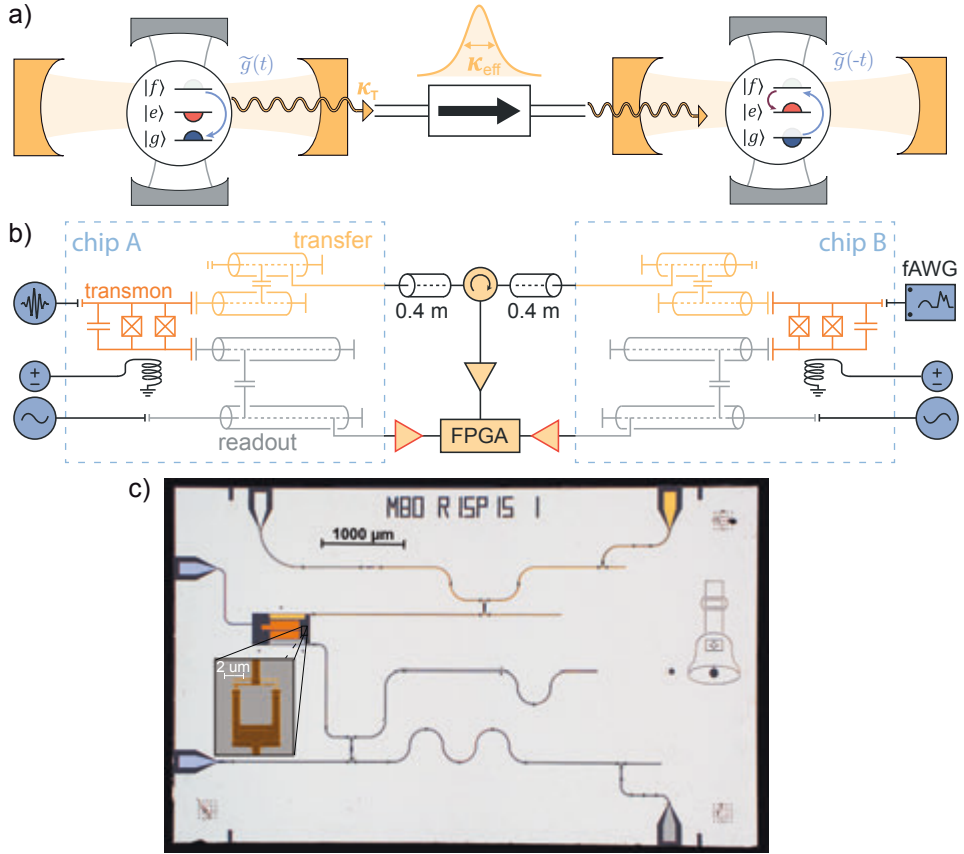


Fig. 2: a) Schematic of a setup for realizing a deterministic unidirectional entanglement protocol between two cavity QED nodes of a quantum network. At the first node, a three level system is prepared in its second excited state $|f\rangle$ (gray half-circle) and coherently driven ($\tilde{g}_{f|g1}(t)$, blue arrow) to $|g\rangle$ (blue half-circle) creating the transfer cavity field (light yellow). The cavity field couples into the directional quantum channel with rate κ_T as a single photon wavepacket with an effective bandwidth κ_{eff}^p (yellow hyperbolic secant shape). In the second quantum node, the time reversed drive $\tilde{g}_{f|g1}(-t)$ transfers the excitation from $|g\rangle$ to $|f\rangle$ in the presence of the transferred photon field $|1\rangle$. Finally, the protocol is completed with a transfer pulse between $|f\rangle$ and $|e\rangle$ (red half-circle) to return to the qubit subspace. Additionally, each three level system is coupled to a readout cavity (gray). b) Implementation of the schematic depicted in (a) in a planar, chip-based, circuit QED architecture. A directional quantum channel between the two nodes is realized using two semi-rigid coaxial cables and a circulator which connects the output ports of the transfer circuit Purcell filters. c) False-colored micrograph of a sample of the same design as the one used in the remote communication experiments. The inset shows a scanning electron microscope (SEM) micrograph of the asymmetric SQUID with a designed ratio of 1 : 5 between the areas of the Josephson junctions. The output transmission lines are galvanically coupled to the corresponding circuit. The input to the transfer circuit is an auxiliary line to perform resonator spectroscopy in transmission.

To verify that the emitted photon envelope has the targeted shape and bandwidth $\kappa_{p,B}^{\text{eff}}/2\pi = 10.6$ MHz, we repeat the emission protocol with an initial transmon state $(|g\rangle + |f\rangle)/\sqrt{2}$ and measure the averaged electric field amplitude $\langle \hat{a}_{\text{out}}(t) \rangle$ of the emitted photon state $(|0\rangle + |1\rangle)/\sqrt{2}$ (Fig. 3c) using heterodyne detection [70]. Instead of preparing the single photon Fock state $|1\rangle$, for this measurement, we prepare the $(|0\rangle + |1\rangle)/\sqrt{2}$ state to conveniently measure its non-vanishing average electric field amplitude [30].

Repeating the emission protocol at node A we observe similar dynamics of the transmon population (see Fig. 3e). We adjust the amplitude and phase of the transfer pulse (Fig. 3d) such that the photons emitted from node A have a bandwidth $\kappa_{p,A}^{\text{eff}}/2\pi = 10.4$ MHz similar to $\kappa_{p,B}^{\text{eff}}/2\pi$. The measured envelope of the photon emitted from node A (Fig. 3f) is slightly distorted by the reflection at node B as determined by the response function of its transfer resonator.

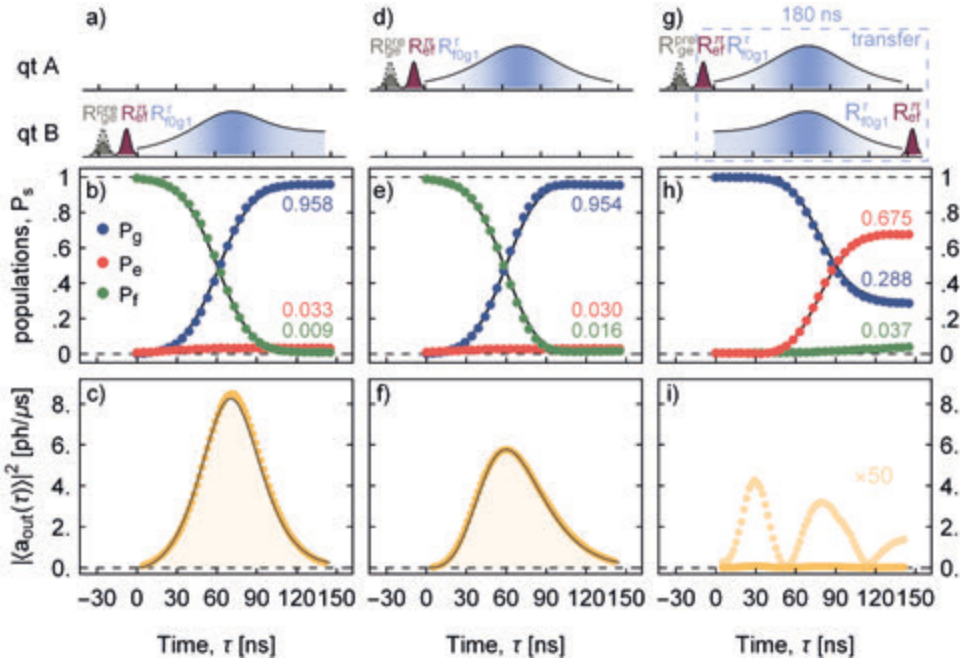


Fig. 3: The transmons at node B (a) and node A (d) are prepared in the state $|f\rangle$. We characterize (dots) the time dependence (τ) of the qutrit populations $P_{g,e,f}$ (b, e) while driving the $|f, 0\rangle$ to $|g, 1\rangle$ transition (f0g1). The phase (white-blue shading) of the f0g1 drive is modulated to compensate the drive-induced quadratic ac Stark shift. The mean field amplitude squared $|\langle \hat{a}_{\text{out}}(\tau) \rangle|^2$ of the traveling photons emitted from node B (c) and node A (f) is obtained for the emitted photon state $(|0\rangle + |1\rangle)/\sqrt{2}$. The effective photon bandwidths are adjusted to be $\kappa_{p,A}^{\text{eff}}/2\pi = 10.4$ MHz and $\kappa_{p,B}^{\text{eff}}/2\pi = 10.6$ MHz. The solid lines in (b, c, e, f, h) are results of master equation simulations. The time dependence of P_s when executing the excitation transfer protocol (pulse shapes in dashed rectangle in g) from qubit A to qubit B (h) are extracted simultaneously with the amplitude of the emitted field from node A. i) shows the residual $|\langle \hat{a}_{\text{out}}(\tau) \rangle|^2$ (light yellow x50) during the absorption process.

We evaluate the ratio of the integrated power of the detected fields $\int |\langle \hat{a}_{\text{out}}^{\text{A}}(t) \rangle|^2 dt / \int |\langle \hat{a}_{\text{out}}^{\text{B}}(t) \rangle|^2 dt$ to extract the photon loss $l_{\text{AB}} = 23.0 \pm 0.5\%$ between node A and B. This is possible, since we detect the field of each of the emitted photons in the same detection line by making use of the circulator between the two nodes (Fig. 2b). The path traveled by the two emitted photons towards the detector differs only by the length of the waveguide separating the two samples from each other. In addition, we estimate the photon loss between nodes A and B based on the electrical specifications of the individual circuit elements connecting the nodes: two printed circuit boards including connectors (each $2.5 \pm 1\%$), two coaxial cables of length 0.4 m (each $4.0 \pm 0.1\%$) [71] and a microwave circulator ($13 \pm 2\%$ according to manufacturer). With these parameters we estimate an accumulated photon loss between nodes A and B of $24 \pm 3\%$, which is in good agreement with the measured value of $l_{\text{AB}} = 23.0 \pm 0.5\%$.

To characterize the absorption of a single time-symmetric photon emitted from node A at the receiving node by time-reversing the emission pulse of node B (Fig. 3a and g), we measure the population of transmon B by interrupting the transfer process at time τ . For this characterization measurement, we apply a π -pulse to transmon B to map the receiving qubit $|f\rangle$ state back to the qubit subspace before performing qubit readout. We observe the population of $|e\rangle$ to smoothly rise and saturate at $P_{\text{e}}^{\text{sat}} = 67.5\%$ (Fig. 3h). This saturation level reflects the efficiency of the protocol for the transfer of a single photon, which is executed in a pulse sequence of 180 ns duration (Fig. 3g). From the ratio of the integrated power of the emitted photon in the absence (Fig. 3i) or presence (Fig. 3f) of the absorption pulse, we determine the absorption efficiency to reach $98.1 \pm 0.1\%$. We estimate that the finite truncation of the f0g1-pulses contributes with $\sim 1.1\%$ to the inefficiency and leakage to the second, weakly-coupled input port of the Purcell filter with $\sim 0.8\%$ (Fig. 2c).

The results of master equation simulations of the excitation transfer (solid lines in Fig. 3), using parameters extracted from independent measurements, display excellent agreement with the measured data. This demonstrates a high level of control over the emission and absorption processes and an accurate understanding of the experimental imperfections dominated by qutrit decoherence and photon loss.

2 Deterministic Protocols

2.1 Deterministic Quantum State Transfer

We use the techniques presented above to deterministically transfer an arbitrary qubit state via a direct quantum channel from node A to node B at a rate of 50 kHz. We prepare the receiving transmon B in state $|g\rangle$, apply a R_{ef}^{π} pulse to the transmitting transmon A, followed by the emission/absorption pulse and finally a rotation R_{ef}^{π} on the receiving transmon B. We characterize the quantum state transfer by reconstructing its process matrix χ_{det} with quantum process tomography [72] (Fig. 4b,c). For that purpose, we prepare each of the six mutually unbiased qubit basis-states $|g\rangle$, $|e\rangle$, $(|g\rangle + |e\rangle)/\sqrt{2}$, $(|g\rangle + i|e\rangle)/\sqrt{2}$, $(|g\rangle - |e\rangle)/\sqrt{2}$, $(|g\rangle - i|e\rangle)/\sqrt{2}$ [73] using the rotations ${}^{\text{x}}R_{\text{ge}}^0$, ${}^{\text{x}}R_{\text{ge}}^{\pi/2}$, ${}^{\text{x}}R_{\text{ge}}^{-\pi/2}$, ${}^{\text{y}}R_{\text{ge}}^{\pi/2}$, ${}^{\text{y}}R_{\text{ge}}^{-\pi/2}$ and ${}^{\text{x}}R_{\text{ge}}^{\pi}$ at node A, transfer the corresponding states to node B, and reconstruct the transferred state using quantum state tomography (QST). QST of a single qutrit [74] is performed by measuring the qutrit state population with single-shot readout, after applying the respective tomography gates: ${}^{\text{x}}R_{\text{ge}}^0$, ${}^{\text{x}}R_{\text{ge}}^{\pi/2}$, ${}^{\text{y}}R_{\text{ge}}^{\pi/2}$, ${}^{\text{x}}R_{\text{ge}}^{\pi}$.

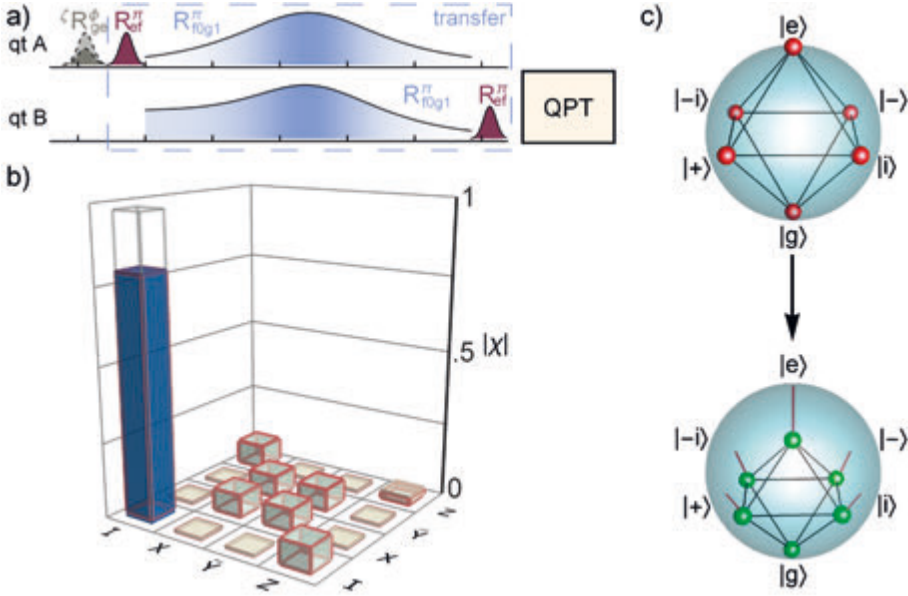


Fig. 4: a) Pulse scheme used to characterize the qubit state transfer between the two nodes. We prepare six mutually unbiased input states at node A (denoted by ${}^{\zeta}R_{\text{ge}}^{\phi}$). b) Experimentally obtain process matrix (absolute value shown as colored bars). The gray and red wire frames show the ideal and the master equation simulation of the absolute values of the process matrix, respectively. c) Bloch sphere representation of the quantum states transfer of the six ideal input states at node A (shown as red dots on the upper sphere) and the corresponding, experimentally obtained output states at node B (shown as green dots on the lower sphere).

$\times R_{\text{ef}}^{\pi/2}$, $y R_{\text{ef}}^{\pi/2}$, $\times R_{\text{ge}}^{\pi} \times R_{\text{ef}}^{\pi/2}$, $\times R_{\text{ge}}^{\pi} y R_{\text{ef}}^{\pi/2}$ and $\times R_{\text{ge}}^{\pi} \times R_{\text{ef}}^{\pi}$. The elements of the density matrix are then reconstructed with a maximum-likelihood method [75]. We obtain the process matrix through linear inversion, from these density matrices. Performing QST of the qutrit subspace is required to characterize residual population in $|f\rangle$ after the qubit state transfer, which is mainly caused by energy relaxation from the $|f\rangle$ level in combination with the single qubit rotation R_{ef}^{π} swapping $|e\rangle$ with $|f\rangle$ populations. The obtained density matrices have a non-unit trace in the qubit subspace and so does the qubit state transfer process matrix. This reduction method gives a conservative estimate of the obtained fidelities and are discussed in more detail in Section 2.2. We determine a process fidelity of $\mathcal{F}_{\text{p}}^{\text{det}} = \text{tr}(\chi_{\text{det}} \chi_{\text{ideal}}) = 80.02 \pm 0.07\%$. The process matrix χ_{sim} calculated using master equation simulations agrees very well with the data (absolute values shown as red wire frames in Fig. 4b) showing the high level of experimental control. This is supported by the small trace distance [35] $\text{tr} |\chi_{\text{det}} - \chi_{\text{sim}}| / 2 = 0.015$, which ideally is 0 for identical process matrices and 1 for orthogonal ones.

2.2 Deterministic Generation of Remote Entanglement

Using a modified excitation transfer protocol we deterministically generate two-qubit remote entangled states between nodes A and B. The protocol starts by preparing transmon A in state

$(|e\rangle + |f\rangle)/\sqrt{2}$ and B in state $|g\rangle$, and applying the emission/absorption pulses followed by a rotation R_{ef}^{π} on transmon B to generate the Bell state $|\psi^+\rangle = (|e, g\rangle + |g, e\rangle)/\sqrt{2}$ (Fig. 5a). Alternatively, a remote entangled state can be generated by preparing the transmon at node A in $|f, 0\rangle$, swapping half of the population to $|g, 1\rangle$ using $R_{f0g1}^{\pi/2}$ and applying the same $|g, 1\rangle \leftrightarrow |f, 0\rangle$ absorption pulse at node B. The $R_{f0g1}^{\pi/2}$ can be used to decrease the emission time, however, the

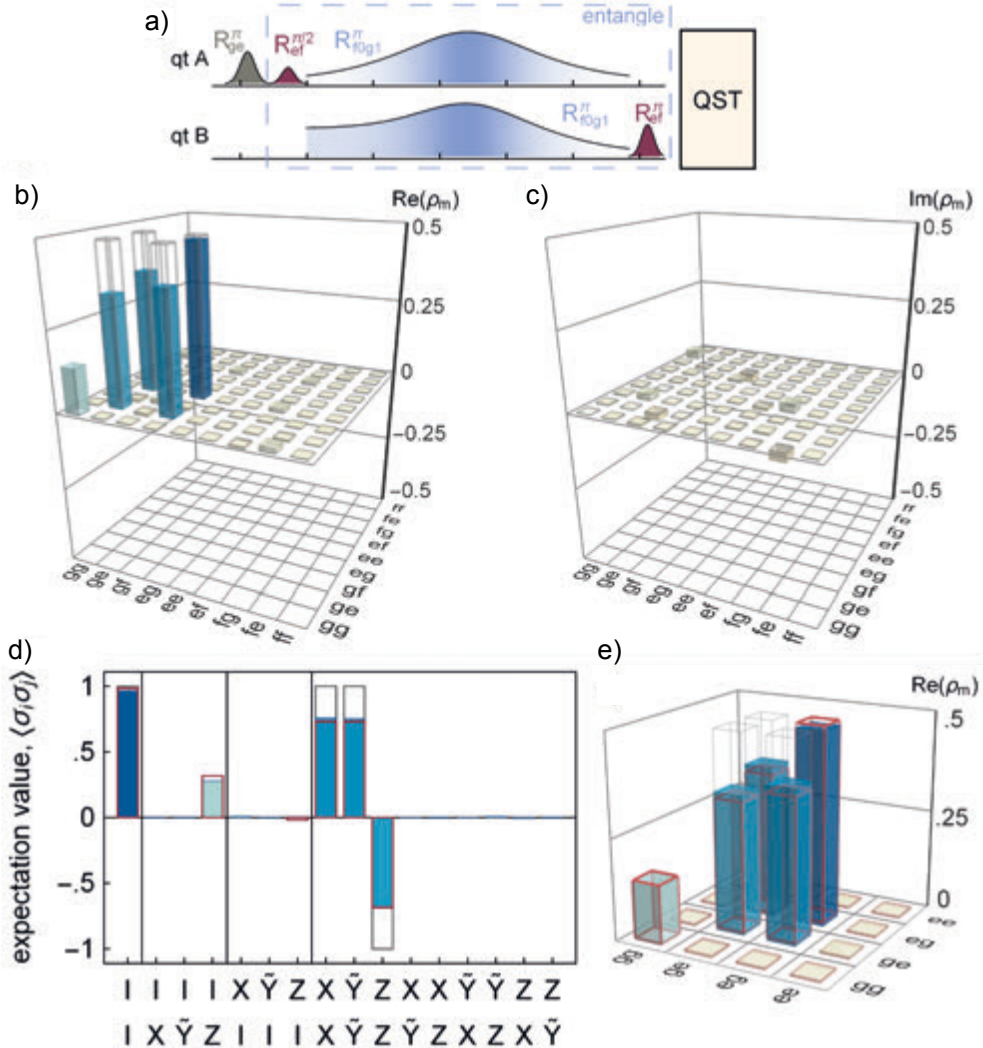


Fig. 5: a) Pulse scheme to generate deterministic remote entanglement between nodes A and B. b) Real and c) imaginary part of the generated density matrix. The ideal Bell state $|\psi^+\rangle$ is depicted with gray wire frames. d) Expectation values of two-qubit Pauli operators and e) reconstructed density matrix after execution of the remote entanglement protocol. The colored bars indicate the measurement results, the ideal expectation values for the Bell state $|\psi^+\rangle = (|e, g\rangle + |g, e\rangle)/\sqrt{2}$ are shown in gray wire frames and the results of a master equation simulation in red.

absorption process requires the same time. Thus, the full alternative protocol is not shorter than our implementation.

Since leakage to the $|f\rangle$ level at both nodes leads to errors in the two-qubit density matrix reconstruction, we extract the full two-qutrit density matrix $\hat{\rho}_{3\otimes 3}^{\text{det}}$ from QST experiments (Fig. 5b,c). To do so, we apply two local tomography gates on transmons A and B, from the 81 pairs of gates that can be formed from the single-qutrit QST gates, and extract the state populations using single shot measurements of the two qutrits. We observe a total residual population of 3.5% of the $|f\rangle$ level of both transmons after the execution of the entanglement protocol showing that the entangled state can not be described rigorously by a two-qubit density matrix. We verify the three-level bipartite entanglement by using the computable cross norm or realignment (CCNR) criterion [76, 77], which states that a state must be entangled if $\text{ccnr} > 1$. We obtain $\text{ccnr} = 1.612 \pm 0.003$ with the measured entangled state $\hat{\rho}_{3\otimes 3}^{\text{det}}$, witnessing unambiguously the entanglement of the two qutrits.

For illustration purposes, we display the two-qubit density matrix $\hat{\rho}_m^{\text{det}}$ (Fig. 5d,e), as the two-qubit elements of $\hat{\rho}_{3\otimes 3}^{\text{det}}$. This choice of reduction from a two-qutrit to a two-qubit density matrix conserves the state fidelity $\mathcal{F}_s^{\text{det}} = \langle \psi^+ | \hat{\rho}_m^{\text{det}} | \psi^+ \rangle = \langle \psi^+ | \hat{\rho}_{3\otimes 3}^{\text{det}} | \psi^+ \rangle$ [35], but $\hat{\rho}_m^{\text{det}}$ has a non-unit trace. In addition, this reduction method gives a conservative estimate of the concurrence $\mathcal{C}(\hat{\rho}_m)$, compared to a projection of $\hat{\rho}_{3\otimes 3}^{\text{det}}$ on the set of physical two-qubit density matrices. Making this choice, we find a state fidelity $\mathcal{F}_s^{\text{det}} = \langle \psi^+ | \hat{\rho}_m^{\text{det}} | \psi^+ \rangle = 78.9 \pm 0.1\%$ relative to the ideal Bell state, and a concurrence $\mathcal{C}(\hat{\rho}_m^{\text{det}}) = 0.747 \pm 0.004$.

The density matrix $\hat{\rho}_{\text{sim}}$ and equivalently the Pauli sets obtained from the master equation simulations of the entanglement protocol (red wireframe in Fig. 5d,e) are in excellent agreement with the experimental results, displaying a small trace distance $\text{tr} |\hat{\rho}_m^{\text{det}} - \hat{\rho}_{\text{sim}}| / 2 = 0.024$. We decompose the infidelity into approximately 10.5% photon loss, 9.5% finite transmon coherence, and 1% photon absorption inefficiency.

3 Time-Bin Encoding Protocols

We extend the deterministic quantum state transfer scheme to a protocol which encodes quantum information as a time-bin superposition of a single photon, making it suitable for heralded quantum communication in a direct quantum channel. The time-bin based transfer protocol relies on encoding the transmitted quantum information in a suitably chosen subspace S such that any error, which may be encountered during transmission, causes the system to leave this subspace. On the receiving end, a measurement which determines whether the system is in S but does not distinguish between individual states within S , can be used to detect if an error occurred. Crucially, when the transfer is successful, this protocol does not disturb the transmitted quantum information. The single photon Fock-state encoding used in Section 2.1 is not suitable to detect errors due to photon loss because the error does not cause a transition out of the code subspace $\{|0\rangle, |1\rangle\}$.

As in the deterministic case, the stationary quantum nodes are transmon qubits coupled to coplanar waveguide resonators. The two lowest energy eigenstates of the transmon, $|g\rangle$ and $|e\rangle$, form the qubit subspace and the second excited state, $|f\rangle$, is used to detect potential errors. A variant of the protocol also allows for remote, heralded entanglement generation (see Section 3.2).

The process for transferring quantum information stored in the transmon qubit into a time-bin superposition state consists of the steps illustrated in Fig. 6a: The transmon qubit at node A is initially prepared in a superposition of its ground and first excited state, $\alpha|g\rangle + \beta|e\rangle$, and the resonator in its vacuum state $|0\rangle$. Next, two pulses are applied to transform this superposition into $\alpha|e\rangle + \beta|f\rangle$. Then, a f0g1-pulse induces the transition from $|f, 0\rangle$ to $|g, 1\rangle$, which is followed by spontaneous emission of a photon from the resonator. The temporal mode profile of the f0g1-pulse is adjusted such that the photon is time-symmetric. After this first step, the system is in the state $\alpha|e, 0\rangle \otimes |0\rangle + \beta|g, 0\rangle \otimes |1_a\rangle$, where $|0\rangle$ and $|1_a\rangle$ denote the vacuum state of the waveguide and the single-photon state in the time-bin mode a . Next, the population of state $|e\rangle$ is swapped into $|f\rangle$ and the photon emission process is repeated, this time to create a single photon in a time-bin mode b . The resulting state of the system is $|g, 0\rangle \otimes (\alpha|1_b\rangle + \beta|1_a\rangle)$.

By reversing both f0g1-pulses in the time-bin encoding scheme (Fig. 6b) an incoming single photon in the time-bin superposition state $\alpha|1_b\rangle + \beta|1_a\rangle$ will cause the receiving transmon-resonator system, initialized in $|g, 0\rangle$, to be driven to the state $\alpha|e, 0\rangle + \beta|g, 0\rangle$ as the photon is absorbed. Thus, this protocol transfers the qubit state encoded as a superposition of $|g\rangle$ and $|e\rangle$ from transmon A to transmon B. In short, the sequence reads

$$\begin{aligned}
 & (\alpha|g\rangle_A + \beta|e\rangle_A) \otimes |g\rangle_B \rightarrow \\
 & |g\rangle_A \otimes (\alpha|1_b\rangle + \beta|1_a\rangle) \otimes |g\rangle_B \rightarrow \\
 & |g\rangle_A \otimes (\alpha|g\rangle_B + \beta|e\rangle_B)
 \end{aligned} \tag{1}$$

where we have omitted the states of the resonators and the propagating field whenever they are in their respective vacuum states.

An important property of this time-bin encoded transfer protocol is its ability to detect photon loss in the communication channel. Indeed, if a photon is lost or not absorbed by the receiver, node B receives a vacuum state at its input instead of the desired single-photon state. This implies that both absorption pulse sequences leave transmon B in its ground state $|g\rangle$ which is

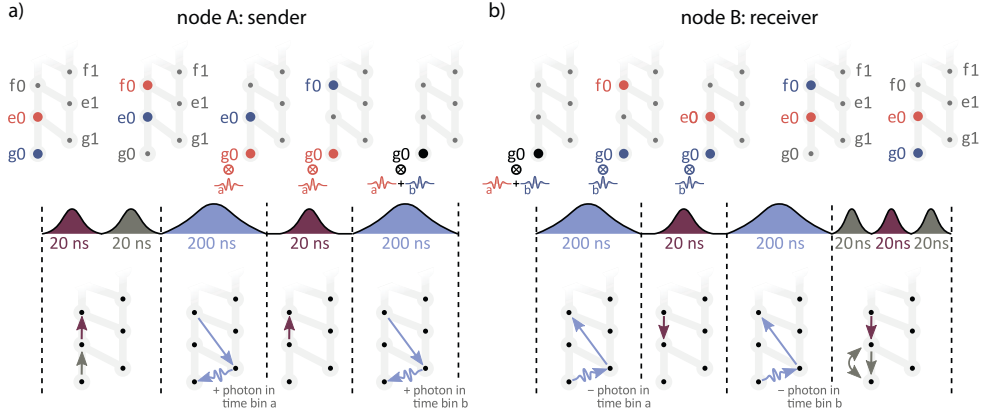


Fig. 6: Schematic representation of the pulse sequence implementing the time-bin encoded transmission process at the sender (a) and the receiver (b). In the level diagrams, the gray vertical lines represent Hamiltonian matrix elements due to the microwave drive between transmon levels while the diagonal ones show the transmon-resonator coupling. a) The qubit state is initially stored as a superposition of $|g, 0\rangle$ (blue dot) and $|e, 0\rangle$ (red dot). The first two pulses map this state to a superposition of $|e, 0\rangle$, $|f, 0\rangle$ and the third pulse transfers $|f, 0\rangle$ into $|g, 0\rangle$ while emitting a photon in time bin a (see red symbol for photon in mode a). Then, after $|e, 0\rangle$ is swapped to $|f, 0\rangle$, the last pulse again transfers $|f, 0\rangle$ into $|g, 0\rangle$ and emits a photon in time bin b . b) Reversing the protocol we reabsorb the photon, mapping the time-bin superposition back onto a superposition of transmon states.

subsequently mapped into $|f\rangle$ by the final three pulses. The successful quantum state transfer can thus be heralded by performing a quantum non-demolition readout of the transmon which distinguishes between $|f\rangle$ and the subspace spanned by $|g\rangle$, $|e\rangle$, but does not measure within this subspace. Such a binary measurement of a qutrit state can, for example, be realized by suppressing the measurement-induced dephasing in the (g,e) -subspace using parametric amplification and feedback [78] or by engineering the dispersive shifts of two transmon states on the readout resonator to be equal [79]. Furthermore, the $|f\rangle$ state population can be swapped to an ancilla qubit B_a which can be read out independently of transmon B in single-shot [80]. The swap operation can be performed using flux pulses [81] or an all-microwave gate driving the $|f_B, g_{B_a}\rangle$ into $|g_B, e_{B_a}\rangle$ (fgge) transition. In addition, the time-bin encoding protocol also detects failures of the state transfer due to energy relaxation of the transmon qubits during the protocol, e.g. if no photon is emitted from A due to energy relaxation to $|g\rangle$ before the first time bin is populated.

3.1 Time-Bin-Encoded Quantum State Transfer

We implement the time-bin encoding protocol in a setup identical to the one used for deterministic quantum communication (Fig. 2), up to an exchange of the cryogenic coaxial circulator in the connection between the two samples with a rectangular waveguide isolator which affects the bandwidth of the transfer resonators due to its different impedance ($\kappa_T^A = 7.4$ MHz, $\kappa_T^B = 12.6$ MHz).

To characterize the time-bin encoding protocol, we perform qutrit single-shot readout and use quantum process tomography. We initialize both transmons in their ground states [68, 69] and subsequently prepare the qubit at node A in one of the six mutually unbiased qubit basis states (Fig. 4b). We then run the time-bin encoded emission and reabsorption protocol (Fig. 6 and Fig. 7a) and implement quantum state tomography at node B for all six input states (Section 2.1). Directly after the tomography pulses, we read out the $|g\rangle$, $|e\rangle$ and $|f\rangle$ states of transmon B in single-shot mode. Based on these single-shot measurements, we postselect experimental runs in

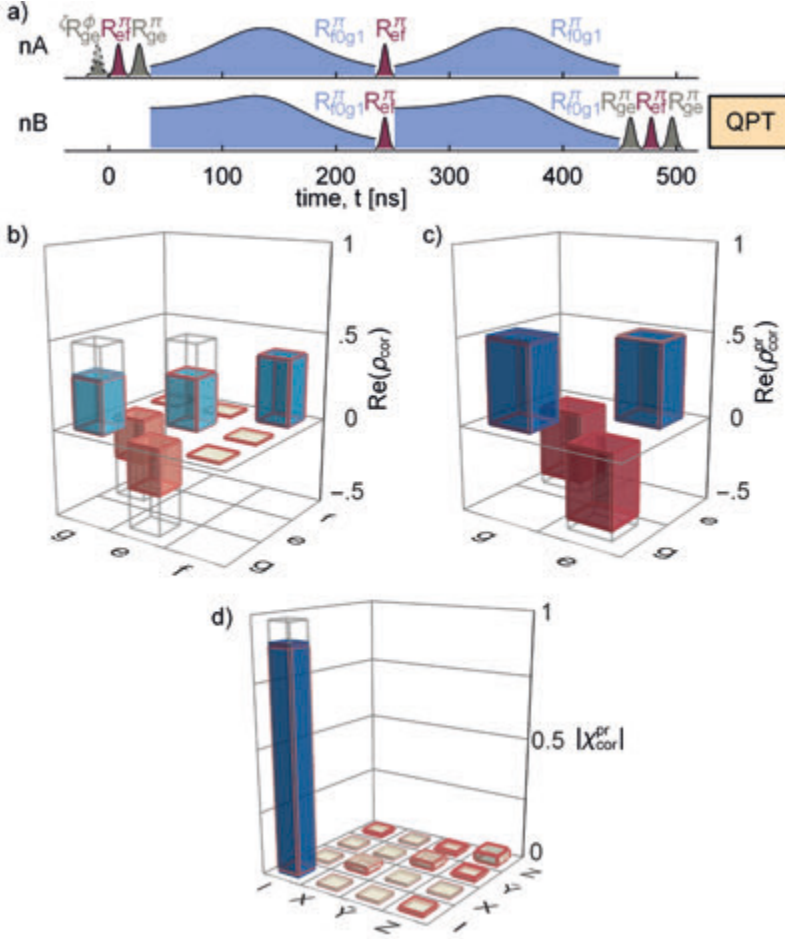


Fig. 7: a) Pulse scheme for characterizing the time-bin encoding protocol using quantum process tomography (QPT). ζR_{ij}^ϕ label DRAG microwave pulses. b) Real part of the qutrit density matrix $\hat{\rho}_{\text{cor}}$ for the input state $|-\rangle = (|g\rangle - |e\rangle)/\sqrt{2}$ after the state transfer protocol reconstructed using measurement-error correction. The magnitude of each imaginary part is < 0.017 . c) Projection of $\hat{\rho}_{\text{cor}}$ (b) onto the (g,e)-subspace $\hat{\rho}_{\text{cor}}^{\text{pe}}$ performed numerically which we use to reconstruct the process matrix $\chi_{\text{cor}}^{\text{pe}}$ (absolute value shown in (d)). The colored bars show the measurement results, the gray wire frames the ideal density or process matrix, respectively. The results of numerical master equation simulations are depicted as red wire frames.

which transmon B was measured in the computational space keeping on average $P_{\text{suc}}^{\text{qst}} = 64.6\%$ of the data, and transferring qubit states at a rate $\Gamma_{\text{qst}}/2\pi = P_{\text{suc}}^{\text{qst}}\Gamma_{\text{exp}}/2\pi \approx 32.3 \text{ kHz}$ based in a bare experimental repetition rate of $\Gamma_{\text{exp}} = 50 \text{ kHz}$. Using the post-selected data, we reconstruct the normalized density matrices $\hat{\rho}_{\text{ps}}$ of the qubit output state at node B using a maximum-likelihood approach without correcting for readout errors. In this post-selected approach, we obtain the process matrix χ_{ps} of the quantum state transfer and compute an averaged state fidelity of $\mathcal{F}_{\text{s}}^{\text{ps}} = \text{avg}(\langle \psi_{\text{in}} | \hat{\rho}_{\text{ps}} | \psi_{\text{in}} \rangle) = 88.2 \pm 0.2\%$ and a process fidelity of $\mathcal{F}_{\text{p}}^{\text{ps}} = \text{tr}(\chi_{\text{ps}} \chi_{\text{ideal}}) = 82.3 \pm 0.2\%$ relative to the ideal input states $|\psi_{\text{in}}\rangle$ and the ideal identity process, respectively.

Detecting errors to herald successful transfers does require a measurement which discriminates between $|f\rangle$ and the subspace spanned by $|g\rangle, |e\rangle$, but does not measure within this subspace, as described in Section 3. To benchmark only the encoding part of the time-bin encoding protocol we assume an ideal heralding measurement. To do so, we reconstruct all six qutrit density matrices $\hat{\rho}_{\text{cor}}$ of the output state at node B using the same dataset as for post-selection but correct for measurement errors in the qutrit subspace. These qutrit density matrices have 39.1% average population of level $|f\rangle$ indicating the detection of errors after the time-bin encoding protocol, which is compatible with $1 - P_{\text{suc}}^{\text{qst}}$ of the post-selected analysis (Fig. 7b). Next, we project these density matrices numerically onto the qubit (g,e)-subspace $\hat{\rho}_{\text{cor}}^{\text{pr}}$ (Fig. 7c), simulating an ideal error detection and reconstruct the process matrix $\chi_{\text{cor}}^{\text{pr}}$ of the quantum state transfer (Fig. 7d). In this way, we find an average state fidelity of $\mathcal{F}_{\text{s}}^{\text{cor}} = \text{avg}(\langle \psi_{\text{in}} | \hat{\rho}_{\text{cor}}^{\text{pr}} | \psi_{\text{in}} \rangle) = 93.5 \pm 0.1\%$ and a process fidelity of $\mathcal{F}_{\text{p}}^{\text{cor}} = \text{tr}(\chi_{\text{cor}}^{\text{pr}} \chi_{\text{ideal}}) = 90.3 \pm 0.2\%$ based on these measurement-error-corrected density matrices. This analysis allows us also to compare the time-bin encoding protocol directly to the deterministic quantum state transfer protocol (Section 2.1), in which we obtained $\mathcal{F}_{\text{p}}^{\text{det}} = 80.02 \pm 0.07\%$. This comparison clearly displays the capability of time-bin encoding to reduce the effect of photon loss, when assuming perfect readout. In addition, we demonstrated with the post-selected experiments that $F_{\text{p}}^{\text{ps}} > F_{\text{p}}^{\text{det}}$ at our current readout fidelities. Such direct comparison of fidelities, however, should be done with caution because it depends on the loss of the quantum channel. We expect the fidelity of the deterministic protocol to decrease linearly with loss, while the fidelity of the time-bin encoding protocol remains constant, with a linear decrease of its success probability. The time-bin encoding protocol thus complements deterministic protocols to perform heralded quantum communication in direct quantum channels.

In addition, we analyze the sources of infidelity by performing numerical master equation simulations of the time-bin encoding protocol which we compare to the measurement-error corrected density and process matrices. We find excellent agreement with the experimental results, indicated by a small trace distance $\text{tr}|\chi_{\text{cor}}^{\text{pr}} - \chi_{\text{sim}}|/2 = 0.03$. The master-equation-simulation results indicate that approximately 5.5% of the infidelity can be attributed to $|f\rangle \rightarrow |e\rangle$ and $|e\rangle \rightarrow |g\rangle$ relaxation at both transmons during the protocol. Pure qutrit dephasing explains the remaining infidelity.

3.2 Time-Bin-Encoded Generation of Remote Entanglement

As in the deterministic case, we use a modification of the state-transfer protocol to generate entanglement between the distant nodes (Fig. 8a). Both transmon-resonator systems at the

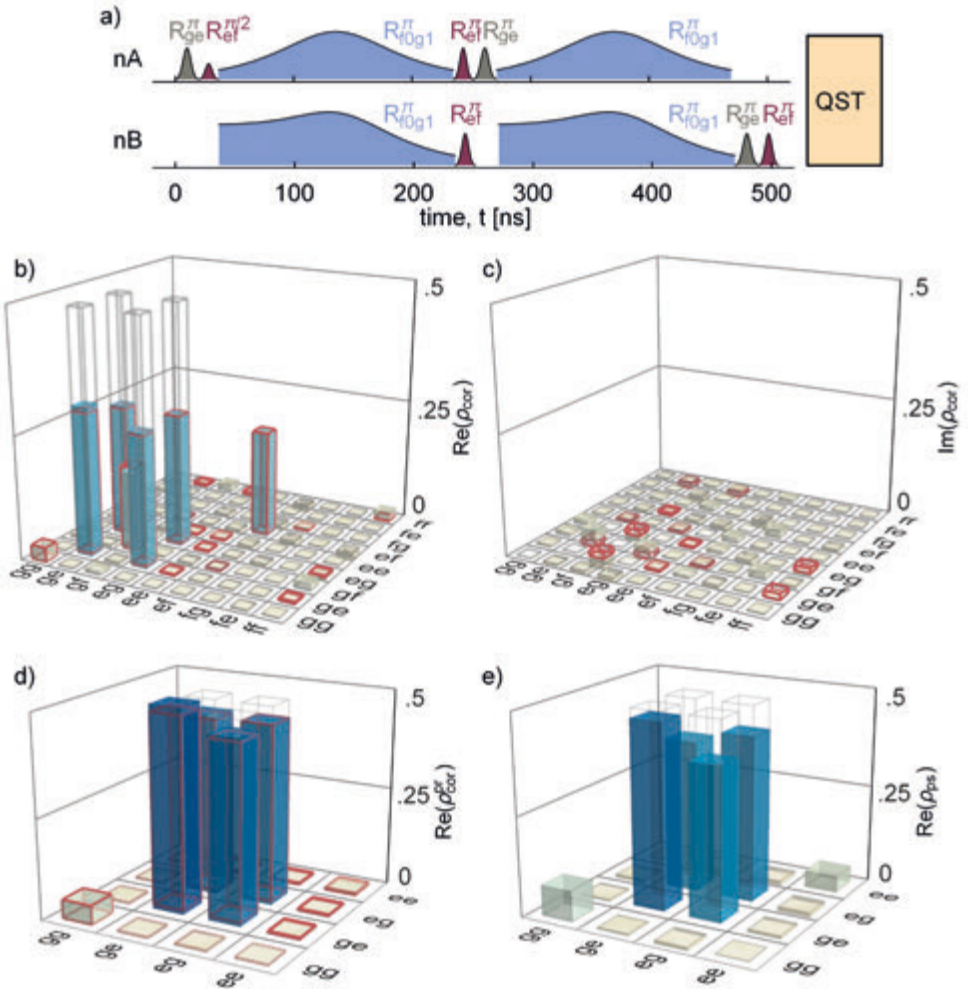


Fig. 8: a) Pulse scheme for generating remote entanglement between nodes A and B (see text for details). Real (b) and imaginary (c) part of the two-qutrit density matrix $\hat{\rho}_{\text{cor}}$ reconstructed after execution of the time-bin remote entanglement protocol correcting for measurement-errors as in the deterministic case. The colored bars indicate the measurement results. The ideal expectation values for the Bell state $|\psi^+\rangle = (|g_A, e_B\rangle + |e_A, g_B\rangle)/\sqrt{2}$ are shown as gray wire frames and the results of a master equation simulation as red wire frames. d) Numerical projection of $\hat{\rho}_{\text{cor}}$ onto the (g,e)-subspace to obtain the two-qubit density matrix $\hat{\rho}_{\text{cor}}^{\text{PR}}$ which is in excellent agreement with our master equation simulations (trace distance of 0.028). e) Post-selected two-qubit density matrix $\hat{\rho}_{\text{ps}}$ without correcting for measurement-errors.

network nodes are first initialized in their ground states. The first two pulses of the remote-entanglement protocol prepare transmon A in an equal superposition state $1/\sqrt{2}(|e\rangle + |f\rangle)$, followed by a pulse sequence which entangles the transmon state $|g\rangle$ and $|e\rangle$ with the time-bin qubit and maps the state of the time-bin qubit to transmon qubit B. This process can be summarized as

$$\begin{aligned} & \frac{1}{\sqrt{2}}(|e\rangle_A + |f\rangle_A) \otimes |g\rangle_B \rightarrow \\ & \frac{1}{\sqrt{2}}(|g\rangle_A \otimes |1_a\rangle + |e\rangle_A \otimes |1_b\rangle) \otimes |g\rangle_B \rightarrow \\ & \frac{1}{\sqrt{2}}(|g\rangle_A \otimes |e\rangle_B + |e\rangle_A \otimes |g\rangle_B). \end{aligned} \quad (2)$$

In case of an error, transmon B ends up in state $|f\rangle_B$.

We execute the entanglement-generation protocol by post-selecting only experimental runs in which neither qutrit is found in the $|f\rangle$ state using individual single-shot readout of both transmon qubits. Under this condition, we retain $P_{\text{suc}}^{\text{ent}} \approx 61.5\%$ of the data and obtain a Bell state fidelity of $\mathcal{F}_s^{\text{ps}} = \langle \psi^+ | \hat{\rho}_{\text{ps}} | \psi^+ \rangle = 82.3 \pm 0.4\%$ compared to an ideal $|\psi^+\rangle$ Bell state (Fig. 8e). In these post-selected experiments, we generate entangled states at rate $\Gamma_{\text{ent}}/2\pi = P_{\text{suc}}^{\text{ent}}\Gamma_{\text{exp}}/2\pi \approx 30.8$ kHz. To benchmark the time-bin encoded entanglement protocol, we use full two-qutrit state tomography of the transmons in which we correct for measurement errors with the same data set as for post-selection. The reconstructed density matrix (Fig. 8b,c) displays a high population of the $|g_A, f_B\rangle$, $|e_A, f_B\rangle$ states, $P_{gf} = 16.0\%$ and $P_{ef} = 21.4\%$, and small population of $|f_A, g_B\rangle$, $|f_A, e_B\rangle$ and $|f_A, f_B\rangle$, $\sum_{i=\{g,e,f\}} P_{fi} = 2.7\%$, which indicates that photon loss is a significant source of error. We project onto the qubit (g,e)-subspace numerically and obtain a two-qubit density matrix, Fig. 8d, showing a fidelity of $\mathcal{F}_s^{\text{cor}} = \langle \psi^+ | \hat{\rho}_{\text{cor}}^{\text{pr}} | \psi^+ \rangle = 92.4 \pm 0.4\%$. F_s^{cor} is the fidelity we expect to obtain with an ideal qubit readout. For finite fidelity in distinguishing between the pair of states $|g\rangle$, $|e\rangle$ and the state $|f\rangle$, we expect a false positive heralding of the entanglement protocol to occur with a probability of approximately 5%, and thus, an overall state fidelity of approximately 87%. Comparing these results to the fully deterministic case (Section 2.2), $F_s^{\text{det}} \approx 79\%$, shows the potential of the presented time-bin encoding protocol to generate high fidelity remote entanglement even in the presence of photon loss. Using a master equation simulation we attribute approximately 6.5% of the infidelity to energy relaxation and the rest to dephasing.

4 Conclusion and Outlook

Quantum networks across remote nodes are based on reliable, high-bandwidth quantum communication, which has been investigated in a variety of physical systems. Here, we have presented the operation of a direct quantum channel in a small network connecting two remote nodes, realized in circuit QED architecture. The direct quantum channel is established using time-symmetric itinerant single photons as information carriers. To emit and absorb these itinerant photons at the nodes we use an all-microwave cavity-assisted Raman process at the f0g1-transition. Using this quantum channel enabled by photon exchange, we employed a Fock-state encoding to realize a deterministic quantum communication protocol between two remote nodes. We achieved a qubit-state-transfer process fidelity of $80.02 \pm 0.07\%$ and generated

Bell states $|\psi^+\rangle$ with a fidelity of $78.9 \pm 0.1\%$. Related work independently achieved similar results with superconducting circuits [24–27]. Our approach demonstrates a high remote quantum communication rate of 50 kHz making the Fock-state encoding particularly interesting for quantum communication in low-loss quantum channels. We are confident that the state transfer rate can be improved to 1 MHz, the inverse of the protocol duration including active reset. The loss in the quantum channel between the two nodes can be reduced by replacing the isolator (insertion loss: $13 \pm 2\%$) in our quantum channel with a low-loss superconducting switch [82] which still allows for thermalization of the waveguide field, or by tunable couplers at both ends of the waveguide [83, 84]. In the second approach, we transferred quantum information in a time-bin superposition of a single photon. The time-bin encoding provides the possibility to detect communication errors making the quantum communication rate linearly proportional to the transmission probability. Using time-bin encoding we improved the qubit-state-transfer process fidelity to $92.4 \pm 0.4\%$ and the Bell-state fidelity to $90.3 \pm 0.2\%$ when considering ideal error detection. Time-bin encoding thus provides an opportunity for realizing heralded quantum communication in lossy, direct quantum channels.

For all remote communication experiments, we find excellent agreement of our experimental results with numerical master equation simulations indicating a high level of control over our experimental system. Using these master equation simulations for the Fock-state encoding, we decompose the Bell-state infidelity into approximately 10.5% photon loss, 9.5% finite transmon coherence times and 1% photon absorption inefficiency for our experimental setup. For the time-bin encoding, we conclude from the master equation simulations that all photon loss is detected in our scheme and that the residual Bell-state infidelity is explained by the finite transmon coherence times. Using time-bin encoding, we also obtain from the master equation simulations that approximately 2/3 of communication errors due to energy relaxation of the transmon are detected, since such decay errors lead to the same final qutrit state as photon loss.

The master equation simulations allow us to predict the performance of improved remote communication experiments. For example, we estimate that state fidelities of 98.3% using the Fock-state encoding and 98.1% using the time-bin encoding can be obtained for coherence times of $30 \mu\text{s}$ for the ge-transitions and of $15 \mu\text{s}$ for the ef-transitions. For these estimates we assume the absence of photon loss and perfect heralding but otherwise circuit parameters identical to the presented experiments. Improving the presented circuit QED systems to obtain higher photon emission and absorption bandwidths, we estimate that our Fock-state encoding protocol can be performed in 100 ns (currently 180 ns), and the time-bin encoding protocol in 230 ns. These short protocol durations and high fidelities indicate that the presented remote communication schemes can potentially be integrated in larger distributed quantum computing algorithms with neither performance degradation nor time overhead, e.g. with state-of-the-art qubit readout times of approximately 60–300 ns [66, 67, 80, 85, 86] and two-qubit gate times of 40–160 ns [87–90]. These advantages make quantum communication protocols such as those implemented using the f0g1-interaction interesting for realizing future large-scale quantum computing architectures.

References

- [1] H. J. Kimble, *The quantum internet*, Nature **453**, 1023 (2008).
- [2] L.-M. Duan, C. Monroe, *Colloquium: Quantum networks with trapped ions*, Rev. Mod. Phys. **82**, 1209 (2010).
- [3] D. P. DiVincenzo, *The Physical Implementation of Quantum Computation*, Fortschritte der Physik **48**, 771 (2000).
- [4] A. S. Cacciapuoti, M. Caleffi, F. Tafuri, F. S. Cataliotti, S. Gherardini, G. Bianchi, *Quantum Internet: Networking Challenges in Distributed Quantum Computing*, arXiv:1810.08421 (2018).
- [5] C. Monroe, R. Raussendorf, A. Ruthven, K. R. Brown, P. Maunz, L.-M. Duan, J. Kim, *Large-scale modular quantum-computer architecture with atomic memory and photonic interconnects*, Phys. Rev. A **89**, 022317 (2014).
- [6] A. Reiserer, G. Rempe, *Cavity-based quantum networks with single atoms and optical photons*, Rev. Mod. Phys. **87**, 1379 (2015).
- [7] L. K. Grover, *Quantum Telecomputation*, arXiv:quant-ph/9704012 (1997).
- [8] R. Cleve, H. Buhrman, *Substituting quantum entanglement for communication*, Phys. Rev. A **56**, 1201 (1997).
- [9] J. I. Cirac, A. K. Ekert, S. F. Huelga, C. Macchiavello, *Distributed quantum computation over noisy channels*, Phys. Rev. A **59**, 4249 (1999).
- [10] L. Jiang, J. M. Taylor, A. S. Sørensen, M. D. Lukin, *Distributed quantum computation based on small quantum registers*, Phys. Rev. A **76**, 062323 (2007).
- [11] J. Yezpez, *Type-II Quantum Computers*, International Journal of Modern Physics C **12**, 1273 (2001).
- [12] A. G. Fowler, D. S. Wang, C. D. Hill, T. D. Ladd, R. Van Meter, L. C. L. Hollenberg, *Surface Code Quantum Communication*, Phys. Rev. Lett. **104**, 180503 (2010).
- [13] K. Fujii, T. Yamamoto, M. Koashi, N. Imoto, *A distributed architecture for scalable quantum computation with realistically noisy devices*, arXiv:1202.6588 (2012).
- [14] C. Horsman, A. G. Fowler, S. Devitt, R. V. Meter, *Surface code quantum computing by lattice surgery*, New Journal of Physics **14**, 123011 (2012).
- [15] N. Nickerson, *Practical fault-tolerant quantum computing*, Ph.D. thesis, Imperial College London (2015).
- [16] S. Kumar, N. Lauk, C. Simon, *Towards long-distance quantum networks with superconducting processors and optical links*, Quantum Science and Technology **4**, 045003 (2019).
- [17] J. I. Cirac, P. Zoller, H. J. Kimble, H. Mabuchi, *Quantum State Transfer and Entanglement Distribution among Distant Nodes in a Quantum Network*, Phys. Rev. Lett. **78**, 3221 (1997).

- [18] S. J. van Enk, J. I. Cirac, P. Zoller, *Ideal Quantum Communication over Noisy Channels: A Quantum Optical Implementation*, Phys. Rev. Lett. **78**, 4293 (1997).
- [19] R. Van Meter, W. J. Munro, K. Nemoto, K. M. Itoh, *Arithmetic on a Distributed-memory Quantum Multicomputer*, J. Emerg. Technol. Comput. Syst. **3**, 2:1 (2008).
- [20] M. Ying, Y. Feng, *An Algebraic Language for Distributed Quantum Computing*, IEEE Transactions on Computers **58**, 728 (2009).
- [21] J. Eisert, K. Jacobs, P. Papadopoulos, M. B. Plenio, *Optimal local implementation of non-local quantum gates*, Phys. Rev. A **62**, 052317 (2000).
- [22] M. Sarvaghad-Moghaddam, M. Zomorodi-Moghaddam, A. Farouk, *A General protocol for Distributed Quantum Gates*, arXiv:1812.07798 (2018).
- [23] I. Cohen, K. Mølmer, *Deterministic quantum network for distributed entanglement and quantum computation*, Phys. Rev. A **98**, 030302 (2018).
- [24] P. Campagne-Ibarcq, E. Zaly-Geller, A. Narla, S. Shankar, P. Reinhold, L. Burkhardt, C. Axline, W. Pfaff, L. Frunzio, R. J. Schoelkopf, M. H. Devoret, *Deterministic Remote Entanglement of Superconducting Circuits through Microwave Two-Photon Transitions*, Phys. Rev. Lett. **120**, 200501 (2018).
- [25] C. Axline, L. Burkhardt, W. Pfaff, M. Zhang, K. Chou, P. Campagne-Ibarcq, P. Reinhold, L. Frunzio, S. M. Girvin, L. Jiang, M. H. Devoret, R. J. Schoelkopf, *On-demand quantum state transfer and entanglement between remote microwave cavity memories*, Nature Physics **14**, 705 (2018).
- [26] P. Kurpiers, P. Magnard, T. Walter, B. Royer, M. Pechal, J. Heinsoo, Y. Salathé, A. Akin, S. Storz, J.-C. Besse, S. Gasparinetti, A. Blais, A. Wallraff, *Deterministic quantum state transfer and remote entanglement using microwave photons*, Nature **558**, 264 (2018).
- [27] N. Leung, Y. Lu, S. Chakram, R. K. Naik, N. Earnest, R. Ma, K. Jacobs, A. N. Cleland, D. I. Schuster, *Deterministic bidirectional communication and remote entanglement generation between superconducting qubits*, npj Quantum Information **5**, 18 (2019).
- [28] P. Kurpiers, M. Pechal, B. Royer, P. Magnard, T. Walter, J. Heinsoo, Y. Salathé, A. Akin, S. Storz, J.-C. Besse, S. Gasparinetti, A. Blais, A. Wallraff, *Quantum Communication with Time-Bin Encoded Microwave Photons*, Phys. Rev. Applied **12**, 044067 (2019).
- [29] C. Eichler, D. Bozyigit, C. Lang, L. Steffen, J. Fink, A. Wallraff, *Experimental State Tomography of Itinerant Single Microwave Photons*, Phys. Rev. Lett. **106**, 220503 (2011).
- [30] M. Pechal, L. Huthmacher, C. Eichler, S. Zeytinoğlu, A. A. Abdumalikov Jr., S. Berger, A. Wallraff, S. Filipp, *Microwave-Controlled Generation of Shaped Single Photons in Circuit Quantum Electrodynamics*, Phys. Rev. X **4**, 041010 (2014).
- [31] S. Zeytinoğlu, M. Pechal, S. Berger, A. A. Abdumalikov Jr., A. Wallraff, S. Filipp, *Microwave-induced amplitude- and phase-tunable qubit-resonator coupling in circuit quantum electrodynamics*, Phys. Rev. A **91**, 043846 (2015).

- [32] J. F. Clauser, M. A. Horne, A. Shimony, R. A. Holt, *Proposed Experiment to Test Local Hidden-Variable Theories*, Phys. Rev. Lett. **23**, 880 (1969).
- [33] S. Hill, W. K. Wootters, *Entanglement of a Pair of Quantum Bits*, Phys. Rev. Lett. **78**, 5022 (1997).
- [34] W. K. Wootters, *Entanglement of formation of an arbitrary state of two qubits*, Phys. Rev. Lett. **80**, 2245 (1998).
- [35] M. A. Nielsen, I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, New York, NY, USA, 2011), 10th edition.
- [36] C. W. Chou, H. de Riedmatten, D. Felinto, S. V. Polyakov, S. J. van Enk, H. J. Kimble, *Measurement-induced entanglement for excitation stored in remote atomic ensembles*, Nature **438**, 828 (2005).
- [37] D. N. Matsukevich, T. Chanelière, S. D. Jenkins, S.-Y. Lan, T. A. B. Kennedy, A. Kuzmich, *Entanglement of Remote Atomic Qubits*, Phys. Rev. Lett. **96**, 030405 (2006).
- [38] C.-W. Chou, J. Laurat, H. Deng, K. S. Choi, H. de Riedmatten, D. Felinto, H. J. Kimble, *Functional Quantum Nodes for Entanglement Distribution over Scalable Quantum Networks*, Science **316**, 1316 (2007).
- [39] J. Laurat, K. S. Choi, H. Deng, C. W. Chou, H. J. Kimble, *Heralded Entanglement between Atomic Ensembles: Preparation, Decoherence, and Scaling*, Phys. Rev. Lett. **99**, 180504 (2007).
- [40] Z.-S. Yuan, Y.-A. Chen, B. Zhao, S. Chen, J. Schmiedmayer, J.-W. Pan, *Experimental demonstration of a BDCZ quantum repeater node*, Nature **454**, 1098 (2008).
- [41] D. L. Moehring, P. Maunz, S. Olmschenk, K. C. Younge, D. N. Matsukevich, L. M. Duan, C. Monroe, *Entanglement of single-atom quantum bits at a distance*, Nature **449**, 68 (2007).
- [42] D. N. Matsukevich, P. Maunz, D. L. Moehring, S. Olmschenk, C. Monroe, *Bell Inequality Violation with Two Remote Atomic Qubits*, Phys. Rev. Lett. **100**, 150404 (2008).
- [43] P. Maunz, S. Olmschenk, D. Hayes, D. N. Matsukevich, L.-M. Duan, C. Monroe, *Heralded Quantum Gate between Remote Quantum Memories*, Phys. Rev. Lett. **102**, 250502 (2009).
- [44] L. Slodička, G. Hétet, N. Röck, P. Schindler, M. Hennrich, R. Blatt, *Atom-Atom Entanglement by Single-Photon Detection*, Phys. Rev. Lett. **110**, 083603 (2013).
- [45] D. Hucul, I. V. Inlek, G. Vittorini, C. Crocker, S. Debnath, S. M. Clark, C. Monroe, *Modular entanglement of atomic qubits using photons and phonons*, Nature Physics **11**, 37 (2014).
- [46] M. Lettner, M. Mücke, S. Riedl, C. Vo, C. Hahn, S. Baur, J. Bochmann, S. Ritter, S. Dürr, G. Rempe, *Remote Entanglement between a Single Atom and a Bose-Einstein Condensate*, Phys. Rev. Lett. **106**, 210503 (2011).

- [47] K. C. Lee, M. R. Sprague, B. J. Sussman, J. Nunn, N. K. Langford, X.-M. Jin, T. Champion, P. Michelberger, K. F. Reim, D. England, D. Jaksch, I. A. Walmsley, *Entangling Macroscopic Diamonds at Room Temperature*, Science **334**, 1253 (2011).
- [48] I. Usmani, C. Clausen, F. Bussieres, N. Sangouard, M. Afzelius, N. Gisin, *Heralded quantum entanglement between two crystals*, Nature Photonics **6**, 234 (2012).
- [49] J. Hofmann, M. Krug, N. Ortegel, L. Gérard, M. Weber, W. Rosenfeld, H. Weinfurter, *Heralded Entanglement Between Widely Separated Atoms*, Science **337**, 72 (2012).
- [50] S. Ritter, C. Nolleke, C. Hahn, A. Reiserer, A. Neuzner, M. Uphoff, M. Mücke, E. Figueroa, J. Bochmann, G. Rempe, *An elementary quantum network of single atoms in optical cavities*, Nature **484**, 195 (2012).
- [51] H. Bernien, B. Hensen, W. Pfaff, G. Koolstra, M. S. Blok, L. Robledo, T. H. Taminiau, M. Markham, D. J. Twitchen, L. Childress, R. Hanson, *Heralded entanglement between solid-state qubits separated by three metres*, Nature **497**, 86 (2013).
- [52] W. Pfaff, B. J. Hensen, H. Bernien, S. B. van Dam, M. S. Blok, T. H. Taminiau, M. J. Tiggelman, R. N. Schouten, M. Markham, D. J. Twitchen, R. Hanson, *Unconditional quantum teleportation between distant solid-state quantum bits*, Science **345**, 532 (2014).
- [53] B. Hensen, H. Bernien, A. E. Dreau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. L. Vermeulen, R. N. Schouten, C. Abellan, W. Amaya, V. Pruneri, M. W. Mitchell, M. Markham, D. J. Twitchen, D. Elkouss, S. Wehner, T. H. Taminiau, R. Hanson, *Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres*, Nature **526**, 682 (2015).
- [54] N. Roch, M. E. Schwartz, F. Motzoi, C. Macklin, R. Vijay, A. W. Eddins, A. N. Korotkov, K. B. Whaley, M. Sarovar, I. Siddiqi, *Observation of Measurement-Induced Entanglement and Quantum Trajectories of Remote Superconducting Qubits*, Phys. Rev. Lett. **112**, 170501 (2014).
- [55] A. Narla, S. Shankar, M. Hatridge, Z. Leghtas, K. M. Sliwa, E. Zolys-Geller, S. O. Mundhada, W. Pfaff, L. Frunzio, R. J. Schoelkopf, M. H. Devoret, *Robust Concurrent Remote Entanglement Between Two Superconducting Qubits*, Phys. Rev. X **6**, 031036 (2016).
- [56] C. Dickel, J. J. Wesdorp, N. K. Langford, S. Peiter, R. Sagastizabal, A. Bruno, B. Criger, F. Motzoi, L. DiCarlo, *Chip-to-chip entanglement of transmon qubits using engineered measurement fields*, Phys. Rev. B **97**, 064508 (2018).
- [57] A. Delteil, Z. Sun, W. Gao, E. Togan, S. Faelt, A. Imamoglu, *Generation of heralded entanglement between distant hole spins*, Nature Physics **12**, 218 (2016).
- [58] R. Stockill, M. J. Stanley, L. Huthmacher, E. Clarke, M. Hugues, A. J. Miller, C. Matthiesen, C. Le Gall, M. Atatüre, *Phase-Tuned Entangled State Generation between Distant Spin Qubits*, Phys. Rev. Lett. **119**, 010503 (2017).
- [59] F. Verstraete, M. M. Wolf, *Entanglement versus Bell Violations and Their Behavior under Local Filtering Operations*, Phys. Rev. Lett. **89**, 170401 (2002).

- [60] B. Julsgaard, A. Kozhekin, E. S. Polzik, *Experimental long-lived entanglement of two macroscopic objects*, Nature **413**, 400 (2001).
- [61] P. C. Humphreys, N. Kalb, J. P. J. Morits, R. N. Schouten, R. F. L. Vermeulen, D. J. Twitchen, M. Markham, R. Hanson, *Deterministic delivery of remote entanglement on a quantum network*, Nature **558**, 268 (2018).
- [62] J. Koch, T. M. Yu, J. Gambetta, A. A. Houck, D. I. Schuster, J. Majer, A. Blais, M. H. Devoret, S. M. Girvin, R. J. Schoelkopf, *Charge-insensitive qubit design derived from the Cooper pair box*, Phys. Rev. A **76**, 042319 (2007).
- [63] A. Wallraff, D. I. Schuster, A. Blais, L. Frunzio, R.-S. Huang, J. Majer, S. Kumar, S. M. Girvin, R. J. Schoelkopf, *Strong coupling of a single photon to a superconducting qubit using circuit quantum electrodynamics*, Nature **431**, 162 (2004).
- [64] A. Wallraff, D. I. Schuster, A. Blais, L. Frunzio, J. Majer, M. H. Devoret, S. M. Girvin, R. J. Schoelkopf, *Approaching Unit Visibility for Control of a Superconducting Qubit with Dispersive Readout*, Phys. Rev. Lett. **95**, 060501 (2005).
- [65] M. D. Reed, B. R. Johnson, A. A. Houck, L. DiCarlo, J. M. Chow, D. I. Schuster, L. Frunzio, R. J. Schoelkopf, *Fast reset and suppressing spontaneous emission of a superconducting qubit*, Appl. Phys. Lett. **96**, 203110 (2010).
- [66] E. Jeffrey, D. Sank, J. Y. Mutus, T. C. White, J. Kelly, R. Barends, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. Megrant, P. J. J. O'Malley, C. Neill, P. Roushan, A. Vainsencher, J. Wenner, A. N. Cleland, J. M. Martinis, *Fast Accurate State Measurement with Superconducting Qubits*, Phys. Rev. Lett. **112**, 190504 (2014).
- [67] T. Walter, P. Kurpiers, S. Gasparinetti, P. Magnard, A. Potočnik, Y. Salathé, M. Pechal, M. Mondal, M. Oppliger, C. Eichler, A. Wallraff, *Rapid, High-Fidelity, Single-Shot Dispersive Readout of Superconducting Qubits*, Phys. Rev. Applied **7**, 054020 (2017).
- [68] P. Magnard, P. Kurpiers, B. Royer, T. Walter, J.-C. Besse, S. Gasparinetti, M. Pechal, J. Heinsoo, S. Storz, A. Blais, A. Wallraff, *Fast and Unconditional All-Microwave Reset of a Superconducting Qubit*, Phys. Rev. Lett. **121**, 060502 (2018).
- [69] D. Egger, M. Werninghaus, M. Ganzhorn, G. Salis, A. Fuhrer, P. Müller, S. Filipp, *Pulsed Reset Protocol for Fixed-Frequency Superconducting Qubits*, Phys. Rev. Applied **10**, 044030 (2018).
- [70] D. Bozyigit, C. Lang, L. Steffen, J. M. Fink, C. Eichler, M. Baur, R. Bianchetti, P. J. Leek, S. Filipp, M. P. da Silva, A. Blais, A. Wallraff, *Antibunching of microwave-frequency photons observed in correlation measurements using linear detectors*, Nat. Phys. **7**, 154 (2011).
- [71] P. Kurpiers, T. Walter, P. Magnard, Y. Salathe, A. Wallraff, *Characterizing the attenuation of coaxial and rectangular microwave-frequency waveguides at cryogenic temperatures*, EPJ Quantum Technology **4**, 8 (2017).
- [72] I. L. Chuang, M. A. Nielsen, *Prescription for experimental determination of the dynamics of a quantum black box*, J. Mod. Opt. **44**, 2455 (1997).

- [73] S. J. van Enk, N. Lütkenhaus, H. J. Kimble, *Experimental procedures for entanglement verification*, Phys. Rev. A **75**, 052318 (2007).
- [74] R. Bianchetti, S. Filipp, M. Baur, J. M. Fink, C. Lang, L. Steffen, M. Boissonneault, A. Blais, A. Wallraff, *Control and Tomography of a Three Level Superconducting Artificial Atom*, Phys. Rev. Lett. **105**, 223601 (2010).
- [75] J. A. Smolin, J. M. Gambetta, G. Smith, *Efficient Method for Computing the Maximum-Likelihood Quantum State from Measurements with Additive Gaussian Noise*, Phys. Rev. Lett. **108**, 070502 (2012).
- [76] K. Chen, L. Wu, *A matrix realignment method for recognizing entanglement*, Quantum Information & Computation **3**, 193 (2003).
- [77] O. Rudolph, *Further Results on the Cross Norm Criterion for Separability*, Quantum Information Processing **4**, 219 (2005).
- [78] G. de Lange, D. Riste, M. J. Tiggelman, C. Eichler, L. Tornberg, G. Johansson, A. Wallraff, R. N. Schouten, L. DiCarlo, *Reversing Quantum Trajectories with Analog Feedback*, Phys. Rev. Lett. **112**, 080501 (2014).
- [79] M. Jerger, P. Macha, A. R. Hamann, Y. Reshitnyk, K. Juliusson, A. Fedorov, *Realization of a Binary-Outcome Projection Measurement of a Three-Level Superconducting Quantum System*, Phys. Rev. Applied **6**, 014014 (2016).
- [80] J. Heinsoo, C. K. Andersen, A. Remm, S. Krinner, T. Walter, Y. Salathé, S. Gasparinetti, J.-C. Besse, A. Potočník, A. Wallraff, C. Eichler, *Rapid High-fidelity Multiplexed Readout of Superconducting Qubits*, Phys. Rev. Applied **10**, 034040 (2018).
- [81] L. DiCarlo, J. M. Chow, J. M. Gambetta, L. S. Bishop, B. R. Johnson, D. I. Schuster, J. Majer, A. Blais, L. Frunzio, S. M. Girvin, R. J. Schoelkopf, *Demonstration of two-qubit algorithms with a superconducting quantum processor*, Nature **460**, 240 (2009).
- [82] M. Pechal, J.-C. Besse, M. Mondal, M. Oppliger, S. Gasparinetti, A. Wallraff, *Superconducting Switch for Fast On-Chip Routing of Quantum Microwave Fields*, Phys. Rev. Applied **6**, 024009 (2016).
- [83] A. Bienfait, K. J. Satzinger, Y. P. Zhong, H.-S. Chang, M.-H. Chou, C. R. Conner, E. Dumur, J. Grebel, G. A. Peairs, R. G. Povey, A. N. Cleland, *Phonon-mediated quantum state transfer and remote qubit entanglement*, Science **364**, 368 (2019).
- [84] Y. P. Zhong, H.-S. Chang, K. J. Satzinger, M.-H. Chou, A. Bienfait, C. R. Conner, E. Dumur, J. Grebel, G. A. Peairs, R. G. Povey, D. I. Schuster, A. N. Cleland, *Violating Bell's inequality with remotely connected superconducting qubits*, Nature Physics (2019).
- [85] D. Ristè, J. G. van Leeuwen, H.-S. Ku, K. W. Lehnert, L. DiCarlo, *Initialization by Measurement of a Superconducting Quantum Bit Circuit*, Phys. Rev. Lett. **109**, 050507 (2012).
- [86] R. Dassonneville, T. Ramos, V. Milchakov, L. Planat, E. Dumur, F. Foroughi, J. Puer-tas, S. Leger, K. Bharadwaj, J. Delaforce, K. Rafsanjani, C. Naud, W. Hasch-Guichard, J. J. Garcia-Ripoll, N. Roch, O. Buisson, *Fast high fidelity quantum non-demolition qubit readout via a non-perturbative cross-Kerr coupling*, arXiv:1905.00271 (2019).

- [87] R. Barends, J. Kelly, A. Megrant, A. Veitia, D. Sank, E. Jeffrey, T. C. White, J. Mutus, A. G. Fowler, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, C. Neill, P. O'Malley, P. Roushan, A. Vainsencher, J. Wenner, A. N. Korotkov, A. N. Cleland, J. M. Martinis, *Superconducting quantum circuits at the surface code threshold for fault tolerance*, Nature **508**, 500 (2014).
- [88] S. Sheldon, E. Magesan, J. M. Chow, J. M. Gambetta, *Procedure for systematically tuning up cross-talk in the cross-resonance gate*, Phys. Rev. A **93**, 060302 (2016).
- [89] C. K. Andersen, A. Remm, S. Lazar, S. Krinner, J. Heinsoo, J.-C. Besse, M. Gabureac, A. Wallraff, C. Eichler, *Entanglement stabilization using ancilla-based parity detection and real-time feedback in superconducting circuits*, npj Quantum Information **5**, 69 (2019).
- [90] M. A. Rol, F. Battistel, F. K. Malinowski, C. C. Bultink, B. M. Tarasinski, R. Vollmer, N. Haider, N. Muthusubramanian, A. Bruno, B. M. Terhal, L. DiCarlo, *Fast, High-Fidelity Conditional-Phase Gate Exploiting Leakage Interference in Weakly Anharmonic Superconducting Qubits*, Phys. Rev. Lett. **123**, 120502 (2019).

A 8 **Quantum Communication**

Christophe Salomon

Ecole Normale Supérieure, Département de Physique, Paris,
France

Please refer to:

http://www.phys.ens.fr/~salomon/anglais/index_an.html

- The Organizers -

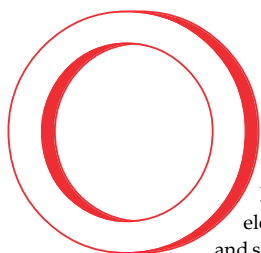
B 1 Quantum Computing with semiconductor spins

Lieven M.K. Vandersypen¹ and Mark A. Eriksson²

¹Delft University of Technology, Delft, The Netherlands

²University of Wisconsin, Dept. of Physics, Wisconsin, USA

Lieven Vandersypen is an Antoni van Leeuwenhoek Professor in the Kavli Institute of Nanoscience Delft and QuTech, a research center of Delft University of Technology in the Netherlands. **Mark Eriksson** is a Vilas Distinguished Achievement Professor in the department of physics at the University of Wisconsin–Madison.



pen any textbook on quantum mechanics, and the two-state system of choice is likely to be a spin- $\frac{1}{2}$ particle, such as an electron. The corresponding states, spin up and spin down, form the prototypical quantum bit (qubit), and rotations of the spin state constitute the simplest quantum logic gates. Because of their negative charge, electrons can be manipulated with voltages applied to nanoscale electrodes, or gates. And the application of appropriate voltages can confine the electrons to small islands called quantum dots (see the article by Marc Kastner, PHYSICS TODAY, January 1993, page 24).

Twenty years ago Daniel Loss and David DiVincenzo proposed that the spin of a single electron in a semiconductor quantum dot could form not just a model but also a real, physical qubit.¹ Their theoretical work predated by four years the first experiments to successfully trap a single electron in a gate-defined quantum dot, and it predated by several more years the first coherent manipulation of a single spin in a semiconductor. Semiconductor spin qubits now come in four distinct flavors, each of which was proposed by theory that set a target for experiments to pursue. Those experiments always brought surprises, and the interplay between theory and experiment makes semiconductor spin qubits a particularly vibrant field of study.

In this article we describe the experimental development and the current state of the art of semiconductor quantum-dot spin qubits. Functional and scalable qubits must meet well-defined criteria.² First, reliably initializing each qubit into one of its two levels must be possible. Second, the final state of each qubit must be knowable by a projective measurement that gives the correct answer with high probability. Third, qubit manipulation must be implementable using high-quality single- and two-qubit gates.

Imagine the spin state as a vector pointing on a sphere, commonly known as the Bloch sphere. Single-qubit gates correspond to rotations of the state vector that are independent of the state of any other qubit in the system. In the case of qubit gates, rotation of one qubit depends on the state of

other. And when the second qubit itself starts off in a superposition of states, the two qubits become entangled with each other. The recent satisfaction of all those requirements with quantum dots led to the demonstration of the first—and at two qubits the smallest possible—quantum semiconductor processor.

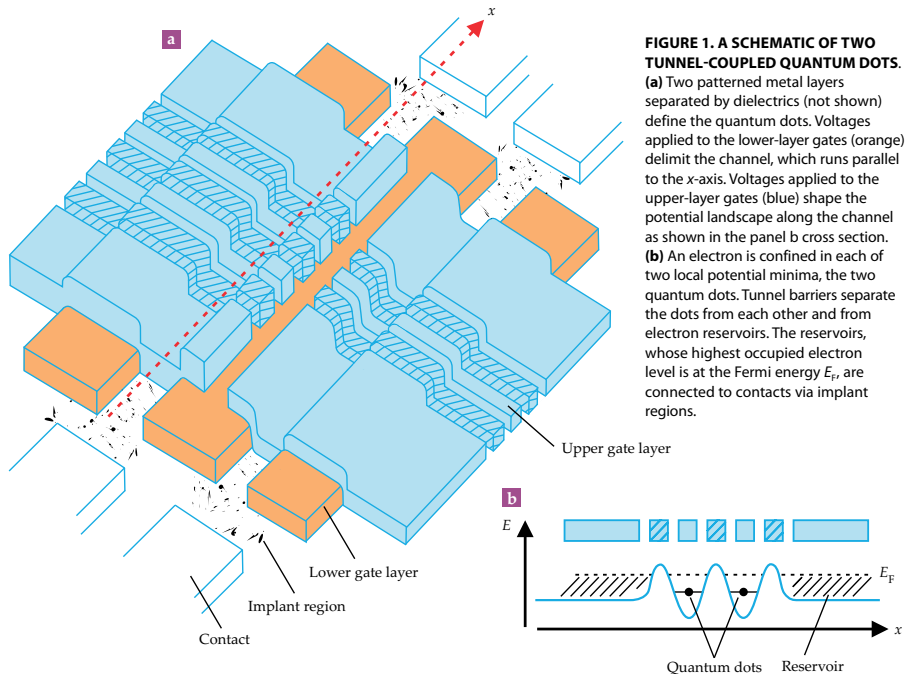
That single-electron spins in a semiconductor chip can act as qubits is remarkable. Unlike atoms or photons in a vacuum, an electron in a semiconductor resides in a noisy, solid-state environment. Engineering that environment so that it doesn't rapidly de-

grade or decohere the spin-qubit states has been a key challenge for our field.

Errors are unavoidable and necessitate quantum error-correction techniques (see PHYSICS TODAY, February 2005, page 19). To be effective, the techniques require that initialization, readout, and single- and two-qubit operations have error rates below 1%. Furthermore, quantum error correction involves an overhead in the number of qubits that can easily reach 1000 physical error-prone qubits to encode one protected error-free qubit. Therefore, a future quantum computer capable of solving relevant problems beyond the reach of a supercomputer will likely contain millions of physical qubits. (See the article by David Weiss and Mark Saffman, PHYSICS TODAY, July 2017, page 44.)

Semiconductor quantum dots have a tiny footprint that offers the prospect of integrating millions of qubits, akin to classical integrated circuits. The corresponding electron density in quantum-dot devices, however, is far smaller than in classical transistors, with each single electron in a qubit typically spread over a region roughly $20\text{ nm} \times 20\text{ nm}$ in size. For such a device to work as intended, the materials and nanofabricated structures must have very little disorder, to ensure that electrons are easy to position and control. Pulling off that achievement entails uniform patterning of the gate electrodes but also having low densities of trapped charges in the substrate, in the dielectrics, and at the interfaces.

SEMICONDUCTOR SPINS



sprint. And research today is motivated by a vision that will take years to bring to fruition. In the case of semiconductor spin qubits, that vision relies on long coherence times and on recent advances in gate fidelity—a common metric to express the quality of quantum gates—fueled by a move to silicon-based devices.

Intriguingly, spin qubits in semiconductors could also be integrated with classical integrated-circuit technology, including processing, memory, and the distribution of signals. Integration on chip is natural, because quantum-dot qubits use gate electrodes just as field-effect transistors do. Integration could also occur at the system level, with clusters of chips communicating with one another.

From transistor to qubit

The field-effect transistor is a good starting point for understanding a quantum dot. In a transistor, the flow of electrons between two contacts (source and drain) is switched on or off via the voltage on a metal gate electrode placed above the space between the contacts (the channel). A positive gate voltage attracts electrons to the channel and produces a conducting path from source to drain. A negative gate voltage, by contrast, empties the channel such that no source-drain current can flow. If one were to replace the gate electrode with three independently biased electrodes, the electronic potential landscape between the contacts could be shaped to create a potential-

energy minimum separated from the contact regions by potential barriers.

At low temperature, typically below 4 K, the thermal energy is lower than the energy needed to add or remove electrons from the potential well. Thus the well is occupied by a discrete number of electrons. When the electrons are confined tightly enough that orbital motion is frozen out quantum mechanically, the device is known as a quantum dot.

Arrays of tunnel-coupled quantum dots can be formed with additional gate electrodes, as shown in figure 1. The voltages on the blue gate electrodes control the depth of the potential minima and thereby the number of electrons on each quantum dot. The voltages on the hatched blue gates control the tunnel barriers between adjacent dots and between the dots and the reservoirs. Nowadays, quantum dots are routinely tuned to the limit in which just a single electron resides on each dot. Researchers can verify the tuning by monitoring the current through an auxiliary nearby quantum dot that acts as an electrometer.

Spin qubits

When one electron resides in each quantum dot in the presence of a magnetic field, each electron spin becomes an appealing qubit. Indeed, that simple configuration, with one electron in one dot, was proposed by Loss and DiVincenzo in 1998. In subsequent years, alternative spin qubits have made their debut.

HOW TO INITIALIZE, MANIPULATE, AND READ OUT A SPIN QUBIT

Reading out the spin state of an electron on a quantum dot involves making a so-called spin-to-charge conversion,¹⁶ whereby the electron is allowed to tunnel from one location to another in a way that depends on its spin state—or more specifically, on whether the qubit is up or down. A nearby charge sensor is sensitive to the dots' electron occupation; the current through the sensor thus indirectly reveals the spin state.

In one scenario, the Pauli exclusion principle provides the spin dependence: Two electrons can reside on the same dot only when they are in a spin-singlet state. For a spin-triplet state, each electron is forced to reside on its own dot. In another scenario, a qubit's two spin states are aligned above and below the reservoir's Fermi level—the highest occupied energy level (see panel a of the figure). That protocol is usually effective for any qubit separated by at least a few times the thermal energy. When the electron in the dot occupies the lower-energy spin, it doesn't have enough energy to leave and no tunneling occurs. But if the higher-energy spin state is occupied, the electron can tunnel out and is detected. Afterwards, another electron tunnels into the dot from the reservoir.

Initialization is commonly the result of

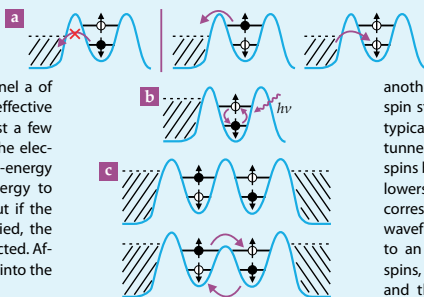
readout, after which an electron with a known spin resides in the dot. Alternatively, initialization can be achieved by allowing the electron spin qubit to thermalize to its ground state.

Resonant control of spin qubits uses magnetic or electric excitation at radio or microwave frequencies. Magnetic excitation can coherently drive spin transitions directly when the excitation is resonant with the energy difference between spin-up and spin-down states (see panel b in the figure).¹⁷ The excitation's amplitude controls the rotation frequency of the spin vector around the Bloch sphere, its phase controls

the rotation axis, and its duration controls the rotation angle.

Resonant electrical excitation, by contrast, can drive single-spin transitions because of spin-orbit coupling.¹⁸ The excitation causes the electron to oscillate back and forth in a quantum dot, and the electron experiences an oscillating effective magnetic field that rotates the electron's spin. Alternatively, in the presence of a suitably engineered magnetic field gradient at the dot location, an electrically driven electron experiences a real, oscillating magnetic field, again allowing for coherent spin rotations. In the case of the quantum-dot hybrid qubit (three electrons in two dots), resonant electric fields alone drive transitions between the qubit states.⁵

Gate-voltage pulses provide another method to controllably manipulate spin states. The basic idea is to abruptly—typically within nanoseconds—turn on the tunnel coupling between two neighboring spins by applying a gate-voltage pulse that lowers the tunneling barrier between their corresponding dots, so that the electron wavefunctions overlap. The overlap leads to an exchange interaction between the spins, as suggested in the figure's panel c, and the two spin states are periodically exchanged.



For instance, a qubit can comprise two collective states of two or three spins that reside in either two or three quantum dots. Those flavors are known as singlet-triplet qubits⁷ (two electrons, one each in two dots), exchange-only qubits⁴ (three electrons in three dots), and quantum-dot hybrid qubits (three electrons in two dots).⁵ The trade-offs between them are many and still under investigation. Ultimately, the various qubit types are initialized, manipulated, and read out using the same physical principles, but their robustness to specific noise sources varies, as does their ease of operation.

The first wave of successful spin-qubit experiments started in the early 2000s and used quantum dots defined by gate electrodes over a gallium arsenide/aluminum gallium arsenide two-dimensional electron gas. That heterostructure technology had been the workhorse of mesoscopic physics for more than a decade and provided a platform in which spin qubits were easy to control. Initial work largely met the important requirements for individual qubits—namely, that they could be initialized, manipulated, and read out.

As outlined in the box above, qubits can be implemented using nanosecond gate-voltage pulses and resonant microwave excitation of gate electrodes or current-carrying wires. Single-shot readout is performed indirectly, by induc-

ing spin-dependent tunneling of an electron while detecting the position of the electron in real time. The groups of Leo Kouwenhoven and one of us (Vandersypen) at Delft University of Technology (TU Delft), Charles Marcus at Harvard University, and Seigo Tarucha at the University of Tokyo were the main players to carry out those early experimental demonstrations. The GaAs work culminated in the creation of entangled states of singlet-triplet qubits by Amir Yacoby and coworkers at Harvard. They reached a fidelity—the extent to which the actual state resembles a two-qubit entangled state—of 72% and later improved it⁶ to greater than 90%.

Relaxation and decoherence

Spin qubits in GaAs benefit from remarkably long energy relaxation times T_1 , the time it takes a qubit to change from a high-energy state to the ground state. For single-spin qubits, T_1 can exceed 1 second at low temperature (100 mK or lower) in a 1 T field. That's three orders of magnitude longer than the longest T_1 in superconducting qubits.

By comparison, T_2^* , the time it takes the qubit phase to randomize, is just tens of nanoseconds in GaAs dots.³ The phase of the electron's spin is randomized through hyperfine coupling to the roughly 1 million nuclear spins of atoms in the

SEMICONDUCTOR SPINS

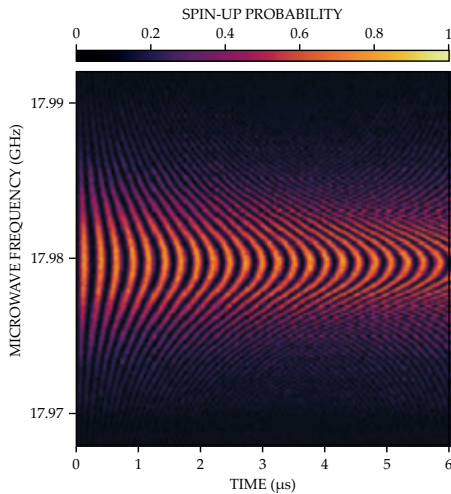


FIGURE 2. DRIVEN EVOLUTION OF A SINGLE QUBIT. When an applied microwave pulse is close to resonance with the spin qubit—that is, with the energy difference between its up and down states (17.98 GHz here)—the qubit undergoes driven rotations, or Rabi oscillation, and the probability of finding it spin-up oscillates as a function of the pulse duration. (Adapted from ref. 10.)

cated pulse schemes,⁶ the random nuclear field has significantly slowed the progress of GaAs-based spin qubits.

Enter silicon

As early as 1998, it was clear that silicon would be preferable to GaAs as a host material for spin qubits. Fewer than 5% of naturally occurring Si atoms carry a nuclear spin, and those nuclear spins can be largely eliminated by using isotopically enriched ²⁸Si. Although Si is the cornerstone of today's semiconductor technology, it has taken many years of materials development and nanofabrication advances to make Si quantum dots suitable for spin-qubit experiments.

Two main quantum-dot platforms have emerged. In the first, pioneered by one of us (Eriksson) and colleagues at the University of Wisconsin–Madison, electrons are confined in Si quantum wells by silicon germanium barriers above and below the well.⁷ In the second, developed by Andrew Dzurak and colleagues at the University of New South Wales (UNSW) in Sydney, electrons are confined against a Si-SiO₂ interface—as in n-doped metal oxide semiconductor technology.⁸ In both cases, gate electrodes on the surface are used to accumulate electrons in quantum dots and to form tunnel barriers between the dots.

The randomization time T_2^* is significantly longer in Si than in GaAs, with T_2^* reaching 1 μs in natural Si and up to 100 μs in purified ²⁸Si. That's an improvement over GaAs by four orders of magnitude,⁹ and it translates directly to single-spin gate

quantum dot, with which the electron wavefunction overlaps. The interaction is impossible to avoid because every Ga and As isotope carries a nuclear spin of ½.

Moreover, despite the low temperatures and strong magnetic fields used with typical spin-qubit measurements, the nuclear spins point in nearly random orientations. The result is a statistically fluctuating and slowly varying collective effect on the electron spin known as the random nuclear or Overhauser field. Although the randomness of the nuclear field can be significantly reduced for singlet-triplet qubits by using sophisti-

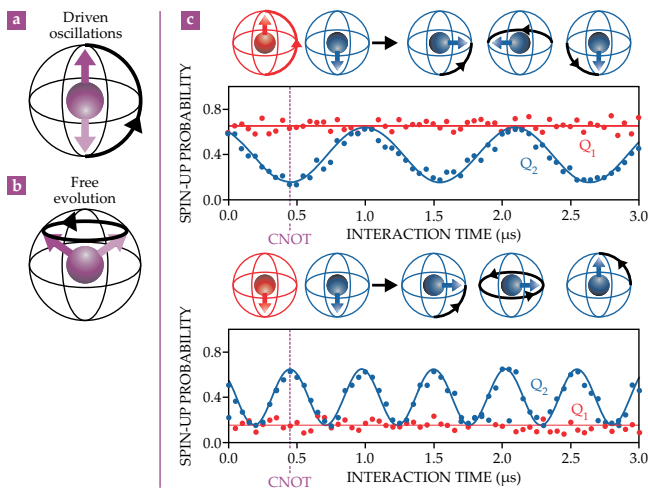


FIGURE 3. A TWO-QUBIT LOGIC GATE. (a) In a Bloch sphere diagram, a qubit rotates along a line of longitude during a resonant microwave pulse (see figure 2). (b) In the absence of a microwave pulse, a state precesses along a latitude line around the vertical axis of the Bloch sphere. (c) A controlled NOT (CNOT) gate is an operation that flips a target qubit (Q2, blue) based on the state of the control qubit (Q1, red). With Q2 initialized spin down, the plots show the time evolution of the spin-up probability of both qubits when Q1 is spin up (top) or spin down (bottom). In each case, two single-qubit $\pi/2$ rotations are applied, separated by free evolution, during which the two qubits interact. For an interaction of 0.5 μs, the sequence flips Q2 if Q1 is down but not if Q1 is up. (Adapted from ref. 11, M. Veldhorst et al.)

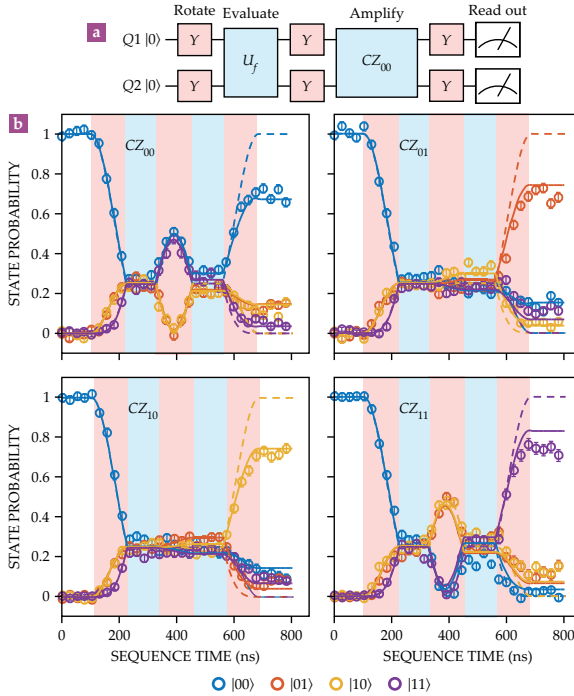


FIGURE 4. A TWO-QUBIT CIRCUIT THAT IMPLEMENTS A QUANTUM SEARCH ALGORITHM. (a) A sequence of operations acts on qubits Q1 and Q2: rotation (Y), interaction (U_f), and amplification (CZ). A detector reads out the final state probabilities of each qubit. (b) The two-spin probabilities of the qubit states' populations are plotted as a function of time; the background colors (white, pink, and blue) correspond to the colors of operations in the circuit. After the first rotation around the y-axis, the qubits are in a superposition $(|00\rangle + |01\rangle + |10\rangle + |11\rangle)/2$, with each term having equal weight. In each panel, U_f is a different interaction (CZ_{*ij*}) that picks out one particular two-qubit state; that state then gets amplified in subsequent steps due to quantum interference. Dashed and solid lines show, respectively, the ideal populations and the results of a model that includes decoherence. (Adapted from ref. 11, T. F. Watson et al.)

prospects for practical Si spin qubits have risen sharply.

Putting it all together

Building on the long-lived coherence in Si quantum-dot spin qubits, several groups have now demonstrated high-fidelity control of two single-spin qubits.¹¹ In 2015 the Dzurak group got a two-qubit gate working with single-qubit control and independent readout of the two spins. The two-qubit gate relied on the interaction between neighboring spins, as outlined in the box. That interaction, in combination with single-qubit rotations, enables a controlled-NOT (CNOT) gate, as illustrated in figure 3. Two years later two teams—a collaboration of our own groups at TU Delft and at the University of Wisconsin–Madison and, independently, the group of Jason Petta at Princeton University—demonstrated entanglement of two single-spin qubits in a Si/SiGe double quantum dot.

To further illustrate the recent progress of Si spin qubits, figure 4 shows the implementation of a simple quantum algorithm on two Si spin qubits. We and our colleagues at TU Delft and the University of Wisconsin–Madison successfully programmed all four instances of Grover's search algorithm for two qubits.¹¹ The algorithm is designed to invert a function $f(x)$ and identify the unique n -bit input value x_0 for which $f(x_0) = 1$. For all other input values, $f(x) = 0$. Without further knowledge of f , there is no more efficient method using a classical computer than exhaustively searching through the space of input values, evaluating $f(x)$ using one input value after another until hitting the input value x_0 .

The quantum case behaves very differently. Figure 4 illustrates how the occupation probabilities of the four basis states $|00\rangle$, $|01\rangle$, $|10\rangle$, and $|11\rangle$ evolve throughout the steps of the quantum algorithm for each of the four possible functions f . Starting off with qubits Q1 and Q2 both in the $|00\rangle$ ground state, the first step is to prepare an equal superposition of the four basis states

fidelities¹⁰ of greater than 99.9% (see figure 2). Furthermore, given that the nuclear-spin bath evolves slowly on the time scale of the electron-spin dynamics, it is possible to extend the coherence times to tens of milliseconds⁹ using dynamic decoupling techniques, extensions of the Hahn spin-echo concept.

Even longer electron-spin coherence times are obtained for electrons bound to phosphorus-31 dopants in ²⁸Si-enriched material. The positively charged ³¹P donor provides the confining potential for the electron. The system is convenient because it avoids the need for bandgap engineering, though actual devices do contain gate electrodes to manipulate the confining potential in time. The group of Andrea Morello at UNSW has shown that individual ³¹P nuclear spins can provide a nuclear-spin qubit with an exceedingly long T_2^* of 0.6 s.

Quantum-dot and donor qubits in ²⁸Si behave in many respects like isolated electrons trapped in a vacuum, and they allow for extremely high single-qubit control fidelity. In contrast to quantum-dot lithography, ion implantation produces an uncertainty that makes it challenging to position multiple donors with respect to each other. The group of Michelle Simmons, also at UNSW, has shown that scanning tunneling microscope lithography can position atoms with much higher precision than is possible through implantation.

With isotopically enriched ²⁸Si now available on wafer scales and at moderate costs, and with several methods available to confine electron spins in electronic devices, the

SEMICONDUCTOR SPINS

via simultaneous 90-degree rotations of each qubit from about 100 ns to 200 ns in the circuit. Next, a unitary transformation U_f is executed that corresponds to calling the function f from about 200 ns to 350 ns.

Because the qubits are in superposition, the function is evaluated for all four of its input values (00, 01, 10, and 11) in a quantum superposition as well. The function call is implemented with a two-qubit gate, which flips the phase of the $|x_0\rangle$ component in the superposition. At that point in the circuit, all probabilities remain $\frac{1}{4}$, as shown in figure 4. Subsequent single-qubit and two-qubit operations, identical for the four cases, boost the amplitude of the term $|x_0\rangle$ using quantum interference at the expense of the other terms.

Networked qubit registers

The two-qubit experiment can be scaled up to a few dozen qubits in linear arrays of quantum dots. Researchers, most notably at CNRS Grenoble, have already gone beyond 1D arrays and reported the first demonstrations of small 2D arrays of quantum dots. But limits exist to the number of tunnel-coupled quantum dots that can be realistically integrated monolithically. To scale up further, it is likely that on-chip quantum links will be required to connect distant quantum registers with each other, forming networks of interconnected multiqubit registers.

Many proposals exist for making such links, and their realization is an active area of research. One heavily pursued approach uses microwave photons stored in on-chip superconducting resonators to indirectly mediate the coupling between distant spins on the chip. Adopting that tack, three groups made a major breakthrough in their recent observation of so-called strong coupling of a single microwave photon and an electron spin qubit (see reference 12 and PHYSICS TODAY, April 2018, page 17). A second promising approach is to apply periodic gate voltages to induce a traveling-wave potential that shuttles electrons through channels across the chip. Initial results on quantum-dot arrays indicate that spin coherence can be preserved during such shuttling.¹³

Challenges in scaling up

Low fabrication yield still slows progress in many labs, and working devices are not all identical. Researchers must compensate for disorder in the form of charged defects and impurities in the semiconductor by tweaking the gate voltages. That's time-consuming, and low-frequency charge noise makes frequent retuning necessary. Furthermore, high-frequency charge noise limits the two-qubit gate fidelity. Nevertheless, the first experiments achieved two-qubit gate fidelities of 92–98% under suboptimal conditions, and 99% fidelity seems within reach.¹⁴

Recent experiments have shown encouraging improvements in charge noise. And yield, qubit uniformity, and charge noise are expected to benefit from industrial efforts to fabricate

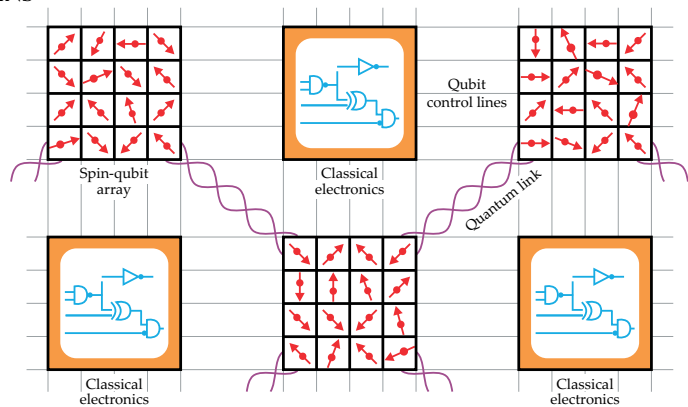


FIGURE 5. A VISION OF FUTURE SILICON QUANTUM ELECTRONICS, containing dense local registers of quantum dots interconnected with quantum links. Classical electronics between the spin-qubit arrays distribute signals on the chip.

quantum-dot arrays using commercial methods. The work is ongoing at the CEA's Leti Institute, an electronics information technology laboratory in Grenoble, France; at Imec, headquartered in Belgium, using electron-beam lithography; and at Intel Corp using all-optical lithography (see page 38).

Another challenge comes from the nature of Si, whose conduction band has six degenerate minima, or valleys, in the bulk. The degeneracy is problematic for spin-qubit operation because the Pauli exclusion principle, which normally forbids two electrons with the same spin to occupy the orbital ground state, gets circumvented and the two-qubit gate fails.

Confined structures such as quantum dots lift that sixfold degeneracy. But the so-called valley splitting—the energy gap to the first excited valley state—depends strongly on atomic-scale details that are locked in during growth and that can vary across a sample. In some of the Si/SiGe quantum dots measured to date, the valley splitting is too small to be useful. In contrast, a metal-oxide semiconductor quantum dot can have large valley splitting because of the hard confinement from the silicon oxide layer. The flip side is that this same oxide interface is a source of disorder that is larger than the disorder at the epitaxial interface of Si/SiGe quantum wells.

Scaling challenges can also arise at higher levels in the system—from the control electronics to the quantum-computer architecture and software layers. For example, every quantum dot (and superconducting qubit) made today requires that at least one wire be connected off-chip, which presents a wiring bottleneck for going beyond a few thousand qubits. To overcome the bottleneck, we envision two solutions that work in tandem: crossbar addressing schemes, like those used in displays and memory chips, and on-chip classical multiplexing circuits to distribute signals.¹⁵

A vision of qubit registers

Imagine a large-scale Si chip consisting of local 2D quantum-dot arrays addressed using crossbars and classical multiplexing electronics that are connected by quantum links.¹⁵ Figure 5


depicts what such a network of quantum and classical electronics might look like.

Si spin qubits are particularly well suited to realize that vision. First, the quantum dots, quantum links, and classical on-chip electronics can all be integrated using the same process steps. Those parts, moreover, can leverage today's transistor technology. Second, with a typical spacing of 100 nm, quantum dots are extremely compact: 1000 dots can fit inside an area of $10\ \mu\text{m}^2$. Third, Si spin-qubit coherence times are extremely long and can accommodate sequential operations on the qubits, which may be needed using crossbar addressing schemes. Fourth, Si spin qubits are resilient to temperature and suffer only modest degradation of charge noise and spin-relaxation times between 20 mK and 1 K.

Those are significant assets for scaling up Si spin qubits into a truly integrated circuit of quantum and classical components on a single chip. Scientific and technological challenges remain, but the prospect is very real that Si spin qubits may be scaled up to the many millions of qubits that will likely be needed to solve real-world problems beyond the reach of any classical machine. For example, a large-scale quantum computer will be capable of efficiently computing the properties of materials and molecules, with possible applications ranging from energy harvesting and storage to the design of drugs and catalysts.

We thank our many colleagues at TU Delft, the University of Wisconsin–Madison, and around the world for numerous collaborations and productive discussions.

REFERENCES

1. D. Loss, D. P. DiVincenzo, *Phys. Rev. A* **57**, 120 (1998).
2. D. P. DiVincenzo, *Fortschr. Phys.* **48**, 771 (2000).
3. J. R. Petta et al., *Science* **309**, 2180 (2005).
4. J. Medford et al., *Phys. Rev. Lett.* **111**, 050501 (2013); K. Eng et al., *Sci. Adv.* **1**, e1500214 (2015).
5. D. Kim et al., *Nature* **511**, 70 (2014); D. Kim et al., *NPJ Quantum Inf.* **1**, 15004 (2015).
6. M. D. Shulman et al., *Science* **336**, 202 (2012); J. M. Nichol et al., *NPJ Quantum Inf.* **3**, 3 (2017).
7. C. B. Simmons et al., *Appl. Phys. Lett.* **91**, 213103 (2007); M. G. Borselli et al., *Appl. Phys. Lett.* **99**, 063109 (2011).
8. S. J. Angus et al., *Nano Lett.* **7**, 2051 (2007); A. Fujiwara et al., *Appl. Phys. Lett.* **88**, 053121 (2006).
9. M. Veldhorst et al., *Nat. Nanotechnol.* **9**, 981 (2014).
10. J. Yoneda et al., *Nat. Nanotechnol.* **13**, 102 (2018).
11. M. Veldhorst et al., *Nature* **526**, 410 (2015); D. M. Zajac, *Science* **359**, 439 (2018); T. F. Watson et al., *Nature* **555**, 633 (2018).
12. N. Samkharadze et al., *Science* **359**, 1123 (2018); X. Mi et al., *Nature* **555**, 599 (2018); A. J. Landig et al., <https://arxiv.org/abs/1711.01932>.
13. H. Flentje et al., *Nat. Commun.* **8**, 501 (2017); T. Fujita et al., *NPJ Quantum Inf.* **3**, 22 (2017).
14. W. Huang et al., *Nature* **569**, 532 (2019); X. Xue et al., *Phys. Rev. X* **9**, 021011 (2019).
15. L. M. K. Vandersypen et al., *NPJ Quantum Inf.* **3**, 34 (2017).
16. K. Ono et al., *Science* **297**, 1313 (2002); J. M. Elzerman et al., *Nature* **430**, 431 (2004).
17. F. H. L. Koppens et al., *Nature* **442**, 766 (2006).
18. K. C. Nowack et al., *Science* **318**, 1430 (2007); M. Pioro-Ladrière et al., *Nat. Phys.* **4**, 776 (2008); R. Maurand et al., *Nat. Commun.* **7**, 13575 (2016); R. Jock et al., *Nat. Commun.* **9**, 1768 (2018). 

B 2 Topological Quantum Computers

Fabian Hassler

JARA-Institute for Quantum Information

RWTH Aachen University

Contents

1	Introduction	2
2	Topological superconductors	3
2.1	Topological quantum number	3
2.2	Kitaev model	5
2.3	Majorana zero modes	7
3	Majorana qubits	7
3.1	Fermionic quantum computation	8
3.2	Encoding of a qubit	9
3.3	Fusion and splitting	10
4	Braiding	11
4.1	Abelian anyons	13
4.2	Ising anyons	14
4.3	Fibonacci anyons	15
5	Conclusion	19

1 Introduction

The concept of *identical particles* is one of the most counterintuitive features of many-body quantum mechanics, maybe second only to the concept of entanglement. In fact, even the bosonic exchange statistics, which is the closest to our classical world, has puzzled researchers time and time again. The hallmark experiment by Hanbury Brown and Twiss [1], showing two photon interference, was initially facing strong criticism: Brannen and Fergusen for example stated that “it would appear to the authors [...] that if such a correlation did exist, it would call for a major revision of some fundamental concepts in quantum mechanics” [2]. As a reply, Purcell did point out that “the electromagnetic field is a classical field after all, which is why the Brown-Twiss effect only appears odd if one looks at it from a particle point of view; its oddness being simply the *peculiarity of bosons*” [3]. Going over to fermions, we do not even have a classical analog. The basic principle of fermions is that they are described by anticommuting numbers. However, all observables, being part of our everyday classical world, are bosonic. The consistency with quantum mechanics then demands ‘*superselection*’, which means that all observables have to be formed by an even number of fermionic operators.

The idea of using identical particles and their exchange statistics as a resource for quantum computation is only a couple of decades old [4, 5]. The reason is that ‘simple’ fermions and bosons, which we are familiar with from the basic physics courses, are not useful for this task. Researchers have been looking at extensions of the concept of exchange statistics of identical particles beyond the notion of fermions and bosons. However, in three dimensions nothing interesting arises as all possible parastatistics can be reduced to bosons and fermions.

In two dimension, we know that charged particles (with charge q) pick up the Aharonov-Bohm phase $q\Phi/\hbar$ when encircling a magnetic flux Φ . The Aharonov-Bohm phase is *topological* in the sense that it does not depend on the concrete trajectory taken by the particles but only on the winding of the particle around the flux. Composite (bosonic) particles consisting of an electric charge q and a magnetic flux Φ acquire a phase $\theta = q\Phi/2\hbar$ when exchanged.¹ The value $\theta = 0$ ($\theta = \pi$) corresponds to a composite boson (fermion). However, also other values are allowed which correspond to Abelian anyons. In three dimension, the construction does not work. In fact, it is a theorem by Dirac [6] that a consistent theory only allows for pointlike magnetic charges (called magnetic monopoles) that produce a flux of size $\Phi_0 = n \times 2\pi\hbar/q$ leading to $\theta = n\pi$ ($n \in \mathbb{Z}$). We can understand Dirac’s argument as follows: since there is no well-defined winding number between two point particles in three dimensions, there can be no topological phase which restricts the charge of potential magnetic monopoles.

The general idea how to use identical particles for topological quantum computation is the following. In the standard gate model of quantum computers, a calculation consists of three steps: (1) initialization in the state $|i\rangle = |0, 0, \dots\rangle$, (2) application of a gate U (a general unitary operator) on $|i\rangle$ which produces $|f\rangle = U|i\rangle$, (3) measurement of the outcomes $o_j \in \{0, 1\}$ in the computational basis with probabilities $P(o_1, o_2, \dots) = |\langle o_1, o_2, \dots | f \rangle|^2$. In a topological quantum computer, the three steps are replaced by operations involving anyons [7]: (1) the initialization is replaced by splitting a pair of anyons out of the vacuum, (2) a gate is done by *braiding* the anyons around each other, (3) the measurement is performed by fusion (pairwise annihilation), see Fig. 4(c).

¹ The factor $\frac{1}{2}$ arises as exchanging the particle only moves them halfway around each other.

To obtain nontrivial gates, braiding should perform a unitary gate on the ground state manifold. Since Abelian anyons only produce phases, they are not useful for this task but rather non-Abelian anyons are needed. Note that the computation on non-Abelian anyons is topological. The trajectories (worldlines) in the process splitting-braiding-fusion form a knot in space-time. The result of the computation does not depend on the concrete form of the worldlines but only on the topology of the knot. It is in this respect, that the computation performed by non-Abelian anyons is robust.

The outline of the lecture is as follows: first, we will introduce the notion of a topological superconductor. Then, we will show that Majorana zero modes appear as zero energy solutions of the Bogoliubov-de Gennes equation describing a spinless p -wave superconductor in one dimension. The zero modes will turn out to be non-Abelian anyons called Ising anyons. However, note that braiding of Ising anyons does not result in the application of any arbitrary unitary gate. Thus, we will introduce an alternative class of anyons called ‘Fibonacci anyons’. We will show that they are in fact universal, i.e., an arbitrary computation can be performed by braiding. There are by now several reviews where further information on these subjects can be found. For Majorana zero modes, see for example Refs. [8–10]. More information about topological quantum computation can be found in Refs. [7, 11–14].

2 Topological superconductors

In the context of quantum mechanics, topology refers to a static system described by a Hamiltonian where the properties of certain eigenstates are insensitive to small perturbations (disorder). In order that these states are not affected by hybridization with other states of the system, the energy necessarily has to be in an energy gap (bandgap) of the system such that direct coupling is forbidden by energy conservation.

2.1 Topological quantum number

A prime example for a topological system in one dimension (1D) is given by the Jackiw-Rebbi (JR) model (1976) with the Hamiltonian [15]

$$H_{\text{JR}} = v_F p \sigma^z + M(x) \sigma^x = \begin{pmatrix} -i\hbar v_F \partial_x & M(x) \\ M(x) & i\hbar v_F \partial_x \end{pmatrix}, \quad (1)$$

where σ^j denote the Pauli matrices which represent an (artificial) spin degree of freedom. The first term describes the motion of a particle with velocity $\pm v_F$ and momentum $p = -i\hbar \partial_x$. The particle is chiral as the direction ‘ \pm ’ of motion depends on the spin. The second term proportional to $M(x)$ leads to backscattering (as it couples the two directions of motion) and the appearance of a ‘mass’, i.e., a gap in the energy spectrum, see below.

Let us first discuss the case where the mass does not depend on position with $M(x) \equiv \bar{M}$. In this case, the system is translation-invariant and we can find eigenstates of the form

$$\psi_k = e^{ikx} \mathbf{v}_k. \quad (2)$$

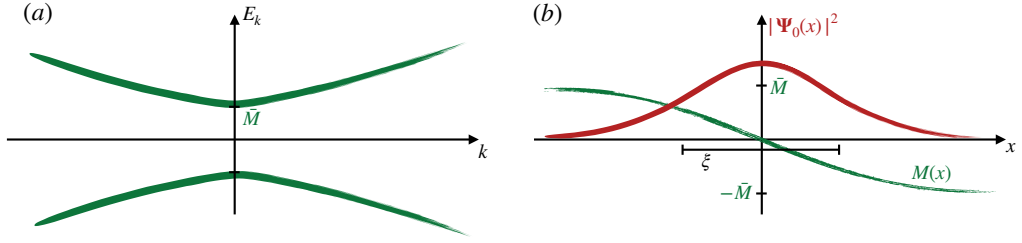


Fig. 1: (a) Spectrum of the Jackiw-Rebbi model. It consists of two bands, one at positive and one at negative energies E_k . For large momentum $\hbar k$, the spectrum assumes the form $E_k = \pm v_F \hbar |k|$. Close to $k \approx 0$, there is an anticrossing with a bandgap of size $2\bar{M}$. (b) Probability distribution $|\Psi_0(x)|^2$ for finding the particle in the bound state with $E = 0$ at the position x . It is peaked at the position $x \approx 0$ where the mass changes sign. The decay happens on the characteristic scale $\xi = \hbar v_F / \bar{M}$ proportional to the inverse gap.

Solving the eigenvalue equation $H_{\text{JR}}\psi_k = E_k\psi_k$ in this case, leads to the result

$$E_k = \pm \sqrt{(\hbar v_F k)^2 + \bar{M}^2}, \quad \mathbf{v}_k = \begin{pmatrix} E_k - k \\ \bar{M} \end{pmatrix}. \quad (3)$$

Looking at the spectrum, cf. Fig. 1(a), we see that the JR Hamiltonian is a model for a 1D semiconductor with a bandgap of $2|\bar{M}|$. At first sight, it looks like the sign of \bar{M} does not matter.

However, let us see what happens if we bring a semiconductor with a positive bandgap $\bar{M} > 0$ in proximity to a semiconductor with a negative bandgap $-\bar{M}$. We model this system by a mass $M(x)$ that depends on position and that assumes the asymptotic values $M(x) \rightarrow \pm\bar{M}$ for $x \rightarrow \mp\infty$, see Fig. 1(b). We claim that there will be in this case a single state with energy $E = 0$. We show this by directly computing the eigenstate, i.e., solving the problem

$$H_{\text{JR}}\Psi_0 = -i\hbar v_F \sigma^z \partial_x \Psi_0 + M(x) \sigma^x \Psi_0 = 0. \quad (4)$$

After a slight rearrangement using $\sigma^z \sigma^x = i\sigma^y$, we obtain

$$\frac{d}{dx} \Psi_0(x) = \kappa(x) \sigma^y \Psi_0(x), \quad \text{with} \quad \kappa(x) = \frac{M(x)}{\hbar v_F}. \quad (5)$$

with the solutions

$$\Psi_0 = \mathcal{N} \exp \left[\pm \int_0^x dx' \kappa(x') \right] \chi_{\pm y}, \quad \chi_{\pm y} = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{i\pi/4} \\ \pm e^{-i\pi/4} \end{pmatrix}, \quad (6)$$

where $\mathcal{N} > 0$ is the normalization constant and $\chi_{\pm y}$ denote the eigenvectors of σ^y to the eigenvalues ± 1 .

For our choice of $M(x)$, we have that

$$\int_0^x dx' \kappa(x') \rightarrow \mp\infty, \quad x \rightarrow \pm\infty. \quad (7)$$

In order that the state Ψ_0 is normalizable, we are only allowed to take the ‘+’ sign in Eq. (6). Concluding, we have found the bound eigenstate $\Psi_0(x) \propto \exp\left[\int_0^x dx' \kappa(x')\right] \chi_{+y}$ of the JR model at energy zero. The rest of the spectrum consists of extended states above the gap with energy $|E| \geq \bar{M}$. The energy of the state Ψ_0 is within the gap.

Recapitulating the argument that has lead to the identification of Ψ_0 , we realize that the presence of the state is independent of the concrete form of $M(x)$ and only depends on the sign of the asymptotic values for $x \rightarrow \pm\infty$. It is in this sense, that the state Ψ_0 is topological and insensitive to disorder. In fact, one calls $\mathcal{Q} = \text{sgn } M$ a (\mathbb{Z}_2) topological charge that can take values ± 1 . Bringing two systems with opposite topological charge in proximity, a bound state within the gap is trapped at the interface. This state is insensitive to disorder as its presence is guaranteed by the properties (topological charge) of the bulk away from the interface.

2.2 Kitaev model

Kitaev (2001) introduced a Hamiltonian (called the Kitaev model) that implements the JR model in a one-dimensional (1D) superconducting system. The model is given by [16]

$$H_K = \sum_p \xi_p c_p^\dagger c_p + \frac{1}{2} \sum_p \Delta p (c_{-p} c_p + c_p^\dagger c_{-p}^\dagger), \quad (8)$$

where c_p are fermionic annihilation operators that obey the canonical anticommutation relation $\{c_p, c_q^\dagger\} = \delta_{p,q}$, $\{c_p, c_q\} = 0$. The first term described spinless electrons with momentum p and energy $\xi_p = p^2/2m - \mu$ relative to the chemical potential μ . The second term proportional to $\Delta > 0$ describes the superconduction pairing.² As we will see below, the fact that the problem is realized in a superconductor has crucial implications on the properties of the bound state Ψ_0 . In fact, it turns out that the resulting excitations at the interface are Majorana zero modes which follow non-Abelian statistics that are useful for quantum computing.

In order to make contact to the JR model, we write the model in Nambu space

$$H_K = \frac{1}{2} \sum_p C_p^\dagger H_{\text{BdG}}(p) C_p \quad (9)$$

with $C_p = (c_p, c_{-p}^\dagger)^T$ and where we have introduced the Bogoliubov-de Gennes Hamiltonian

$$H_{\text{BdG}}(p) = \xi_p \tau^z + \Delta p \tau^x \quad (10)$$

with τ^j Pauli matrices acting on the Nambu space. Due to the fact, that the Hamiltonian H_{BdG} is derived from H_K by doubling the degrees of freedom when going over from c_p to C_p , it enjoys the relation $\tau^x H_{\text{BdG}}(-p) \tau^x = -H_{\text{BdG}}(p)$. As a result, each eigenmode of H_K leads to two eigenstates: for each eigenstate $\Psi_E(p)$ of H_{BdG} at energy E , there is a state $\tau^x \Psi_E(-p)$ at energy $-E$.

As before, we are interested in states close to $E = 0$. The Hamiltonian H_{BdG} allows for such a state only when $p = 0$ (such that the second term vanishes) and $\mu = 0$ (such that the first term

² The pairing is of p -type as there is no s -wave pairing for spinless electrons.

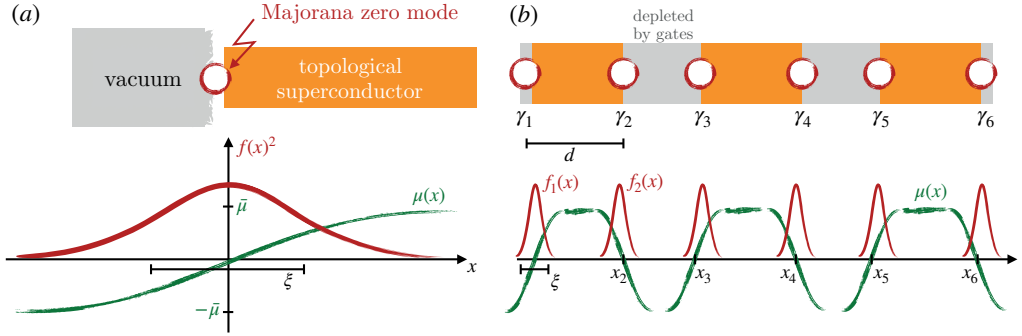


Fig. 2: (a) A Majorana zero mode appears at the interface between a 1D topological superconductor and the vacuum. The vacuum is modeled by a negative chemical potential μ such that there are no electrons present. (b) Segments of a topological superconductor can be depleted by gates. In our effective model, the gates lead to a negative chemical potential. At each domain wall (interface between the ‘vacuum’ and the superconductor) at position x_j a Majorana zero mode emerges. Due to the overlap proportional to $\exp(-d/\xi)$ with $\xi = \hbar\Delta/\bar{\mu}$, these modes are not exactly at zero energy but at an energy $\epsilon \simeq \bar{\mu}e^{-d/\xi}$.

vanishes). To describe the physics near the band closing at $\mu = 0$ and $p = 0$, we expand around $p = 0$ ³

$$H_{\text{BdG}} \approx -\mu\tau^z + \Delta p\tau^x. \quad (11)$$

We see that H_{BdG} is in fact the JR model with the replacement

$$\sigma^z \mapsto \tau^x, \quad v_F \mapsto \Delta, \quad (12)$$

$$\sigma^x \mapsto -\tau^z, \quad M \mapsto \mu. \quad (13)$$

With that, we can identify the sign of μ with the topological charge. Indeed, the Kitaev model H_K describes a topological superconductor when $\mu > 0$. For $\mu < 0$, the electrons are depleted from the wire and the state is a conventional insulator.

The mapping of H_{BdG} onto the JR-model allows to predict that an end state will appear at the interface of a topological superconductor ($\mu > 0$) to vacuum ($\mu < 0$). In particular, we model the situation of Fig. 2 by a chemical potential $\mu(x)$ that changes sign at $x = 0$. The Hamiltonian H_{BdG} has the bound state

$$\Psi_0(x) = f(x)\chi_{+y}, \quad f(x) = \mathcal{N} \exp \left[\int_0^x dx' \mu(x')/\hbar\Delta \right] \quad (14)$$

at energy $E = 0$ that is localized close to the interface at $x = 0$. It is separated by a gap $\bar{\mu}$ from the extended states.

³ Treating the rest as a perturbation that can be included in principle later on.

2.3 Majorana zero modes

As we have seen before, a single eigenmode of H_K with energy $E > 0$ corresponds to two states of H_{BdG} with energy $\pm E$. It is thus interesting to understand what the single state of H_{BdG} at $E = 0$ corresponds to. To this end, we calculate the second-quantized operator

$$\gamma = \sqrt{2} \sum_p \int dx \Psi_0(x)^* \cdot C_p e^{ipx/\hbar} = \int dx f(x) \left[e^{i\pi/4} \psi(x) + e^{-i\pi/4} \psi^\dagger(x) \right] \quad (15)$$

that corresponds to the eigenstate Ψ_0 ; here, $\psi(x) = \sum_p e^{ipx/\hbar} c_p$ is the field operator that annihilates a particle at the position x and fulfills the anticommutation relations $\{\psi(x), \psi^\dagger(x')\} = \delta(x - x')$, $\{\psi(x), \psi(x')\} = 0$.

The fermionic mode γ is special in that it is Hermitian with $\gamma^\dagger = \gamma$. In physical terms, this means that annihilating an excitation in the mode γ is the same as creating one. Such an excitation is only possible at zero energy which corresponds to the chemical potential of the superconductor. Due to the aforementioned Hermiticity, the excitation described by γ is called *Majorana zero mode*. With a straightforward calculation, one can show that $\gamma^2 = \gamma^\dagger \gamma = 1$. Generalizing these properties to multiple points x_j where the chemical potential μ changes sign, we obtain a Majorana zero mode γ_j at each interface (domain wall) at x_j . These modes fulfill the Clifford algebra

$$\{\gamma_i, \gamma_j\} = 2\delta_{i,j} \quad (16)$$

which encodes the normalization $\gamma_j^2 = 1$ together with the fermionic statistics $\gamma_i \gamma_j = -\gamma_j \gamma_i$.

As we have seen before, a single Majorana zero mode γ that is formed at the position where the chemical potential changes sign is pinned at zero energy. We can understand this in algebraic terms by noting that the effective low energy Hamiltonian H_{eff} can only involve the excitation γ as the other excitations are at the energy $\bar{\mu}$. Due to superselection (see below), each term of a Hamiltonian has to involve an even power of fermionic operators. However, all such terms are trivial as $\gamma^{2n} = (\gamma^2)^n = 1$ so the effective Hamiltonian vanishes.

This argument does not work any more as soon as two Majorana zero modes γ_1 and γ_2 are present. In fact, there is only one Hermitian combination $i\gamma_1\gamma_2$ that involves an even number of fermionic operators. So the effective Hamiltonian has to be of the form

$$H_{\text{eff}} = \epsilon i\gamma_1\gamma_2. \quad (17)$$

In fact, the energy ϵ is approximately given by the overlap

$$\epsilon \approx \langle f_1 | H_{\text{BdG}} | f_2 \rangle \simeq \bar{\mu} \int dx f(x - x_1)^* f(x - x_2) \simeq \bar{\mu} e^{-d/\xi} \quad (18)$$

with $d = |x_2 - x_1|$ the distance between the zero modes and $\xi = \hbar\Delta/\bar{\mu}$ the decay length that is proportional to the inverse gap.

3 Majorana qubits

We have seen that the low energy properties of a topological superconductor are given by the Majorana zero modes γ_j that are pinned at the domain walls where the sign of μ changes. We

are interested in the Hilbert space that can be accessed by acting with the modes γ_j on the ground state $|0\rangle$. It is an easy exercise in algebra to show that given a set of $2N$ Majorana zero modes, we can construct N Dirac fermions c_j via

$$c_j = \frac{1}{2}(\gamma_{2j-1} + i\gamma_{2j}), \quad c_j^\dagger = \frac{1}{2}(\gamma_{2j-1} - i\gamma_{2j}). \quad (19)$$

The Dirac fermions fulfill the canonical anticommutation relations $\{c_i, c_j^\dagger\} = \delta_{ij}$ and $\{c_i, c_j\} = 0$.

The Hilbert space of a single fermionic mode ($N = 1$) is two-dimensional: the mode is either filled or empty distinguished by the eigenvalue of the number operators $n_j = c_j^\dagger c_j$ which have eigenvalues 0 or 1.⁴ Operators which will turn out to be important in the following discussion are the fermion parity operators $\mathcal{P}_j = 1 - 2n_j = (-1)^{n_j}$ which have the eigenvalue $+1$ if the number of fermions is even and -1 if the number of fermions is odd. In terms of the Majorana operators, the parity operators assume the simple form

$$\mathcal{P}_j = -i\gamma_{2j-1}\gamma_{2j}. \quad (20)$$

If we think about an implementation for a quantum computer, we are used to the example of a spin- $\frac{1}{2}$ particle which is a model system for a generic two-level system [17]. However, we can ask ourself the question whether we can also use the many-body Fock space for quantum computation purposes. We know that the occupation states $|n_1, n_2, \dots, n_N\rangle$ with $n_j \in \{0, 1\}$ form a basis for the N -mode fermionic Fock space generated by the creation operators c_j^\dagger , $j \in \{1, \dots, N\}$, starting from the vacuum state denoted by $|0\rangle$. The Fock space has dimension 2^N (each mode can be either occupied or empty). Thus counting the degrees of freedom, we are tempted to conclude that a fermionic system with N -modes emulates N -qubits. In the next section, we will see that this naïve counting argument is not completely correct as it violates the so-called superselection rule.

3.1 Fermionic quantum computation

Expressing a Hamiltonian H or in fact any physical observable A which are bosonic operators in terms of fermionic creation and annihilation operators, we are bound to only include terms where an even number of fermion operators appear.⁵ The result is that the total fermion parity $\mathcal{P} = \prod_j \mathcal{P}_j = (-1)^{\sum_j n_j}$ is strictly conserved in a closed system; the reason for this is the fact that

$$\mathcal{P}A\mathcal{P} = A \quad (21)$$

which follows from $\mathcal{P}c_j\mathcal{P} = -c_j$ and the fact that each term in A involves an even number of fermionic operators. Note that the superconducting Hamiltonian (8) conserves the total fermion parity even so the number of fermions is not conserved. Due to this constraint, we have the

⁴ Note that n_j is idempotent as $n_j^2 = c_j^\dagger c_j c_j^\dagger c_j = c_j^\dagger (1 - c_j^\dagger c_j) c_j = n_j$ which proves the fact that the eigenvalues of n_j are 0 or 1.

⁵ From the correspondence principle, we know that for large quantum numbers the expectation values of operators for physical observables should behave like (real) numbers. Due to the anticommutation relation of fermionic operators, the correspondence principle for a potential fermionic observable would instead lead to anticommuting Grassmann numbers on the classical level.

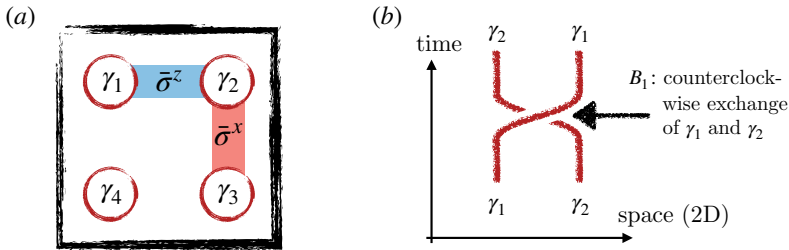


Fig. 3: (a) Sketch of the parity Majorana qubit: Two Majorana zero modes together form a single Dirac fermionic mode whose Hilbert space is two-dimensional as the mode can either be empty or filled. Both states have the same energy. Four Majorana zero modes thus form a four-dimensional Hilbert space of which due to the conservation of the total fermion parity only a two-dimensional subspace can be accessed. This degenerate two-dimensional subspace is the Majorana qubit. Gates on the qubit can be either performed by braiding or by coupling two Majorana zero modes. As indicated in the figure, coupling γ_1 to γ_2 implements a $\bar{\sigma}^z$ -operation whereas coupling γ_2 to γ_3 leads to a $\bar{\sigma}^x$ -operation. Given the fact that the Majorana zero modes are sufficiently far apart from each other and that the environment only acts locally on the system, these operations are not performed ‘accidentally’ by the environment and the Majorana qubit is protected from both sign flip and bit flip errors. As these protection originates from the conservation of the total fermion parity, the qubit is called parity-protected. (b) Elementary operation of the braid group. The geometric representation of the braid group is in space-time; the horizontal axis is the spacial axis whereas the vertical one is temporal. The counterclockwise exchange of Majorana zero modes γ_1 and γ_2 in space-time forms the braid B_1 .

following superselection rule: given two states in a fermionic Fock space $|\psi_+\rangle$ and $|\psi_-\rangle$ with different fermion parity, $\mathcal{P}|\psi_\pm\rangle = \pm|\psi_\pm\rangle$ we have

$$\langle\psi_-|A|\psi_+\rangle = \langle\psi_-|\mathcal{P}A\mathcal{P}|\psi_+\rangle = -\langle\psi_-|A|\psi_+\rangle = 0 \quad (22)$$

for all observables A . Thus, there is no point in making superpositions between states of different parity as there will be no effect on any observable. We can thus restrict ourselves to one superselection sector and keep the total fermion parity fixed with either $\mathcal{P} = +1$ or $\mathcal{P} = -1$. The conclusion of this argument is that out of the 2^N states in a fermionic Fock space, only 2^{N-1} can be effectively used for quantum computation purposes.

A further restriction to quantum computation using fermions arises from the fact that noninteracting fermions subject to beam splitters, phase-shifters (delay lines), measurements of the state of a single electron (so-called fermionic linear optics) does in fact not lead to any entanglement [18]. In order to generate entanglement, we need to add parity measurement of two electrons which effectively involves interactions between different electrons [19, 20].

3.2 Encoding of a qubit

We have seen in the last section that due to the parity-conservation, we need to have two fermionic modes to encode a single qubit. For concreteness, we will work in the even par-

ity superselection sector and have the single logical qubit encoded as $|\bar{0}\rangle = |00\rangle$ and $|\bar{1}\rangle = |11\rangle$. Thinking about a possible implementation in terms of Majorana modes, we encode each fermionic mode in a pair of Majorana zero modes which are localized states sufficiently far separated from each other, see Fig. 3. We denote the Majorana zero modes on the top as γ_1 and γ_2 and the one on the bottom as γ_3 and γ_4 , correspondingly. The Majorana modes are at zero energy thus the two states $|\bar{0}\rangle$ and $|\bar{1}\rangle$ are degenerate in energy. The parity of the number of electrons on the superconducting segments are given by $\mathcal{P}_1 = -i\gamma_1\gamma_2$ and $\mathcal{P}_2 = -i\gamma_3\gamma_4$. Due to the parity constraint, we have $\mathcal{P}_1 = \mathcal{P}_2$ and the action of both operators on the logical qubit emulates the σ^z Pauli-operator,

$$\bar{\sigma}^z = -i\gamma_1\gamma_2 = -i\gamma_3\gamma_4. \quad (23)$$

In order to have a complete qubit, we are left with the task to find a logical $\bar{\sigma}^x$, an operator which anticommutes with $\bar{\sigma}^z$. It is easy to see that

$$\bar{\sigma}^x = -i\gamma_2\gamma_3 = -i\gamma_1\gamma_4 \quad (24)$$

anticommutes with $\bar{\sigma}^z$ due to the fact that the single Majorana fermions shared by both operators anticommute with each other. In the situation where all the Majorana modes are sufficiently far separated from each other, either gate on the logical qubit is a nonlocal operator. Due to this nonlocality, it is highly unlikely that uncontrolled, random fluctuations in the environment will execute a gate and thus act as an error on the logical qubit. This protection of the Majorana qubit is called symmetry-protected topological order [21, 22] or simply parity-protection [23]. The decisive difference to full topological order, as it is for example present in Kitaev's toric code [24], is the fact that logical Pauli operators are only required to be nonlocal as long as the parity symmetry is conserved. Having a reservoir tunneling single electrons on the superconducting island is a local process which violates the parity-conservation and immediately brings the Majorana qubit out of its computational subspace.

The requirement for operating the Majorana qubit successfully in a protected manner is that the environment does not provide single unpaired electrons. This sounds on the first sight very stringent. However, the physical implementation of the system does only involve superconductors where most of the electrons are paired up into Cooper pairs and where at temperature T only a exponentially small fraction proportional to the Boltzmann factor $e^{-\Delta/k_B T}$ remains unpaired. The storage time of quantum information in a Majorana qubit thus is expected to increase exponentially when lowering the electron temperature.

3.3 Fusion and splitting

We have seen before that if you bring the Majorana zero modes γ_1 and γ_2 close together, there is a finite energy splitting $\epsilon \simeq \bar{\mu}e^{-d/\xi}$ due to the overlap of the modes, see Fig. 4(a). This splitting breaks the parity protection and as a result the two states $|0\rangle$ and $|1\rangle$ are no longer degenerate in energy. It is then possible to detect in which state the system is [12, 23, 25]. The process of bringing two Majorana modes together is called fusion. Graphically, we denote the fusion experiment as the red box in Fig. 4(b). The fusion experiment implements the measurement in the computational basis and is the last step in a quantum computation.

The reverse process is called splitting and is depicted as the blue box in Fig. 4(b). If one starts with the vacuum state $|0\rangle$ when the Majorana zero modes are overlapping, the Majorana states

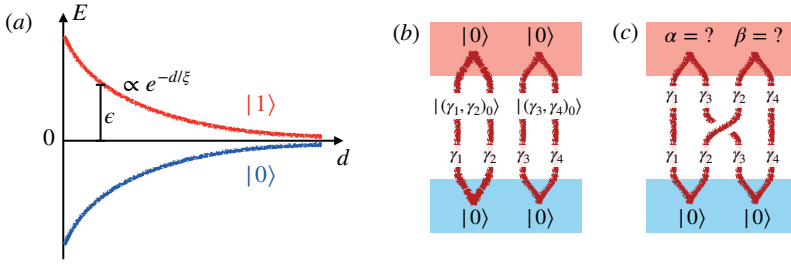


Fig. 4: (a) Bringing two Majorana zero modes in close proximity (called fusion) such that their distance d is of the decay length ξ , the energy difference $\propto \exp(-d/\xi)$ between the state $|0\rangle$ and $|1\rangle$ breaks the parity protection. By measuring whether or not the system is in the ground state, a projective measurement in the computational basis states $|0\rangle$ and $|1\rangle$ is performed. Similarly, we can split the Majorana zero modes that are initialized in the ground state $|0\rangle$ and produce a parity protected state. (b) The fusion outcome (red box) of Majorana zero modes that have been initialized to the vacuum (by splitting, blue box) is well-defined. (c) If the Majorana zero modes have been interchanged (braided), the fusion outcomes α, β fluctuate.

can be separated from each other such that the initial state $|0\rangle$ becomes degenerate with $|1\rangle$ and parity protection is achieved. It is clear that if one fuses the zero modes again immediately after the splitting, the outcome of the fusion experiment is the state $|0\rangle$ with certainty, see Fig. 4(b). On the other hand, if the pairing is changed such that at the fusion different Majorana zero modes are paired up as in the splitting, the result is unclear, see Fig. 4(c).

As we are now thinking about different ways of pairing the Majorana zero modes to produce Dirac fermion that can be either empty or occupied, it is useful to introduce the notation $|(\gamma_1, \gamma_2)_j\rangle$. It denotes whether the state formed by γ_1 and γ_2 is occupied ($j = 1$) or empty ($j = 0$). With this new notation, we can write for the logical state

$$|\bar{0}\rangle = |(\gamma_1, \gamma_2)_0 (\gamma_3, \gamma_4)_0\rangle, \quad |\bar{1}\rangle = |(\gamma_1, \gamma_2)_1 (\gamma_3, \gamma_4)_1\rangle. \quad (25)$$

The fusion experiment in Fig. 4(c) produces the outcome $\alpha, \beta \in \{0, 1\}$ with probability

$$P(\alpha, \beta) = |\langle (\gamma_1, \gamma_3)_\alpha (\gamma_2, \gamma_4)_\beta | (\gamma_1, \gamma_2)_0 (\gamma_3, \gamma_4)_0 \rangle|^2. \quad (26)$$

Due to the conservation of the total parity \mathcal{P} , the only possible outcomes are $\alpha = \beta = 0$ and $\alpha = \beta = 1$. We will calculate the probabilities in the next section. However, we have to first introduce the part of Fig. 4(c) in between the splitting and the fusion of the Majorana zero modes. The exchange of the zero modes is called braiding.

4 Braiding

Exchange statistics is introduced in the basic physics courses as the action of permutations of the symmetric group \mathcal{S}_N with $N!$ elements on the wavefunction of N identical particles. There are two possibilities: either the wavefunction remains invariant ('bosons') or the wavefunction acquires a minus sign ('fermions') under the exchange of two particles. On a mathematical

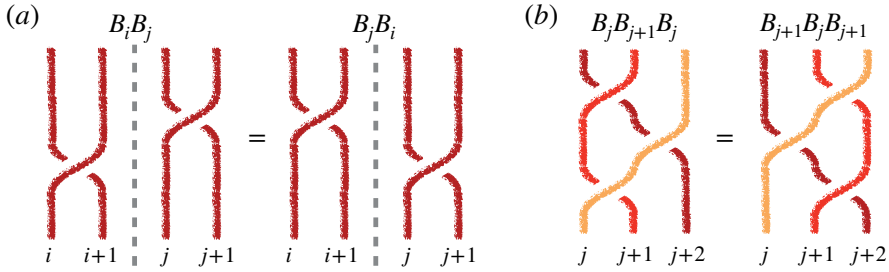


Fig. 5: (a) Braiding operators commute if their index i and j are more than 2 apart. The reason is that B_i and B_j do not have any strand in common so either one can be executed first. (b) The Yang-Baxter equation $B_j B_{j+1} B_j = B_{j+1} B_j B_{j+1}$ provides a nontrivial relation between the generators B_j and B_{j+1} . That the two braids are topologically equivalent can be seen as follows: in both braids the dark strand can be considered to lie in the very back and connects the initial position $j + 2$ to the final position j . Similarly, the bright strand lies in front and connects j to $j + 2$. The middle strand starts at $j + 1$ and ends at the same place. The braids are equivalent as they can be deformed into each other by sliding the middle strand $j + 1$ in between the two other strands from the left to the right.

level, the origin of this distinction lies in the fact that the Hamiltonian of identical particles commutes with an arbitrary element of the symmetric group S_N . Thus, it is possible to classify the eigenstates of the Hamiltonian in terms of irreducible representations of the permutation group. There are only two one-dimensional representations: the trivial representation (corresponding to bosons) and the sign representation (corresponding to fermions). Any representation whose dimension is larger than one leads to a degeneracy, which is called exchange degeneracy as it originates simply from the fact that particles are indistinguishable.⁶ In fact, the spin-statistic theorem can be proven in the context of relativistic field theory in 3+1 dimensions, which states that particles with integer spin are bosons whereas particles with half-integer spin are fermions.

It has been pointed out that the physical process of exchanging two particles is important. In fact, the exchange of two identical particles has to be viewed as a (slow) process that occurs in space-time [26]. In 2 + 1 dimensions, the relevant group is the braid group \mathcal{B}_N of N strands as trajectories in space-time for exchanging two particles clock- or counterclockwise are topological distinct. The braid group \mathcal{B}_N consists of N strands. The generators of the group B_j denote the braiding strand j and $j + 1$ in the counterclockwise direction (in Fig. 3(b), B_1 braids strand 1 and 2 counterclockwise). The inverse operation braids the strands clockwise. Two elements of the group are equivalent if the corresponding braids can be smoothly deformed into each other by keeping their ends fixed. Note that different from the symmetric group $B_j \neq B_j^{-1}$. The braid group of N strands is generated by B_1, \dots, B_{N-1} where the generators fulfill the following two

⁶ The statistics of identical particles which transforms according to higher dimension representations of the permutation groups is called parastatistics. However, even if particles with parastatistics were to exist they would offer nothing new as a set of Klein transformations could be used to map particles with parastatistics onto bosons or fermions with a set of internal quantum numbers (like spin, ...). Later, we will see that such a mapping is not possible in 2 + 1 dimensions and that higher dimensional representations of the braid group are truly different from the one-dimensional representations.

relations (Artin [27])

$$B_i B_j = B_j B_i, \quad |i - j| \geq 2 \quad \text{and} \quad B_j B_{j+1} B_j = B_{j+1} B_j B_{j+1}; \quad (27)$$

the latter is also called Yang-Baxter equation, see Fig. 5. Different from the symmetric group \mathcal{S}_N the group order is infinity which makes the classification of all irreducible representation difficult.

In $3 + 1$ dimensions, clock- and counterclockwise depends on the observer (coordinate system) and thus the exchanges B_j and B_j^{-1} are topologically equivalent. As a result, we obtain $B_j^2 = B_j^{-1} B_j = 1$. Under this additional constraint, the braid group \mathcal{B}_N reduces to the symmetric group \mathcal{S}_N . The dependence of the exchange statistic on the dimension of space is as follows:

- 1D: no exchange of particles is possible
- 2D: the exchange is described by the braid group \mathcal{B}_N
- 3D: we have $B_j^2 = 1$ and the symmetric group \mathcal{S}_N characterizes the exchange (only bosons and fermions are possible)

We see that 2D is special and we will study some examples of nontrivial exchange statistics in $2 + 1$ dimensions in the following.

4.1 Abelian anyons

The one-dimensional (unitary) representations of the braid group are simple to construct. In general, the action of B_j onto a wavefunction is given by a phase factor $\rho(B_j) = e^{i\theta_j}$ with $\theta_j \in [0, 2\pi)$. The Yang-Baxter equation demands

$$\rho(B_j)\rho(B_{j+1})\rho(B_j) = \rho(B_{j+1})\rho(B_j)\rho(B_{j+1}) \Rightarrow e^{2i\theta_j + i\theta_{j+1}} = e^{i\theta_j + 2i\theta_{j+1}} \Rightarrow \theta_j = \theta_{j+1}$$

for any 1D representation. As a result, all the angles are equal and the representation

$$\rho_\theta(B_j) = e^{i\theta} \quad (28)$$

is characterized by a single angle θ .

Note that for $\theta = 0$, we get the customary result for bosons that interchanging two particles does nothing to the wavefunction whereas for $\theta = \pi$ interchanging introduces a minus sign which is the result for fermions. In $2 + 1$ dimension, all angles in between 0 and π are allowed and particles with $\theta \neq 0$ or π are called (Abelian) anyons. As an example, we note that quasiparticles in the fractional quantum Hall effect at filling fraction $\nu = \frac{1}{n}$ with n an odd integer are anyons with $\theta = \nu\pi$. Even though the Abelian anyons are interesting from a physical point of view, they do not offer any resources that can be used for quantum computing as braiding only introduces simple phase factors.

Particles whose wavefunctions transform according to higher dimensional irreducible representations of the braid group are called non-Abelian anyons. A necessary ingredient is a ground state degeneracy (which grows exponentially with the number of particles). The effect of B_j ,

the counterclockwise exchange of two particles j and $j + 1$, is then represented by a unitary matrix $\rho(B_j)$ on the ground state manifold. As different unitary matrices do not commute, the representation is non-Abelian which is the reason for their name.

The usefulness of non-Abelian anyons for topological quantum computation relies on the fact that the degeneracy of the ground state manifold is protected and the gates implemented by the exchange of particles are exact (up to an unimportant global phase) [12]. A specific species of non-Abelian anyons is called universal for quantum computation, if, for any given gate, a braid can be found which approximates the gate with arbitrary accuracy.

4.2 Ising anyons

We have seen that Majorana zero modes in topological superconductors lead to a degenerate ground state of size 2^{N-1} that grows exponentially with the number of zero modes. It is thus a natural question to ask if the process of braiding the zero modes leads to a nontrivial operator $\rho_\gamma(B_j)$ on the ground state manifold. It can be shown, see Appendix, that braiding of Majorana zero modes in topological superconductors is described by the unitary representation [28–32]

$$\rho_\gamma(B_j) = \exp\left(-\frac{\pi}{4}\gamma_j\gamma_{j+1}\right) = \frac{1}{\sqrt{2}}(1 - \gamma_j\gamma_{j+1}) \quad (29)$$

of the braid group. The representation ρ_γ is also called Ising anyons. The clockwise exchange of the strands j and $j + 1$ is implemented by $\rho_\gamma(B_j^{-1}) = \rho_\gamma(B_j)^\dagger = \rho_\gamma^\dagger(B_j) = \exp\left(\frac{\pi}{4}\gamma_j\gamma_{j+1}\right)$.

Under the action of $\rho_\gamma(B_j)$, the Majorana zero modes are mapped onto each other with

$$\gamma'_j = \rho_\gamma^\dagger(B_j)\gamma_j\rho_\gamma(B_j) = -\gamma_{j+1}, \quad \gamma'_{j+1} = \rho_\gamma^\dagger(B_j)\gamma_{j+1}\rho_\gamma(B_j) = \gamma_j \quad (30)$$

while the other modes remain unaffected. That this is the correct expression is reinforced by the fact that the (local) parity

$$\mathcal{P}'_j = \rho_\gamma^\dagger(B_j)\mathcal{P}_j\rho_\gamma(B_j) = -i\gamma'_j\gamma'_{j+1} = -i\gamma_j\gamma_{j+1} = \mathcal{P}_j \quad (31)$$

is conserved [33].

To check that ρ_γ is a representation of the braid group, we have to check $\rho_\gamma(B_i)\rho_\gamma(B_j) = \rho_\gamma(B_j)\rho_\gamma(B_i)$, $|i - j| \geq 2$, and the Yang-Baxter equation, see Eq. (27). The first relation follows easily as $\rho_\gamma(B_i)$ and $\rho_\gamma(B_j)$ commute due to the fact that they act on different Majorana zero modes and consist of an even number of Majorana operators. The Yang-Baxter equation can be directly evaluated

$$\begin{aligned} \rho_\gamma(B_j)\rho_\gamma(B_{j+1})\rho_\gamma(B_j) &= \frac{1}{2^{3/2}}(1 - \gamma_j\gamma_{j+1})(1 - \gamma_{j+1}\gamma_{j+2})(1 - \gamma_j\gamma_{j+1}) \\ &= -\frac{1}{\sqrt{2}}(\gamma_j\gamma_{j+1} + \gamma_{j+1}\gamma_{j+2}) \end{aligned} \quad (32)$$

and similarly

$$\rho_\gamma(B_{j+1})\rho_\gamma(B_j)\rho_\gamma(B_{j+1}) = -\frac{1}{\sqrt{2}}(\gamma_j\gamma_{j+1} + \gamma_{j+1}\gamma_{j+2}); \quad (33)$$

as a result, the unitary gate performed by braiding Majorana zero modes only depends on the braid and not on the concrete paths taken. In particular, both ways depicted in Fig. 5(b) of

moving the Majorana zero modes around each other produces the same operation on the ground state manifold.

Thus, braiding can be used to perform topologically protected gates onto the encoded Majorana qubit states $|\bar{0}\rangle$ and $|\bar{1}\rangle$. Indeed for $N = 4$, the interchange of Majorana zero modes implements the operations

$$\rho_\gamma(B_1) = \rho_\gamma(B_3) = \exp\left(-i\frac{\pi}{4}\bar{\sigma}^z\right), \quad \rho_\gamma(B_2) = \exp\left(-i\frac{\pi}{4}\bar{\sigma}^x\right) \quad (34)$$

that correspond to rotations by 90° on the Bloch sphere. We have now all the ingredients to calculate the probability $P(\alpha, \beta)$ of Eq. (26), see Fig. 4(c). Performing the braid B_2 acts as $\rho_\gamma(B_2) = 2^{-1/2}(1 - i\bar{\sigma}^x)$ on the ground state manifold. In particular, the initial state $|\bar{0}\rangle = |(\gamma_1, \gamma_2)_0 (\gamma_3, \gamma_4)_0\rangle$ gets transformed into

$$\rho_\gamma(B_2)|\bar{0}\rangle = \frac{1}{\sqrt{2}}(|\bar{0}\rangle - i|\bar{1}\rangle) = \frac{1}{\sqrt{2}}(|00\rangle - i|11\rangle). \quad (35)$$

The probability $P(\alpha, \beta)$ for the outcomes $\alpha = \beta \in \{0, 1\}$ after the braid are given by

$$P(0, 0) = |\langle 00 | \rho_\gamma(B_2) |\bar{0}\rangle|^2 = \frac{1}{2}, \quad P(1, 1) = |\langle 11 | \rho_\gamma(B_2) |\bar{0}\rangle|^2 = \frac{1}{2}. \quad (36)$$

As the result with $P(\alpha, \alpha) = 50\%$ for both $\alpha = 0$ and $\alpha = 1$ relies on the way Majorana zero modes braid, the fusion experiment in Fig. 4(c) has been proposed recently as a first test of the non-Abelian nature of Majorana zero modes [34].

Even though these operations are protected, braiding of Ising anyons is not enough to perform arbitrary unitary operations on the ground state manifold. The single qubit rotations are not complete as only rotations by (multiples of) 90° around the coordinate axes x, y, z can be implemented. In particular, a rotation by 45° is missing (called $\frac{\pi}{8}$ -phase or T-gate) and an entangling gate needs to be added [12, 35]. In concrete realizations, entanglement can be obtained by a joint parity measurement of two qubits [20, 23]. In fact, it is enough if the T-gate is implemented with a fidelity of 90% as a distillation protocol using the exact Clifford gates called the Magic state distillation can be employed to purify the state [35].

4.3 Fibonacci anyons

Even though the Ising anyons provided by the Majorana modes have the nice property that they allow for noise-insensitive operations on a parity-protected qubit by braiding, the group of operations that can be obtained in this way is not enough for universal quantum computation. However, luckily, there exist other anyons, in particular Fibonacci anyons that allow for a universal set of operations by braiding [4, 12]. That Majorana zero modes do not offer universality is connected to the fact that the Hilbert space can be locally assigned to Majorana modes: any pair of Majorana zero modes can either be filled or empty. For Fibonacci anyons the relation between the anyons and the Hilbert space of the ground state thus has to be more complicated to overcome this issue.

We denote Fibonacci anyons with the letter τ . A pair of Fibonacci anyons has two possible fusion outcomes: it can either fuse to the vacuum 0 or to a Fibonacci anyon τ . In particular,

	$ i\rangle = f_{j-1}, f_j, f_{j+1}\rangle$	$\rho_\tau(B_j) i\rangle = \sum_{f'_j} c_{f'_j} f_{j-1}, f'_j, f_{j+1}\rangle$
	$ 0, \tau, \tau\rangle$	$\omega^{-1} 0, \tau, \tau\rangle$
	$ \tau, \tau, 0\rangle$	$\omega^{-1} \tau, \tau, 0\rangle$
	$ 0, \tau, 0\rangle$	$\omega^{-2} 0, \tau, 0\rangle$
	$ \tau, 0, \tau\rangle$	$\phi^{-1}\omega^2 \tau, 0, \tau\rangle + \phi^{-1/2}\omega \tau, \tau, \tau\rangle$
	$ \tau, \tau, \tau\rangle$	$\phi^{-1/2}\omega \tau, 0, \tau\rangle - \phi^{-1} \tau, \tau, \tau\rangle$

Table 1: Result of the elementary braid B_j of Fibonacci anyons on the basis formed by the fusion outcomes. Note that the effect of braiding the strands j and $j+1$ is local in the sense that it only changes the fusion outcome f_j . The parameters are the phase $\omega = -q = -e^{2\pi i/5}$ and the golden ratio $\phi = (1 + \sqrt{5})/2 = q + \bar{q} + 1$.

we have that $|(\tau_1, \tau_2)_0\rangle, |(\tau_1, \tau_2)_\tau\rangle$ forms a basis of the two-dimensional ground state manifold of two Fibonacci anyons. However, akin to the Ising anyons, there is a superselection rule forbidding us to use these two states as a qubit. The ‘Fibonacci qubit’ thus has to be formed from a subspace of three Fibonacci anyons τ_1, τ_2, τ_3 with the possible fusion outcomes⁷

$$|\bar{0}\rangle = |((\tau_1, \tau_2)_0, \tau_3)_\tau\rangle, \quad |\text{nc}\rangle = |((\tau_1, \tau_2)_\tau, \tau_3)_0\rangle, \quad |\bar{1}\rangle = |((\tau_1, \tau_2)_\tau, \tau_3)_\tau\rangle. \quad (37)$$

Note that the first and the last state are in the same superselection sector (as they fuse to τ) and thus can be used as a genuine Fibonacci qubit. The state of the qubit is then determined by the fusion outcome of the first two Fibonacci anyons while the fusion with the last Fibonacci anyon is constraint to be τ . The third state with total fusion outcome 0 is then a noncomputational (nc) state.

For the Hilbert space of the degenerate ground state of the N -Fibonacci anyons τ_1, \dots, τ_N , we introduce the following notation

$$|0, \tau, f_2, \dots, f_N\rangle = |(\dots((\tau_1, \tau_2)_{f_2}, \tau_3)_{f_3}, \dots)_{f_N}\rangle, \quad f_j \in \{0, \tau\}, \quad (38)$$

where we added the fusion outcomes 0 and τ of the zeroth and first Fibonacci anyons for future convenience. A particular state is then labelled by writing the fusion outcome f_j in between strand j and $j+1$. In order to determine the dimension of the ground state, we have to find the number of states of the form in Eq. (38). The fusion outcomes $f_0 = 0, f_1, \dots, f_N$ are only constraint by the fact that the fusion of the vacuum $f_j = 0$ (at step j) with a Fibonacci anyon necessarily gives a Fibonacci anyon as fusion outcome and thus $f_{j+1} = \tau$. Thus, we need to count the number of states $|0, \tau, f_1, \dots, f_N\rangle$ without two 0s in a row. Denoting by $Z_j(O_j)$ the number of states with $f_j = 0$ ($f_j = 1$), we have to solve the recurrence relations

$$Z_{j+1} = O_j, \quad O_{j+1} = O_j + Z_j, \quad O_0 = 0, \quad Z_0 = 1 \quad (39)$$

with the result $Z_{j+1} = O_j = F_j$ with F_j the j -th Fibonacci number. Note that this relation is the reason why the anyons are called Fibonacci anyons.

⁷ Note that the fusion of the vacuum 0 with τ can only give τ .

The unitary representation $\rho_\tau(B_j)$ of braiding the Fibonacci anyons τ_j and τ_{j+1} can be expressed as a local relation of the fusion outcomes f_{j-1}, f_j, f_{j+1} , see Table 1 and [36, 37].⁸ It can be directly checked that ρ_τ is unitary. To check that it is in fact a representation, we have to prove Eq. (27). As before, the first relation in (27) follows from the fact that $\rho_\tau(B_j)$ only depends on f_{j-1}, f_j, f_{j+1} and only changes the value of f_j . As a result, $\rho(B_i)$ trivially commutes with $\rho(B_j)$ for $|i - j| \geq 2$.

In the last step, we have to test whether ρ_τ satisfies the Yang-Baxter equation. As for the Ising anyons, we do this directly by calculating UVU and VUV with $U = \rho_\tau(B_j)$ and $V = \rho_\tau(B_{j+1})$ separately and subsequently verify that $UVU = VUV$. The Yang-Baxter equation acts on the three strand $j, j+1, j+2$. In a first step, we determine the representation of $\rho_\tau(B_j)$ and $\rho(B_{j+1})$ in the Hilbert space $|f_{j-1}, f_j, f_{j+1}, f_{j+2}\rangle$. Note that we suppress the labels of the other fusion outcomes as they remain unchanged. Moreover, we see from Table 1 that U and V cannot change the value of f_{j-1} and f_{j+2} (superselection) and thus we can verify the relation for each value of $a = f_{j-1}$ and $b = f_{j+2}$ separately.

With the rules of Table 1, we find [U_{ab} is the matrix $\rho(B_j)$ in the subspace with fixed a and b]

$$\begin{aligned} U_{00} = V_{00} &= (\omega^{-1}); & U_{0\tau} &= \begin{pmatrix} \omega^{-2} & 0 \\ 0 & \omega^{-1} \end{pmatrix}, V_{0\tau} = \begin{pmatrix} \phi^{-1}\omega^2 & \phi^{-1/2}\omega \\ \phi^{-1/2}\omega & -\phi^{-1} \end{pmatrix}; \\ U_{\tau 0} = V_{\tau 0}, V_{\tau 0} &= U_{0\tau}; & U_{\tau\tau} &= \begin{pmatrix} \phi^{-1}\omega^2 & 0 & \phi^{-1/2}\omega \\ 0 & \omega^{-1} & 0 \\ \phi^{-1/2}\omega & 0 & -\phi^{-1} \end{pmatrix}, V_{\tau\tau} = \begin{pmatrix} \omega^{-1} & 0 & 0 \\ 0 & \phi^{-1}\omega^2 & \phi^{-1/2}\omega \\ 0 & \phi^{-1/2}\omega & -\phi^{-1} \end{pmatrix}. \end{aligned} \quad (40)$$

Here, we have ordered the basisstates as follows: $\{|0, \tau, \tau, 0\rangle\}_{00}$, $\{|0, \tau, 0, \tau\rangle, |0, \tau, \tau, \tau\rangle\}_{0\tau}$, $\{|\tau, 0, \tau, 0\rangle, |\tau, \tau, \tau, 0\rangle\}_{\tau 0}$, and $\{|\tau, 0, \tau, \tau\rangle, |\tau, \tau, 0, \tau\rangle, |\tau, \tau, \tau, \tau\rangle\}_{\tau\tau}$. It is now an easy exercise in matrix multiplication to verify $UVU = VUV$ and thus to show that ρ_τ is a unitary representation of the braid group.

We obtain more insights in the operations performed by braiding by choosing $j = 1, f_0 = 0$ and looking at the states $|((\tau_1, \tau_2)_a, \tau_3)_b\rangle$ of the three Fibonacci anyons in Eq. (38). Performing B_1 by brading the first two anyons (corresponding to the matrices U_{0b}), the states simply acquire the phases $\omega^{-2} = e^{-4\pi i/5}$ (if $a = 0$) and $\omega^{-1} = e^{3\pi i/5}$ (if $a = \tau$) depending only on their fusion outcome a irrespective of the value of b .⁹ On the other hand, braiding the anyons τ_2 and τ_3 with B_2 (corresponding to V_{0b}) is not diagonal in this basis as the two anyons do not have a well-defined fusion outcome; in the alternative basis $|(\tau_1, (\tau_2, \tau_3)_{a'})_b\rangle$ the operation B_2 would be the simple phase factors ω^{-1}, ω^{-2} depending on a' as before. It is easy to check that $V_{0\tau} = F^{-1}U_{0\tau}F$ with the transformation matrix

$$F = F^{-1} = \begin{pmatrix} \phi^{-1} & \phi^{-1/2} \\ \phi^{-1/2} & -\phi^{-1} \end{pmatrix}, \quad (41)$$

which corresponds to the basis transformation

$$|(\tau_1, (\tau_2, \tau_3)_\tau)_0\rangle = |((\tau_1, \tau_2)_\tau, \tau_3)_0\rangle, \quad (42)$$

$$\begin{aligned} |(\tau_1, (\tau_2, \tau_3)_0)_\tau\rangle &= \phi^{-1}|((\tau_1, \tau_2)_0, \tau_3)_\tau\rangle + \phi^{-1/2}|((\tau_1, \tau_2)_\tau, \tau_3)_\tau\rangle \\ |(\tau_1, (\tau_2, \tau_3)_\tau)_\tau\rangle &= \phi^{-1/2}|((\tau_1, \tau_2)_0, \tau_3)_\tau\rangle - \phi^{-1}|((\tau_1, \tau_2)_\tau, \tau_3)_\tau\rangle \end{aligned} \quad (43)$$

⁸ Note that the references denote our $0, \tau$ by $*, p$.

⁹ The corresponding phases are known as elements of the R-matrix.

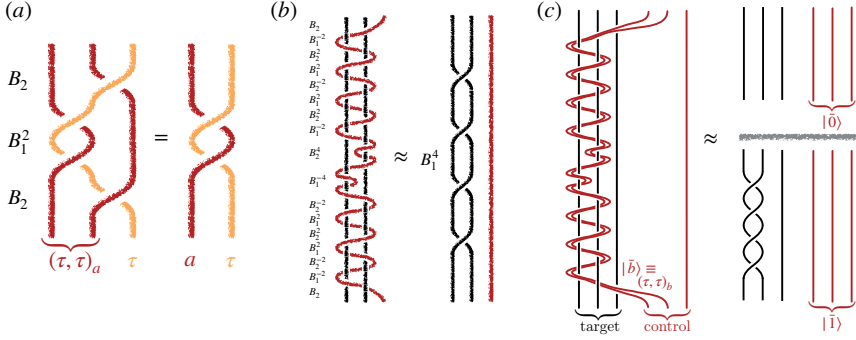


Fig. 6: (a) The braid $B_2 B_1^2 B_2$ moves the combined object $(\tau_1, \tau_2)_a$ (dark) around the anyon τ_3 (light). If the fusion outcome is the vacuum ($a = 0$), the state of the system remains unchanged. On the other hand, if the fusion outcome is an anyon ($a = \tau$), the result is equivalent to $(U_{0\tau})^2$ obtained by simply moving an elementary anyon around the light braid. (b) The braid where the red anyon is ‘weaved’ around the black anyons is approximately equivalent to B_1^4 with an error of 1 %. (c) The approximate braid of (b) can be used to implement a two gate (a controlled unitary gate) on 6 anyons (the black anyons form the target and the red anyons the control qubit). The state of the control qubit is encoded in the fusion outcome of the two leftmost red braids. If the fusion outcome is the vacuum (the control qubit is in state $|\bar{0}\rangle$), the result of the braid is the identity operation. If the fusion outcome is an anyon (the control qubit is in state $|\bar{1}\rangle$), the net result is $U_4 = (U_{0\tau})^4$ according to (b).

also known as the F -move.

A single Fibonacci qubit is realized by three Fibonacci anyons τ_1, τ_2, τ_3 in the state $|0, \tau, f, \tau\rangle$. The states with $f \in \{0, \tau\}$ form the logical qubit $|\bar{0}\rangle, |\bar{1}\rangle$. From the calculation in Eq. (40) we know that the gate $U_{0\tau}$ with the phases $e^{-\pi i/10 \pm 7\pi i/10}$ is performed while braiding the first two anyons. Up to the irrelevant (Abelian) factor $e^{-\pi i/10}$, the braid corresponds to a rotation by $\frac{7}{5}\pi \equiv 252^\circ$ around the z -axis of the Bloch sphere. Similarly, braiding anyons τ_2 and τ_3 yields the gate $V_{0\tau}$ in a protected fashion. This gate is another rotation by 252° around an axis which corresponds to the basis transformation given by F . In terms of the Bloch sphere, the rotation $V_{0\tau}$ is a rotation around the axis $\mathbf{v} = (2\phi^{-3/2}, 0, \phi^{-2} - \phi^{-1})^T$. Braiding the anyons τ_1, τ_2, τ_3 , arbitrary products of $U_{0\tau}$ and $V_{0\tau}$ can be implemented and all single qubit gates can be performed in a protected manner.

To understand the idea of Ref. [38] of how to implement two qubit gates, we first have to appreciate another crucial property of anyon braiding: braiding the fusion product $(\tau_1, \tau_2)_a$ as a composite object results in the same net effect as braiding the elementary object a . As a simple example, let us consider the states $|0, \tau, a, b\rangle = |((\tau_1, \tau_2)_a, \tau_3)_b\rangle$ as before. As seen in Fig. 6(a), the braid $B = B_2 B_1^2 B_2$ moves the composite object $(\tau_1, \tau_2)_a$ around τ_3 and is represented by

$W = \rho_\tau(B)$. From Eq. (40), we obtain after a straightforward calculation the result¹⁰

$$W_{ab} = \begin{cases} 1, & a = 0, b = \tau, \\ \omega^{-4}, & a = \tau, b = 0, \\ \omega^{-2}, & a = \tau, b = \tau. \end{cases} \quad (44)$$

We observe that as expected moving two anyons which fuse to the vacuum ($a = 0$) around another anyon does not change the state of the system. However, when the fusion outcome is another anyon ($a = \tau$), the result is $(U_{0\tau})^2$.

The following procedure allows to obtain a controlled two-qubit gate by braiding: first, we find a ‘weave’, i.e., a braid where only a single anyon is moved around two static anyons that is equivalent to a braid which only involves the static anyons. It can be checked that the weave of the third anyon shown in Fig. 6(b) is equivalent to B_1^4 of the first two anyons up to an error that is smaller than a percent. Having found such a weave by brute force search, a controlled gate can be obtained with the idea of the composite object explained above, see Fig. 6(c). As before, we encode two qubits in the states $|\bar{a}, \bar{b}\rangle = |((\tau_1, \tau_2)_a, \tau_3)_\tau\rangle \otimes |((\tau_4, \tau_5)_b, \tau_6)_\tau\rangle$ of six Fibonacci anyons.¹¹ We use $|\bar{b}\rangle$ (formed by anyons τ_4, τ_5, τ_6 in red) as the control qubit and $|\bar{a}\rangle$ (formed by anyons τ_1, τ_2, τ_3 in black) as the target. As the state of the control qubit is given by the fusion outcome b of τ_4, τ_5 moving the two anyons together around the anyons of the target qubit does nothing as long as $b = 0$ (in this case, they are equivalent to the vacuum that braids trivially). In order to obtain a controlled gate that does not change the state of $|\bar{b}\rangle$, we let the pair of anyons τ_4, τ_5 perform the weave found above. In this case, the control anyons (up to a small error) remain unaffected and, provided that $b = \tau$, the operation $U_4 = (U_{0\tau})^4 = \text{diag}(e^{2\pi i/5}, e^{-4\pi i/5})$ is performed on the target. As a result, the braid of Fig. 6(c) acting on two qubits entangles them by a controlled- U_4 gate. Together with the universal set of single qubit gates found before, all gates can be performed by braiding Fibonacci anyons and the anyons are *universal* for quantum computation. More information on Fibonacci anyons can be found, e.g., in Refs. [7, 12, 39, 40].

5 Conclusion

We have shown how non-Abelian anyons can be used for topological quantum computers. The computation is performed by the wordlines of the anyons forming knots in space-time. The outcome of the computation is the fusion outcomes at the last step of the computation. The computation is topologically protected in the sense that the gates performed do not depend on the concrete trajectories of the anyons but only on the overall topology of the braid.

The discussion so far has been only concerned with zero temperature. At any finite temperature there is a finite fraction $\propto \exp(-\bar{M}/k_B T)$ of thermal anyons present in the system (\bar{M} is the gap protecting the topological phase). Those anyons braid in an unwanted, uncontrollable fashion around the anyons forming the qubit. As a result, the computation is dephased and fails. If

¹⁰ It follows from $V_{0\tau}(U_{0\tau})^2 V_{0\tau} = \text{diag}(1, \omega^{-2})$ and $V_{00}(U_{00})^2 V_{00} = \omega^{-4}$.

¹¹ In terms of our states $|f_0, f_1, \dots\rangle$ where f_j denotes the fusion outcome of all the anyons to the left, we have to add another passive anyon, which fuses with τ to the vacuum, in between the three anyons forming qubit a and the three anyons forming qubit b . As a result, we have the encoding $|\bar{a}, \bar{b}\rangle = |0, \tau, a, \tau, 0, \tau, b, \tau\rangle$. Without the additional fusion outcome 0 the value b would not denote the fusion outcome of the two anyons to the left of it.

topological quantum computation is so fragile with respect to temperature, why did we bother to discuss the ideas at length? The answer to this question has many facets. First and most importantly, non-Abelian anyons are an interesting new form of quantum matter. The study of the potential ways in which (quasi-)particles may braid depending on the dimension of space is an interesting and important question of basic physics research even without applications to quantum computation.

Regarding the potential for quantum computation, the argument given above simply means that we have to employ error correction at some level. Non-Abelian anyons have inspired error correction for a long time and have lead to the concept of topological codes [5]. In particular, even if there is no passive system which features Fibonacci anyons at zero temperature, we can imagine keeping a system actively in a topological state with non-Abelian anyons by constantly performing syndrome measurements [41,42].

Regarding the parity-protected quantum computation with Majorana zero modes; even though the protocol strictly only works at zero temperatures, we may expect the Majorana qubit to have a rather long lifetime at finite temperature. Moreover, it has recently been shown that the non-Abelian nature of the Majorana zero modes allows for dedicated error protection protocols relying on their non-Abelian nature [43–46]. The study of systems where the elementary excitations are particles with exotic exchange statistics is still a very active field with many surprises to come. Presently, there is a big experimental push to realize Ising-anyons in the form of Majorana zero modes in superconductor-semiconductor heterostructures, see Ref. [47]. However, so far, no compelling evidence of particles with non-Abelian exchange statistics has ever been observed.

I want to thank Lisa Arndt and Alex Ziesen for carefully reading the manuscript and proposing changes that helped to considerably increase the readability of the text.

Appendix

Non-Abelian Berry phase and braiding

In this appendix, we want to derive the unitary operation that is performed when exchanging two Majorana zero modes. We discuss this in a system consisting of 4 Majorana zero modes which are aligned in a Y-junction [32,48]. We call the Majorana modes at the three ends of the Y-junction $\gamma_1, \gamma_2, \gamma_3$. The Majorana mode in the middle is denoted by γ_0 . The Hamiltonian of the system assumes the form

$$H = i \sum_{j=1}^3 \epsilon_j \gamma_0 \gamma_j = i \gamma_0 (\boldsymbol{\epsilon} \cdot \boldsymbol{\gamma}) \quad (45)$$

with ϵ_j three parameters which gives the coupling strength of the two Majorana modes on each leg of the junction. Slowly tuning these couplings in a controlled fashion makes it possible to interchange Majorana modes and thus we can observe the non-Abelian statistics. In particular, we would like to study the situation in Fig. 7.

We start at time (a) with $\epsilon_3 = \bar{\epsilon}$ and $\epsilon_1 = \epsilon_2 = 0$. Then we increase ϵ_1 to $\bar{\epsilon}$ and afterwards

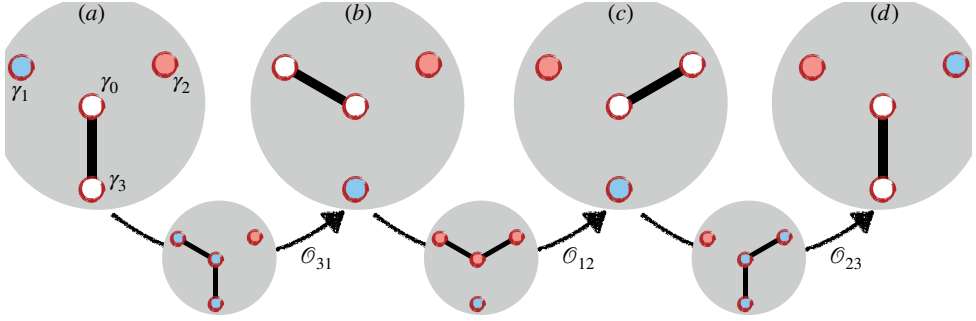


Fig. 7: Steps to interchange the blue and red Majorana zero mode that are initially located at position 1 and 2. The first move from (a) to (b), moves the blue Majorana zero mode from 1 to 3. In the first step (small circle below), the coupling between γ_0 and γ_1 is increased which leads to the delocalization of the blue Majorana zero mode as a superposition of $\gamma_1, \gamma_0, \gamma_3$. In the second step, the coupling between γ_0 and γ_3 is reduced which moves the blue Majorana mode to position 3. In the step from (b) to (c), the red Majorana mode is moved from position 2 to position 1. In the last step, the blue Majorana mode is moved from 3 to 2.

decrease ϵ_3 to 0 such that at time (b) we have the blue Majorana zero mode moved to the bottom (from position 1 to position 3). To finish the exchange, we do the same procedure from position 2 to 1 and finally from position 3 to 2.

Non-Abelian Berry phase

The Berry phase is a geometric phase which arises when a parameter of the Hamiltonian is changed slowly and we want to observe the change in the ground state wavefunction. In the case where the ground state is degenerate, the change in the parameter might induce transitions between the different degenerate states and we have a non-Abelian Berry phase which is generated by the non-Abelian gauge field

$$A_{ab}^k = i\langle\psi_a(\epsilon)|\partial_{\epsilon_k}|\psi_b(\epsilon)\rangle; \quad (46)$$

here, $\psi_a(\epsilon)$ denote the different degenerate ground state wavefunctions for the Hamiltonian with parameters ϵ .

Changing the parameter along a contour $\epsilon(s)$, the ground states are transformed according to the unitary matrix

$$U = \mathcal{P} \exp \left(i \int \mathbf{A} \cdot d\epsilon \right) \quad (47)$$

where \mathcal{P} denotes the path ordering.

A basis independent way to obtain the evolution of the ground state due to Kato is the following [49]: given the projector $P(\epsilon)$ onto the ground state manifold (which depends on the parameters ϵ), we define the Kato Hamiltonian

$$K^k = i[P(\epsilon), \partial_{\epsilon_k} P(\epsilon)]. \quad (48)$$

with which the unitary evolution in Eq. (47) can be written as

$$U_K = \mathcal{P} \exp \left(i \int \mathbf{K} \cdot d\epsilon \right). \quad (49)$$

It can be shown that $U_K = U$ for any closed contour [50].

Calculating the non-Abelian Berry phase

The Hamiltonian H has the eigenenergies $\pm\epsilon$ ($\epsilon = |\epsilon|$) where both of them are doubly degenerate. For the Berry phase, we need to project onto the ground state sector on which $H = -\epsilon$. The projector is given by

$$P(\epsilon) = \frac{1}{2\epsilon}(\epsilon - H) = \frac{1}{2\epsilon}[\epsilon - i\gamma_0(\epsilon \cdot \gamma)] \quad (50)$$

which leads to

$$\partial_{\epsilon_k} P(\epsilon) = \frac{i\gamma_0}{2\epsilon^3} \sum_{j \neq k} (\epsilon_k \epsilon_j \gamma_j - \epsilon_j^2 \gamma_k). \quad (51)$$

Now it is straightforward to calculate the non-Abelian gauge field

$$\begin{aligned} K^k &= \frac{i}{4\epsilon^4} \sum_i \sum_{j \neq k} \epsilon_i \left(\epsilon_k \epsilon_j \overbrace{[\gamma_0 \gamma_i, \gamma_0 \gamma_j]}^{2(\delta_{ij}-1)\gamma_i \gamma_j} - \epsilon_j^2 \overbrace{[\gamma_0 \gamma_i, \gamma_0 \gamma_k]}^{2(\delta_{ik}-1)\gamma_i \gamma_k} \right) \\ &= \frac{i}{2\epsilon^4} \left[\sum_{j \neq k} \epsilon_j^2 \sum_{i \neq k} \epsilon_i \gamma_i \gamma_k - \epsilon_k \sum_{i \neq j} \sum_{j \neq k} \epsilon_i \epsilon_j \gamma_i \gamma_j \right] \\ &= \frac{i}{2\epsilon^4} \left[(\epsilon^2 - \epsilon_k^2)(\epsilon \cdot \gamma - \epsilon_k \gamma_k) \gamma_k + \epsilon_k^2 (\epsilon \cdot \gamma - \epsilon_k \gamma_k) \gamma_k \right] \\ &= \frac{i}{2\epsilon^2} (\epsilon \cdot \gamma - \epsilon_k \gamma_k) \gamma_k. \end{aligned} \quad (52)$$

The result when changing a single parameter reads

$$\begin{aligned} i \int_{\epsilon_{\text{start}}}^{\epsilon_{\text{stop}}} d\epsilon_k K^k &= \overbrace{(\epsilon_k \gamma_k - \epsilon \cdot \gamma) \gamma_k}^{-\sum_{j \neq k} \epsilon_j \gamma_j \text{ (indep. of } \epsilon_k)} \times \int_{\epsilon_{\text{start}}}^{\epsilon_{\text{stop}}} \frac{d\epsilon_k}{2\epsilon^2} \\ &= \frac{1}{2\epsilon_{\perp}} \arctan(\epsilon_k / \epsilon_{\perp}) \Big|_{\epsilon_{\text{start}}}^{\epsilon_{\text{stop}}} \end{aligned} \quad (53)$$

with $\epsilon_{\perp} = \left(\sum_{j \neq k} \epsilon_j^2 \right)^{1/2}$.

A braid is constructed from elementary moves, see Fig. 7. We exemplify the calculation of \mathcal{O}_{31} . Starting with $\epsilon_3 = \bar{\epsilon}$ and $\epsilon_1 = \epsilon_2 = 0$, we first turn on ϵ_1 from 0 to $\bar{\epsilon}$ which yields ($k = 1, \epsilon_{\text{start}} = 0, \epsilon_{\text{stop}} = \bar{\epsilon}$)

$$U_a = \exp \left(-\bar{\epsilon} \gamma_3 \gamma_1 \times \frac{1}{2\bar{\epsilon}} \frac{\pi}{4} \right) = \exp \left(\frac{\pi}{8} \gamma_1 \gamma_3 \right).$$

In the second step, we start with $\epsilon_1 = \epsilon_3 = \bar{\epsilon}$ and $\epsilon_2 = 0$ and reduce ϵ_3 down to 0. This step yields a non-Abelian Berry phase ($k = 3$, $\epsilon_{\text{start}} = \bar{\epsilon}$, $\epsilon_{\text{stop}} = 0$)

$$U_b = \exp \left(-\bar{\epsilon} \gamma_1 \gamma_3 \times -\frac{1}{2\bar{\epsilon}} \frac{\pi}{4} \right) = \exp \left(\frac{\pi}{8} \gamma_1 \gamma_3 \right).$$

So all together, we have the transformation

$$\mathcal{O}_{31} = U_b U_a = \exp \left(\frac{\pi}{4} \gamma_1 \gamma_3 \right) = \frac{1}{\sqrt{2}} (1 + \gamma_1 \gamma_3) \quad (54)$$

from (a) to (b) that moves the Majorana zero mode from position 1 to 3. The whole operation of exchanging the particle 1 and 2 is given by

$$\rho_\gamma(B_1) = \mathcal{O}_{23} \mathcal{O}_{12} \mathcal{O}_{31} = \frac{1}{2^{3/2}} (1 + \gamma_3 \gamma_2) (1 + \gamma_2 \gamma_1) (1 + \gamma_1 \gamma_3) = \frac{1}{\sqrt{2}} (1 - \gamma_1 \gamma_2) = e^{-\frac{\pi}{4} \gamma_1 \gamma_2}. \quad (55)$$

The braiding operation $\rho(B_1)$ transforms the Majorana zero modes as

$$\gamma_1 \mapsto \rho_\gamma(B_1) \gamma_1 \rho_\gamma(B_1)^\dagger = \gamma_2, \quad \gamma_2 \mapsto \rho_\gamma(B_1) \gamma_2 \rho_\gamma(B_1)^\dagger = -\gamma_1. \quad (56)$$

It is a natural question to ask which of the two modes γ_1 or γ_2 gets a minus sign. It turns out that in our setting of Fig. 7, γ_2 obtains a minus sign (it is transformed to $-\gamma_1$) as it is only moved once in the second step (\mathcal{O}_{12}). On the other hand, γ_1 is moved twice, in the first step (\mathcal{O}_{31}) and in the last step (\mathcal{O}_{32}). Indeed, in a single step, we have

$$\mathcal{O}_{31} \gamma_1 \mathcal{O}_{31}^\dagger = -\gamma_3 \quad (57)$$

and thus γ_1 is moved to $-\gamma_3$. As a result, we would obtain the conjugate representation $\rho_\gamma(B_1) = e^{\frac{\pi}{4} \gamma_1 \gamma_2}$, that is often employed in the literature, if the geometry in Fig. 7 would look like a Λ rather than a Y .

References

- [1] R. Hanbury Brown and R. Q. Twiss, A test of a new type of stellar interferometer on Sirius, *Nature* **178**, 1046 (1956).
- [2] E. Brannen and H. I. S. Ferguson, The question of correlation between photons in coherent light rays, *Nature* **178**, 481 (1956).
- [3] I. Silva and O. Freire, The concept of the photon in question, *Hist. Stud. Nat. Sci.* **43**, 453 (2013).
- [4] M. H. Freedman, M. Larsen, and Z. Whang, A modular functor which is universal for quantum computation, *Commun. Math. Phys.* **227**, 605 (2002).
- [5] A. Yu. Kitaev, Fault-tolerant quantum computation by anyons, *Ann. Phys. (N. Y.)* **303**, 2 (2003).
- [6] P. A. M. Dirac, Quantised singularities in the electromagnetic field, *Proc. R. Soc. A* **133**, 60 (1931).
- [7] J. Preskill, *Lecture notes: Quantum computation*, chapter 9 (2004).
- [8] J. Alicea, New directions in the pursuit of Majorana fermions in solid state systems, *Rep. Prog. Phys.* **75**, 076501 (2012).
- [9] M. Leijnse and K. Flensberg, Introduction to topological superconductivity and Majorana fermions, *Semicond. Sci. Technol.* **27**, 124003 (2012).
- [10] C. W. J. Beenakker, Search for Majorana fermions in superconductors, *Annu. Rev. Con. Mat. Phys.* **4**, 113 (2013).
- [11] J. K. Pachos, *Introduction to topological quantum computation* (Cambridge University Press, Cambridge, 2012).
- [12] C. Nayak, S. H. Simon, A. Stern, M. Freedman, and S. Das Sarma, Non-Abelian anyons and topological quantum computation, *Rev. Mod. Phys.* **80**, 1083 (2008).
- [13] A. Stern and N. H. Lindner, Topological quantum computation—from basic concepts to first experiments, *Science* **339**, 1179 (2013).
- [14] S. Das Sarma, M. Freedman, and C. Nayak, Majorana zero modes and topological quantum computation, *npj Quantum Inf.* **1**, 15001 (2015).
- [15] R. Jackiw and C. Rebbi, Solitons with fermion number $1/2$, *Phys. Rev. D* **13**, 3398 (1976).
- [16] A. Yu. Kitaev, Unpaired Majorana fermions in quantum wires, *Phys.-Usp.* **44** (suppl.), 131 (2001).
- [17] R. P. Feynman, R. B. Leighton, and M. Sands, *Feynman lectures on physics: Quantum mechanics*, vol. 3 (Addison-Wesley, Reading, 1965).
- [18] B. M. Terhal and D. P. DiVincenzo, Classical simulation of noninteracting-fermion quantum circuits, *Phys. Rev. A* **65**, 032325 (2002).

- [19] C. W. J. Beenakker, D. P. DiVincenzo, C. Emary, and M. Kindermann, Charge detection enables free-electron quantum computation, *Phys. Rev. Lett.* **93**, 020501 (2004).
- [20] F. Hassler, Majorana qubits, in *Quantum Information Processing. Lecture Notes of the 44th IFF Spring School*, edited by D. P. DiVincenzo (Verlag des Forschungszentrums Jülich, 2013).
- [21] Z.-C. Gu and X.-G. Wen, Tensor-entanglement-filtering renormalization approach and symmetry protected topological order, *Phys. Rev. B* **80**, 155131 (2009).
- [22] F. Pollmann, E. Berg, A. M. Turner, and M. Oshikawa, Symmetry protection of topological order in one-dimensional quantum spin systems, *Phys. Rev. B* **85**, 075125 (2012).
- [23] F. Hassler, A. R. Akhmerov, and C. W. J. Beenakker, Top-transmon: Hybrid superconducting qubit for parity-protected quantum computation, *New J. Phys.* **13**, 095004 (2011).
- [24] A. Kitaev, Anyons in an exactly solved model and beyond, *Ann. Phys. (N. Y.)* **321**, 2 (2006).
- [25] A. Stern, Anyons and the quantum Hall effect—a pedagogical review, *Ann. Phys. (N. Y.)* **323**, 204 (2008).
- [26] F. Wilczek, Magnetic flux, angular momentum, and statistics, *Phys. Rev. Lett.* **48**, 1144 (1982).
- [27] E. Artin, Theory of braids, *Ann. Math.* **48**, 101 (1947).
- [28] N. Read and D. Green, Paired states of fermions in two dimensions with breaking of parity and time-reversal symmetries, and the fractional quantum Hall effect, *Phys. Rev. B* **61**, 10267 (2000).
- [29] D. A. Ivanov, Non-Abelian statistics of half-quantum vortices in p-wave superconductors, *Phys. Rev. Lett.* **86**, 268 (2001).
- [30] J. Alicea, Y. Oreg, G. Refael, F. von Oppen, and M. P. A. Fisher, Non-Abelian statistics and topological quantum computation in 1D wire networks, *Nat. Phys.* **7**, 412 (2011).
- [31] D. J. Clarke, J. D. Sau, and S. Tewari, Majorana fermion exchange in quasi-one-dimensional networks, *Phys. Rev. B* **84**, 035120 (2011).
- [32] B. van Heck, A. R. Akhmerov, F. Hassler, M. Burrello, and C. W. J. Beenakker, Coulomb-assisted braiding of Majorana fermions in a Josephson junction array, *New J. Phys.* **14**, 035019 (2012).
- [33] B. I. Halperin, Y. Oreg, A. Stern, G. Refael, J. Alicea, and F. von Oppen, Adiabatic manipulations of Majorana fermions in a three-dimensional network of quantum wires, *Phys. Rev. B* **85**, 144501 (2012).
- [34] D. Aasen, M. Hell, R. V. Mishmash, A. Higginbotham, J. Danon, M. Leijnse, T. S. Jespersen, J. A. Folk, C. M. Marcus, K. Flensberg, and J. Alicea, Milestones toward Majorana-based quantum computing, *Phys. Rev. X* **6**, 031016 (2016).

- [35] S. Bravyi and A. Yu. Kitaev, Universal quantum computation with ideal Clifford gates and noisy ancillas, *Phys. Rev. A* **71**, 022316 (2005).
- [36] L. H. Kauffman and S. J. Lomonaco, Jr., q — Deformed spin networks, knot polynomials and anyonic topological quantum computation, *J. Knot Theory Ramif.* **16**, 267 (2007).
- [37] P. W. Shor and S. P. Jordan, Estimating Jones polynomials is a complete problem for one clean qubit, *Quantum Inf. Comput.* **8**, 681 (2008).
- [38] N. E. Bonesteel, L. Hormozi, G. Zikos, and S. H. Simon, Braid topologies for quantum computation, *Phys. Rev. Lett.* **95**, 140503 (2005).
- [39] L. Hormozi, G. Zikos, N. E. Bonesteel, and S. H. Simon, Topological quantum compiling, *Phys. Rev. B* **75**, 165310 (2007).
- [40] S. Trebst, M. Troyer, Z. Wang, and A. W. W. Ludwig, A short introduction to Fibonacci anyon models, *Prog. Theor. Phys. Supp.* **176**, 384 (2008).
- [41] R. Koenig, G. Kuperberg, and B. W. Reichardt, Quantum computation with Turaev-Viro codes, *Ann. Phys. (N. Y.)* **325**, 2707 (2010).
- [42] N. E. Bonesteel and D. P. DiVincenzo, Quantum circuits for measuring Levin-Wen operators, *Phys. Rev. B* **86**, 165113 (2012).
- [43] B. M. Terhal, F. Hassler, and D. P. DiVincenzo, From Majorana fermions to topological order, *Phys. Rev. Lett.* **108**, 260504 (2012).
- [44] S. Vijay, T. H. Hsieh, and L. Fu, Majorana fermion surface code for universal quantum computation, *Phys. Rev. X* **5**, 041038 (2015).
- [45] S. Plugge, L. A. Landau, E. Sela, A. Altland, K. Flensberg, and R. Egger, Roadmap to Majorana surface codes, *Phys. Rev. B* **94**, 174514 (2016).
- [46] D. Litinski, M. S. Kesselring, J. Eisert, and F. von Oppen, Combining topological hardware and topological software: Color code quantum computing with topological superconductor networks, *Phys. Rev. X* **7**, 031048 (2017).
- [47] E. Prada, P. San-Jose, M. W. A. de Moor, A. Geresdi, E. J. H. Lee, J. Klinovaja, D. Loss, J. Nygård, R. Aguado, and L. P. Kouwenhoven, From Andreev to Majorana bound states in hybrid superconductor-semiconductor nanowires, *arXiv:1911.04512* (2019).
- [48] J. D. Sau, D. J. Clarke, and S. Tewari, Controlling non-Abelian statistics of Majorana fermions on Majorana dimer lattices, *Phys. Rev. B* **84**, 094505 (2011).
- [49] T. Kato, On the adiabatic theorem of quantum mechanics, *J. Phys. Soc. Jap.* **5**, 435 (1950).
- [50] J. E. Avron, L. Sadun, J. Seger, and B. Simon, Chern numbers, quaternions, and Berry's phases in Fermi systems, *Commun. Math. Phys.* **124**, 595 (2009).

B 3 Majorana Qubits

Charles M. Marcus

Niels Bohr Institute, Condensed Matter Physics,
Copenhagen, Denmark

Please refer to:

<https://qdev.nbi.ku.dk/>

- The Organizers -

B 4 An introduction into optimal control for quantum technologies

Frank K. Wilhelm, S. Kirchhoff, S. Machnes, N. Wittler, and D. Sugny¹

Theoretical Physics, Saarland University, Saarbrücken

¹Laboratoire Interdisciplinaire Carnot de Bourgogne (ICB) UMR 6303 CNRS-Université Bourgogne-Franche Comté, 21078 Dijon Cedex, France

Contents

1	Introduction	3
2	Elementary optimal control	4
2.1	Optimal control of a classical harmonic oscillator	4
2.2	Optimal control for a classical system	5
2.3	Gradient-based optimal quantum control with the GRAPE algorithm	8
2.4	The Krotov algorithms	12
2.5	Modern numerical issues	13
3	Applied optimal quantum control	17
3.1	Closing the loop for pulse calibration	18
3.2	CRAB	19
3.3	GOAT	20
3.4	Evaluating fidelity with randomized benchmarking.	21
3.5	Approximating time evolutions with the Magnus expansion	26

3.6	Real-world limitations	27
4	Examples	29
4.1	Optimal control of a qubit	29
4.2	Exploring the speed limit with high parameter counts	29
4.3	Open systems	31
4.4	DRAG and its derivatives	33
5	Summary and outlook	37
	References	38

1 Introduction

Control is a key component in turning science into technology [1], [2]. Broadly and colloquially speaking, control looks at providing the user / experimenter with external parameters to steer a given dynamical system to her liking ¹ rather than simply observing its internal dynamics. Control is in this sense ubiquitous to modern technology. In this colloquial sense, quantum control is transferring that idea to quantum systems and thus contains both hard- and software of many kinds.

The ubiquity of control has given rise to the field of control theory. This is a field of applied mathematics that looks at how to choose said external parameters in order to drive the dynamical system to one's liking. It has spawned ideas of open-loop control, i.e., the pre-determination of controls given the laws of nature (that were a key ingredient to, e.g., the Apollo program) as well as closed-loop-control, interleaving of observation and adjustment as we know it in our daily lives from thermostats. This type of optimal control theory takes the hardware setup as a given, however ideally, these are developed in tandem. The mathematical procedures of open-loop-control typically involve optimizing a cost function, hence the name optimal control.

The application of optimal control is not an entirely new idea. Pioneering applications were primarily chemistry, such as the laser control of chemical reactions and magnetic resonance. By now, quantum optimal control is also applied to a large spectrum of modern quantum technologies (Quantum 2.0) [2]. This implies a certain tradition of fragmentation - quantum optimal control researchers tend to be in departments of mathematics, chemistry, computer science, and physics and follow their specific idiosyncrasies [1]. Modern efforts have gone very far in overcoming this fragmentation which is fruitful in learning from each other and respecting the different goals – quantum control of complex reactions does for example deal with large Hilbert spaces whereas control in quantum computing aims at sufficiently low errors in order to meet error correction thresholds.

In this series of lectures, we would like to introduce the audience to quantum optimal control. The first lecture will cover basic ideas and principles of optimal control with the goal of demystifying its jargon. The second lecture will describe computational tools (for computations both on paper and in a computer) for its implementation as well as their conceptual background. The third chapter will go through a series of popular examples from different applications of quantum technology.

These are lectures notes. Other than a textbook, it makes a significant difference to attend the lectures it goes with rather than use it to self-study. Other than a review, it is not complete but rather serves to introduce clarify the concepts of the field. This also means that the choice of references is certainly not complete, rather, it is the subjective choice of what the authors find most suitable and got inspired by.

¹ where female attributions are the default, male is considered included

2 Elementary optimal control

We start with classical examples of control, which lay a lot of foundations for quantum systems.

2.1 Optimal control of a classical harmonic oscillator

In order to understand the basic concept and structure of quantum optimal control, let us start with a simple classical example: control of the harmonic oscillator.

The equation of motion of a harmonic oscillator driven by force $F(t) = mf(t)$ where m is the mass and eigenfrequency Ω is given by

$$\ddot{x} + \Omega^2 x(t) = f(t).$$

Its general solution is parameterized through the Green's function

$$G(\tau) = \frac{\theta(\tau)}{\Omega} \sin(\Omega\tau)$$

(where θ is the Heaviside function) as

$$x(t) = x(0) \cos \Omega t + \frac{\dot{x}(0)}{\Omega} \sin \Omega t + \int_0^t dt' \frac{\sin \Omega(t-t')}{\Omega} f(t')$$

Readers not familiar with Green's functions can easily verify that this expression does indeed solve the equation of motion of the driven oscillator.

From this we get the velocity

$$\dot{x}(t) = \dot{x}(0) \cos \Omega t - \Omega x(0) \sin \Omega t + \int_0^t dt' \cos \Omega(t-t') f(t')$$

Thus, imposing target values $x(T)$ and $\dot{x}(T)$ we find the conditions

$$\begin{aligned} x(T) - x(0) \cos \Omega T - \frac{\dot{x}(0)}{\Omega} \sin \Omega T &= \int_0^T dt' \frac{\sin \Omega(T-t')}{\Omega} f(t') \\ \dot{x}(T) - \dot{x}(0) \cos \Omega T + \Omega x(0) \sin \Omega T &= \int_0^T dt' \cos \Omega(T-t') f(t'). \end{aligned}$$

These equations allow a few observations that have analogies all over quantum control: Firstly, the control $f(t)$ is needed to push the system away from its natural dynamics (the terms on the left) – it is redirecting the natural drift of the system. Secondly, there are two constraints for a function given through an integral – so we can expect many solutions. As an example, let's look at the case that we move a particle by a fixed distance $x(0) = 0$ and $x(T) = X$ from rest to rest $\dot{x}(T) = \dot{x}(0) = 0$. We thus need to satisfy

$$\int_0^T dt' \sin[\Omega(t-t')] f(t') = \Omega X \quad \int_0^T dt' \cos[\Omega(t-t')] f(t') = 0$$

and we can easily show that this fixes low Fourier components of $f(t')$ but leaves higher ones open.

The situation changes, if we impose, e.g., an energetic constraint to the control. This typically leads to constraints of the form

$$\int_0^T dt f^2(t) = \int_{-\infty}^{\infty} d\omega f(\omega) f^*(\omega) \leq A$$

where A is the imposed maximum and we have used symmetry properties of the Fourier transform of real-valued functions. Thus, the sum of Fourier components needs to be bounded and if the constraint is too close, there may not even be any solution. This is an example showing that constraints clearly influence the number of accessible solutions and their potential performance, which is commonly seen in practice.

2.2 Optimal control for a classical system

The previous section hinged on having a closed-form Green's function solution of the equation of motion, which is not always available. This follows chapter 2.3 of Bryson and Ho [3].

Suppose we have a dynamical system that can be controlled by a control parameter u that enters a dynamic equation for the state variable x in the form

$$\dot{x} = f[x(t), u(t), t] \quad 0 \leq t \leq T \quad (1)$$

with a given $x(0)$. Both x and u can be single variables or vectors of variables. We wish to optimize a cost function at the end of the process $J[x(T), T]$. We recall classical Lagrangian mechanics and introduce a Lagrange multiplier function λ and can thus state based on the constrained calculus of variations that we need to find a stationary point of

$$\bar{J} = J[x(T), T] + \int_0^T dt \lambda^T(t) (f[x(t), u(t), t] - \dot{x})$$

where we have allowed for the complex of coupled equations and thus vector-valued Lagrange multipliers.

The introduction of the Lagrange multiplier allows for the optimization of J , while satisfying the equation of motion (1) at specified times. As such, this means λ has to be time-dependent as well.

We introduce the associated Hamilton's function (which has a similar mathematical origin in the calculus of variations as the Hamiltonians of mechanics yet a very different physical motivation)

$$H[x(t), u(t), \lambda(t), t] = \lambda^T(t) f[x(t), u(t), t] \quad (2)$$

and rewrite our constrained cost function by integrating the last term by parts

$$\bar{J} = J[x(T), T] + \lambda^T(T)x(T) - \lambda^T(0)x(0) + \int_0^T dt \left\{ H[x(t), u(t), \lambda(t), t] + \dot{\lambda}^T x(t) \right\}.$$

Now let's consider the variation in \bar{J} based on variations in $u(t)$ recalling that the times as well as the initial state variable are given. We find

$$\delta\bar{J} = \left(\frac{\partial J}{\partial x} - \lambda^T \right) \delta x \Big|_{t=T} + \lambda^T \delta x \Big|_{t=0} + \int_0^T dt \left[\left(\frac{\partial H}{\partial x} + \dot{\lambda}^T \right) \delta x + \frac{\partial H}{\partial u} \delta u \right].$$

Note that in general we choose the variation at the beginning to be $\delta x(0) = 0$, since we know the exact initial state of the dynamics.

Now the variations of x and u are not independent, they are linked by the equation of motion. Were we not to work with the Lagrange multiplier, we would need to tediously solve the equation of motion for different control functions and then work out how these variations are related. Fortunately, the Lagrange multiplier method allows us to circumvent that problem. Our goal is for $\delta\bar{J}$ to vanish to first order. Choosing a specific Lagrange multiplier to realize this, we finally arrive at

$$\dot{\lambda}^T = -\frac{\partial H}{\partial x} = -\lambda^T \frac{\partial f}{\partial x} \quad \lambda^T(t_f) = \frac{\partial J}{\partial x(t_f)}. \quad (3)$$

These are the Euler-Lagrange equations pertaining to the system. That being satisfied, we are left with the total variation

$$\delta\bar{J} = \underbrace{\lambda^T(0)\delta x(0)}_{=0} + \int_0^T dt \frac{\partial H}{\partial u} \delta u$$

For an extremum to be reached under any variation of the control, we need

$$\frac{\partial H}{\partial u} = \lambda^T \frac{\partial f}{\partial u} = 0 \quad 0 \leq t \leq T. \quad (4)$$

We have shown the ingredients to what can be formalized as the Pontryagin Maximum Principle (PMP). More pragmatically, these equations give us a recipe on how to solve the thus formulated optimal control problem by a coupled gradient search: From a suitable initial guess for $u(t)$

1. Solve the equation of motion eq. (1) to find $x(t)$ using the initial value $x(0)$ that is part of the control problem
2. Find the Lagrange multiplier by solving eq. (3). Note that there is a definite value given at the *end* time T , i.e., we have a final value problem – that is solved like an initial value problem but propagating backwards in time. This back-propagation is typical when we consider this cost functional.
3. With these, compute the effective gradient in eq. (4) and update the values of u following the direction of the gradient. Adjust the step size as needed.

Iterating these three steps will get us to a local solution, depending on the initial conditions, if the control landscape admits one.

2.2.1 Example: Driven harmonic oscillator

Let us get back to formulating these steps for the optimal control problem of the driven harmonic oscillator described above in section 2.1. We identify the control as the dimensionless force $u \equiv f$ and write the equation of motion as a coupled system

$$\begin{aligned}\frac{dx}{dt} &= \dot{x} \\ \frac{d\dot{x}}{dt} &= -\Omega^2 x + u \\ x(0) &= 0 \quad \dot{x}(0) = 0\end{aligned}$$

In order to have a differentiable performance index that forces the particle to end at a at time T and in rest we can write

$$J = \Omega^2 (x - a)^2 + \dot{x}^2. \quad (5)$$

This leads us to Hamilton's function following the prescription of 2

$$H = \lambda_1 \dot{x} + \lambda_2 (u - \Omega^2 x).$$

So the Euler-Lagrange equations 3 describing the Lagrange Multiplier

$$\dot{\lambda}_1 = \lambda_2 \Omega^2 \quad \dot{\lambda}_2 = -\lambda_1 \quad (6)$$

which remarkably describes a free harmonic oscillator. It is such interpretations that lead to the Lagrange multiplier to be called the adjoint system. The final conditions from eq. (3) are

$$\lambda_1(T) = 2\Omega^2(x(T) - a) \quad \lambda_2(T) = 2\dot{x}(T) \quad (7)$$

which are of course both zero if the final conditions are met (thus, for the optimal solution, the adjoint system vanishes at T). The gradient flow for the control is given by eq. (4)

$$\frac{\partial H}{\partial u} = \lambda_2.$$

Again, iterating these equations will give us a suitable control.

We could guess as a first control that $u_0(t) = \Omega^2 a$ (which is the force needed to keep the particle at rest at the final position, so at least a motivated guess) thus leading to the equation of motion

$$\ddot{x}_0 + \Omega^2(x_0 - a) = 0$$

with the solution $x_0(t) = a(1 - \cos \Omega t)$ and thus $\dot{x}_0 = a\Omega \sin \Omega t$. This clearly does not solve the control problem, we have from eq. (5) $J = \Omega^2 a^2$. In fact, the final conditions eq. 7 for the adjoint system are $\lambda_1(T) = -2\Omega^2 a \cos \Omega T$ and $\lambda_2(T) = 2a\Omega \sin \Omega T$ leading us, by solving eq. (6)

$$\begin{aligned}\lambda_1 &= -2\Omega^2 a \cos \Omega T \cos [\Omega(t - T)] + 2a\Omega^2 \sin \Omega T \sin [\Omega(t - T)] \\ &= -2a\Omega^2 \cos \Omega t\end{aligned}$$

and $\lambda_2 = -2 \sin \Omega t$. This means that the gradient suggests introducing a resonant drive – as we have seen from the exact solution above.

For further treatment of the classical Harmonic oscillator, see [4].

2.3 Gradient-based optimal quantum control with the GRAPE algorithm

These principles can be transferred to the control of quantum systems in a straightforward way. This is easily illustrated with the GRAdient Ascent Pulse Engineering (GRAPE) algorithm [5].

2.3.1 State-to-state control

We start with a simple state preparation problem. Suppose WLOG that our system is described by a Hamiltonian

$$\hat{H}(t) = \hat{H}_0 + \sum_{i=1}^n u_i(t) \hat{H}_i.$$

We call the time-independent part of the Hamiltonian \hat{H}_0 the drift, the fields u_i are the controls and \hat{H}_i are the control Hamiltonians. In atomic physics, say, \hat{H}_0 describes the energy level structure of the atom, u_i are laser or microwave fields and \hat{H}_i are dipole operators describing the different field modes including polarization. Our task is now to start at an initial state $|\psi_0\rangle$ at time $t = 0$ and find controls such that we reach state $|\psi_1\rangle$ at time $t = T$. As in quantum physics the global phase is meaningless, this corresponds to maximizing the overlap $J = |\langle \psi_1 | \psi(T) \rangle|^2$. The dynamics of our system is, of course, subject to the Schrödinger equation

$$i\hbar \partial_t |\psi(t)\rangle = \hat{H}(t) |\psi(t)\rangle.$$

Mathematically we got ourselves a system of the exact same structure as the previous one. We give its derivation in the form of Ref. [5].

Many practical generators for u_i such as standard arbitrary wave form generators (AWGs or Arbs) used in superconducting qubits represent² the pulse in a piecewise constant fashion, so it is natural³ to represent the $u_i(t)$ in that same way: We chop the total time into N intervals of length $\delta t = T/N$ and write

$$u_i(t) = u_i(j) \quad \text{for} \quad (j-1)\delta t \leq t < j\delta t.$$

This allows us to write down the formal solution of the Schrödinger equation as

$$\hat{U}(T) = \hat{U}_N \hat{U}_{N-1} \cdots \hat{U}_2 \hat{U}_1$$

² but not necessarily output, as the output is typically smoothed and filtered

³ although not always optimal, see below

with

$$\hat{U}_k = \exp \left(-\frac{i}{\hbar} \delta_t \left(\hat{H}_o + \sum_i u_i(j) \hat{H}_i \right) \right) \quad (8)$$

which we can introduce into the performance index as

$$J = \left| \langle \psi_1 | \hat{U}(T) \psi_0 \rangle \right|^2 = \left| \langle \psi_1 | \hat{U}_N \cdots \hat{U}_1 \psi_0 \rangle \right|^2.$$

We are at liberty to move some of the factors into the adjoint state, giving us

$$J = \left| \langle U_{m+1}^\dagger \cdots U_N^\dagger \psi_1 | \hat{U}_m \cdots \hat{U}_1 \psi_0 \rangle \right|^2$$

or $J = |\langle \lambda_m | \rho_m \rangle|^2$ with $|\rho_m\rangle = \hat{U}_m \cdots \hat{U}_1 |\psi_0\rangle$, $|\lambda_m\rangle = \hat{U}_N \cdots \hat{U}_{m+1} |\psi_1\rangle$. Here, the partially propagated state $|\rho_m(t)\rangle$ is overlapped with the partially back-propagated adjoint state $|\lambda_m(t)\rangle$ – both states are overlapped at time t_m . We thus sweep the time at which we calculate the overlap based on the actual pulse that we apply. Now the final ingredient we need is the derivative of an exponential proven in Theorem 4.5 of [6] (see also [7])

$$\left. \frac{d}{dt} \right|_{t=0} e^{X+tY} = e^X \left\{ Y - \frac{[X, Y]}{2!} + \frac{[X, [X, Y]]}{3!} - \cdots \right\} \quad (9)$$

Both of these together allow us to determine all the gradients needed to compute an update at any time step as shown in the left column of figure 1.

We can rewrite this as

$$\left. \frac{d}{dt} \right|_{t=0} e^{X+tY} = e^X \int_0^1 d\tau e^{\tau X} Y e^{-\tau X}$$

by simple power counting. This allows us to analytically compute the derivative of the propagator across one time step by identifying $\hat{X} = \hat{H}(t)$ (the Hamiltonian including the current values of the control) and $\hat{Y} = \hat{H}_i$, one of the control Hamiltonians. In order to simplify the right-hand side, we define $\hat{U}_k(j) = \hat{U}_k^j$ (taking the exponential here simply means to stretch time and study the integral on the right

$$\begin{aligned} & \int_0^1 d\tau \hat{U}_k(j) \hat{H}_i \hat{U}_k^\dagger(j) \\ &= \int d\tau \left(1 - i\tau \delta_t \hat{H} - \tau^2 \delta_t^2 \hat{H}^2 + \dots \right) \hat{H}_i \left(1 + i\tau \delta_t \hat{H} - \tau^2 \delta_t^2 \hat{H}^2 + \dots \right) \\ &= \int d\tau \left(\hat{H}_i - i\tau \delta_t [\hat{H}, \hat{H}_i] + \dots \right) \\ &\simeq \hat{H}_i \end{aligned}$$

where we assume that the time steps chosen are so small that the integral over the commutator can be neglected⁴. A self-contained derivation is presented later in 2.5.4. Restoring all the units leads us to the closed gradient formula

$$\frac{\partial J}{\partial u_i(j)} = -i\delta_t \left\langle \lambda_j \left| \hat{H}_i \right| \rho_j \right\rangle \quad (10)$$

⁴ we will later, under the Magnus expansion, study related steps more carefully

meaning that we can expect, with an appropriate value of ϵ compute a gradient-based update

$$u_i(j) \mapsto u_i(j) + \epsilon \frac{\partial J}{\partial u_i(j)} \quad (11)$$

This allows us to extremalize J hence to find controls that best approximate the final state with the following algorithm. Starting from an initial guess for the controls:

1. Compute the propagated initial state $|\rho_m\rangle = \hat{U}_m \cdots \hat{U}_1 |\psi_0\rangle$ for all $m \leq N$ by iterative matrix multiplication.
2. Compute the back-propagated final state $|\lambda_m\rangle = \hat{U}_N \cdots \hat{U}_{m+1} |\psi_1\rangle$ by iterative matrix multiplication
3. Compute the gradient of the performance index and update the controls following eqs. (10), (11)
4. Iterate until the value of J is satisfactory or the updates are below a threshold

There are a lot of practical improvements that were found beyond this which we will describe below.

One must not underestimate the importance of this analytical derivation of a gradient. Whenever a gradient is available, it greatly improves the convergence of a search specifically when going from a rough initial guess that can often be obtained by solving an approximate version of the problem at hand to a solution that has the very high precision generally demanded by quantum technologies. If a gradient is available, its analytical and exact derivation is also paramount – numerical gradients are very hard to control numerically as they involve a small difference between two potentially large numbers. In pioneering, pre-GRAPE work [8] this was rather obvious – even with large computational effort, only few parameters could be optimized.

2.3.2 An alternative, direct derivation

An alternative derivation of the variational approach to quantum optimal control is as follows:

Let us again look at the state transfer task. We shall construct a functional, J , to be maximized, and utilize Lagrange multipliers to enforce both the initial condition and the equation of motion. We shall parameterize our control fields, $u(t)$ using a vector of scalar real parameters $\vec{\alpha}$.

Our aim is to maximize the overlap of the goal state $|\psi_{\text{goal}}\rangle$ and the state at final time T , $|\psi(T)\rangle$,

$$J_{\text{goal}} = |\langle \psi(T) | \psi_{\text{goal}} \rangle|^2. \quad (12)$$

We need to impose an initial condition, utilizing a Lagrange multiplier λ_{init}

$$J_{\text{init}} = \lambda_{\text{init}} (|\langle \psi(0) | \psi_{\text{init}} \rangle|^2 - 1). \quad (13)$$

Next, we must guarantee the Schrödinger equation, $(i\hbar\partial_t - H(\bar{\alpha}, t))|\psi(t)\rangle = 0$ is upheld at all times. To do that, at each point in time, t , we must multiply the equation of motion by the Lagrange multiplier $\langle\chi(t)|$, and we must add the contributions for all points in time:

$$J_{\text{e.o.m}} = \int_0^T \langle\chi(t)| i\hbar\partial_t - H(\bar{\alpha}, t) |\psi(t)\rangle \quad (14)$$

Note that in $J_{\text{e.o.m}}$, $\langle\chi(t)|$ can be interpreted as a conjugate state, propagating backwards in time, as the term can be rewritten as $\langle(-i\hbar\partial_t - H(\bar{\alpha}, t))\chi(t)|\psi(t)\rangle$.

The functional to be minimized is then

$$J = J_{\text{init}} + J_{\text{e.o.m}} + J_{\text{goal}} \quad (15)$$

We then proceed in the standard variational approach, taking the gradient of this functional with respect to $\bar{\alpha}$ and requiring

$$\partial_{\bar{\alpha}} J = 0. \quad (16)$$

2.3.3 Synthesis of unitary gates

We will now go to the topic of finding controls that best approximate a quantum gate. This can be viewed as a generalization of the state preparation problem to rotating a full basis of the Hilbert space into a desired new basis. This first begs the question of how to find an appropriate performance index. It can be accomplished by starting with a distance measure between the desired and the actual final unitary $\|\hat{U}_{\text{target}} - \hat{U}(T)\|$. The most common choice is based on the 2-norm

$$\begin{aligned} \|\hat{U}_{\text{target}} - \hat{U}(T)\|_2^2 &= \text{Tr} \left[\left(\hat{U}_{\text{target}}^\dagger - \hat{U}^\dagger(T) \right) \left(\hat{U}_{\text{target}} - \hat{U}(T) \right) \right] \\ &= \text{Tr} \left[\hat{U}_{\text{target}}^\dagger \hat{U}_{\text{target}} + \hat{U}^\dagger(T) \hat{U}(T) - \hat{U}_{\text{target}}^\dagger \hat{U}(T) - \hat{U}^\dagger(T) \hat{U}_{\text{target}} \right] \\ &= 2 \left(d - \text{ReTr} \hat{U}_{\text{target}}^\dagger \hat{U}(T) \right) \end{aligned}$$

where d is the underlying Hilbert space dimension. Thus, we see that minimizing the error corresponds to maximizing the overlap $\text{ReTr} \hat{U}_{\text{target}}^\dagger \hat{U}(T)$.

Now the real part looks suspicious – if we have the gate right up to a global phase, $\hat{U}(T) = e^{i\phi} \hat{U}_{\text{target}}$ this overlap indicates a non-perfect result. In fact, numerical experimentation shows that this would be a serious drawback. We can trace this error back to the original distance measure. The high-brow step to take now would be to elevate the description to full quantum

channels. Pragmatically, we move from real part to absolute square and thus the most common performance index for gates is

$$J = \left| \text{Tr} \left(\hat{U}_{\text{target}}^\dagger \hat{U}(T) \right) \right|^2.$$

This quantity can be interpreted in a somewhat operational fashion: First apply the gate you have, then undo the gate you want. If everything goes right you have but a global phase – the same one on all vectors of the standard basis. If not, you measure the deviation from unity for the complete standard basis. There are other possible choices (and good reasons to think about them), which we will discuss later. With this quantity, we can proceed in a way similar to state transfer, only that now we of course start at the unit matrix. We again use piecewise constant controls and define both the intermediate propagator and the intermediate back-propagated target

$$\hat{X}_j = \hat{U}_j \cdots \hat{U}_1 \quad \hat{P}_j = \hat{U}_{j+1}^\dagger \cdots \hat{U}_N^\dagger \hat{U}_{\text{target}}$$

allowing us to rewrite $J = \left| \text{Tr} \hat{P}_j^\dagger \hat{X}_j \right|^2$ for all values of j . We can now apply the same identities as before and find

$$\begin{aligned} \frac{\partial J}{\partial u_i(j)} &= \frac{\partial}{\partial u_i(j)} \left(\text{Tr} \hat{P}_j^\dagger \hat{X}_j \right) \left(\text{Tr} \hat{P}_j^\dagger \hat{X}_j \right)^* \\ &= 2\text{Re} \left[\left(\frac{\partial}{\partial u_i(j)} \text{Tr} \hat{P}_j^\dagger \hat{X}_j \right) \left(\text{Tr} \hat{P}_j^\dagger \hat{X}_j \right) \right] \\ &= -2i\delta t \text{Re} \left[\left(\text{Tr} \hat{P}_j^\dagger \hat{H}_i \hat{X}_j \right) \left(\text{Tr} \hat{P}_j^\dagger \hat{X}_j \right) \right]. \end{aligned}$$

With this analytical gradient, the GRAPE algorithm can be applied as above.

2.4 The Krotov algorithms

The Krotov algorithm [9–12] has been formulated before the GRAPE algorithm. Some of its presentations are historically based on applications in chemistry and emphasizes constraints more than its core. Looking back on how GRAPE is applied, we are blessed with an analytical gradient formula which in each iteration allows us to calculate the gradient of the cost function(al) with respect to all controls at all times and then by walking against it look for improved controls. Notably, the gradient is always computed at a point in parameter space given by the controls computed in the *previous* iteration.

There are two different algorithms which carry the name “Krotov” – a fact which can be quite confusing, even for experts in the field.

The first Krotov, prides itself with its monotonic convergence, which is achieved by propagating the forward state using the old control field, while the backward-propagating state makes use of the new field. A detailed description, with Python implementation, can be found in [13].

The second Krotov can be considered a greedy version of GRAPE, and is described in detail in [14]: In this version of the Krotov algorithm, all previously computed knowledge is used, i.e., once an entry to the gradient is computed, it is applied right away and the next element of the gradient is computed with that correction already applied. This approach of not leaving any information behind in general lowers the number of iterations needed to reach convergence and it comes with proven monotonic convergence. On the other hand, each iteration step takes more time.

The various update strategies are visualized in figure 1.

Benchmarking of the various optimal control algorithms is a topic of ongoing research.

2.5 Modern numerical issues

2.5.1 Control landscapes

A gradient search with an analytical gradient as outlined is the best way to find a local extremum of an optimization landscape. If the optimization landscape has multiple local minima, it can get stuck in a local minimum and needs to be enhanced.

In a seminal series of papers, Rabitz has shown (see e.g., [15]) that there is indeed only one extremum in the control landscape and that it is global. This theorem is a correct derivation of its assumptions – one of which is the absence of constraints in pulse amplitude and temporal resolution. In practice, these constraints exist and multiple local extrema exist – the more constrained the optimization, the more local extrema. Specifically in situations close to the quantum speed limit (see below), with low control resolution (Ref [16] looks at a single bit of amplitude resolution and required genetic algorithms to converge) or with complex many-body dynamics and only few controls, these call for more advanced methods.

If one has a good intuition about the optimal pulse say, by solving a model that is very close to the desired model or by rescaling a solution that works at a longer gate duration, one can often stay close to the global extremum and otherwise requires a gradient search. If that is not the case, one needs to first start with a more global search method covering a large parameter space. Known systems for such gradient-free approaches are GROUP [17], genetic algorithms ([16, 18]), they are part of CRAB (see section 3.2) and simulated annealing [19].

2.5.2 Fidelities

We would like to come back to the choice of fidelity based on the 2-norm described above. It has been argued that the most appropriate way to characterize quantum processes is the use of the diamond norm [20]. It can be expressed for a quantum operation \mathcal{E} compared to an ideal

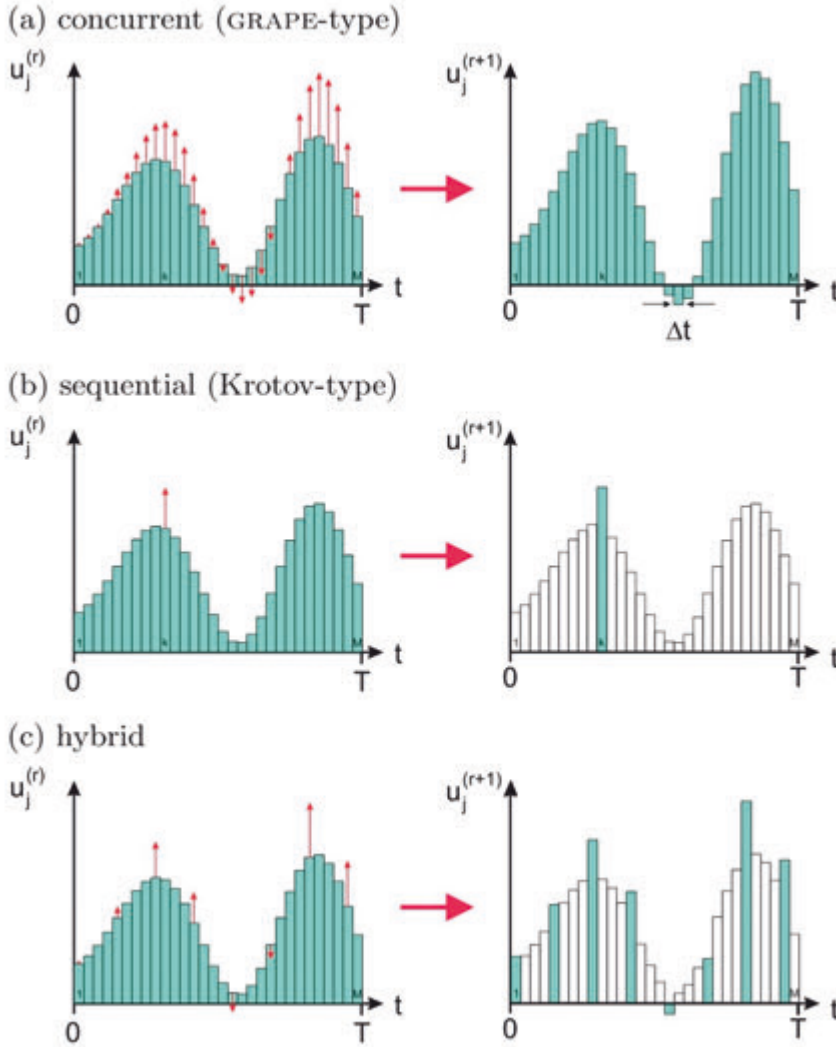


Fig. 1: Overview on the update schemes of gradient-based optimal control algorithms in terms of the set of time slices $T(q) = \{k(q), k(q), \dots, k(q)\}$ for which the control amplitudes are concurrently updated in each iteration. Subspaces are enumerated by q , gradient-based steps within each subspace by s , and r is the global step counter. In grape (a) all the M piecewise constant control amplitudes are updated at every step, so $T(1) = \{1, 2, \dots, M\}$ for the single iteration $q \equiv 1$. Sequential update schemes (b) update a single time slice once, in the degenerate inner-loop $s \equiv 1$, before moving to the subsequent time slice in the outer loop, q ; therefore here $T(q) = \{q \bmod M\}$. Hybrid versions (c) follow the same lines: for instance, they are devised such as to update a (sparse or block) subset of p different time slices before moving to the next (disjoint) set of time slices.

operation \mathcal{U} as

$$\|\mathcal{U}_{\text{ideal}} - \mathcal{E}\|_{\diamond} = \sup_q \max_{\psi} |\text{Tr} [\mathcal{U}_{\text{ideal}}(|\psi\rangle\langle\psi|) - \mathcal{E}(|\psi\rangle\langle\psi|)]| \quad (17)$$

This involves two generalizations of the 2-norm: On the one hand, rather than taking the 2-norm distance which is equivalent to averaging over all possible input states to the operation, we are taking the maximum over $|\psi\rangle$, i.e., we choose the input state that maximizes the distance. On the other hand, rather than directly using the unitary operation, we enhance the Hilbert space by adding another space of dimension q on which the identity operation is performed. The diamond norm is then the supremum over q . The latter may sound rather academic, but it is not if, e.g., the initial state is entangled between the original and the auxiliary system.

For the purposes of quantum optimal control, the diamond norm is rather impractical – it is hard to compute (as it contains a supremum) and it can be non-differentiable (as it contains taking a maximum over states, the state at which it has reached can jump in state space). What does this mean for the applicability of quantum optimal control in the context of fault tolerance?

There are two answers to this question. On the one hand, one can at least find performance indices that emphasize the worst case more strongly while being differentiable. A straightforward option is [21]

$$\begin{aligned} J_q &= \max_{\alpha \in [0, 2\pi)} \left\| \hat{U}_{\text{target}} - e^{i\alpha} \hat{U}(T) \right\|_{2q}^{2q} \\ &= \max_{\alpha \in [0, 2\pi)} \text{Tr} \left[\left(\hat{U}_{\text{target}}^\dagger - e^{-i\alpha} \hat{U}^\dagger(T) \right) \left(\hat{U}_{\text{target}} - e^{i\alpha} \hat{U}(T) \right) \right]^q \end{aligned}$$

which can be implemented in a straightforward fashion yet does not have a known extension that avoids optimizing the global phase.

On the other hand, it is pragmatically not very crucial to go through these steps as long as the algorithm converges properly: Our goal is to get the error as close to zero as possible and, as these norms can be continuously mapped onto each other, one pragmatically searches for controls that reduce the error in the 2-norm to an extremely low value which guarantees that even in the desired norm the error is low enough – using the paradigm to control and verify with two different measures.

2.5.3 Increasing precision of GRAPE

The GRAPE algorithm above defines a straightforward gradient algorithm for optimal control. There are a few known measures to speed up its convergence.

One measure is the improvement of the use of the gradient by moving to a quasi-Newton method, the Broyden, Fletcher, Goldfarb, and Shanno (BFGS) method [22]. Newton's method, as the reader may have learned in an elementary introduction to numerical mathematics, rely on approximating the function whose zero we desire to find by its tangent – in our case, we desire

to find the zero of the gradient, i.e., we need to approximate the functions up to its second derivative. As we are optimizing a scalar that depends on many parameters – all the controls taken at all the times of interest – the matrix of second derivatives is a high-dimensional object. In order to approximate the zero of the gradient, one would have to invert that matrix, which is numerically hard and would likely negate the potential computational advantage. The BFGS method instead relies on directly approximating the inverse Hessian

2.5.4 The gradient of a matrix exponential

Expanding on the discussion surrounding eq. (8), (9), any gradient-driven optimal control optimization, such as GRAPE or Krotov, which treats the control fields as piecewise constant, will describe the coherent propagator of time slice m as

$$U_m = \exp \left(-\frac{i}{\hbar} \delta_t H(\bar{\alpha}, t_m) \right) \quad (18)$$

where $\bar{\alpha}$ parameterizes the control functions $u(t)$. We are searching for the value of $\bar{\alpha}$ which will minimize the infidelity. At step j of the optimization, to compute the gradient of the goal function with respect to $\bar{\alpha}$, we must compute $\partial_{\bar{\alpha}} U_m(\bar{\alpha})|_{\bar{\alpha}=\bar{\alpha}_j}$. At this point we can rewrite eq. (18) as in eq. (9),

$$U_m = \exp \left(-\frac{i}{\hbar} \delta_t \left(H_{m,j} + \epsilon_{\bar{\alpha}} \tilde{H}_{m,j} \right) \right)$$

where $\epsilon_{\bar{\alpha}}$ is small and we seek $\partial_{\epsilon_{\bar{\alpha}}} U_m(\epsilon_{\bar{\alpha}})|_{\epsilon_{\bar{\alpha}}=0}$. Following [23, 24], and their summary in Appendix A of [14], we denote the eigenvalues and eigenvectors of H_j by e_k and $|e_k\rangle$, respectively, then using the spectral theorem

$$\langle e_l | \partial_{\epsilon_{\bar{\alpha}}} U_m | e_k \rangle = \begin{cases} -\frac{i}{\hbar} \delta_t \langle e_l | \tilde{H}_{m,j} | e_k \rangle \exp \left(-\frac{i}{\hbar} \delta_t e_l \right) & \text{if } e_l = e_k \\ -\frac{i}{\hbar} \delta_t \langle e_l | \tilde{H}_{m,j} | e_k \rangle \frac{\exp \left(-\frac{i}{\hbar} \delta_t e_l \right) - \exp \left(-\frac{i}{\hbar} \delta_t e_k \right)}{-\frac{i}{\hbar} \delta_t (e_l - e_k)} & \text{if } e_l \neq e_k \end{cases}$$

one may invoke the spectral theorem in a standard way and calculate matrix functions via the eigendecomposition.

To simplify notation, we shall look at $\partial_x e^{A+xB}$, with A, B being an arbitrary pair of Hermitian (non-commuting) matrices and $x \in \mathbb{R}$. As previously $\{|e_l\rangle\}$ as the orthonormal eigenvectors to the eigenvalues $\{e_l\}$ of A . We then obtain the following straightforward, if somewhat lengthy, derivation:

$$\begin{aligned}
D &= \left. \langle e_l | \partial_x e^{A+xB} | e_k \rangle \right|_{x=0} \\
&= \left. \left\langle e_l \left| \partial_x \sum_{n=0}^{\infty} \frac{1}{n!} (A+xB)^n | e_k \right\rangle \right|_{x=0} \\
&= \left. \left\langle e_l \left| \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{q=1}^n (A+xB)^{q-1} B (A+xB)^{n-q} | e_k \right\rangle \right|_{x=0} \\
&= \left\langle e_l \left| \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{q=1}^n A^{q-1} B A^{n-q} | e_k \right\rangle \right. \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{q=1}^n e_l^{q-1} \langle e_l | B | e_k \rangle e_k^{n-q} \\
&= \langle e_l | B | e_k \rangle \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{q=1}^n e_l^{q-1} e_k^{n-q}
\end{aligned}$$

This provides the answer for in the case where $e_l = e_k$. For $e_l \neq e_k$ a bit more work is needed:

$$\begin{aligned}
D &= \langle e_l | B | e_k \rangle \sum_{n=0}^{\infty} \frac{1}{n!} e_k^{n-1} \sum_{q=1}^n \left(\frac{e_l}{e_k} \right)^{q-1} \\
&= \langle e_l | B | e_k \rangle \sum_{n=0}^{\infty} \frac{1}{n!} e_k^{n-1} \frac{(e_l/e_k)^n - 1}{(e_l/e_k) - 1} \\
&= \langle e_l | B | e_k \rangle \sum_{n=0}^{\infty} \frac{1}{n!} \frac{e_l^n - e_k^n}{e_l - e_k} \\
&= \langle e_l | B | e_k \rangle \frac{e^{e_l} - e^{e_k}}{e_l - e_k}
\end{aligned}$$

Note that we have explicitly made use of the orthogonality of eigenvectors to different eigenvalues in normal matrices.

3 Applied optimal quantum control

While quantum optimal control is a well-developed field and has been very successful in atomic and molecular systems, its track record in solid-state quantum technologies is somewhat less developed. The reason has to do with the accuracy of the models, i.e., the precision at which we know every ingredient of the Hamiltonian. First of all, a quantum-technological device (specifically, but not exclusively, in the solid state) has human-made components which contain some fabrication uncertainty. This affects the drift Hamiltonian – even if its eigenvalues can be accurately determined using spectroscopy, it is much more involved to find its eigenvectors.

These naturally also affect the matrix elements of the control Hamiltonians. On top of that, some solid-state quantum devices need to be extremely well isolated from their environments including high-temperature black-body radiation. This means, that an applied control signal will get distorted on its way to the sample in a way that can be measured only to a limited degree, see fig. 2 for a summary. While one can improve hardware and characterization to meet these challenges, it is hard to get this to the precision required by, say, fault-tolerant quantum computing. Thus, other approaches are called for.

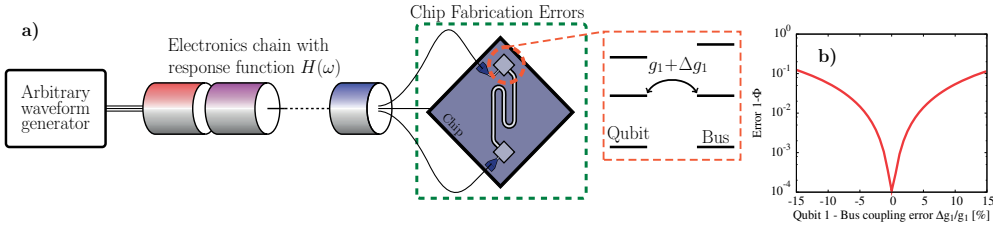


Fig. 2: Typical sources of inaccuracy in quantum control for superconducting qubits including the transmission from the generator to the sample and inaccuracy of the Hamiltonian model. Right: Typical error sensitivity for a gate between superconducting qubits.

3.1 Closing the loop for pulse calibration

One possible approach to handle uncertainties would be to use a robust control methodology inspired by magnetic resonance in ensembles. While this method can be useful, it slightly misses the point: It still requires a good estimate for the uncertainty and then it improves performance across the relevant parameter interval. Here, the situation is different, we do not have a parameter distribution but a single set of parameters – we just cannot find it or even the relevant model *a priori*.

One way to still find good pulses are hybrid control methods such as Adaptive Hybrid Optimal Control (AdHOC, [25]), Optimized Randomized Benchmarking for Immediate Tuneup (ORBIT, [26]), and Adaptive Control via Randomized Optimization Nearly Yielding Maximization (ACRONYM, [27]). The idea of these methods is rather similar: After an initial design phase that may or may not contain traditional optimal control, a set of pulses is constructed based on models that are believed to approximate the actual system but its parameterization is left open to some corrections. These corrections are then determined in a closed loop – the fidelity is measured and the pulses are updated based on these fidelity measurements.

In the example of AdHOC, the pulse measurement is based on randomized benchmarking (described below) and the optimization that determines the corrections is based on the Nelder-Mead simplex algorithm, which is available in most numerical mathematics toolboxes. What is crucial is that this is a gradient-free algorithm in order to avoid issues with taking gradients of measurement data. Is that as a simplex algorithm, the search for a pulse described by n

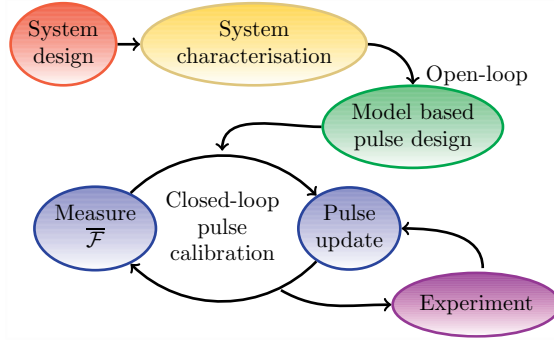


Fig. 3: Typical two-stage control workflow with an open loop modeling stage

parameters needs to be initialized using $n + 1$ initial guesses. This raises the important question how the number n can be kept as small as possible (but not smaller, see below) by finding an efficient parameterization.

This is not an easy problem. So far, we have always assumed that the pulses are parameterized in piecewise constant fashion and have argued that this is naturally compatible with arbitrary wave form generators. However, this parameterization does not naturally lend itself to reduction of the number of parameters – simple, sparse controls in quantum physics are typically sine and cosine functions with smooth, Gaussian-derived envelopes. On the other hand, the piecewise constant parameterization was instrumental in deriving the gradient formula in an analytical way and cannot be easily removed.

3.2 CRAB

Albeit originally developed from a different motivation, the optimization of many-body dynamics, the Chopped RAndom Basis (CRAB)⁵ algorithms serves that purpose, [28]. It introduced the concept of simple and sparse pulse parameterizations, i.e., finding a pulse parameterization that is not necessarily piecewise constant but rather can be written as

$$H(\bar{\alpha}, t) = H_0 + \sum_{k=1}^C c_k(\bar{\alpha}, t) H_k, \quad (19)$$

where the functions c_k can e.g., be harmonic functions characterized by amplitude, frequency and phase or a sequence of Gaussians

$$c_k(\bar{\alpha}, t) = \sum_{j=1}^m A_{k,j} \exp\left(-(t - \tau_{k,j})^2 / \sigma_{k,j}^2\right). \quad (20)$$

In complex systems that were the initial motivation for CRAB, one has very little prior knowledge about a suitable basis and it is at best chosen random, hence the name. CRAB utilizes a gradient-

⁵ pronounced with a rolling 'r' and a voiced 'b'

free search, specifically Nelder-Mead (although other algorithms could be used), similarly to what we have already described for AdHOC⁶.

The fact CRAB is model-free, with the gradient-free search treating the quantity to be optimized as a black box, provides a distinct advantage in situations a precise model is unknown or when the model is known, but the gradient cannot be computed due to numerical complexity or other reasons. This makes CRAB appropriate for closed-loop experimental calibration of control fields in systems ranging from nitrogen vacancy centers in nano-diamonds [29] and cancer treatment formulations [30], to DMRG-based simulations [31]. Further, CRAB enjoys huge success in studying quantum phase transitions, preparing large Schrödinger cat states, sensing and many more.

A variant of CRAB, known as dCRAB [32], deals with a situation where the control parameterization has a higher dimensionality than can be optimized by Nelder-Mead, by iteratively optimizing different subsets (or low-dimension projections) of the high-dimension full parameter space.

3.3 GOAT

Gradient Optimization of Analytic conTrols (GOAT) is a recently [33] proposed optimal control algorithm which does not derive from the variational formulation of optimal control, defined earlier. Rather, GOAT finds the equations of motion for the gradient of the goal function with respect to the control parameters, integrating as you would the Schrödinger equation (as piecewise-constant approximation, or using standard ODE tools such as Runge-Kutta optimizers).

For our purpose, the goal function to minimize is defined as the projective SU distance (infidelity) between the desired gate, U_{goal} , and the implemented gate, $U(T)$, [34] (also [35])

$$g(\bar{\alpha}) := 1 - \frac{1}{\dim(U)} \left| \text{Tr} \left(U_{\text{goal}}^\dagger U(T) \right) \right|, \quad (21)$$

where $U(t)$ is the time ordered (\mathbb{T}) evolution operator

$$U(\bar{\alpha}, T) = \mathbb{T} \exp \left(\int_0^T -\frac{i}{\hbar} H(\bar{\alpha}, t) dt \right). \quad (22)$$

GOAT's ability to use any control ansatz makes it feasible to find drive shapes described by a small number of parameters, suitable for closed-loop calibration.

A gradient-based optimal control algorithm requires two ingredients: an efficient computation of $\partial_{\bar{\alpha}} g(\bar{\alpha})$ and a gradient-based search method over parameter space. GOAT presents a novel method for the former, while using any standard algorithm for the latter, such as BFGS.

Consider the gradient of the goal function eq. (21) with respect to $\bar{\alpha}$,

$$\partial_{\bar{\alpha}} g(\bar{\alpha}) = -\text{Re} \left(\frac{g^*}{|g|} \frac{1}{\dim(U)} \text{Tr} \left(U_{\text{goal}}^\dagger \partial_{\bar{\alpha}} U(\bar{\alpha}, T) \right) \right). \quad (23)$$

⁶ note that CRAB was proposed before AdHOC

Neither $U(\bar{\alpha}, T)$ nor $\partial_{\bar{\alpha}}U(\bar{\alpha}, T)$ can be described by closed form expressions. U evolves under the equation of motion $\partial_t U(\bar{\alpha}, t) = -\frac{i}{\hbar} H(\bar{\alpha}, t) U(\bar{\alpha}, t)$. By taking the derivative of the U equation of motion with respect to $\bar{\alpha}$ and swapping derivation order, we arrive at a coupled system of equations of motion for the propagator and its gradient,

$$\partial_t \begin{pmatrix} U \\ \partial_{\bar{\alpha}}U \end{pmatrix} = -\frac{i}{\hbar} \begin{pmatrix} H & 0 \\ \partial_{\bar{\alpha}}H & H \end{pmatrix} \begin{pmatrix} U \\ \partial_{\bar{\alpha}}U \end{pmatrix}. \quad (24)$$

As $\bar{\alpha}$ is a vector, $\partial_{\bar{\alpha}}U$ represents multiple equations of motion, one for each component of $\bar{\alpha}$. $\partial_{\bar{\alpha}}H$ is computed using the chain rule.

GOAT optimization proceeds as follows: Starting at some initial $\bar{\alpha}$ (random or educated guess), initiate a gradient driven search (e.g., L-BFGS [22]) to minimize eq. (21). The search algorithm iterates, requesting evaluation of eqs. (21,23) at various values of $\bar{\alpha}$, and will terminate when the requested infidelity is reached or it fails to improve g further. Evaluation of $g(\bar{\alpha})$, $\partial_{\bar{\alpha}}g(\bar{\alpha})$ requires the values of $U(\bar{\alpha}, T)$ and $\partial_{\bar{\alpha}}U(\bar{\alpha}, T)$. These are computed by numerical forward integration of eq. (24), by any mechanism for integration of ordinary differential equations that is accurate and efficient for time-dependent Hamiltonians, such as adaptive Runge-Kutta. Initial conditions are $U(t=0) = \mathcal{I}$ and $\partial_{\bar{\alpha}}U(t=0) = 0$. Note that no back propagation is required.

Experimental constraints can be easily accommodated in GOAT by mapping the optimization from an unconstrained space to a constrained subspace, and computing the gradient of the goal function using the chain rule. For example, $\bar{\alpha}$ components may be constrained by applying bounding functions, e.g., $\alpha_k \rightarrow \frac{1}{2}(v_{\max} - v_{\min}) \sin(\bar{\alpha}_k) + \frac{1}{2}(v_{\max} + v_{\min})$ which imposes $\alpha_k \in [v_{\min} \dots v_{\max}]$. Amplitude constraints and a smooth start and finish of the control pulse can be enforced by passing the controls through a window function which constrains them to a time-dependent envelope. Gradients for $\partial_{\bar{\alpha}}H$ flow via the chain rule.

3.4 Evaluating fidelity with randomized benchmarking.

The closed-loop approaches mentioned above crucially rely on a measurement of success. While in state-transfer problems, e.g., creating an ordered state quickly or steering a chemical reaction, there may be generic tools to determine this success with a given experimental apparatus. In the case of a quantum gate, this is not so simple. While classic textbooks like first label quantum process tomography, this has a number of drawbacks, and is now replaced by more efficient methods.

3.4.1 The trouble with tomography

To understand this, let's first take a look at quantum state tomography [36]. This is, in a nutshell, the reconstruction of a quantum state (characterized by its density matrix) by performing a complete set of observable measurements. Next to some practical drawbacks having to do

with guaranteeing a positive density matrix [37], this is also impractical: A typical quantum device can be read out with a single machine – an electric or optical measurement. Formally this corresponds to measuring in one basis (we will assume that we are dealing with qubits, so recording the expectation value completely characterizes the output distribution). In order to measure a complete set of operators, one has to first perform a basis change in the shape of performing a coherent operation. As this operation itself is prone to error, this will falsify the result. Together with the intrinsic imperfection of the readout device this constitutes measurement error.

From state tomography, it is another step to process tomography, i.e., the reconstruction of a quantum channel – linear map from input to output density matrices – from measured. Formally, one can using the Choi-Jamiołkowski isomorphism [36] map the process matrix of the channel onto the density matrix of a state and treat the problem of process tomography as one of state tomography. Practically, process tomography involves to now measure complete sets of both initial and final states that undergo the channel. Similar to measurement, also state preparation is usually possible only in one distinct basis – if state preparation is performed by measurement it is the measurement basis, if state preparation is performed via thermalization or optical pumping it is the drift Hamiltonian’s eigenbasis – and it is imperfect – both of these give rise to state preparation errors. Thus, in total, the quantum channel that one would like to characterize is masked by state preparation and measurement (SPAM) errors.



Fig. 4: SPAM errors in process tomography: The state that can be prepared and measured needs to be transferred into the basis that has to be prepared and measured, introducing additional errors obfuscating the channel.

On top of that, full process tomography is also forbiddingly labourious. The state of a d dimensional quantum system is characterized by a d^2 entries in a density matrix that, accounting for hermiticity and norm boil down to $d^2 - d + 1$ real numbers. This has to be squared again to describe a quantum channel, leading to $O(d^4)$ numbers – which then are recombined to compute a single fidelity. In an n - Qubit system, we have $d = 2^n$ making full tomography forbiddingly data intensive. On top of that, we would like to ensure complete positivity of the measured channel, which gives rise to inequality constraints that are practically hard to meet specifically when the map is close to unitary. Now there are several methods such as compressed sensing and Monte Carlo sampling [38, 39] that reduce that problem, but with SPAM still included, there is strong motivation to look for an independent method to evaluate fidelity in an experiment. Here, randomized benchmarking and its descendants (RB+) have appeared as a quasi-standard. A comprehensive review of RB+ has currently not been published. We are going to mention

key papers on the way and otherwise refer to the work of J. Emerson.

3.4.2 Randomization of quantum channels

Let's first lay the foundation of how we describe a quantum channel [36]: A linear map that takes any valid density matrix onto another valid density matrix, i.e., with

$$\rho \mapsto \mathcal{E}[\rho]$$

we demand that if ρ is hermitian, positively semidefinite, and has a normalized trace, so is $\mathcal{E}[\rho]$. This is satisfied by the Kraus representation

$$\mathcal{E}[\rho] = \sum_k \hat{A}_k \rho \hat{A}_k^\dagger \quad \sum_k \hat{A}_k^\dagger \hat{A}_k = \mathbb{I}.$$

The (non-unique) *Kraus operators* A_k characterize the channel. It can be easily shown that the Kraus representation leads to a valid channel and it takes a bit more attention to show that the validity of the channel also requires the Kraus representation.

Now to estimate the average fidelity over a channel relative to a desired unitary \hat{U} we apply the channel to a pure initial state, then undo the ideal channel, compute the overlap with the pure state and average over all pure inputs

$$F = \int d\psi \langle \psi | U^\dagger \mathcal{E} [| \psi \rangle \langle \psi | U] | \psi \rangle$$

where the integral runs over a suitable uniform distribution of all states called the Haar measure. We now aim at replacing the average in this formula by another randomization procedure [40]. We now decompose the real operation into an ideal operation followed by an error channel and Kraus-decompose the error channel

$$\mathcal{E} = \Lambda \circ \mathcal{U} \quad \Lambda = \sum_k A_k \rho A_k^\dagger.$$

Plugging this into the expression for the average gets us

$$F = \int d\psi \langle \psi | U^\dagger \Lambda [U | \psi \rangle \langle \psi | U^\dagger] U | \psi \rangle.$$

We can read this expression as implementing the motion-reversal transformation $U^\dagger \cdot U$ with an error Λ occurring in the middle.

Now instead of going for F directly, let us average the fidelity over all unitaries that can enter the motion-reversal map – assuming tacitly that we have the same Λ at all times. We now compute a at first glance very different average – we keep a single initial state $\rho = |\psi\rangle\langle\psi|$ and instead average over all unitaries

$$E = \int dU \operatorname{Tr} [\rho U^\dagger \Lambda (U \rho U^\dagger) U].$$

Now we exchange the order of integration and change the order under the trace and write this as

$$E = \text{Tr} \left(\rho \left[\int dU U \Lambda U^\dagger \right] \rho \right)$$

We can now read this exchanged expression at face value – in the center is noise averaged over all unitaries

$$\Lambda_{\text{ave}} = \int dU U \Lambda U^\dagger.$$

Building on the operations of unitary maps as generalized rotations, this is called a twirled channel. It can be mathematically shown what is physically rather obvious – this channel must be highly symmetric, it cannot prefer any basis over the other. The only channel compatible with this is the depolarizing channel

$$\Lambda_{\text{ave}}[\rho] = p\rho + \frac{1-p}{d}\mathbb{I}$$

which has a single error probability p . With this the error averaged over all unitaries equals the fidelity of the twirled channel computed for a single input state

$$E = \text{Tr}(\rho \Lambda_{\text{ave}} \rho) = F.$$

where the last equality requires some more involved math to show that this is also the same as the average fidelity of a unitary averaged over all states. The fact that a single input state is enough – we have delegated the need for averaging from all states to twirling the channel – addresses the problem of SPAM errors.

Now what is needed is an efficient way to implement Λ_{ave} . We need to replace the integral over all unitaries by a sum over random elements that converges to this integral. This brings in the concept of a unitary 2-design: a set that correctly reproduces the full unitary set in polynomials of degree 2. It can be shown [41, 42] (in a rather pedestrian way) that the Clifford group is sufficient. The Clifford group [36] is formally defined as the normalizer of the Pauli group. For n qubits, this Pauli $P_n = \{\sigma_n\}$ group consists of all direct products of Pauli matrices $\sigma_n = \otimes_{j=1}^n \sigma_{i_j}$, $i_j \in \{0, 1, 2, 3\}$ so the corresponding Clifford group is the set of all unitaries that map all n -qubit Pauli matrices onto Pauli matrices

$$C_n = \{U \in SU(2^n) : \forall \sigma_n \in P_n \exists \sigma_m \in P_m : \sigma_m = U \sigma_n U^\dagger\}.$$

For a single qubit, this group is generated by all quarter-turns around the Bloch sphere. The Clifford group is a discrete group and quantum algorithms consisting of only Clifford gates can be efficiently classically simulated. These together lead to the remarkably simple protocol of randomized benchmarking.

3.4.3 Randomized Benchmarking

Let's pull all of these ingredients together into a handy protocol:

1. Repeat for a few representative sequences
 - (a) Draw a random set of Clifford gates
 - (b) Compute the resulting operation and its inverse. Add the inverse to the end of the sequence
 - (c) Repeat the following to establish an estimate for the final probability for survival of the initial state
 - i. initialize the system in a convenient state
 - ii. run the sequence
 - iii. measure if the outcome is the same state or not
 - (d) Average to estimate the survival probability for the given sequence
2. Average to estimate the survival probability averaged of the Clifford groups. As a function of sequence length, the result will have the form

$$p(n) = p_0 + \lambda^n.$$

Here, λ is the average Clifford gate fidelity and can be determined by fitting, whereas p_0 is the SPAM error.

It turns out practically and can be reasoned analytically that the need for averaging is acceptable, artifacts of ensemble sizes vanish quickly [43].

In this basic version of RB, there are a lot of assumptions that can be questioned. The theory of randomized benchmarking has been extended to adapt most of the demands resulting from weakening these assumptions. We cannot do the vast literature full justice here but mention a few highlights.

First of all, standard RB finds the fidelity averaged over the whole Clifford group. If one instead desires to characterize a single Clifford gate, the technique of interleaved randomized benchmarking (IRB) [44] can be applied. There, one first performs regular RB. Then, one takes the sequences used for RB and interleaves the desired Clifford gate between any two of the gates from the sequence. The inverse to the resulting sequence needs to be re-computed. The comparison between the interleaved and the regular frequencies gives the average fidelity of that special Clifford gate.

In a similar vein, issues like leakage out of the computational subspace, gate-dependent error and others can be taken into account [43], leading to the modern concept of cycle benchmarking. Including non-Clifford gates, however, can only be done at the cost of significant overhead, as the inverting operation is hard to compute as well as hard to invert - it is an arbitrarily quantum gate encompassing the whole system and not part of the Clifford group. A combination of RB with Monte Carlo sampling can be applied to still keep parts of the benefits of IRB [39].

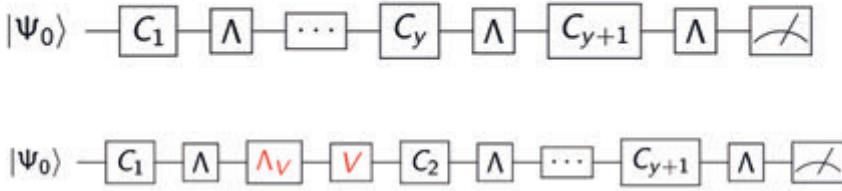


Fig. 5: Quantum channels for randomized benchmarking. Top: Randomized Benchmarking consists of a sequence of random (perfect) Clifford C_i and Errors Λ inverted by the last Clifford gate C_{y+1} . Bottom: Interleaved randomized Benchmarking interleaves a particular Clifford gate V into this sequence.

That being said, in many practical architectures, the only non-Clifford gate is the T-gate, a $\pi/4$ z-axis rotation which can be done in software to high precision, so it is not crucial to calibrate it with optimal control. Also, as the two-qubit CNOT gate is a Clifford gate, one cannot claim that natively and without error correction Clifford gates are easier than non-Clifford.

3.5 Approximating time evolutions with the Magnus expansion

Control calculations involve solving the time-dependent Schrödinger equation. While this can be done analytically in, e.g., rotating wave situations or approximations, this can quickly become hard – even for a system as simple as a harmonically driven two-state-system this is a daunting task [45]. If we would like to proceed analytically with optimal control as far as possible, computing the final gate analytically is a key ingredient to which the Magnus expansion is an important ingredient. Numerically, techniques for coupled ordinary differential equations like Runge-Kutta can be used as well as split-operator techniques. For analytical calculations, one can use the Dyson series familiar from regular advanced quantum mechanics as systematic perturbation theory. In many cases, it is however more effective to use the Magnus expansion, an asymptotic expansion that used the number of nested commutators as a small parameter. It is exact but usually truncated at low order. Our treatment mostly follows [46].

The problem at hand is to start from a Hamiltonian that has a (hopefully) large but solvable component and a perturbation $\hat{H} = \hat{H}_0(t) + \hat{V}(t)$. A clever choice of this division is key and there is no need for the former to be time-independent. We can transfer to the interaction picture with respect to \hat{H}_0 . The resulting transformed perturbation $\hat{V}^I(t)$ will then acquire additional time-dependence, often in the form of large oscillating terms. The objective is now to approximately calculate the time evolution

$$\hat{U}^I(t) = \mathbb{T} \exp \left(-\frac{i}{\hbar} \int_0^t d\tau \hat{V}^I(\tau) \right)$$

where \mathbb{T} is the usual time ordering operator. The Dyson expansion of this term starts as

$$\hat{U}^I(t) = 1 - \frac{i}{\hbar} \int_0^t d\tau \hat{V}^I(\tau) - \frac{1}{\hbar^2} \int_0^t d\tau \int_0^\tau d\tau' \hat{V}^I(\tau) \hat{V}^I(\tau') + \dots$$

which we can expect to converge quickly if the perturbation combined with oscillations are so small that the integration over (potentially) long times does not hinder convergence. If this is not the case, one could resort to self-energy techniques as they are known in quantum field theory. For these time-dependent systems, the Magnus expansion is a related route. It provides an expansion

$$\hat{U}^I(t) = e^{-i \sum_{n=0}^{\infty} \bar{H}_n(t)} \quad (25)$$

thus truncating this series happens in the exponent and maintains unitarity and is compatible with going to long times. Its lowest orders can be understood as follows: We start with the average Hamiltonian

$$\bar{H}_0(t) = \int_0^t d\tau \hat{V}^I(\tau)$$

i.e. the expression that collects the classical part and ignores all commutators. The next order contains one commutator

$$\bar{H}_1(t) = -\frac{i}{2} \int_0^t d\tau_1 d\tau_2 \left[\hat{V}^I(\tau_2), \hat{V}^I(\tau_1) \right]$$

but as it is in the exponent, it collects terms from all orders of the Dyson series (you can convince yourself by expanding the exponential in eq. (25). The next order of the expansion is

$$\bar{H}_2(t) = -\frac{1}{6} \int_0^t d\tau_1 d\tau_2 d\tau_3 \left\{ \left[\hat{V}^I(\tau_3), \left[\hat{V}^I(\tau_2), \hat{V}^I(\tau_1) \right] \right] + \left[\hat{V}^I(\tau_1), \left[\hat{V}^I(\tau_2), \hat{V}^I(\tau_3) \right] \right] \right\}$$

i.e. it contains two nested commutators. We will only be able to appreciate this expansion when we apply it, but we can already see that the different orders will inherit different operator structures from the different commutators and that stacking on more integrals will create ever more demanding resonance conditions, so higher orders likely oscillate out. That notwithstanding, the Magnus expansion is asymptotic in nature: Its formal radius of convergence is zero hence adding higher orders does not always improve the accuracy.

3.6 Real-world limitations

When applying (quantum) optimal control to real-world systems, we have to contend with the fact that all parameters under our control have practical limitations: power, frequency, timing, etc. are all constrained by the capabilities of the equipment through which we apply said control. Moreover, any feedback scheme (such as Ad-HOC), must account for experimental noise, uncertainties in the experimental system (both gaps in system characterization, and "random walk"-like drifts of experimental parameters) and imperfections in both control and readout.

These issues above are complex and have to be dealt with simultaneously in real-world scenarios. There is no known textbook solution to these problems, and they are subject to ongoing research. We shall therefore limit ourselves to a very brief review of some of the approaches currently available:

Constraints on applicable controls: Two approaches can be taken: Either the space of possible controls can be defined such all points in the search space are valid, applicable, controls, or the optimization space is defined more liberally, and we penalize controls which fail to conform.

For the first approach, limiting the control subspace, a partial solution is to choose and fix some parameters, such as control field frequency, ahead of time. This is the solution suggested by the CRAB optimal control algorithm [28]. A more general approach is to use bounded functions, such as cosine or inverse tangent, to transform an unconstrained physical parameter to a constrained one. For example, the search parameter α may be unconstrained and $\mathcal{O}(1)$, and we transform it to a constrained field amplitude via $A := 500\text{MHz} \times \cos(\alpha)$, which is subsequently used in the system Hamiltonian.

Sometimes, the approaches above are insufficient as constraints are complex and include multiple parameters; or perhaps such substitutions are not a good fit to the optimal control problem. In such cases, we can impose a penalty term which will modify the functional for which we seek a minimum. For example, if we wish to impose a low-bandwidth solution on the control field $c(t)$, we may add a penalty term proportional to $\int_0^T |\partial_t c(t)|^2 dt$, which will be significant for highly oscillatory functions and zero for the DC component.

Robust controls: Experiments are often noisy environments, which noise appearing both on control fields and on the underlying system Hamiltonians. To provide a control scheme which provides consistently good performance, one must add the robustness requirement of the optimization requirements. This can be done using "ensemble optimization", where each optimization step averages over multiple manifestations of the dynamics, each with a different noise realization. The specific noise manifestations can be either fixed for the duration of the optimization or varied with each iteration step. The former approach is simpler to implement, but runs the risk of the optimization solving the problem only to the small subset of noises it encountered. The latter approach tends to result in more robust controls, but introduces a noisy goal function, which is harder to optimize reliably. In either case, ensemble optimization tends to be expensive in terms of computational resources. In some cases, it is possible to replace it with the a penalty term which is proportional to the absolute value of the gradient of the standard optimization goal with respect to the noisy variable (i.e. require that the control's performance will be weakly dependent on the noisy parameter). In all cases, robust controls often exhibit the "no free lunch" rule of control theory – robust controls often require more time, more bandwidth, or provide a worst average-case performance than their non-robust counterparts [47].

4 Examples

4.1 Optimal control of a qubit

Let's start with a really elementary analytical example: A single qubit with Hamiltonian $\hat{H}(t) = u(t)\hat{\sigma}_x$ looking at the fastest state transfer possible from $|0\rangle$ to $e^{i\phi}|1\rangle$. We can parameterize the state as $|\psi\rangle(t) = (x_0 + iy_0)|0\rangle + (x_1 + iy_1)|1\rangle$. The Schrödinger equation can be expressed in these real parameters as

$$\dot{x}_0 = uy_1 \quad \dot{y}_0 = -ux_1 \quad \dot{x}_1 = uy_0 \quad \dot{y}_1 = -ux_0$$

which are coupled in two sets of two that do not talk to the other components, already telling us that $\phi = \pm\pi/2$. Keep in mind, however, that u can be time-dependent. Now we clearly see that the speed of evolutions scales with the control amplitude u so our initial question was not even well-posed. We need to at least limit the amplitude of the control field. We make this dimensionless $|u| \leq u_{\max}$. The optimal solution exhausts that amplitude and, indeed, plugging in $u = u_{\max}$ we find

$$\ddot{x}_0 + u_{\max}^2 x_0 = 0$$

the harmonic oscillator equation of motion which leads to the desired solution $x_0 = 0$ after time $t_{\min} = \pi/2u_{\max}$. Solutions of this kind are called “bang” solutions. More generally, in strictly bilinear control problems like this one, the optimal solution jumps between its boundaries (which in the case of multiple controls can be quite intricate), then called “bang-bang”-control.

It is interesting to study the physical significance of this result. A real system in its laboratory frame always has an attached drift

$$\hat{H}_1(t) = \frac{E}{2}\hat{\sigma}_z + u(t)\hat{\sigma}_x$$

Now if $u_{\max} \gg |E|$ we can expect the previous solution to still hold approximately. If this condition is violated, the situation is different: The vectors $(\pm u_{\max}, y, z)^T$ define two non-collinear axes on the Bloch sphere and a given initial state can reach all final states that are on the circle around that axis including that state. In general, we will need up to three “bangs” to reach out goal. The limitation of u_{\max} may

4.2 Exploring the speed limit with high parameter counts

The quantum speed limit (QSL) is defined as the minimal time that is needed to evolve a system from a given state ρ_0 to another state $\rho(t)$ with a specific fidelity $\Phi(\rho_0, \rho(t))$ [48]. This is relevant e.g., for qubit gate implementations, because it limits the minimal gate time (for unrestricted controls). When the control bandwidth is restricted, then the dimension of the set of reachable states $D_{\mathcal{W}}$ and the available bandwidth $\Delta\Omega$ give a lower bound for the evolution time [49]:

$$T \geq \frac{D_{\mathcal{W}}}{\Delta\Omega}$$

This is a continuous version of the Solovay-Kitaev theorem.

The set of reachable states consists of all states that can be written as

$$|\psi(t)\rangle = U(t_0, t)|\psi_0\rangle \quad (26)$$

where $U(t_0, t)$ is the propagation operator of the system. A system is called completely controllable if one can choose the control parameters in such a way that the propagation operator is equal to any specific operator [50].

A method to explore the QSL for a gate is the following [51]: For different given gate times one optimizes the gate and plots the fidelity Φ_{goal} or the error $g(T) = 1 - \Phi_{\text{goal}}$ (see equation (21)) of the optimized gates against the gate times. If a QSL exists, there will be minimal time for which the error is small. For shorter gate times the error is significantly larger. This time is the QSL.

The result depends on the chosen optimization method, concretely we show an example:

In fig. 6 and fig. 7 the error g is plotted against gate duration for two different parameterizations. The system is a CR-gate implementation of a CNOT gate [52].

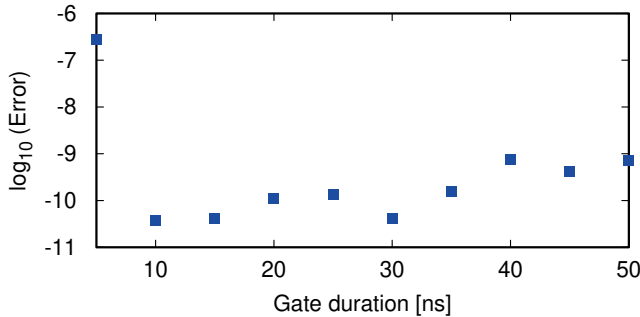


Fig. 6: Gate error as a function of gate time. The optimization was done using GRAPE with a PWC parameterization with 500 pieces. The QSL is around 10ns.

In fig. 6 the QSL is shown for an optimization using GRAPE with a piecewise constant (PWC) parameterization with 500 pieces and unconstrained controls. One can see that there is a jump around 10ns which indicates that this is the QSL in this case. Fig. 7 shows the same, but for an optimization using GOAT with a Fourier decomposition into 167 components. The QSL is here around 40ns and is reached more slowly.

The difference is related to optimally have the controls interact with redirecting the drift. A key step to a theory of this phenomenon has been undertaken in [49].

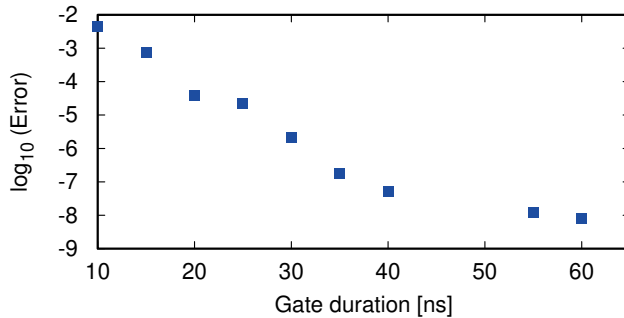


Fig. 7: Gate error as a function of gate time. The optimization was done using GOAT with a Fourier decomposition into 167 pieces. The QSL is around 40ns.

4.3 Open systems

In these notes we have mostly concentrated on the optimal control for closed quantum systems. One can ask related questions for open quantum systems as well. A treatment of this situation would go way beyond the scope of these lecture notes. Here, the space of potentially reachable states / of reachable time evolutions is much larger than in the unitary case. The theory of controllability and reachability is thus more involved, it is for example not at all clear, if the impact of decoherence can be reduced to zero, i.e., if the subset of unitary time evolutions is reachable. We would thus like to describe a pragmatic approach and refer the reader to the literature. For a Lindblad equation, it can be shown that the control fields cannot cancel dissipation effect and the system is not completely controllable. This is still an open question in the non-Markovian regime.

As a first rule of thumb, there are situations when the decoherence experienced by the quantum subsystem has no or very little structure – e.g., in the case of uniform decoherence leading to a fully depolarizing channel and, at least for the synthesis of gates, for most Markovian decoherence models. These do not give an open system optimal control algorithm any space to actually exploit the structure of the decoherence to perform an optimization, rather, we can expect that the fastest solution of the closed system also is close to an optimal solution for the open system. Thus, running a closed-system version of optimal control and benchmarking it in a realistic open system is a good initial approach.

If one suspects that the decoherence mechanism contains exploitable structure, or if one tries to accomplish a task that actively uses decoherence – such as tasks changing the entropy of the state, e.g., cooling, it is possible to generalize the aforementioned methods of optimal control. More specifically, e.g., in OpenGRAPE, one simply replaces the Schrödinger equation as the dynamical constraint by a suitable description of open systems dynamics, such as a master equation. One caveat lies in the need for backwards-in-time propagation: Open system dynamics is asymptotically irreversible, which can make back-propagation unstable. Practically,

this can be handled by either focusing on decoherence rates that are not too large or by suitable initial guesses.

As a well-defined example, let us consider a single qubits perturbed by a two-level fluctuator, i.e., a second two-state system that is coupled to a heat bath. This is a common situation in superconducting qubits [53].

We specifically model a qubit coupled to a single TLF by $H = H_S + H_I + H_B$. H_S consists of the qubit and the coupled two-state system, i.e.

$$H_S = E_1(t)\sigma_z + \Delta\sigma_x + E_2\tau_z + \Lambda\sigma_z\tau_z$$

where σ_i and τ_i are the usual Pauli matrices operating in qubit and fluctuator Hilbert space respectively. $E_1(t)$ is time-dependent and serves as an external control. The source of decoherence is the coupling of the fluctuator to the heat bath, which leads to incoherent transitions between the fluctuator eigenstates, $H_I = \sum_i \lambda_i(\tau^+ b_i + \tau^- b_i^\dagger)$, $H_B = \sum_i \hbar\omega_i b_i^\dagger b_i$. We introduce an Ohmic bath spectrum $J(\omega) = \sum_i \lambda_i^2 \delta(\omega - \omega_i) = \kappa\omega\Theta(\omega - \omega_c)$ containing the couplings λ_i , the dimensionless damping κ , and a high-frequency cutoff ω_c (which we assume to be the largest frequency in the system). Now depending on the bath damping constant κ the fluctuator can flip fast or slow – and in the limit of slow flipping, the qubit sees noise with strong temporal correlation leading to highly non-Markovian qubit dynamics.

To formally treat this system, we can on the other hand still set up a Markovian master equation for the augmented system of qubit *and* fluctuator and only after its solution trace over the fluctuator to get the effective density matrix of the qubit alone. We formulate the control approach by rewriting the master equation as $\dot{\rho}(t) = -(i\mathcal{H}(E_1(t)) + \Gamma(E_1(t)))\rho(t)$ with the Hamiltonian commutator superoperator $\mathcal{H}(E_1(t))(\cdot) = [H(E_1(t)), \cdot]$ and the relaxation superoperator Γ , both time-dependent via the control $E_1(t)$. The formal solution to the master equation is a linear quantum map operating on a physical initial state according to $\rho(t) = F(t)\rho(0)$. Thus F itself follows the operator equation of motion

$$\dot{F} = -(i\mathcal{H} + \Gamma) F \quad (27)$$

with initial condition $F(0) = \mathbb{I}$, as in ref. [54].

Here, multiplication of quantum maps denotes their concatenation. The task is to find control amplitudes $E_1(t)$ with $t \in [0, t_g]$, t_g being a fixed final time, such that the difference $\delta F = F_U - F(t_g)$ between dissipative time evolution $F(t_g)$ obeying eqn. (27) and a target unitary map F_U is minimized with respect to the Euclidean distance $\|\delta F\|_2^2 \equiv \text{tr} \{ \delta F^\dagger \delta F \}$. Clearly, this is the case, when the trace fidelity

$$\phi = \text{Re tr} \left\{ F_U^\dagger F(t_g) \right\} \quad (28)$$

is maximal. Note, that in an open system, one cannot expect to achieve zero distance to a unitary evolution F_U [54]. The goal is to come as close as possible. On this setting, we find optimal pulses by gradient search.

It is interesting to investigate the resulting pulses and performance limits. We see in figure ... that optimal control pulses allow to reach great gate performance after overcoming a quantum speed limit. Remarkably, the dependence on gate duration is non-monotonic at least in the regime of low κ when the two settings of the TLS can be resolved. At some magic times, the frequency split from the TLS naturally refocuses, constraining the optimization much less than at other times.

More remarkable, the maximally attainable fidelity also has a non-monotonic dependence on κ . At hindsight, this can be understood as follows: At low κ there is no randomness of the system, it is fully reversible. The optimal control algorithm just has to deal with the fact that the setting of the TLS is unknown, which it perfectly accomplishes. On the other hand, at high κ , the phenomenon of *motional narrowing* occurs: Fast motion of the impurity broadens its spectrum thus reducing its spectral weight at low frequencies.

4.4 DRAG and its derivatives

In general a quantum system will contain additional states outside of a subspace, we want to operate in. If our control couples also to transitions out of the subspace we will leak population and degrade the performance of our operation. The Derivative Removal with Adiabatic Gate (DRAG) method provides a framework to identify these leakages and to modify the control signals to counteract them.

We will review the basic idea along the procedure shown in [55]. Consider a 3-level-system that is controlled by a signal $u(t) = u_x(t) \cos(\omega_d t) + u_y(t) \sin(\omega_d t)$. The first two levels make up the computational subspace $|0\rangle, |1\rangle$ with transition frequency ω_1 that we want to operate in and $|2\rangle$ accounts for the leakage. It is modeled by the Hamiltonian

$$H/\hbar = \omega_1 |1\rangle \langle 1| + (2\omega_1 + \Delta) |2\rangle \langle 2| + u(t) \hat{\sigma}_{0,1}^x + \lambda u(t) \hat{\sigma}_{1,2}^x \quad (29)$$

where the Pauli operators are $\hat{\sigma}_{j,k}^x = |k\rangle \langle j| + |j\rangle \langle k|$, λ describes the coupling of the drive to the 1-2 transition. We expressed the second transition frequency by the anharmonicity $\Delta = \omega_2 - 2\omega_1$.

Let's say we want to implement a simple Rabi pulse by choosing $u_x(t) = \Omega(t)$ and $u_y(t) = 0$. This gives rise to unwanted leakage out of the computational subspace with the term $\lambda \Omega(t) \hat{\sigma}_{1,2}^x$. The DRAG idea shows how we can counteract this leakage by choosing $u_y(t)$ appropriately.

We first express the Hamiltonian in the rotating frame with $R = \exp(i\omega_d |1\rangle \langle 1| + 2i\omega_d |1\rangle \langle 1|)$ following the rule $H^R = RHR^\dagger + i\hbar \dot{R}R^\dagger$ which gives

$$H^R/\hbar = \delta_1 |1\rangle \langle 1| + \delta_2 |2\rangle \langle 2| + \sum_{\alpha=x,y} \frac{u_\alpha}{2}(t) \hat{\sigma}_{0,1}^\alpha + \lambda \frac{u_\alpha}{2}(t) \hat{\sigma}_{1,2}^\alpha,$$

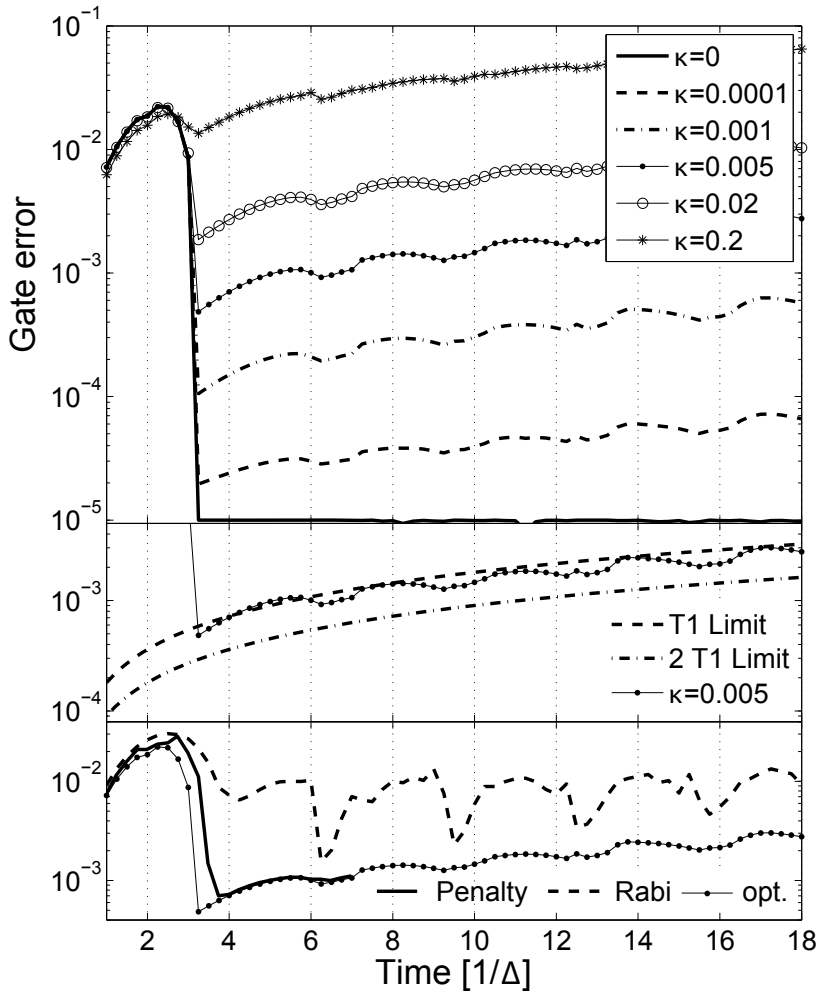


Fig. 8: Top: Gate error versus pulse time t_g for optimal Z-gate pulses in the presence of a non-Markovian environment with dissipation strength κ . A periodic sequence of minima at around $t_n = n\pi/\Delta$, where $n \geq 1$, is obtained. Middle: The gate error of optimized pulses approaches a limit set by T_1 and $2T_1$, as shown with $\kappa = 0.005$. Bottom: Optimized pulses reduce the error rate by approximately one order of magnitude compared to Rabi pulses for $\kappa = 0.005$. Pulses starting from zero bias and with realistic rise times (penalty) require only a small additional gate time. In all figures the system parameters are $E_2 = 0.1\Delta$, $\Lambda = 0.1\Delta$ and $T = 0.2\Delta$.

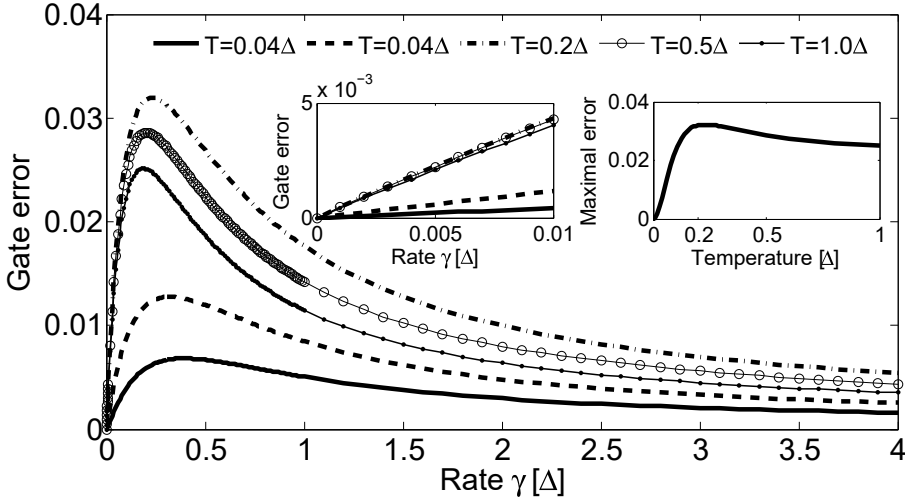


Fig. 9: Gate error versus TLF rate γ for various temperatures for an optimized pulse with $t_g = 5.0/\Delta$. The left inset is a magnification of the low- γ part of the main plot and reveals the linear behaviour. The right inset shows the maximum of the curves of the main plot versus temperature. ($E_2 = 0.1\Delta$ and $\Lambda = 0.1\Delta$)

using the detunings $\delta_1 = \omega_1 - \omega_d$ and $\delta_2 = \Delta + 2\delta_1$ between the drive and transition frequencies.

Applying an adiabatic transformation $V(t)$ by calculating $H^V = VHV^\dagger + i\hbar\dot{V}V^\dagger$ allows us to look at the system in a frame where the leakage and the y -component necessary to counteract it are visible. We take

$$V(t) = \exp \left[-i \frac{u_x(t)}{2\Delta} (\hat{\sigma}_{0,1}^y + \lambda \hat{\sigma}_{1,2}^y) \right],$$

a transformation that depends on our intended signal u_x , and apply it to first order in u_x/Δ to find

$$\begin{aligned} H^V/\hbar = & \left(\delta_1 - \frac{(\lambda^2 - 4)u_x^2}{4\Delta} \right) |1\rangle\langle 1| + \left(\delta_2 + \frac{(\lambda^2 + 2)u_x^2}{4\Delta} \right) |2\rangle\langle 2| \\ & + \frac{u_x}{2} \hat{\sigma}_{0,1}^x + \lambda \frac{u_x^2}{8\Delta} \hat{\sigma}_{0,2}^x + \left[\frac{u_y}{2} + \frac{\dot{u}_x}{2\Delta} \right] (\hat{\sigma}_{0,1}^y + \lambda \hat{\sigma}_{1,2}^y) \end{aligned}$$

From this expression we can see that our intended drive is unchanged $u_x/2\hat{\sigma}_{0,1}^x$ but if we also choose $u_y = -\dot{u}_x/\Delta$ we cancel the last term that is responsible for driving out of the computational subspace $\propto \lambda \hat{\sigma}_{1,2}^y$. The transformation also suggest detuning the drive by $\delta_1 = (\lambda^2 - 4)u_x^2/4\Delta$ to avoid stark shifting of the 0-1 transition. This example illustrates the main working principle of DRAG which can be generalized to account for more than just leakage to a third level. By modifying $V(t)$, for example adding terms $\propto \hat{\sigma}_{0,2}^y$, or iteratively performing transformations $V_j(t)$ the inertial terms, the inertial terms $i\hbar\dot{V}_jV_j^\dagger$ generate more conditions on the control signals and its derivatives.

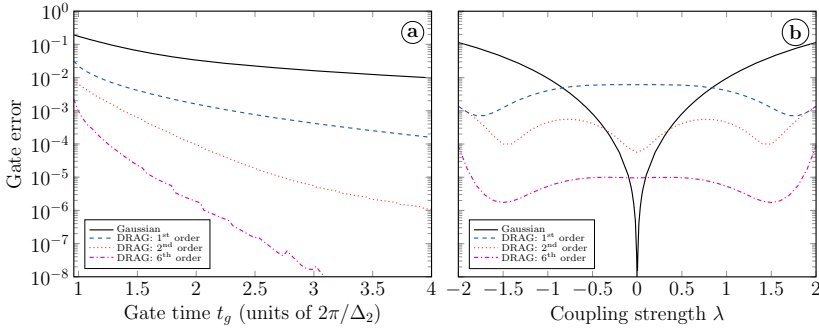


Fig. 10: (a): Performance of non-optimized DRAG variants as a function of gate time, derived from an iterative Schrieffer-Wolff expansion to higher orders. Target : $\hat{\sigma}_x$ rotation of a single qubit described by the lowest three levels of Hamiltonian (29). – (b): Performance of the DRAG pulses used in (a) for a fixed gate time $t_g = 4\pi/\Delta_2$ as a function of coupling strength λ to the leakage level.

The performance of solutions to different orders, obtained via iterative transformations, is depicted in Fig.10a as a function of pulse length, and in Fig.10b as a function of coupling strength λ for a fixed gate time $t_g = 4\pi/\Delta_2$. Higher order solutions are taken from [56]. Note also that when the $|0\rangle \leftrightarrow |2\rangle$ transition is controlled via an additional corresponding frequency component, exact solutions to the three-level system exist (cf. chapter 8 in [57]).

Turning to the experimental implementation [58,59] of DRAG pulses: In practice, actual system parameters differ somewhat from those assumed in theory due to characterization gaps, system drift, or unknown transfer functions affecting the input field shapes [60]. As a simplification, we assume the low order terms in DRAG are easier to implement as their shape will be mostly maintained on entry into the dilution fridge. Even so, many different low-order variants of DRAG have been found in the literature for third-level leakage [55, 56, 58, 61]. This reduced functional form can further be optimized theoretically [62] and/or through a closed-loop process experimentally [25, 26] to account for the effect of higher order terms and experimental uncertainties (preferably using more advanced gradient-free algorithms such as CMA-ES [63]). A systematic experimental study of the tune-up of the prefactors in front of the functional forms for the control operators was performed in [64]. In writing up these optimizations and adapting them, the Magnus expansion, see chapter 3.5 is typically used.

For instance, let us denote the Gaussian pulse implementing a $\hat{\sigma}_x$ gate for the qubit by $G(t)$. Then the first order solutions described in [55,56,61] are parameterized by the limited functional basis $u_x \propto G$, $u_y \propto \partial_t G$ and $\delta \propto G^2$, which mimics the limited shaping control that can exist in experiment. None of the reported solutions are optimal within this functional basis: For typical example parameters, infidelities may be further reduced from $10^{-5.28}$ to $10^{-6.63}$ by slightly adjusting the prefactors of the control fields. For example, [55]’s first order DRAG solution may be transformed according to $u_x \rightarrow (1 + \alpha_x)u_x$ and similarly for u_y and δ , and

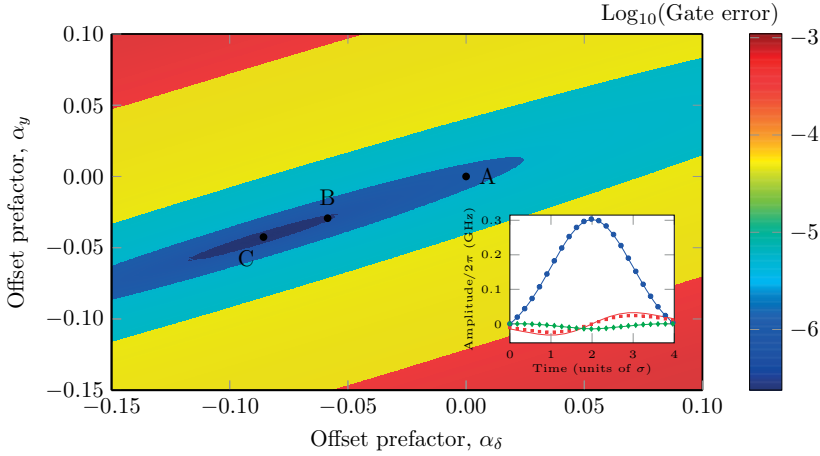


Fig. 11: A slice of the 3D calibration landscape for DRAG solution up to the first order in the small parameter to the qubit σ_x -gate leakage problem. Point A and B denote [55]’s and [61]’s first-order solutions, respectively. Point C is the optimum for this control function subspace (here $\alpha_x = -0.0069$), with infidelity of $10^{-6.63}$. A successful calibration process will typically start at a known DRAG solution, i.e. points A or B, and conclude in point C. The inset illustrates the associated pulse shapes: markers represent the unoptimized shapes (u_x : \bullet , u_y : \blacksquare , δ : \blacklozenge) whereas solid lines depict the corresponding optimal solution (C).

then the constants α_x , α_y and α_δ are optimized. A discussion for why optimization within a severely restricted functional subspace may often be sufficient is given in [65] and follow-up publications. A schematic of the optimization task involved in the calibration, as well as the shape of the associated controls, is shown in Fig.11.

5 Summary and outlook

Optimal control is a mature discipline of theoretical physics and related fields. In experimentation, it has remarkable success in situations in which physical systems are well characterized. Reaching out to engineered systems requires a close integration with characterization and benchmarking.

Experimentalists and users of quantum control should have taken home an introduction of concepts, jargon, and results of the field. Theorists should feel motivated to embrace these challenges and to fashion their results into tools that can be used efficiently and scalably so quantum control and quantum technology applications can mutually benefit from their potential.

Acknowledgements

We acknowledge collaboration with the optimal control group at Saarland University (and its previous locations), including Daniel Egger, Likun Hu, Kevin Pack, Federico Roy, Ioana Serban, and Lukas Theis as well as continuous collaboration with Tommaso Calarco, Steffen Glaser, Christiane Koch, Simone Montangero, and Thomas Schulte-Herbrüggen. Some of this work is sponsored by the Intelligence Advanced Research Projects Activity (IARPA) through the LogiQ Grant No. W911NF-16-1-0114, by the European Union under OpenSuperQ and the ITN Qusco.

References

- [1] S. J. Glaser, U. Boscain, T. Calarco, C. P. Koch, W. Köckenberger, R. Kosloff, I. Kuprov, B. Luy, S. Schirmer, T. Schulte-Herbrüggen, D. Sugny, F. K. Wilhelm, *Training Schrödinger's cat: quantum optimal control*, Eur. Phys. J. D **69** (2015).
- [2] A. Acín, I. Bloch, H. Buhrman, T. Calarco, C. Eichler, J. Eisert, D. Esteve, N. Gisin, S. J. Glaser, F. Jelezko, S. Kuhr, M. Lewenstein, M. F. Riedel, P. O. Schmidt, R. Thew, A. Wallraff, I. Walmsley, F. K. Wilhelm, *The quantum technologies roadmap: a European community view*, New J. Phys. **20**, 080201 (2018).
- [3] A. E. Bryson, Y.-C. Ho, *Applied optimal control, revised printing*, Hemisphere, New York (1975).
- [4] B. Andresen, K. H. Hoffmann, J. Nulton, A. Tsirlin, P. Salamon, *Optimal control of the parametric oscillator*, European journal of physics **32**, 827 (2011).
- [5] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, S. J. Glaser, *Optimal control of coupled spin dynamics: design of NMR pulse sequences by gradient ascent algorithms*, Journal of Magnetic Resonance **172**, 296 (2005).
- [6] B. C. Hall, *An Elementary Introduction to Groups and Representations*, arXiv:math-ph/0005032.
- [7] B. C. Hall, *Lie Groups, Lie Algebras, and Representations, Graduate Texts in Mathematics*, volume 222 (Springer, New York, 2003).
- [8] A. Niskanen, J. Vartiainen, M. Salomaa, *Optimal Multiqubit Operations for Josephson Charge Qubits*, Phys. Rev. Lett. **90**, 197901 (2003).
- [9] D. M. Reich, M. Ndong, C. P. Koch, *Monotonically convergent optimization in quantum control using Krotov's method*, The Journal of Chemical Physics **136**, 104103 (2012).

- [10] R. Kosloff, S. Rice, P. Gaspard, S. Tersigni, D. Tannor, *Wavepacket dancing: Achieving chemical selectivity by shaping light pulses*, Chem. Phys. **139**, 201 (1989).
- [11] V. Krotov, I. Feldman, *Iteration method of solving the problems of optimal control*, Eng. Cybern **21**, 123 (1983).
- [12] A. P. Peirce, Mohammed, A. Dahleh, H. Rabitz, *Optimal control of quantum-mechanical systems: Existence, numerical approximation, and applications*, Phys. Rev. A **37**, 4950 (1988).
- [13] M. H. Goerz, D. Basilewitsch, F. Gago-Encinas, M. G. Krauss, K. P. Horn, D. M. Reich, C. P. Koch, *Krotov: A Python implementation of Krotov's method for quantum optimal control*, SciPost Phys. **7**, 80 (2019).
- [14] S. Machnes, U. Sander, S. J. Glaser, P. de Fouquières, A. Gruslys, S. Schirmer, T. Schulte-Herbrüggen, *Comparing, optimizing, and benchmarking quantum-control algorithms in a unifying programming framework*, Phys. Rev. A **84**, 022305 (2011).
- [15] H. A. Rabitz, M. M. Hsieh, C. M. Rosenthal, *Quantum Optimally Controlled Transition Landscapes*, Science **303**, 1998 (2004).
- [16] P. J. Lieberman, F. K. Wilhelm, *Optimal Qubit Control Using Single-Flux Quantum Pulses*, PHYSICAL REVIEW APPLIED **6** (2016).
- [17] J. J. W. H. Sørensen, M. O. Aramburu, T. Heinzl, J. F. Sherson, *Quantum optimal control in a chopped basis: Applications in control of Bose-Einstein condensates*, Phys. Rev. A **98**, 022119 (2018).
- [18] R. S. Judson, H. Rabitz, *Teaching Lasers to Control Molecules*, Phys. Rev. Lett. **68**, 1500 (1992).
- [19] Z.-C. Yang, A. Rahmani, A. Shabani, H. Neven, C. Chamon, *Optimizing Variational Quantum Algorithms Using Pontryagin's Minimum Principle*, Phys. Rev. X **7**, 021027 (2017).
- [20] Y. R. Sanders, J. J. Wallman, B. C. Sanders, *Bounding quantum gate error rate based on reported average fidelity*, New Journal of Physics **18**, 012002 (2016).
- [21] L. Hu, *Optimal control with an arbitrary performance index*, honors Thesis, University of Waterloo (2008).
- [22] J. Nocedal, *Updating quasi-Newton matrices with limited storage*, Mathematics of Computation **35**, 773 (1980).
- [23] K. Aizu, *Parameter Differentiation of Quantum-Mechanical Linear Operators*, Journal of Mathematical Physics **4**, 762 (1963).

- [24] T. Levante, T. Bremi, R. Ernst, *Pulse-sequence optimization with analytical derivatives. Application to deuterium decoupling in oriented phases*, Journal of Magnetic Resonance, Series A **121**, 167 (1996).
- [25] D. J. Egger, F. K. Wilhelm, *Adaptive hybrid optimal quantum control for imprecisely characterized systems*, Physical review letters **112**, 240503 (2014).
- [26] J. Kelly, R. Barends, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, I.-C. Hoi, E. Jeffrey, A. Megrant, J. Mutus, C. Neill, P. J. J. O'Malley, C. Quintana, P. Roushan, D. Sank, A. Vainsencher, J. Wenner, T. C. White, A. N. Cleland, J. M. Martinis, *Optimal Quantum Control Using Randomized Benchmarking*, Phys. Rev. Lett. **112**, 240504 (2014).
- [27] C. Ferrie, O. Moussa, *Robust and efficient in situ quantum control*, Phys. Rev. A **91**, 052306 (2015).
- [28] P. Doria, T. Calarco, S. Montangero, *Optimal Control Technique for Many-Body Quantum Dynamics*, Phys. Rev. Lett. **106**, 190501 (2011).
- [29] J. M. Binder, A. Stark, N. Tomek, J. Scheuer, F. Frank, K. D. Jahnke, C. Müller, S. Schmitt, M. H. Metsch, T. Unden, et al., *Qudi: A modular python suite for experiment control and data processing*, SoftwareX **6**, 85 (2017).
- [30] F. Angaroni, A. Graudenzi, M. Rossignolo, D. Maspero, T. Calarco, R. Piazza, S. Montangero, M. Antoniotti, *Personalized therapy design for liquid tumors via optimal control theory*, bioRxiv p. 662858 (2019).
- [31] P. Silvi, E. Rico, T. Calarco, S. Montangero, *Lattice gauge tensor networks*, New Journal of Physics **16**, 103015 (2014).
- [32] N. Rach, M. M. Müller, T. Calarco, S. Montangero, *Dressing the chopped-random-basis optimization: A bandwidth-limited access to the trap-free landscape*, Phys. Rev. A **92**, 062343 (2015).
- [33] S. Machnes, E. Assémat, D. Tannor, F. K. Wilhelm, *Tunable, Flexible, and Efficient Optimization of Control Pulses for Practical Qubits*, Phys. Rev. Lett. **120**, 150401 (2018).
- [34] J. P. Palao, R. Kosloff, *Quantum Computing by an Optimal Control Algorithm for Unitary Transformations*, Phys. Rev. Lett. **89**, 188301 (2002).
- [35] C. M. Tesch, L. Kurtz, R. de Vivie-Riedle, *Applying optimal control theory for elements of quantum computation in molecular systems*, Chemical Physics Letters **343**, 633 (2001).
- [36] I. L. C. M. A. Nielsen, *Quantum Computation and Quantum Information* (Cambridge University Press, 2010).

- [37] M. Paris, J. Rehacek, *Quantum State Estimation* (Springer, Berlin, 2004).
- [38] D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, J. Eisert, *Quantum State Tomography via Compressed Sensing*, Phys. Rev. Lett. **105**, 150401 (2010).
- [39] T. Chasseur, D. M. Reich, C. P. Koch, F. K. Wilhelm, *Hybrid benchmarking of arbitrary quantum gates*, PHYSICAL REVIEW A **95** (2017).
- [40] J. Emerson, R. Alicki, K. Życzkowski, *Scalable noise estimation with random unitary operators*, Journal of Optics B: Quantum and Semiclassical Optics **7**, S347 (2005).
- [41] C. Dankert, R. Cleve, J. Emerson, E. Livine, *Exact and approximate unitary 2-designs and their application to fidelity estimation*, Phys. Rev. A **80**, 012304 (2009).
- [42] E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, D. J. Wineland, *Randomized benchmarking of quantum gates*, Phys. Rev. A **77**, 012307 (2008).
- [43] T. Chasseur, F. K. Wilhelm, *Complete randomized benchmarking protocol accounting for leakage errors*, Phys. Rev. A **92**, 042333 (2015).
- [44] E. Magesan, J. M. Gambetta, B. R. Johnson, C. A. Ryan, J. M. Chow, S. T. Merkel, M. P. da Silva, G. A. Keefe, M. B. Rothwell, T. A. Ohki, M. B. Ketchen, M. Steffen, *Efficient Measurement of Quantum Gate Error by Interleaved Randomized Benchmarking*, Phys. Rev. Lett. **109**, 080505 (2012).
- [45] A. Gangopadhyay, M. Dzero, V. Galitski, *Exact solution for quantum dynamics of a periodically driven two-level system*, Phys. Rev. B **82**, 024303 (2010).
- [46] W. S. Warren, *Effects of arbitrary laser or NMR pulse shapes on population inversion and coherence*, J. Chem. Phys. **81**, 5437 (1984).
- [47] B. Khani, S. T. Merkel, F. Motzoi, J. M. Gambetta, F. K. Wilhelm, *High-fidelity quantum gates in the presence of dispersion*, Phys. Rev. A **85**, 022306 (2012).
- [48] V. Giovannetti, S. Lloyd, L. Maccone, *Quantum limits to dynamical evolution*, Phys. Rev. A **67**, 052109 (2003).
- [49] S. Lloyd, S. Montangero, *Information Theoretical Analysis of Quantum Optimal Control*, Phys. Rev. Lett. **113**, 010502 (2014).
- [50] S. G. Schirmer, A. I. Solomon, J. V. Leahy, *Degrees of controllability for quantum systems and application to atomic systems*, Journal of Physics A: Mathematical and General **35**, 4125 (2002).

- [51] J. J. W. Sørensen, M. K. Pedersen, M. Munch, P. Haikka, J. H. Jensen, T. Planke, M. G. Andreassen, M. Gajdacz, K. Mølmer, A. Lieberoth, et al., *Exploring the quantum speed limit with computer games*, Nature **532**, 210 (2016).
- [52] S. Kirchhoff, T. Keßler, P. J. Liebermann, E. Assémat, S. Machnes, F. Motzoi, F. K. Wilhelm, *Optimized cross-resonance gate for coupled transmon systems*, Phys. Rev. A **97**, 042348 (2018).
- [53] P. Rebentrost, I. Serban, T. Schulte-Herbrüggen, F. Wilhelm, *Optimal control of a qubit coupled to a two-level fluctuator*, quant-ph/0612165.
- [54] T. Schulte-Herbrüggen, A. Spörl, N. Khaneja, S. J. Glaser, *Optimal control for generating quantum gates in open dissipative systems*, Journal of Physics B: Atomic, Molecular and Optical Physics **44**, 154013 (2011).
- [55] F. Motzoi, J. M. Gambetta, P. Rebentrost, F. K. Wilhelm, *Simple Pulses for Elimination of Leakage in Weakly Nonlinear Qubits*, Phys. Rev. Lett. **103**, 110501 (2009).
- [56] F. Motzoi, F. K. Wilhelm, *Improving frequency selection of driven pulses using derivative-based transition suppression*, Phys. Rev. A **88**, 062318 (2013).
- [57] F. Motzoi, *Controlling Quantum Information Devices*, Ph.D. thesis, U. Waterloo (2012).
- [58] E. Lucero, J. Kelly, R. C. Bialczak, M. Lenander, M. Mariantoni, M. Neeley, A. D. O'Connell, D. Sank, H. Wang, M. Weides, J. Wenner, T. Yamamoto, A. N. Cleland, J. M. Martinis, *Reduced phase error through optimized control of a superconducting qubit*, Phys. Rev. A **82**, 042339 (2010).
- [59] J. M. Chow, L. DiCarlo, J. M. Gambetta, F. Motzoi, L. Frunzio, S. M. Girvin, R. J. Schoelkopf, *Optimized driving of superconducting artificial atoms for improved single-qubit gates*, Phys. Rev. A **82**, 040305 (2010).
- [60] F. Motzoi, J. M. Gambetta, S. T. Merkel, F. K. Wilhelm, *Optimal control methods for rapidly time-varying Hamiltonians*, Phys. Rev. A **84**, 022307 (2011).
- [61] Gambetta, J. M., Motzoi, F., Merkel, S. T., Wilhelm, F. K., *Analytic control methods for high-fidelity unitary operations in a weakly nonlinear oscillator*, Phys. Rev. A **83**, 012308 (2011).
- [62] L. S. Theis, F. Motzoi, F. K. Wilhelm, *Simultaneous gates in frequency-crowded multilevel systems using fast, robust, analytic control shapes*, Phys. Rev. A **93**, 012324 (2016).
- [63] N. Hansen, S. D. Müller, P. Koumoutsakos, *Reducing the Time Complexity of the Derandomized Evolution Strategy with Covariance Matrix Adaptation (CMA-ES)*, Evol. Comput. **11**, 1 (2003).

-
- [64] Z. Chen, J. Kelly, C. Quintana, R. Barends, B. Campbell, Y. Chen, B. Chiaro, A. Dunsworth, A. Fowler, E. Lucero, E. Jeffrey, A. Megrant, J. Mutus, M. Neeley, C. Neill, P. O'Malley, P. Roushan, D. Sank, A. Vainsencher, J. Wenner, T. White, A. Korotkov, J. M. Martinis, *Measuring and Suppressing Quantum State Leakage in a Superconducting Qubit*, Phys. Rev. Lett. **116**, 020501 (2016).
- [65] T. Caneva, T. Calarco, S. Montangero, *Chopped random-basis quantum optimization*, Phys. Rev. A **84**, 022326 (2011).

B 5 Programming a quantum computer: the difficulty in designing optimal quantum algorithms

Harry Buhrman, Subhasree Patro, Florian Speelman
Research Center for Quantum Software (QuSoft)
Center for Mathematics and Computer Science (CWI)
University of Amsterdam

P.O. Box 94079
1090 GB Amsterdam
The Netherlands

Contents

1	Introduction	3
2	Computability and Complexity Theory	4
2.1	Gödel's Theorem	5
2.2	Time Bounded Complexity Theory	6
2.3	Edit Distance & Orthogonal Vectors	7
2.4	P versus NP	8
2.5	Strong Exponential Time Hypothesis	9
3	Quantum Query Complexity	9
4	Revisiting the Exponential Time Hypothesis (definitions)	11
5	Defining a Quantum Strong Exponential-Time Hypothesis	13

5.1	The basic QSETH	14
5.2	Extending QSETH to general properties	15
5.3	Observations on the set of compression oblivious properties	17
6	QSETH lower bounds for Orthogonal Vectors and Useful Proofs of Work	18
6.1	Quantum time lower bounds based on DEPTH2-QSETH	19
6.2	Quantum Proofs of Useful Work	20
7	Lower bound for edit distance using NC-QSETH	20
7.1	Reduction from ORTHOGONAL VECTORS to EDIT DISTANCE	21
7.2	Observations	22
	References	24

1 Introduction

Quantum computers are heralded as being a game changer in computation and communication tasks. This is illustrated by Google's claim of reaching quantum supremacy [AAB⁺19] and the many national and international programs (e.g. US, China, UK, Canada, Germany, the Netherlands, and the EU-quantum flagship) for quantum technologies as one of the main focal points for research and development. Besides Google many of the other big tech companies are getting involved as well, for example IBM, Microsoft, Intel, Amazon, Alibaba, and many others.

Quantum computers are based on the laws of quantum mechanics. These laws describe Nature at the smallest scale and have some counterintuitive features like *superposition*, *interference*, and *entanglement*. The superposition principle suggests that a particle (or larger system like e.g. a molecule) can be in two or more *different* states at the *same* time. A famous example by Schrödinger is that of a cat that is in a superposition of being both alive and dead. Although the theory is counterintuitive, it has been verified many times in laboratories around the world and to date it is the most accurate description of Nature we have. A system in superposition can also *interfere* with itself. Superposition together with this interference forms the magic of quantum computing. A qubit is the quantum analog of a classical bit and can be in a superposition of 0 and 1.

Quantum information processing gained a lot of momentum after the breakthrough result in 1994 of Peter Shor [Sho97] who demonstrated that factoring numbers into their prime-factors can be efficiently solved on a quantum computer, whereas no classical algorithm is known that solves this problem quickly. The relevance of this result becomes apparent when one realizes that the security of most of modern cryptography is based on the fact that no fast algorithms for the factorization problem are known. Shor's quantum algorithm will thus break all of these cryptographic protocols, provided that we have a quantum computer that operates on a few thousand qubits.

Shor's result gives some indication that quantum computers can solve certain computational problems more efficiently than classical computers. But it is not true that every computation can be sped-up exponentially and one can even show that for a broad class of problems no speed-up whatsoever is possible. Moreover, it is mathematically still possible that only a very limited speed-up is possible at all! This comes from the fact that we don't understand computation very well. It could be that quantum computers can be efficiently simulated on *classical* computers. This would imply an efficient classical version of Shor's algorithm that factors numbers. It could even be that a much broader class of problems is efficiently computable on a classical computer: PSPACE. This is the class of problems that can be solved using only a polynomial amount of space, but potentially run for exponential time. We stress that this seems unlikely but can not be ruled out with our current (limited) knowledge of the power of (classical) computation. The running hypothesis is that quantum computers can solve certain problems more efficiently than classical problems.

Quantum algorithms are hard to construct and developing them requires new ideas and insights that are completely different from classical algorithm design. The main ingredients are always the superposition principle coupled with smart use of interference. Since the first algorithms of Deutsch [Deu85], Deutsch and Jozsa [DJ92], Simon [Sim97], leading up to Shor's algorithm [Sho97], and Grover's algorithm [Gro96], quite a few more quantum algorithms have

been developed. See for example the overview paper of Montanaro [Mon16] and the quantum algorithms zoo page [Jor20]. It is more accurate to describe these algorithms in terms of techniques: For example the quantum Fourier transform, amplitude amplification, quantum walks, quantum linear system of equations solver (HHL), Span Programs, QAOA etc. These are general quantum algorithmic techniques (and some are heuristics) akin to the classical situation where there also exist a handful of general algorithmic techniques, like branch and bound, search, dynamic programming, deep learning etc.

One often hears the complaint that only very few quantum algorithms have been developed. This could equally well be said about classical algorithms! Therefore the number of algorithms is not a meaningful measure. It is all about the techniques and heuristics that are available and whether they are applicable. The main challenge is to apply them to a specific computational problem at hand, often requiring additional and problem-specific tricks and ideas.

What problems can be solved more efficiently on a quantum computer than on a classical computer? This is the main driving question in the field. A follow-up question is which problems admit a quantum advantage if the implementation has to be done on only a few number of noisy qubits? We will mostly focus on the first question, as it is difficult enough to answer!

In order to get a grip on things we will first have to revisit some ideas from classical computation and computational complexity theory.

2 Computability and Complexity Theory

Complexity theory deals precisely with the question raised above. Determine the amount of resources, e.g. time, memory, randomness, bits of communication, in order to perform a certain computational or communication task. In order to make this precise one first needs a precise model of computation. This was done in the 1930's (around the same time that quantum mechanics was developed) independently by Post [Pos43], already developed in 1920 but only first published in 1943, Church [Chu36], and Turing [Tur36]. The notion of computation that is nowadays used is called a Turing machine and precisely formalizes what a memory cell is and what a computation step is. See [AB09] for precise definitions. For now just think of a Turing machine as a program in your favorite programming language and fix that for the rest of these lectures. You should have a precise idea of what a single computation step is and what a single memory cell is. If you are in doubt, look up the precise definitions. Another important feature is that you can identify programs with natural numbers by interpreting the text of your program as number. So we can talk about the program (or Turing machine) that has number i . Note that we can also interpret each natural number as a program – if it is not a number that comes from a valid program then interpret it as the trivial program that does nothing and stops. This means we can talk about the i^{th} Turing machine (or program).

Definition 1. A function $f : \{0, 1\}^* \rightarrow \{0, 1\}^*$ is **computable** if there is a Turing machine (program) M_f such that for every input $x \in \{0, 1\}^*$, program M_f with input x (we will write $M_f(x)$) will run for a finite number of steps and then output a string $y \in \{0, 1\}^*$ if and only if $f(x) = y$. Note that $\{0, 1\}^*$ indicates the set of all finite binary strings.

We will call a set S **computable** if and only if its characteristic sequence $\chi : \{0, 1\}^* \rightarrow \{0, 1\}$ is computable. Here $\chi(x) = 1 \leftrightarrow x \in S$.

Turing showed that there are functions that are *not* computable. He showed that determining whether a program M_i on input x halts after a finite number of steps or goes on forever is not computable, a problem which is captured by the halting set H , defined as follows.

$$H = \{\langle i, x \rangle \mid M_i(x) \text{ stops}\}$$

Turing [Tur36] showed that there is no Turing machine that computes H . He used a diagonalization argument invented originally by Cantor [Can74] who used it to show that the cardinality of \mathbb{R} is strictly larger than that of \mathbb{Q} . The argument goes along the following lines. Suppose, for a contradiction, that H is computable. Then the following variant will also be computable:

$$H' = \{i \mid M_i(i) \text{ stops}\}$$

This implies there is a machine that halts on all inputs: $M_{j_0}(i) = 1$ if $M_i(i)$ stops and 0 otherwise. Then we can construct the following machine M_{j_1} :

$$M_{j_1}(i) = \begin{cases} \text{halts} & \text{if } M_{j_0}(i) = 0 \\ \text{does not halt}^1 & \text{if } M_{j_0}(i) = 1 \end{cases}$$

In other words $M_{j_0}(i)$ halts if and only if $M_i(i)$ does not halt. Now for the contradiction, let's feed M_{j_1} its own code j_1 as input. We now have that $M_{j_1}(j_1)$ halts if and only if $M_{j_1}(j_1)$ does not halt.

2.1 Gödel's Theorem

Using the same ideas Gödel proved his famous incompleteness Theorem [Göd30] that informally states that there are theorems that are true but that have no proof. Using the language of computability introduced here, there is a simple proof of this.

Consider the following mathematical statements: for any i and x , Turing machine $M_i(x)$ stops. It is easy to see that if $M_i(x)$ stops there is a proof of this: just give the sequence of computation steps that $M_i(x)$ makes until it halts. So for all such halting i and x we can prove that $M_i(x)$ stops.

Next consider for i and x the opposite statements: $M_i(x)$ does *not* halt. The statement is true for those i and x for which $M_i(x)$ does not hold. Assume that for all those i and x there *is* a proof that machine $M_i(x)$ does not halt. Now consider the following algorithm A that can decide for any i and x whether $M_i(x)$ stops or not. Test all proofs one by one and examine whether they are a proof of whether $M_i(x)$ halts or not hold. Once such a proof is found stop and output whether $M_i(x)$ halts or not. This would show that the Halting set is decidable, in contradiction with what we proved above. We conclude now that it can not be the case that for all i and x for which $M_i(x)$ does not halt there is a proof. So there must be some i and x such that $M_i(x)$ does not halt but there is no proof for this.

It is interesting to see that the recent breakthrough results [JNV⁺20] of Ji, Natarajan, Vidick, Wright, and Yuen that disprove Connes' embedding conjecture using quantum information also hinges on the undecidability of the Halting set.

¹ This can be achieved by letting $M_{j_1}(i)$ go into an infinite loop if $M_{j_0}(i) = 1$

2.2 Time Bounded Complexity Theory

We are in practice not just interested in whether a computational problem is computable. We also need the computation to finish in a reasonable amount of time. Problems that are in theory computable but for which no *efficient* algorithm exists are in practice uncomputable. In the following we will capture the notion of efficiently computable. We start by defining, for any reasonable function $t : \mathbb{N} \rightarrow \mathbb{N}$, the class of problems that are computable in time t .

Definition 2. For any nice² time bound $t(n)$ let $DTIME(t(n))$ be the class of sets A such that there is a Turing machine M_A , such that:

1. $\forall x \in \{0, 1\}^* : x \in A$ if and only if $M_A(x) = 1$.
2. $\forall n \in \mathbb{N}$: the number of steps $M_A(x)$ makes on inputs x of length n is bounded by $t(n)$.

For any $x \in \{0, 1\}^*$, we will write $|x|$ to denote the length of x , that is if $x = x_1 \dots x_n$ then $|x| = n$.

An important result in this area shows that if you allow the computation to use more time, then more sets and functions become computable. More formally we have the following time hierarchy theorem.

Theorem 1. $DTIME(o(\frac{f(n)}{\log f(n)})) \subsetneq DTIME(f(n))$ for time constructible functions $f(n)$.

The *little-o* notation $o(g(n))$ indicates the class of functions $g'(n)$ such that $\lim_{n \rightarrow \infty} g'(n)/g(n) = 0$. Theorem 1 for example states that there are sets/problems A computable in time $O(n^2)$ which can not be computed in time $O(n)$. The proof of Theorem 1 is a refinement of the diagonalization argument above. See [Ž83] for more details. In essence one constructs a set $D \in DTIME(f(n))$ analogous to the Halting set above:

$$D = \{ \langle i, x \rangle \mid M_i(\langle i, x \rangle) = 0 \text{ and } M_i(\langle i, x \rangle) \text{ makes } \leq \frac{f(|\langle i, x \rangle|)}{\log f(|\langle i, x \rangle|)} \text{ many steps} \}$$

It is not hard to see that $D \in DTIME(f(x))$. The $\log f(|x|)$ factor is due to the overhead in time in simulating the i^{th} Turing machine. On the other hand one can show that $D \notin DTIME(o(\frac{f(n)}{\log f(n)}))$. For suppose it is, then there is a Turing machine running in time $o(\frac{f(n)}{\log f(n)})$, lets call it M_{j_0} . As before one can now argue that $M_{j_0}(\langle j_0, x \rangle) = 1 \iff M_{j_0}(\langle j_0, x \rangle) = 0$, a contradiction.

Next we define the class of problems that can be solved in polynomial time.

Definition 3. P is the class of problems $A \subset \{0, 1\}^*$ such that there is a c such that $A \in DTIME(n^c)$.

Strictly speaking P deals with sets but we can also define an analogous class for functions $f : \{0, 1\}^* \rightarrow \{0, 1\}^*$ that can be computed by Turing machines whose running time is bounded by a polynomial in the length of the input.

² Formally the function has to be time constructible. Throughout the text we will use very simple functions like polynomials and exponential functions.

The class P is also called the class of feasible problems and indeed many of the problems that are in practice solved on a computer, like for example the shortest path problem used in routing software in GPS systems, or computing the edit distance when comparing DNA sequences, fall into this class P .

Complexity theory tries to determine the minimum time bound for which a problem A can be solved.

Definition 4. *The complexity of a problem A is the time bound $t(n)$ such that $A \in DTIME(t(n))$ but $A \notin DTIME(o(t(n)))$.*

In order to determine that the complexity of A is $t(n)$ one needs to establish the following two parts.

1. Upper bound. $A \in DTIME(t(n))$. This is done by designing an algorithm that establishes this.
2. Lower bound. $A \notin DTIME(o(t(n)))$. This is often more difficult and requires a proof that no algorithm exists that runs in time bound $o(t(n))$.

Computer scientist have been quite resourceful with item 1 above and many algorithmic techniques have been employed to design algorithms in order to get upper bounds. However, the area of lower bounds is seriously lagging behind. The best lower bound for an *explicitly defined* function [IM02] is $5n - o(n)$. We note that this result is in the stronger computational model of Boolean circuits, but omit this technicality in these notes.

Open Problem 1. *Establish a superlinear lower bound for an explicitly defined function.*

2.3 Edit Distance & Orthogonal Vectors

We now define properly the EDIT DISTANCE and ORTHOGONAL VECTORS problems which we will study during the lectures in more detail. See also Section 7 for more details.

Definition 5 (The EDIT DISTANCE problem). *Given two strings a and b over an alphabet set Σ , the edit distance between a and b is the minimum number of operations (insertions, deletions, substitutions) on the symbols required to transform string a to b (or vice versa).*

For example if $a = \text{BEAGLE}$ and $b = \text{BADGE}$ then the edit distance between a and b is 3. It is possible to change a into b if in a we replace the first E with an A, replace the A with a D, and delete the L, BEAGLE becomes BADGE. It is not hard to see that 3 is the minimum number of operations needed.

Interestingly, there is a beautiful algorithm [WF74] that solves the EDIT DISTANCE problem in time $O(n^2)$. It makes use of a programming technique called dynamic programming [CLRS01]. The algorithm exploits recurrence relations that define the EDIT DISTANCE problem and computes these in a clever way. Essentially no faster algorithm is known and it is conjectured that this is not possible.

Another computational problem that will figure prominently is the ORTHOGONAL VECTORS problem.

Definition 6 (ORTHOGONAL VECTORS problem). *Given two sets U and V of n vectors, each over $\{0, 1\}^d$ where $d = \omega(\log n)$, determine whether there exists a $u \in U$ and a $v \in V$ such that $\langle u|v \rangle = 0$ (i.e. $\sum_{l \in [d]} u_l v_l = 0$).*

What is the computational complexity of the ORTHOGONAL VECTORS problem? The trivial, and best known, algorithm runs in time $O(n^2 d)$ by testing all pairs of vectors. It is conjectured that this is essentially optimal, i.e., there is no $O(n^{2-\epsilon})$ algorithm for $d = \omega(\log n)$.

2.4 P versus NP

Another problem that has a prominent place in complexity theory is the satisfiability problem:

$$\text{SAT} = \{\phi(x_1, \dots, x_n) \mid \exists \alpha_1, \dots, \alpha_n \in \{T, F\} : \phi(x_1 = \alpha_1, \dots, x_n = \alpha_n) = T\}$$

The formula ϕ on n variables x_1, \dots, x_n is usually in conjunctive normal form (CNF) which means that it is written as:

$$\bigwedge_{i=1}^m \bigvee_{j \in C_i} x_j$$

Where $C_i \subset \{1, \dots, 2n\}$, which is called a *clause*. We will use the convention that x_{n+i} is the negation of variable x_i . When the cardinality of all the clauses is at most k we will call the problem k -SAT.

What is the computational complexity of SAT? It is not hard to see that one can try all possible assignments of T and F to the n variables obtaining a $O(2^n)$ algorithm.³ On the other hand no better lower bound than n is known. Whether there is a faster algorithm for SAT or whether there is a better lower bound is one of the biggest open problems in theoretical computer science and mathematics. Is known by the P versus NP problem. Where NP stands for the class of problems that solutions that can be efficiently checked (i.e. in polynomial time). More formally:

Definition 7. *A is in NP if there is a constant c and a set $B \in \text{P}$:*

$$x \in A \iff \exists y \in \{0, 1\}^{|x|^c} \langle x, y \rangle \in B$$

We call y the witness that $x \in A$. For example for $\phi \in \text{SAT}$, a witness will be a satisfying assignment α such that $\phi(x = \alpha) = T$.

The question whether $\text{P} \stackrel{?}{=} \text{NP}$ can be shown equivalent [Lev73, Co071] to whether $\text{SAT} \stackrel{?}{\in} \text{P}$. It is widely believed that P is not equal to NP and hence that there does *not* exist a polynomial time algorithm for SAT. But as mentioned above showing even a superlinear lower bound is elusive.

³ This is not quite precise since evaluating the formula on a specific assignment α costs time linear in the number of clauses and variables.

2.5 Strong Exponential Time Hypothesis

Conjectures even stronger than $P \neq NP$ have been put forward [IP01]. The Strong Exponential Time Hypothesis (SETH) informally states that essentially the best algorithm for SAT is the one that tests all possible 2^n assignments and declares the formula satisfiable if and only if it found a satisfying assignment. This implies that no algorithm that runs significantly faster than 2^n exists. We will explore this in Section 4, see Conjectures 1 and 2 for precise statements.

There is a surprising consequence of SETH concerning the complexity of the EDIT DISTANCE and ORTHOGONAL VECTORS problems. See [Vas18] for more of such consequences.

Theorem 2. [Wil05] *SETH implies for any $\epsilon > 0$ there is no $O(n^{2-\epsilon})$ time algorithm for ORTHOGONAL VECTORS.*

Proof. Let $\phi(x_1, \dots, x_n)$ be a k -CNF formula with $m = O(n)$ clauses⁴ C_1, \dots, C_m . We will generate two sets of Boolean vectors U and V of dimension d . The size of $|U| = |V| = 2^{n/2}$. We split the variables of ϕ in two groups of equal size. Let α be an assignment to the first group and β to the second group. Let $(u_\alpha)_j$ denote the j^{th} entry of vector u_α . We set $(u_\alpha)_j = 0 \iff \alpha$ satisfies C_j ($1 \leq j \leq m$). Likewise we define $(v_\beta)_j = 0 \iff \beta$ satisfies C_j . We define U the set of all u_α and V the set of all v_β . It is not hard to see that assignment α, β satisfies ϕ iff for every j , α or β satisfies C_j . This is the case iff for every j $(u_\alpha)_j = 0$ or $(v_\beta)_j = 0$, hence α, β satisfies ϕ if and only if $\langle u_\alpha | v_\beta \rangle = 0$.

Note that for ϕ with n variables and $O(n)$ clauses, the size of the OV-problem is $cn2^{n/2} = N$ for some constant c . Now suppose we have an algorithm for ORTHOGONAL VECTORS problem that runs in time $N^{2-\epsilon}$ for some $\epsilon > 0$. This means we get an algorithm for SAT that runs in time $O(2^{n/2} + (n2^{n/2})^{2-\epsilon}) < O(2^{n-\epsilon'n})$ for some $\epsilon' > 0$ which contradicts SETH. \square

A much more intricate argument (see also Section 4 and Section 7.1) shows that similar results can be obtained for the EDIT DISTANCE problem.

This beautiful connection shows that SETH has great explanatory power in order to establish (conditional) optimal algorithms for computational problems in P . This field is called *Fine-Grained Complexity*. In what follows we will try to set up such a Fine-Grained Complexity theory for quantum algorithms. This was done first in [BPS19, ACL⁺19]. We will now turn our attention to quantum algorithms and quantum query complexity.

3 Quantum Query Complexity

We will use the standard notation introduced in for example the textbook [NC00]. Many of the quantum algorithms developed, including the best known algorithms of Shor [Sho97] and Grover [Gro96], can be cast in the black-box or query formalism [BBC⁺01].

For an excellent treatment on the subject, quantum algorithms, and also some techniques, see the lecture notes of de Wolf [dW20].

⁴ We may assume that $m = O(n)$ due to the sparsification Lemma [IPZ01], that we don't discuss here.

Quantum algorithms are expressed in the gate-model. A universal set of gates is a fixed set of 1 or 2 qubit gates that can be used to approximate any unitary on n qubits. Such a universal set could be the Hadamard and Toffoli (CCNOT) gates or CNOT, Hadamard and the phase-gate $= R_\pi/4$.

A quantum algorithm is a unitary operation U that is implemented with only gates from a universal set of gates. The quantum algorithm starts with the register $|x\rangle|0^m\rangle$, where x is the input and the second register has m auxiliary qubits. Then the gates are applied to this initial register resulting in a final state

$$\sum_{y \in \{0,1\}^{n+m}} \alpha_y |y\rangle \text{ and } \sum_y |\alpha_y|^2 = 1$$

Next we measure the final state in the computational basis and obtain the output of the quantum algorithm. Note this could also be a partial measurement when we are interested in fewer than n bits, or just a single bit. We will observe y with probability $|\alpha_y|^2$. Usually we require this outcome to be correct with probability $> 2/3$.

The time that the quantum algorithm takes is the number of gates that are used in the quantum circuit. Questions we would like to answer are: what is the quantum time needed for solving problems like EDIT DISTANCE and ORTHOGONAL VECTORS? It should not come as a surprise that even less is known for the quantum case than for the classical. We do have a result, analogous (but weaker) to the classical one, that gives a linear $((1 + \epsilon)n)$ lower bound for an explicit problem [BSV19].

In order to obtain matching upper and lower bounds for certain problems we look at a weaker model of computation, called the black-box or query model. In this model we want to compute a function $f : \{0, 1\}^N \rightarrow \{0, 1\}$ on a given input $x = x_0 \dots x_{N-1} \in \{0, 1\}^N$. We have access to x in a random access manner. That is we can ask/query the i^{th} bit of x and obtain the value of x_i . This costs a query. The goal is to compute f using as few queries as possible. Note that we do not charge for any computation that is needed in order to compute which queries we want to query.

Formally we define a reversible quantum query as follows:

$$O_x : |i, b\rangle \mapsto |i, b \oplus x_i\rangle$$

where $b \in \{0, 1\}$. In particular $|i, 0\rangle \mapsto |i, x_i\rangle$. Since we are in the quantum setting an algorithm is allowed to ask queries in superposition. A T-query algorithm starts in a fixed state, say the all zeros state $|0, \dots, 0\rangle$, and then interleaves fixed unitary transformations U_0, U_1, \dots, U_T with queries resulting in the final state that can be written as follows:

$$U_T O_x U_{T-1} O_x \dots O_x U_1 O_x U_0 |0, \dots, 0\rangle$$

After this the state is measured in the computational basis. If we are interested in a single bit only a single qubit is measured which should result in the value of $f(x)$. For example Grover's algorithm can be seen to compute the following OR-function: $f_{OR}(x) = 1 \iff \exists i : x_i = 1$. Grover's algorithm establishes a sequence of $O(\sqrt{N})$ unitaries⁵ such that after measuring the final state we find an i such that $x_i = 1$ (if one exists) with probability $> 2/3$.

⁵ In fact in the case of Grover's algorithm, U_0 is the Hadamard transform on all the qubits, and the remaining U 's are all the same.

Definition 8. For a function f the quantum query complexity of f , $Q_\epsilon(f)$ is defined as the minimum number of queries such that f can be computed with probability $1/2 + \epsilon$.

Most of the known quantum algorithms can be cast in this framework. For many functions f , tight upper and lower bounds are known. For example for the OR-function: $Q_\epsilon(f_{\text{OR}}) = \theta(\sqrt{N})$. In order to obtain the lower bounds several techniques have been developed. Notably the polynomial method [BBC⁺01] and the adversary method [Amb00]. In the next part of these lecture notes we will see how we can connect these query lower bounds to time lower bounds via a quantum version of the SETH.

4 Revisiting the Exponential Time Hypothesis (definitions)

The material in this section is taken from the paper by Buhrman, Speelman, and Patro [BPS19]. There is a rich diversity of problems that can be solved in polynomial time, some that have surprisingly fast algorithms, such as the computation of Fourier transforms or solving linear programs, and some for which the worst-case run time has not improved much for many decades. Of the latter category EDIT DISTANCE is a good example: this is a problem with high practical relevance, and an $O(n^2)$ algorithm using dynamic programming, which is simple enough to be taught in an undergraduate algorithms course, has been known for many decades. Even after considerable effort, no algorithm has been found that can solve this problem in fewer than $O(n^2/\log^2 n)$ steps [MP80], still a nearly quadratic run time.

Traditionally, the field of complexity theory has studied the time complexity of problems in a relatively coarse manner – the class P, the problems solvable in polynomial time, is one of the central objects of study in complexity theory.

Consider CNF-SAT, the problem of whether a formula, input in conjunctive normal form, has a satisfying assignment. What can complexity theory tell us about how hard it is to solve this problem? For CNF-SAT, the notion of NP-completeness gives a convincing reason why it is hard to find a polynomial-time algorithm for this problem: if such an algorithm is found, all problems in the complexity class NP are also solvable in polynomial time, showing $P = NP$.

Not only is no polynomial-time algorithm known, but (if the clause-length is arbitrarily large) no significant speed-up over the brute-force method of trying all 2^n assignments is known. Impagliazzo, Paturi, and Zane [IP01, IPZ01] studied two ways in which this can be conjectured to be optimal. The first of which is called the *Exponential-Time Hypothesis* (ETH).

Conjecture 1 (Exponential-Time Hypothesis). *There exists $\alpha > 0$ such that CNF-SAT on n variables and m clauses can not be solved in time $O(m2^{\alpha n})$ by a (classical) Turing machine.*

This conjecture can be directly used to give lower bounds for many natural NP-complete problems, showing that if ETH holds then these problems also require exponential time to solve. The second conjecture, most importantly for the current work, is the *Strong Exponential-Time Hypothesis* (SETH).

Conjecture 2 (Strong Exponential-Time Hypothesis). *There does not exist $\delta > 0$ such that CNF-SAT on n variables and m clauses can be solved in $O(m2^{n(1-\delta)})$ time by a (classical) Turing machine.*

The strong exponential-time hypothesis also directly implies many interesting exponential lower bounds within NP, giving structure to problems within the complexity class. A wide range of problems (even outside of just NP-complete problems) can be shown to require strong exponential time assuming SETH: for instance, recent work shows that, conditioned on SETH, classical computers require exponential time for so-called *strong simulation* of several models of quantum computation [HNS18, MT19].

Surprisingly, the SETH conjecture is not only a very productive tool for studying the hardness of problems that likely require exponential time, but can also be used to study the difficulty of solving problems within P, forming a foundation for the field of *fine-grained complexity*. The first of such a SETH-based lower bound was given in [Wil05], via a reduction from CNF-SAT to the ORTHOGONAL VECTORS problem, showing that a truly subquadratic algorithm that can find a pair of orthogonal vectors among two lists would render SETH false.

The ORTHOGONAL VECTORS problem became one of the central starting points for proving SETH-based lower bounds, and conditional lower bounds for problems such as computing the Frechet distance between two curves [Bri14], sequence comparison problems such as the string alignment problem [AVW14], Longest Common Subsequence and Dynamic Time Warping [ABV15], can all be obtained via a reduction from ORTHOGONAL VECTORS. Also the EDIT DISTANCE problem [BI15] can be shown to require quadratic time conditional on SETH, implying that any super-logarithmic improvements over the classic simple dynamic programming algorithm would also imply better algorithms for satisfiability – a barrier which helps explain why it has been hard to find any new algorithms for this problem.

Besides CNF-SAT, the conjectured hardness of other key problems like 3SUM and APSP is also commonly used to prove conditional lower bounds for problems in P. See the recent surveys [Vas15, Vas18] for an overview of the many time lower bounds that can be obtained when assuming only the hardness of these key problems.

All these results give evidence for the hardness of problems relative to classical computation, but interestingly SETH does not hold relative to *quantum* computation. Using Grover’s algorithm [Gro96, BV97], quantum computers are able to solve CNF-SAT (and more general circuit satisfiability problems) in time $2^{n/2}$, a quadratic speedup relative to the limit that SETH conjectures for classical computation.

Even though this is in violation of the SETH bound, it is not in contradiction to the concept behind the strong exponential-time hypothesis: the input formula is still being treated as a black box, and the quantum speedup comes ‘merely’ from the general quadratic improvement in unstructured search⁶.

It could therefore be natural to formulate the quantum exponential time hypothesis as identical to its classical equivalent, but with an included quadratic speedup, as a ‘basic QSETH’. For some problems, such as ORTHOGONAL VECTORS, this conjecture would already give tight results, since these problems are themselves amenable to a speedup using Grover’s algorithm. See for instance the Master’s thesis [Ren19] for an overview of some of the SETH-based lower bounds that are violated in the quantum setting.

⁶ For unstructured search this bound is tight [BBBV97, BBHT98]. Bennett, Bernstein, Brassard, and Vazirani additionally show that with probability 1 relative to a random oracle all of NP cannot be solved by a bounded-error quantum algorithm in time $o(2^{n/2})$.

On the other hand, since the conditional lower bound for all problems are a quadratic factor lower than before, such a ‘basic QSETH’ lower bound for EDIT DISTANCE would be merely linear. Still, the best currently-known quantum algorithm that computes edit distance takes quadratic time, so we would lose some of the explanatory usefulness of SETH in this translation to the quantum case.

In this work, we present a way around this limit. Realize that while finding a single marked element is quadratically faster for a quantum algorithm, there is no quantum speedup for many other similar problems. For instance, computing whether the number of marked elements is odd or even can not be done faster when allowing quantum queries to the input, relative to allowing only classical queries [BBC⁺01, FGGS98].

Taking the edit distance again as an illustrative example, after careful inspection of the reductions from CNF-SAT to EDIT DISTANCE [BK15, BI15, AHVWW16], we show that the result of such a reduction encodes more than merely the existence of an a satisfying assignment. Instead, the result of these reductions also encodes whether *many* satisfying assignments exist (in a certain pattern), a problem that could be harder for quantum computers than unstructured search. The ‘basic QSETH’ is not able to account for this distinction, and therefore does not directly help with explaining why a linear-time quantum algorithm for EDIT DISTANCE has not been found.

We present a framework of conjectures, that form an analogue of the strong exponential-time hypothesis: QSETH. In this framework, we account for the complexity of computing various properties on the set of satisfying assignments, giving conjectured quantum time lower bounds for variants of the satisfiability problem that range from $2^{n/2}$ up to 2^n .

5 Defining a Quantum Strong Exponential-Time Hypothesis

Almost all known lower bounds for quantum algorithms are defined in terms of *query* complexity, which measures the number of times any quantum algorithm must access the input to solve an instance of a given problem. There are two main methods in the field for proving lower bounds on quantum query complexity: The first one is the *polynomial method*, based on the observation that the (approximate) degree of the unique polynomial representing a function is a lower bound on the number of queries any bounded-error quantum algorithm needs to make [BBC⁺01]. The second main method is the *adversary method* [Amb00] which analyzes a hypothetical quantum adversary that runs the algorithm with a superposition of inputs instead of considering a classical adversary that runs the algorithm with one input and then modifies the input.

Despite the success of quantum query complexity and the fact that we know tight query lower bounds for many problems, the model does not take into account the computational efforts required after querying the input. In particular, it is not possible to use query complexity to prove any lower bound greater than linear, since any problem is solvable in the query-complexity model after all bits are queried. In general we expect the time needed to solve most problems to be much larger than the number of queries required for the computation, but it still seems rather difficult to formalize methods to provide unconditional quantum time lower bounds for explicit problems. We overcome these difficulties by providing a framework of conjectures that

can assist in obtaining *conditional* quantum time lower bounds for many problems in BQP. We refer to this framework as the QSETH framework.

Variants of the classical SETH. The Strong Exponential-Time Hypothesis (SETH) was first studied [IP01, IPZ01], who showed that the lack of a $O(2^{n(1-\delta)})$ for a $\delta > 0$ algorithm to solve CNF-SAT is deeply connected to other open problems in complexity theory. Despite it being one of the most extensively studied problems in the field of (classical) complexity theory, the best known classical algorithms for solving k -SAT run in $2^{n-n/O(k)}m^{O(1)}$ time [PPSZ05], while the best algorithm for the more-general CNF-SAT is $2^{n-n/O(\log \Delta)}m^{O(1)}$ [CIP06], where m denotes the number of clauses and $\Delta = m/n$ denotes the clause to variable ratio.

Even though no refutation of SETH has been found yet, it is plausible that the CNF structure of the input formulas does allow for a speed-up. Therefore, if possible, it is preferable to base lower bounds on the hardness of more general kinds of (satisfiability) problems, where the input consists of wider classes of circuits. For example, lower bounds based on NC-SETH, satisfiability with NC-circuits as input,⁷ have been proven for EDIT DISTANCE, LONGEST COMMON SUBSEQUENCE and other problems [AHVWW16], in particular all the problems that fit the framework presented in [BK15].

Additionally, a different direction in which the exponential-time hypothesis can be weakened, and thereby made more plausible, is requiring the computation of different properties of a formula than whether at least one satisfying assignment exists. For example, hardness of *counting* the number of satisfying assignments is captured by #ETH [DHM⁺14]. Computing existence is equivalent to computing the OR of the set of satisfying assignments, but it could also conceivably be harder to output, e.g., whether the number of satisfying assignments is odd or even or whether the number of satisfying assignments is larger than some threshold. In the quantum case, generalizing the properties to be computed is not only a way to make the hypothesis more plausible: for many of such tasks it is likely that the quadratic quantum speedup, as given by Grover's algorithm, no longer exist.

5.1 The basic QSETH

To build towards our framework, first consider what would be a natural generalization of the classical SETH.

Conjecture (Basic QSETH). *No bounded error quantum algorithm solves CNF-SAT on n variables and m clauses in $O(2^{\frac{n}{2}(1-\delta)}m^{O(1)})$ time, for any $\delta > 0$.*

This conjecture is already a possible useful tool in proving conditional quantum lower bounds, as we present an example of this in Section 6.1.⁸

We first extend this conjecture with the option to consider wider classes of circuits. Let γ denote a class of representations of computational models. Such a representation can for example

⁷ NC circuits are of polynomial size and polylogarithmic depth consisting of fan-in 2 gates.

⁸ Additional examples of implications from such a version of QSETH can be found in the recent independent work of [ACL⁺19].

be polynomial-size CNF formulas, polylog-depth circuits NC, polynomial-size branching programs BP, or the set of all polynomial-size circuits. The complexity of the latter problem is also often studied in the classical case, capturing the hardness of CircuitSAT.

Conjecture (Basic γ -QSETH). *A quantum algorithm cannot, given an input C from the set γ , decide in time $O(2^{\frac{n}{2}(1-\delta)})$ whether there exists an input $x \in \{0, 1\}^n$ such that $C(x) = 1$ for any $\delta > 0$.*

We also define DEPTH2 for the set of all depth-2 circuits consisting of unbounded fan-in, consisting only of AND and OR gates. This definition is later convenient when considering wider classes of properties, and it can be easily seen that ‘basic DEPTH2-QSETH’ is precisely the ‘basic QSETH’ as defined above.

Since both these basic QSETH variants already contain a quadratic speedup relative to the classical SETH, conditional quantum lower bounds obtained via these assumptions will usually also be quadratically worse than any corresponding classical lower bounds for the same problems. For some problems, lower bounds obtained using the basic QSETH, or using γ -QSETH for a wider class of computation, will be tight. However, for other problems no quadratic quantum speedup is known.

5.2 Extending QSETH to general properties

We now extend the ‘basic γ -QSETH’ as defined in the previous section, to also include computing different properties of the set of satisfying assignments. By extending QSETH in this way, we can potentially circumvent the quadratic gap between quantum and classical lower bounds for some problems.

Consider a problem in which one is given some circuit representation of a boolean function $f : \{0, 1\}^n \rightarrow \{0, 1\}$ and asked whether a property $P : \{0, 1\}^{2^n} \rightarrow \{0, 1\}$ on the truth table of this function evaluates to 1, that is, given a circuit C the problem is to decide if $P(\text{tt}(C)) = 1$, where $\text{tt}(C)$ denotes the truth table of the boolean function computed by the circuit C . If one can only access C as a black box then it is clear that the amount of time taken to compute $P(\text{tt}(C))$ is lower bounded by the number of queries made to the string $\text{tt}(C)$. However, if provided with the description of C , which we denote by $\text{desc}(C)$, then one can analyze C to compute $P(\text{tt}(C))$ possibly much faster.

For example, take the representation to be polynomial-sized CNF formulas and the property to be OR. Then for polynomial-sized CNF formulas this is precisely the CNF-SAT problem. Conjecturing quantum hardness of this property would make us retrieve the ‘basic QSETH’ of the previous section. Do note that we cannot simply conjecture that any property is hard to compute on CNF formulas: Even though the query complexity of AND on a string of length 2^n is $\Omega(2^n)$ classically and $\Omega(2^{n/2})$ in the quantum case, this property can be easily computed in polynomial time both classically and quantumly when provided with the description of the $n^{O(1)}$ sized CNF formula.

To get around this problem, we can increase the complexity of the input representation: If we consider inputs from DEPTH2, the set of all depth-2 circuits consisting of unbounded fan-in AND and OR gates, we now have a class that is closed under complementation. For this class,

it is a reasonable conjecture that both AND, the question whether the input is a tautology and all assignments are satisfying, and OR, the normal SAT problem, are hard to compute.

After this step we can look at further properties than AND and OR. For instance, consider the problem of computing whether there exists an even or an odd number of satisfying assignments. This task is equivalent to computing the PARITY of the truth table of the input formula. How much time do we expect a quantum algorithm to need for such a task?

The quadratic speedup for computing the OR is already captured in the model where the quantum computation only tries possible assignments and then performs Grover's algorithm in a black box way. If PARITY is also computed in such a way, then we know from query complexity [BBC⁺01] that there is no speedup, and the algorithm will have to use $\Omega(2^n)$ steps. Our QSETH framework will be able to consider more-complicated properties, like PARITY.

Finally, observe that such a correspondence, i.e., between the query complexity of a property and the time complexity of computing this property on the set of satisfying assignments, cannot hold for *all* properties, even when we consider more complicated input classes besides CNF formulas. For instance, consider a property which is 0 on exactly the strings that are truth tables of polynomial-sized circuits, and is PARITY of its input on the other strings. Such a property has high quantum query complexity, but is trivial to compute when given a polynomial-sized circuit as input. We introduce the notion of *compression oblivious* below to handle this problem.

Defining QSETH. We formalize the above intuitions in the following way. Let the variable γ denote a class of representation at least as complex as the set DEPTH2, where DEPTH2 denotes the set of poly sized depth-2 circuits consisting of only OR and AND gates of unbounded fan-in. We define a meta-language L_P such that $L_P = \{\text{desc}(C) \mid C \text{ is an element from the set } \gamma \text{ and } P(\text{tt}(C)) = 1\}$. We now define the following terms:

Definition 9 (White-box algorithms). *An algorithm A decides the property P in **white-box** if A decides the corresponding meta-language L_P . That is, given an input string $\text{desc}(C)$, A accepts if and only if $P(\text{tt}(C)) = 1$. We use $\text{qTimeWB}_\epsilon(P)$ to denote the time taken by a quantum computer to decide the language L_P with error probability ϵ .*

Definition 10 (Black-box algorithms). *An algorithm A decides the property P in **black-box** if the algorithm $A^f(1^n, 1^m)$ accepts if and only if $P(\text{tt}(f)) = 1$. Here, f is the boolean function computed by the circuit C and m is the upper bound on $|\text{desc}(C)|$ which is the size of the representation⁹ that describes f , and A^f denotes that the algorithm A has oracle access to the boolean function f . We use $\text{qTimeBB}_\epsilon(P)$ to denote the time taken by a quantum computer to compute the property P in the black-box setting with error probability ϵ .*

We define the set of *compression oblivious* properties corresponding to γ as the set of properties where the time taken to compute this property in the black-box setting is lower bounded by the quantum query complexity of this property on all strings. Formally,

$$\mathcal{CO}(\gamma) = \{\text{properties } P \text{ such that } \text{qTimeBB}_\epsilon(P|_{S_\gamma}) \geq \Omega(Q_\epsilon(P))\},$$

where $Q_\epsilon(P)$ denotes the quantum query complexity of the property P in a ϵ -bounded error query model and $S_\gamma = \{\text{tt}(C) \mid C \text{ is an element of the set } \gamma\}$. For example, the properties AND and OR are in $\mathcal{CO}(\text{DEPTH2})$ because the adversarial set that gives the tight query

⁹ For instance a CNF/DNF formula, an NC circuit, or a general circuit.

bound for the property AND (OR) are truth tables of functions that can be represented by $n^{O(1)}$ sized DNF (CNF) formulas. As $Q_\epsilon(\text{AND}|_{S_{\text{DEPTH2}}}) = Q_\epsilon(\text{AND})$ and $\text{qTimeBB}_\epsilon(\text{AND}|_{S_{\text{DEPTH2}}}) \geq Q_\epsilon(\text{AND}|_{S_{\text{DEPTH2}}})$. Therefore, we have $\text{AND} \in \mathcal{CO}(\text{DEPTH2})$. The same result holds for the property OR as well.

For each class of representation γ we now define the corresponding γ -QSETH, which states that computing any compression-oblivious property P in the *white-box* setting is at least as hard as computing P in the *black-box* setting. More formally,

Conjecture 3 (γ -QSETH). *For every class of representation γ , such as the class of depth-2 circuits DEPTH2 or poly-sized circuits of a more complex class, for all properties $P \in \mathcal{CO}(\gamma)$, we have $\text{qTimeWB}_\epsilon(P|_\gamma) \geq \Omega(Q_\epsilon(P))$.*

5.3 Observations on the set of compression oblivious properties

As the class γ gets more complex, the corresponding γ -QSETH becomes more credible. The set of compression oblivious properties is an interesting object of study by itself. First consider the following facts about sets of compression-oblivious properties relate, relative to different computational models.

Fact 1. *Given two classes of representations A and B , if $A \subseteq B$ then for every property P , we have $P \in \mathcal{CO}(B)$ whenever $P \in \mathcal{CO}(A)$.*

Fact 2. *Given two classes of representations A and B , if $A \subseteq B$ then A -QSETH implies B -QSETH.*

Proof. For Fact 1. If $A \subseteq B$ then also for the corresponding sets of truth tables it holds that $S_A \subseteq S_B$. If a property $P \in \mathcal{CO}(A)$, then $\text{qTimeBB}_\epsilon(P|_{S_A}) \geq \Omega(Q_\epsilon(P))$ also implies $\text{qTimeBB}_\epsilon(P|_{S_B}) \geq \text{qTimeBB}_\epsilon(P|_{S_A})$ as S_B is a superset of S_A . Therefore, $P \in \mathcal{CO}(B)$.

For Fact 2: Whenever some property $P \in \mathcal{CO}(A)$ is hard to compute for inputs coming from A , this property is also $P \in \mathcal{CO}(B)$ by Fact 1. Therefore, it is also hard to compute on an even wider range of inputs. \square

Given an explicit property P and a class of representation γ , it would be desirable to unconditionally prove that the property P is γ -compression oblivious¹⁰. This is possible for some simple properties that have query complexity $\Theta(\sqrt{N})$ like OR, corresponding to ordinary satisfiability, and AND. Unfortunately, for more complicated properties, like computing the parity of the number of satisfying assignments, it turns out to be hard to find an unconditional proof that such a property is compression oblivious. The following theorem shows a barrier to finding such an unconditional proof: proving that such a property is compression oblivious implies separating P from PSPACE.

Theorem 3. *If there exists a property P such that $Q_\epsilon(P) = \tilde{\omega}(\sqrt{N})$ and P is γ -compression oblivious, and for all $\epsilon > 0$ we have $P \in \text{SPACE}(N^\epsilon)$, then $P \neq \text{PSPACE}$.*

¹⁰ We call a property P a γ -compression oblivious property if $P \in \mathcal{CO}(\gamma)$.

Proof (sketch). By way of contradiction, assume $P = \text{PSPACE}$. We are given a promise that the circuit to whom we have black-box access to is in the set γ . Using a simplified version of the algorithm for the oracle identification problem [AIK⁺04, Kot14] and assuming $P = \text{PSPACE}$, we can extract a compressed form of the entire input using only $\tilde{O}(\sqrt{N})$ quantum time.

As the property $P \in \text{SPACE}(N^\epsilon)$, using the $P = \text{PSPACE}$ assumption again, we can directly compute P in time $O(N^\epsilon)$ for any arbitrarily small ϵ . Therefore, the total number of (quantum) steps taken is $\tilde{O}(\sqrt{N}) + O(N^\epsilon)$, which for an $\epsilon < \frac{1}{2}$ is in contradiction to the assumption that P is γ -compression oblivious. \square

Note that SETH is already a much stronger assumption than $P \neq \text{PSPACE}$, therefore this observation leaves open the interesting possibility of proving that properties are compression oblivious assuming that the (Q)SETH holds for simpler properties. (For instance, these simpler properties could include OR and AND, for which it is possible to unconditionally prove that they are compression oblivious.)

6 QSETH lower bounds for Orthogonal Vectors and Useful Proofs of Work

Recall that DEPTH2 denotes the set of polynomial-sized depth-2 circuits consisting of only OR and AND gates of unbounded fan-in. Because of its simple structure, the DEPTH2-QSETH conjecture is therefore closest to the classical SETH, and implies the ‘basic QSETH’ as introduced in Section 5.1:

Corollary 1. *If DEPTH2-QSETH is true then there is no bounded error quantum algorithm that solves CNF-SAT on n variables, m clauses in $O(2^{(1-\delta)n/2} m^{O(1)})$ time, for any $\delta > 0$.*

Proof. Consider the property OR: $\{0, 1\}^{2^n} \rightarrow \{0, 1\}$. Using the fact that $\text{OR} \in \mathcal{CO}(\text{DEPTH2})$, as shown in the previous section, we get $\text{qTimeWB}_\epsilon(\text{OR}|_{\text{DEPTH2}}) \geq \Omega(Q_\epsilon(\text{OR})) = \Omega(2^{n/2})$. Due to the structure of the DNF formulas one can compute the property OR on DNF formulas on n variables, m clauses in $n^{O(1)} m^{O(1)}$ time. This implies that the hard cases in the set DEPTH2 for the OR property are the CNF formulas. Therefore, $\text{qTimeWB}_\epsilon(\text{OR}|_{\text{CNF}}) \geq \Omega(2^{n/2})$ where the set CNF denotes all the polynomial sized CNF formulas. \square

In this section we present several immediate consequences of the DEPTH2-QSETH conjecture, including:

1. For some problems, classical SETH-based $\Omega(T)$ time lower bounds carry over to the quantum case, with DEPTH2-QSETH-based $\Omega(\sqrt{T})$ quantum time lower bounds using (almost) the same reduction.
2. The *Proofs of Useful Work* of Ball, Rosen, Sabin and Vasudevan [BRSV17] require time $\tilde{O}(n^2)$ to solve on a quantum computer, equal to their classical complexity, under DEPTH2-QSETH.

6.1 Quantum time lower bounds based on DEPTH2-QSETH

The statement of DEPTH2-QSETH along with Corollary 1 can give quantum time lower bounds for some problems for which we know classical lower bounds under SETH (Conjecture 2).

Corollary 2. *Let P be a problem with an $\Omega(T)$ time lower bound modulo SETH. Then, P has an $\tilde{\Omega}(\sqrt{T})$ quantum time lower bound conditioned under DEPTH2-QSETH if there exists a classical reduction from CNF-SAT to the problem P taking $O(2^{\frac{n}{2}(1-\alpha)})$ (for $\alpha > 0$) time or if there exists an efficient reduction that can access a single bit of the reduction output.¹¹*

As examples we will consider the ORTHOGONAL VECTORS and the EDIT DISTANCE problem. The ORTHOGONAL VECTORS (OV) problem is defined as follows. Given two sets U and V of n vectors, each over $\{0, 1\}^d$ where $d = \omega(\log n)$, determine whether there exists a $u \in U$ and a $v \in V$ such that $\sum_{l \in [d]} u_l v_l = 0$. In [Wil05], Williams showed that SETH implies the non-existence of a sub-quadratic classical algorithm for the OV problem. In the quantum case the best-known query lower bound is $\Omega(n^{2/3})$, which can be achieved by reducing the 2-TO-1 COLLISION problem to the ORTHOGONAL VECTORS problem; however, the known quantum time upper bound is $\tilde{O}(n)$ [Ren19]. First note that we cannot use Williams' classical reduction directly, since a hypothetical quantum algorithm for OV expects quantum access to the input, and writing down the entire reduction already takes time $2^{n/2}$. Instead, observe that the reduction produces a separate vector for each partial assignment: let $t(n)$ be the time needed to compute a single element of the output of the reduction, then $t(n) = \text{poly}(n)$, which is logarithmic in the size of the total reduction. Let $N = O^*(2^{n/2})$ be the size of the output of the reduction of [Wil05], for some CNF formula with n variables. Any quantum algorithm that solves OV in time N^α , can solve CNF-SAT in time $t(n)O^*(2^{\alpha n/2}) = O^*(2^{\alpha n/2})$.¹² Assuming DEPTH2-QSETH, this implies that a quantum algorithm requires time $\tilde{\Theta}(N)$ to solve OV for instances of size N .

See the recent results by Aaronson, Chia, Lin, Wang, and Zhang [ACL⁺19] for more examples of reductions from (a variant of) QSETH, that also hold for the basic QSETH of our framework. Additionally, there the authors define the notion of *Quantum Fine-grained Reductions* more generally, and present a study of OV that also includes the case of constant dimension.

The next example we consider is the EDIT DISTANCE problem. The EDIT DISTANCE problem is defined as follows. Given two strings a and b over an alphabet set Σ , the edit distance between a and b is the minimum number of operations (insertions, deletions, substitutions) on the symbols required to transform string a to b (or vice versa). A reduction by [BI15] shows that if the edit distance between two strings of length n can be computed in time $O(n^{2-\delta})$ for some constant $\delta > 0$, then satisfiability on CNF formulas with n variables and m clauses can be computed in $O(m^{O(1)} \cdot 2^{(1-\frac{\delta}{2})n})$ which would imply that SETH (Conjecture 2) is false. Just like in the ORTHOGONAL VECTORS case, we observe that the classical reduction from CNF-SAT

¹¹ Note that SETH talks about solving CNF-SAT as opposed to bounded k -SAT problems. One could also define a quantum hardness conjecture for k -CNF or k -DNF, for an arbitrary constant k , in the same way as the original SETH. This variant is required for reductions that use the fact that k is constant, which can occur through usage of the sparsification lemma [IP01]. For examples where this is necessary within fine-grained complexity, see the *Matching Triangles* problem mentioned in [AVY15] or reductions like in [CDL⁺16].

¹² We use O^* to denote asymptotic complexity ignoring polynomial factors.

to EDIT DISTANCE is local, in the sense that accessing a single bit of the exponentially-long reduction output can be done in polynomial time: Every segment of the strings that are an output of the reduction, depend only on a single partial satisfying assignment, out of the $2^{n/2}$ possible partial assignments.

This observation directly lets us use the reduction of [BI15] to give a quantum time lower bound of $\tilde{\Omega}(n)$ for the EDIT DISTANCE problem, where n here is the length of the inputs to EDIT DISTANCE, conditioned on DEPTH2-QSETH. However, an unconditional quantum query lower bound of $\Omega(n)$ can also be easily achieved by embedding of a problem with high query complexity, such as the majority problem, in an edit distance instance.

We witness that with the DEPTH2-QSETH conjecture, the SETH-based fine-grained lower bounds at best transfer to a square root lower complexity in the quantum case. This is definitely interesting on its own, but we are aiming for larger quantum lower bounds, which is why we focus on our more general framework.

6.2 Quantum Proofs of Useful Work

Other applications of DEPTH2-QSETH include providing problems for which *Proofs of Useful Work* (uPoW) can be presented in the quantum setting. The paper [BRSV17] proposes uPoW protocols that are based on delegating the evaluation of low-degree polynomials to the prover. They present a classical uPoW protocol for the ORTHOGONAL VECTORS problem (OV) whose security proof is based on the assumption that OV needs $\Omega(n^{2-o(1)})$ classical time in the worst case setting, implying that the evaluation of a polynomial that encodes the instance of OV has average-case hardness. At the end of this protocol, the verifier is able to compute the number of orthogonal vectors in a given instance.

Therefore, the same protocol also works to verify the solutions to $\oplus\text{OV}$, where $\oplus\text{OV}$ denotes the parity version of OV, i.e., given two sets U, V of n vectors from $\{0, 1\}^d$ each, output the parity of number of pairs (u, v) such that $u \in U, v \in V$ and $\sum_{l \in [d]} u_l v_l = 0$, where d is taken to be $\omega(\log n)$. Assuming DEPTH2-QSETH and assuming $\text{PARITY} \in \mathcal{CO}(\text{DEPTH2})$ we get that $\oplus\text{CNF-SAT}$ takes $\Omega(2^n)$ quantum time. Due to the reduction¹³ given in [Wil05], this protocol then implies a conditional quantum time lower bound of $\Omega(n^2)$ for the $\oplus\text{OV}$ problem. Therefore, the uPoW protocol by [BRSV17] also requires quantum provers to take time $\tilde{\Omega}(n^2)$.

7 Lower bound for edit distance using NC-QSETH

We now discuss a consequence of the NC-QSETH conjecture: A quantum time lower bound for the Edit Distance problem. As mentioned in Section 2.3, edit distance is a measure of dissimilarity between two strings. For input strings of length n , the well known Wagner–Fischer algorithm (based on dynamic programming) classically computes the edit distance in $O(n^2)$ time. Unfortunately, all the best known classical (and quantum) algorithms to compute the edit distance are also nearly quadratic. The result by [BI15] proves that these near quadratic clas-

¹³ Note that here one can use the classical reduction from CNF-SAT to ORTHOGONAL VECTORS that runs in $\tilde{O}(2^{n/2})$.

sical time bounds might be tight. Using a simple reduction from OV to Edit Distance, they show that a sub-quadratic classical algorithm for computing Edit Distance would imply a sub-quadratic algorithm for OV which would make SETH (refer to Conjecture 2) false. SETH also implies quadratic lower bounds for many other string comparison problems like Longest Common Subsequence and Dynamic Time Warping whose trivial algorithms are also based on dynamic programming [BI15, BK15]. Bouroujeni *et al.* in [BEG⁺18] give a sub-quadratic quantum algorithm for approximating edit distance within a constant factor which was followed by a better classical algorithm in [CDG⁺18] by Chakraborty *et al.* However, no quantum improvements over the classical algorithms in the exact case are known. Investigating why this is the case is an interesting open problem, which can be addressed in two directions. We formulate the following questions for the example of the Edit Distance problem.

1. Is there a bounded-error quantum algorithm for EDIT DISTANCE that runs in a sub-quadratic amount of time?
2. Can we use a different reduction to raise the linear lower bound for EDIT DISTANCE that we achieve under ‘basic QSETH’?

While the first question still remains open, we address the second question in the following subsection. Independently from [BPS19], Ambainis *et al.* [ABI⁺19] present a quantum query lower bound of $\Omega(n^{1.5-o(1)})$ for solving Edit Distance when solved using the most natural approach by reducing Edit Distance to connectivity on a 2D grid. However, it doesn’t rule out the possibility of other $\tilde{O}(n^{1.5-\alpha})$ quantum algorithms for Edit Distance, for $\alpha > 0$. In this section, we will sketch a conditional quantum time lower bound of $\Omega(n^{1.5})$ for the Edit Distance problem using the NC-QSETH conjecture.

7.1 Reduction from ORTHOGONAL VECTORS to EDIT DISTANCE

For the sake of simplicity we will first give an overview of the reduction from OV to Edit Distance presented by [BI15]. Recall that OV is defined as follows: Given two sets U and V of n vectors, each over $\{0, 1\}^d$ where $d = \omega(\log n)$, determine whether there exists a $u \in U$ and a $v \in V$ such that $\langle u | v \rangle = 0$ (i.e. $\sum_{i \in [d]} u_i v_i = 0$). In the reduction, for each vector $u \in U$ and $v \in V$, we construct a *vector gadget*. Then, the vector gadgets for all $u \in U$ are concatenated together to form the first *input sequence* a . Similarly, vector gadgets for all $v \in V$ are concatenated together to form the second input sequence b . These gadget sequences a and b are constructed in such a way that edit distance between a and b is equal to a constant Y if there are no orthogonal pairs in the list U and V , and, is strictly less than Y if there exists an orthogonal pair.

We will now briefly describe all the gadgets and provide some intuition behind the construction. The alphabet set over which the sequence is constructed is $\Sigma = \{0, 1, 2, 3, 4, 5, 6\}$. Let $k_0 = 1000d$, where d is the dimension of the vectors in OV. Each vector gadget will be built up from *coordinate gadgets*, that are defined as follows:

Given $x \in \{0, 1\}$,

$$CG_1(x) = \begin{cases} 2^{k_0} 01112^{k_0}, & \text{if } x = 0, \\ 2^{k_0} 00012^{k_0}, & \text{otherwise.} \end{cases} \quad CG_2(x) = \begin{cases} 2^{k_0} 00112^{k_0}, & \text{if } x = 0, \\ 2^{k_0} 11112^{k_0}, & \text{otherwise.} \end{cases}$$

The coordinate gadgets are so designed that for any two numbers $x_1, x_2 \in \{0, 1\}$,

$$\text{Edit-Dist}(CG_1(x_1), CG_2(x_2)) = \begin{cases} 1, & \text{if } x_1 \cdot x_2 = 0, \\ 3, & \text{if } x_1 \cdot x_2 = 1. \end{cases}$$

For any two vectors $u, v \in \{0, 1\}^d$, and defining $w = 10^{d-1}$, the vector gadgets are constructed as follows:

$$\begin{aligned} VG_1(u) &= 4^{k_1} (\bigcirc_{i \in [d]} CG_1(u[i])) 3^{k_1} (\bigcirc_{i \in [d]} CG_1(w[i])) 4^{k_1}, \\ VG_2(v) &= 3^{k_1} (\bigcirc_{i \in [d]} CG_2(v[i])) 3^{k_1}, \end{aligned}$$

here \bigcirc stands for concatenation, $k_1 = (1000 \cdot d)^2$ and $u[i]$ denotes the i^{th} coordinate of the vector u while $v[i]$ denotes the i^{th} coordinate of v . Consequently, this design of the vector gadgets ensures that,

$$\text{Edit-Dist}(VG_1(u), VG_2(v)) = \begin{cases} E_s := 2k_1 + k + d, & \text{if } u \cdot v = 0, \\ E_u := 2k_1 + k + d + 2, & \text{if } u \cdot v \geq 1, \end{cases}$$

where $k = d(4 + 2k_0)$ and $u \cdot v = \sum_{i \in [d]} u_i v_i$.¹⁴

Finally, proceed by concatenating the vector gadgets into two sequences. Given the two lists of vectors U and V consisting of n vectors each over $\{0, 1\}^d$, define:

$$P'_1 = 6^{|P'_2|} (\bigcirc_{u \in U} 5^T VG_1(u) 5^T) 6^{|P'_2|},$$

$$P'_2 = (\bigcirc_{i=1}^{n-1} 5^T VG_2(f) 5^T) (\bigcirc_{v \in V} 5^T VG_2(v) 5^T) (\bigcirc_{i=1}^{n-1} 5^T VG_2(f) 5^T),$$

where f is 1^d and $T = \Theta(d^3)$. Again, the construction at this final level ensures that there exists a constant $Y = 2P'_2 + nE_u$ (only dependent on the size of OV), such that,

$$\text{Edit-Dist}(P'_1, P'_2) = \begin{cases} Y, & \text{if there are no orthogonal vectors in } U \text{ and } V, \\ \leq Y - 2, & \text{if there exists a pair of orthogonal vectors in } U, V. \end{cases}$$

Note that the size of the sequences P'_1 and P'_2 are of $O(nd^3)$. Any sub-quadratic algorithm for Edit Distance would imply a sub-quadratic algorithm for OV. In Section 2.5 we discussed a simple yet elegant reduction from CNF-SAT to OV. Therefore, it is easy to see that there doesn't exist a sub-quadratic algorithm for Edit Distance unless SETH is false.

7.2 Observations

Combining the two reductions from CNF-SAT to OV and OV to Edit Distance, we notice that computing the edit distance between the two final sequences P'_1 and P'_2 gives more than just the existence of a satisfying assignment. Instead, the result of these reductions encodes whether *many* satisfying assignments exist (in a certain pattern), a property on the satisfying assignments that might not be amenable to a speedup using Grover's algorithm. We call this property PP_{edit} . The 'basic QSETH' is not able to account for this distinction, and therefore

¹⁴ Note that, in the proof of the Theorem 2 by [BI15] it is assumed that $v \in \{0, 1\}^d$ is any binary vector that starts with 1. This can be easily achieved by just increasing the dimensions of OV to $(d + 1)$.

does not directly help with explaining why a linear-time quantum algorithm for Edit Distance has not been found. Using the more general DEPTH2-QSETH and conjecturing that PP_{edit} is a DEPTH2-compression oblivious property, we are able to get a better than linear quantum lower bound for the Edit Distance problem.

The DEPTH2 circuits (CNF and DNF formulas) have a relatively simple structure, therefore, the result would be strengthened if the hardness can be based on computing such a property of a more-complicated class of formulas, such as NC circuits.

With this in mind, we are able to modify the first part of this reduction to mimic the approach in [AHVWW16] while constructing the final sequences in the same way as in [BI15]. After careful inspection of the reduction(s) we are able to prove that computing edit distance between the final sequences, also computes the property PP_{edit} on the satisfying assignments to an NC circuit. Simultaneously, it is possible to prove a $\Omega(2^{0.75n})$ quantum query lower bound for PP_{edit} in the black-box setting by using the quantum adversary method given by [Amb02]. We then use this query lower bound as a time lower bound in the white-box setting assuming that NC-QSETH is true and PP_{edit} is NC-compression oblivious. Which then, because of this combined reduction, gives us a quantum time lower bound of $\Omega(n^{1.5})$ for the Edit Distance problem.

References

- [AAB⁺19] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C. Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando G. S. L. Brandao, David A. Buell, Brian Burkett, Yu Chen, Zijun Chen, Ben Chiaro, Roberto Collins, William Courtney, Andrew Dunsworth, Edward Farhi, Brooks Foxen, Austin Fowler, Craig Gidney, Marissa Giustina, Rob Graff, Keith Guerín, Steve Habegger, Matthew P. Harrigan, Michael J. Hartmann, Alan Ho, Markus Hoffmann, Trent Huang, Travis S. Humble, Sergei V. Isakov, Evan Jeffrey, Zhang Jiang, Dvir Kafri, Kostyantyn Kechedzhi, Julian Kelly, Paul V. Klimov, Sergey Knysh, Alexander Korotkov, Fedor Kostritsa, David Landhuis, Mike Lindmark, Erik Lucero, Dmitry Lyakh, Salvatore Mandrà, Jarrod R. McClean, Matthew McEwen, Anthony Megrant, Xiao Mi, Kristel Michielsen, Masoud Mohseni, Josh Mutus, Ofer Naaman, Matthew Neeley, Charles Neill, Murphy Yuezhen Niu, Eric Ostby, Andre Petukhov, John C. Platt, Chris Quintana, Eleanor G. Rieffel, Pedram Roushan, Nicholas C. Rubin, Daniel Sank, Kevin J. Satzinger, Vadim Smelyanskiy, Kevin J. Sung, Matthew D. Trevithick, Amit Vainsencher, Benjamin Villalonga, Theodore White, Z. Jamie Yao, Ping Yeh, Adam Zalcman, Hartmut Neven, and John M. Martinis. Quantum supremacy using a programmable superconducting processor. *Nature*, 574(7779):505–510, 2019.
- [AB09] Sanjeev Arora and Boaz Barak. *Computational Complexity: A Modern Approach*. Cambridge University Press, New York, NY, USA, 1st edition, 2009.
- [ABI⁺19] Andris Ambainis, Kaspars Balodis, Janis Iraids, Krisjanis Prūsis, and Juris Smotrovs. Quantum Lower Bounds for 2D-Grid and Dyck Language. 2019.
- [ABV15] Amir Abboud, Arturs Backurs, and Virginia Vassilevska Williams. Tight hardness results for LCS and other sequence similarity measures. In *Proceedings of the 2015 IEEE 56th Annual Symposium on Foundations of Computer Science (FOCS)*, FOCS '15, pages 59–78, Washington, DC, USA, 2015. IEEE Computer Society.
- [ACL⁺19] Scott Aaronson, Nai-Hui Chia, Han-Hsuan Lin, Chunhao Wang, and Ruizhe Zhang. On the quantum complexity of closest pair and related problems. *arXiv preprint arXiv:1911.01973*, 2019.
- [AHVWW16] Amir Abboud, Thomas Dueholm Hansen, Virginia Vassilevska Williams, and Ryan Williams. Simulating Branching Programs with Edit Distance and Friends Or: a Polylog Shaved is a Lower Bound Made. In *Proceedings of the Forty-eighth Annual ACM Symposium on Theory of Computing*, STOC '16, pages 375–388, New York, NY, USA, 2016. ACM.
- [AIK⁺04] Andris Ambainis, Kazuo Iwama, Akinori Kawachi, Hiroyuki Masuda, Raymond H. Putra, and Shigeru Yamashita. Quantum identification of boolean oracles. In Volker Diekert and Michel Habib, editors, *STACS 2004*, pages 105–116, Berlin, Heidelberg, 2004. Springer Berlin Heidelberg.

- [Amb00] Andris Ambainis. Quantum lower bounds by quantum arguments. In *Proceedings of the Thirty-second Annual ACM Symposium on Theory of Computing*, STOC '00, pages 636–643, New York, NY, USA, 2000. ACM.
- [Amb02] A. Ambainis. Quantum lower bounds by quantum arguments. *Journal of Computer and System Sciences*, 64(4):750–767, 2002. Earlier version in STOC'00. quant-ph/0002066.
- [AVW14] Amir Abboud, Virginia Vassilevska Williams, and Oren Weimann. Consequences of faster alignment of sequences. In *ICALP*, 2014.
- [AVY15] Amir Abboud, Virginia Vassilevska Williams, and Huacheng Yu. Matching triangles and basing hardness on an extremely popular conjecture. In *Proceedings of the Forty-seventh Annual ACM Symposium on Theory of Computing*, STOC '15, pages 41–50, New York, NY, USA, 2015. ACM.
- [BBBV97] Charles H. Bennett, Ethan Bernstein, Gilles Brassard, and Umesh Vazirani. Strengths and weaknesses of quantum computing. *SIAM J. Comput.*, 26(5):1510–1523, October 1997.
- [BBC⁺01] Robert Beals, Harry Buhrman, Richard Cleve, Michele Mosca, and Ronald de Wolf. Quantum lower bounds by polynomials. *J. ACM*, 48(4):778–797, July 2001.
- [BBHT98] Michel Boyer, Gilles Brassard, Peter Høyer, and Alain Tapp. Tight bounds on quantum searching. *Fortschritte der Physik: Progress of Physics*, 46(4-5):493–505, 1998.
- [BEG⁺18] Mahdi Boroujeni, Soheil Ehsani, Mohammad Ghodsi, Mohammad Taghi Hajiaghayi, and Saeed Seddighin. Approximating edit distance in truly subquadratic time: Quantum and mapreduce. *CoRR*, abs/1804.04178, 2018.
- [BI15] Arturs Backurs and Piotr Indyk. Edit distance cannot be computed in strongly subquadratic time (unless SETH is false). *STOC*, 2015.
- [BK15] Karl Bringmann and Marvin Kunnemann. Quadratic conditional lower bounds for string problems and dynamic time warping. In *Proceedings of the 2015 IEEE 56th Annual Symposium on Foundations of Computer Science (FOCS)*, FOCS '15, pages 79–97, Washington, DC, USA, 2015. IEEE Computer Society.
- [BPS19] Harry Buhrman, Subhasree Patro, and Florian Speelman. The quantum strong exponential-time hypothesis, 2019.
- [Bri14] Karl Bringmann. Why walking the dog takes time: Frechet distance has no strongly subquadratic algorithms unless seth fails. In *Proceedings of the 2014 IEEE 55th Annual Symposium on Foundations of Computer Science*, FOCS '14, pages 661–670, Washington, DC, USA, 2014. IEEE Computer Society.
- [BRSV17] Marshall Ball, Alon Rosen, Manuel Sabin, and Prashant Nalini Vasudevan. Proofs of useful work. Cryptology ePrint Archive, Report 2017/203, 2017.

- [BSV19] Harry Buhrman, Mario Szegedy, and Thomas Vyskocil. Lower Bounds for Quantum Circuits. manuscript, 2019.
- [BV97] E. Bernstein and U. Vazirani. Quantum complexity theory. *SIAM Journal on Computing*, 26(5):1411–1473, 1997.
- [Can74] G. Cantor. Über eine Eigenschaft des Inbegriffes aller reellen algebraischen Zahlen. *Crelle's Journal für Mathematik*, 77:258–263, 1874.
- [CDG⁺18] Diptarka Chakraborty, Debarati Das, Elazar Goldenberg, Michal Koucký, and Michael E. Saks. Approximating edit distance within constant factor in truly sub-quadratic time. *CoRR*, abs/1810.03664, 2018.
- [CDL⁺16] Marek Cygan, Holger Dell, Daniel Lokshtanov, Dániel Marx, Jesper Nederlof, Yoshio Okamoto, Ramamohan Paturi, Saket Saurabh, and Magnus Wahlström. On problems as hard as cnf-sat. *ACM Trans. Algorithms*, 12(3):41:1–41:24, May 2016.
- [Chu36] Alonzo Church. An unsolvable problem of elementary number theory. *American Journal of Mathematics*, 1936.
- [CIP06] Chris Calabro, Russell Impagliazzo, and Ramamohan Paturi. A duality between clause width and clause density for SAT. In *Proceedings of the 21st Annual IEEE Conference on Computational Complexity*, CCC '06, pages 252–260, Washington, DC, USA, 2006. IEEE Computer Society.
- [CLRS01] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms*. MIT Press & McGraw–Hill, 2nd edition, 2001.
- [Coo71] Stephen A. Cook. The complexity of theorem-proving procedures. In *STOC*, pages 151–158, 1971.
- [Deu85] David Deutsch. Quantum theory, the church-turing principle and the universal quantum computer. *Proc. Roy. Soc. London Ser. A*, 1985.
- [DHM⁺14] Holger Dell, Thore Husfeldt, Dániel Marx, Nina Taslaman, and Martin Wahlén. Exponential time complexity of the permanent and the tutte polynomial. *ACM Transactions on Algorithms (TALG)*, 10(4):21, 2014.
- [DJ92] David Deutsch and Richard Jozsa. Rapid solution of problems by quantum computation. *Proc. Roy. Soc. London Ser. A*, 493(1907):553–558, December 1992.
- [dW20] Ronald de Wolf. Quantum Computing: Lecture Notes. <https://homepages.cwi.nl/~rdewolf/qcnotes.pdf>, 2020. [Online; accessed January-2020].
- [FGGS98] Edward Farhi, Jeffrey Goldstone, Sam Gutmann, and Michael Sipser. Limit on the speed of quantum computation in determining parity. *Phys. Rev. Lett.*, 81:5442–5444, Dec 1998.

- [Göd30] K. Gödel. Über formal unentscheidbare sätze der principia mathematica und verwandter systeme i. *Monatsh. f. Mathematik und Physik*, 38:173–198, 1930.
- [Gro96] Lov K. Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of the Twenty-eighth Annual ACM Symposium on Theory of Computing*, STOC '96, pages 212–219, New York, NY, USA, 1996. ACM.
- [HNS18] Cupjin Huang, Michael Newman, and Mario Szegedy. Explicit lower bounds on strong quantum simulation. *arXiv preprint arXiv:1804.10368*, 2018.
- [IM02] Kazuo Iwama and Hiroki Morizumi. An explicit lower bound of $5n - o(n)$ for boolean circuits. In Krzysztof Diks and Wojciech Rytter, editors, *Mathematical Foundations of Computer Science 2002*, pages 353–364, Berlin, Heidelberg, 2002. Springer Berlin Heidelberg.
- [IP01] Russell Impagliazzo and Ramamohan Paturi. On the complexity of k -SAT. *Journal of Computer and System Sciences*, 62(2):367 – 375, 2001.
- [IPZ01] Russell Impagliazzo, Ramamohan Paturi, and Francis Zane. Which problems have strongly exponential complexity? *Journal of Computer and System Sciences*, 63(4):512 – 530, 2001.
- [JNV⁺20] Zheng-Feng Ji, Anand Natarajan, Thomas Vidick, John Wright, and Henry Yuen. $MIP^* = RE$. *ArXiv*, abs/2001.04383, 2020.
- [Jor20] Stephen Jordan. Quantum Algorithm Zoo. <http://quantumalgorithmzoo.org>, 2020. [Online; accessed January-2020].
- [Kot14] Robin Kothari. An optimal quantum algorithm for the oracle identification problem. In Ernst W. Mayr and Natacha Portier, editors, *31st International Symposium on Theoretical Aspects of Computer Science (STACS 2014)*, volume 25 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 482–493, Dagstuhl, Germany, 2014. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.
- [Lev73] Leonid Levin. Universal search problems. *Problems of Information Transmission*, 9(3):265–266, 1973.
- [Mon16] Ashley Montanaro. Quantum algorithms: an overview. *npj Quantum Information*, 2(1):15023, 2016.
- [MP80] William J. Masek and Michael S. Paterson. A faster algorithm computing string edit distances. *Journal of Computer and System Sciences*, 20(1):18 – 31, 1980.
- [MT19] Tomoyuki Morimae and Suguru Tamaki. Fine-grained quantum computational supremacy. *Quantum Information & Computation*, 19(13&14):1089–1115, 2019.
- [NC00] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*, volume 70. Cambridge University Press, 2000.
- [Pos43] Emil Post. Formal reductions of the general combinatorial decision problem. *American Journal of Mathematics*, 2(65):197–215, 1943.

- [PPSZ05] Ramamohan Paturi, Pavel Pudlák, Michael E. Saks, and Francis Zane. An improved exponential-time algorithm for k-SAT. *J. ACM*, 52(3):337–364, May 2005.
- [Ren19] Jorg Van Renterghem. The implications of breaking the strong exponential time hypothesis on a quantum computer. Master’s thesis, Ghent University, 2019.
- [Sho97] P. W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26(5):1484–1509, 1997. Earlier version in FOCS’94. quant-ph/9508027.
- [Sim97] Daniel R. Simon. On the power of quantum computation. *SIAM J. Comput.*, 26(5):1474–1483, 1997.
- [Tur36] Alan Turing. On computable numbers, with an application to the entscheidungsproblem. *Proceedings of the London Mathematical Society*, 2(42):230—265, 1936. Adendum 1937.
- [Vas15] Virginia Vassilevska Williams. Hardness of easy problems: Basing hardness on popular conjectures such as the strong exponential time hypothesis. *IPEC*, 2015.
- [Vas18] Virginia Vassilevska Williams. On some fine-grained questions in algorithms and complexity. In *Proceedings of the ICM*, 2018. To appear.
- [Ž83] Stanislav Žák. A turing machine time hierarchy. *Theoretical Computer Science*, 26(3):327 – 333, 1983.
- [WF74] Robert A. Wagner and Michael J. Fischer. The string-to-string correction problem. *J. ACM*, 21(1):168?173, January 1974.
- [Wil05] Ryan Williams. A new algorithm for optimal 2-constraint satisfaction and its implications. *Theor. Comput. Sci.*, 348(2):357–365, December 2005.

The IFF Spring School and Quantum Information at Jülich

web: *www.iff-springschool.de*
email: *springschool@fz-juelich.de*
phone: *++49 2461 61-4750*

The annual IFF Spring School has a long-standing tradition dating back to 1969 when the Institut für Festkörperforschung (IFF) was founded. Since then, the Schools have made it possible for students and young researchers to gain a two-week insight into current topics related to condensed matter research. Since the restructuring in 2011, research in the area of electronic systems and information technology became part of the Peter Grünberg Institute (PGI) named after IFF's Nobel Prize winner (Physics in 2007). Biophysics and soft matter research has found a home in the Institute of Complex Systems (ICS). Two additional method-oriented institutes which emerged from the IFF are the Institute for Advanced Simulation (IAS) and the Jülich Center for Neutron Science (JCNS).

Extensive research is underway in Jülich in the areas of quantum technology and quantum information science. Much of it is hosted in the Institute for Quantum Information (IQI), a joint enterprise of PGI and Aachen University. In close cooperation with several other PGI divisions (Theoretical Nanoelectronics, Quantum Control, and Semiconductor Nanoelectronics), work is underway to develop both the theoretical and experimental underpinnings of future quantum computers. Our experimental groups work on the realization of quantum computing components in single-spin quantum dots in semiconducting systems (both gallium-arsenide and silicon-germanium based), and in hybrid superconducting-semiconducting structures that realize Majorana qubits. In cooperation with our technical department, the Central Institute for Engineering and Analytics (ZEA-2), we are working towards the development of electronic control systems that will make scalable quantum devices possible.

Quantum Technology

51st IFF Spring School **2020**

23.03.-03.04.2020

H. Bluhm, T. Calarco, D. DiVincenzo

ISBN: 978-3-95806-449-2

Scattering!Soft, Functional and Quantum Materials

50th IFF Spring School **2019**

11. - 22.03.2019

M. Angst, T. Brückel, S. Förster, K. Frieze, R. Zorn

ISBN: 978-3-95806-380-8

Physics of Life

49th IFF Spring School **2018**

26.02.-09.03.2018

G. Gompper, J. Dhont, J. Elgeti, C. Fahlke, D.

Fedosov, S. Förster, P. Lettinga, A. Offenhäusser

ISBN: 978-3-95806-286-3

Topological Matter - Topological Insulators, Skyrmions and Majoranas

48th IFF Spring School **2017**

27.03.-07.04.2017

S. Blügel, Y. Mokrousov, T. Schäpers, Y. Ando

ISBN: 978-3-95806-202-3

Memristive Phenomena - From Fundamental Physics to Neuromorphic Computing

47th IFF Spring School **2016**

22.02.-04.03.2016

R. Waser, M. Wuttig

ISBN: 978-3-95806-091-3

Functional Soft Matter

46th IFF Spring School **2015**

23.02.-06.03.2015

J. Dhont, G. Gompper, G. Meier, D. Richter, G.

Vliegenthart, R. Zorn

ISBN: 978-3-89336-999-7

Computing Solids - Models, ab-initio methods and supercomputing

45th IFF Spring School **2014**

10.-21.03.2014

S. Blügel, N. Helbig, V. Meden, D. Wortmann

ISBN: 978-3-89336-912-6

Quantum Information Processing

44th IFF Spring School **2013**

25.02.-08.03.2013

D. DiVincenzo

ISBN: 978-3-89336-833-4

Scattering Methods for Condensed Matter Research: Towards Novel Applications at Future Sources

43rd IFF Spring School **2012**

05.-16.03.2012

M. Angst, T. Brückel, D. Richter, R. Zorn

ISBN: 978-3-89336-759-7

Macromolecular Systems in Soft- and Living-Matter

42nd IFF Spring School **2011**

14.-25.02.2011

J. K.G. Dhont, G. Gompper, P. R. Lang, D. Richter,

M. Ripoll, D. Willbold, R. Zorn

ISBN: 978-3-89336-688-0

Electronic Oxides - Correlation Phenomena, Exotic Phases and Novel Functionalities

41st IFF Spring School **2010**

08.-19.03.2010

S. Blügel, T. Brückel, R. Waser, C.M. Schneider

ISBN: 978-3-89336-609-5

Spintronics – From GMR to Quantum Information

40th IFF Spring School **2009**

09.-20.03.2009

S. Blügel, D. Bürgler, M. Morgenstern, C. M.

Schneider, R. Waser

ISBN: 978-3-89336-559-3

Soft Matter - From Synthetic to Biological Materials

39th IFF Spring School **2008**

03.-14.03.20108

J.K.G. Dhont, G. Gompper, G. Nägele, D. Richter,

R.G. Winkler

ISBN: 978-3-89336-517-3

Probing the Nanoworld - Microscopies, Scattering and Spectroscopies of the Solid State

38th IFF Spring School **2007**

12.-23.03.2007

K. Urban, C. M. Schneider, T. Brückel, S. Blügel,

K. Tillmann, W. Schweika, M. Lentzen, L.

Baumgarten

ISBN: 978-3-89336-462-6

Computational Condensed Matter Physics

37th IFF Spring School **2006**

06.-17.03.2006

S. Blügel, G. Gompper, E. Koch, H. Müller-

Krumbhaar, R. Spatschek, R. G. Winkler

ISBN: 978-3-89336-430-5

**Magnetism goes Nano - Electron Correlations,
Spin Transport, Molecular Magnetism**

36th IFF Spring School **2005**

14.-25.02.2005

S. Blügel, T. Brückel, C. M. Schneider

ISBN: 3-89336-381-5

**Physics meets Biology - From Soft Matter of
Cell Biology**

35th IFF Spring School **2004**

22.03.-02.04.2004

G. Gompfer, U. B. Kaupp, J. K. G. Dhont, D.

Richter, R. G. Winkler

ISBN: 3-89336-348-3

Fundamentals of Nanoelectronics

34th IFF Spring School **2003**

10.-21.03.2003

S. Blügel, M. Luysberg, K. Urban, R. Waser

ISBN: 3-89336-319-X

**Soft Matter - Complex Materials on Mesoscopic
Scales**

33rd IFF Spring School **2002**

04.-15.03.2002

J.K.G. Dhont, G. Gompfer, D. Richter

ISBN: 3-89336-297-5

Neue Materialien für die Informationstechnik

32. IFF-Ferienkurses **2001**

05.-16.03.2001

R. Waser

ISBN: 3-89336-279-7

Band / Volume 197

Topological transport in non-Abelian spin textures from first principles

P. M. Buhl (2019), vii, 158 pp

ISBN: 978-3-95806-408-9

Band / Volume 198

Shortcut to the carbon-efficient microbial production of chemical building blocks from lignocellulose-derived D-xylose

C. Brüsseler (2019), X, 62 pp

ISBN: 978-3-95806-409-6

Band / Volume 199

Regulation and assembly of the cytochrome *bc*₁-aa₃ supercomplex in *Corynebacterium glutamicum*

C.-F. Davoudi (2019), 135 pp

ISBN: 978-3-95806-416-4

Band / Volume 200

Variability and compensation in Alzheimer's disease across different neuronal network scales

C. Bachmann (2019), xvi, 165 pp

ISBN: 978-3-95806-420-1

Band / Volume 201

Crystal structures and vibrational properties of chalcogenides: the role of temperature and pressure

M. G. Herrmann (2019), xi, 156 pp

ISBN: 978-3-95806-421-8

Band / Volume 202

Current-induced magnetization switching in a model epitaxial Fe/Au bilayer

P. Gospodarič (2019), vi, 120, XXXVIII pp

ISBN: 978-3-95806-423-2

Band / Volume 203

Network architecture and heme-responsive gene regulation of the two-component systems HrrSA and ChrSA

M. Keppel (2019), IV, 169 pp

ISBN: 978-3-95806-427-0

Band / Volume 204

Spin-orbitronics at the nanoscale: From analytical models to real materials

J. Bouaziz (2019), 228 pp

ISBN: 978-3-95806-429-4

Band / Volume 205

Advanced methods for atomic scale spin simulations and application to localized magnetic states

G. P. Müller (2019), xx, 194 pp

ISBN: 978-3-95806-432-4

Band / Volume 206

Different growth modes of molecular adsorbate systems and 2D materials investigated by low-energy electron microscopy

J. E. Felter (2019), vi, 114, XXXIV pp

ISBN: 978-3-95806-434-8

Band / Volume 207

NADPH-related studies performed with a SoxR-based biosensor in *Escherichia coli*

A. Spielmann (2019), IV, 73 pp

ISBN: 978-3-95806-438-6

Band / Volume 208

Chemisorption aromatischer Moleküle auf Übergangsmetalloberflächen: Bildung molekularer Hybridmagnete

S. Schleicher (2019), 109 pp

ISBN: 978-3-95806-442-3

Band / Volume 209

Regulatory interactions between *Corynebacterium glutamicum* and its prophages

M. Hünnefeld (2019), IV, 209 pp

ISBN: 978-3-95806-445-4

Band / Volume 210

Quantum Technology

Lecture Notes of the 51st IFF Spring School 2020

23 March – 03 April 2020, Jülich, Germany

ed. by H. Bluhm, T. Calarco, D. DiVincenzo (2020), ca. 700 pp

ISBN: 978-3-95806-449-2

Weitere **Schriften des Verlags im Forschungszentrum Jülich** unter
<http://wwwzb1.fz-juelich.de/verlagextern1/index.asp>

Schlüsseltechnologien / Key Technologies
Band / Volume 210
ISBN 978-3-95806-449-2