

Czech Technical University in Prague
Faculty of Electrical Engineering
Department of Computer Science and Engineering



Quantum computing, phase estimation and applications

by

Miroslav Dobšíček

A doctoral thesis submitted to
the Faculty of Electrical Engineering, Czech Technical University in Prague,
in partial fulfilment of the requirements for the degree of Doctor.

Ph.D. Programme: Electrical Engineering and Information Technology
Branch of study: Information Science and Computer Engineering

January 2008

Thesis Supervisor:

Josef Kolář
Department of Computer Science and Engineering
Faculty of Electrical Engineering
Czech Technical University in Prague
Karlovo náměstí 13
121 35 Prague 2
Czech Republic

Thesis Co-Supervisor:

Róbert Lórencz
Department of Computer Science and Engineering
Faculty of Electrical Engineering
Czech Technical University in Prague
Karlovo náměstí 13
121 35 Prague 2
Czech Republic

Abstract

Recently, the field of unconventional computing has witnessed a huge research effort to solve the problem of the assumed power of computers operating purely according to the laws of quantum physics. Quantum computing can be seen as a special intermediate case between digital and real analog computing. Importantly, there is a threshold theorem for error correction, as opposed to the pure analog case. Alternatively, quantum computing can be viewed as generalized probabilistic computing, where non-negative real probabilities are replaced with complex amplitudes. The main new resources are quantum mechanical phenomena such as state superposition, interference and entanglement. Superposition together with interference provide a special kind of parallelism, while entanglement, especially when spatially shared, supports unique means of communication.

One of the most important theoretical result is a proof by Bernstein and Vazirani (1993) that there is an oracle relative to which there is a language that can be efficiently accepted by a quantum Turing machine, but cannot be efficiently accepted by a bounded-error probabilistic Turing machine. The problem which was considered is called Recursive Fourier Sampling and the proposed quantum algorithm gives a quasipolynomial speedup, $O(n)$ versus $O(n^{\log n})$. Next, Abrams and Lloyd showed that a quantum computer can efficiently simulate many-body quantum systems having a local Hamiltonian. An additional (informal) evidence of the assumed power of a quantum computer is a bounded-error quantum polynomial time algorithm for large integer factoring. The Abrams-Lloyd algorithm is potentially the most useful quantum algorithm known so far, and if a quantum computer is ever built, it will revolutionize quantum chemical calculations. Thus there is a growing consensus regarding investments into the experimental quantum computing.

In this thesis, attention is paid to small experimental testbed applications with respect to the quantum phase estimation algorithm, the core approach for finding energy eigenvalues. An iterative scheme for quantum phase estimation (IPEA) is derived from the Kitaev phase estimation, a study of robustness of the IPEA utilized as a few-qubit testbed application is performed, and an improved protocol for phase reference alignment is presented. Additionally, a short overview of quantum cryptography is given, with a particular focus on quantum steganography and authentication.

The work was supported by SSF Nanodev Consortium, by the European Commission through the IST-015708 EuroSQIP project and by the Czech Technical University through the grant CTU0507213.

Keywords: quantum computing, quantum phase estimation, iterative quantum protocols, quantum cryptography, quantum steganography

Publications of the author

Parts of the thesis have been presented in the following journal papers and presentations:

Journal papers

- [A.1] Miroslav Dobšíček, Göran Johansson, Vitaly Shumeiko, and Göran Wendin. Arbitrary accuracy iterative quantum phase estimation algorithm using a single ancillary qubit: A two-qubit benchmark. *Physical Review A*, 76:030306(R), 2007.
- [A.2] Miroslav Dobšíček. Simulation on Quantum Authentication. *Physics of Particles and Nuclei Letters*, 4 (2):158-161, 2007.

Presentation list

- [A.3] Göran Johansson, Lars Tornberg, Miroslav Dobšíček, Margareta Wallquist, Vitaly Shumeiko, and Göran Wendin. Superconducting Qubits and Few Qubit Experiments. Invited talk. In *International Conference on Quantum Information Processing and Communication*, 15-19th of October 2007, Barcelona, Spain, 2007.
- [A.4] Miroslav Dobšíček, Göran Johansson, Vitaly Shumeiko, and Göran Wendin. Arbitrary accuracy iterative phase estimation algorithm as a two qubit benchmark. Contributed poster. In *7th QIPC Cluster Meeting*, 13-14th of October 2006, London, UK, 2006.
- [A.5] Miroslav Dobšíček, Josef Kolář, and Róbert Lórencz. A theoretic-framework for quantum steganography. Contributed poster. In *CTU Workshop 2006*, Prague, Czech Republic, 2006.
- [A.6] Miroslav Dobšíček. Commercial products for quantum cryptography (in Czech). Contributed talk. In *Cryptofest 2005*, Prague, Czech Republic, 2005.

Other publications

- [A.7] Miroslav Dobšíček. Extended steganographic system. Contributed poster. In *8th International Student Conference on Electrical Engineering*, Prague, Czech Republic, 2004.

- [A.8] Miroslav Dobšíček and Radim Ballner. Linux - security and exploits (in Czech). Book. Kopp Publishing, České Budějovice, 2004.

Citations

- Paper [A.1] has been cited in:
 1. Ignacio García-Mata and Dima L. Shepelyansky. Quantum phase estimation algorithm in presence of static imperfections. *The European Physical Journal D*, 47:151-156, 2008.
 2. Liu Xiu-Mei, Luo Jun, and Sun Xian-Ping. Experimental realization of arbitrary accuracy iterative phase estimation algorithms on ensemble quantum computers. *Chinese Physics Letters*, 24:3316–3319, 2007.
 3. Péter Varga and Barnabás Apagyi. Phase estimation procedure to solve quantum-mechanical eigenvalue problems. *Physical Review A*, 78:022337, 2008.

- Paper [A.7] has been cited in:
 1. Nameer N. El.Emam. Hiding a large amount of data with high security using steganography algorithm. *Journal of Computer Science*, 3 (4):223–232, 2007.
 2. Nameer N. El.Emam. Embedding a large amount of information using high secure neural based steganography algorithm. *International Journal of Signal Processing*, 4 (1):95–106, 2007.

Acknowledgements

Many people have contributed in one or another way to my PhD thesis. I would like to express my gratitude to all of them.

First of all, I want to give special thanks to my PhD supervisor Josef Kolář. Without his constant help, advises and support I would never be able to finish my thesis. Also, he provided me with numerous opportunities for travelling, meeting new people and professional advances. Thank you a lot. Many thanks goes to my co-supervisor Róbert Lórencz. I am very thankful to him for introducing me to his group of Applied Numerics and Cryptography, and for being such a supportive colleague and friend. It was thanks to him I talked with a top class cryptographer such as Tomáš Rosa and meet new friends Jiří Buček, Tomáš Zahradnický and Tomáš Brabec in the group.

In the same way, I want to deeply thank Vitaly Shumeiko and Göran Wendin from the Chalmers University of Technology in Sweden, where I spent almost a half of my PhD studies. Vitaly and Göran gave me a great opportunity to work with them in the Laboratory of Applied Quantum Physics. The laboratory provides a very flexible environment and the research performed there, regarding superconducting quantum computing, is one of the best in the world. It was a great honor, pleasure and fun to be a member of the lab. I gratefully acknowledge a financial support from the laboratory.

I am especially thankful to Göran Johansson for helping me with my research. His teaching, personal passion for science and easy going attitude have been the main reasons for my achievements during my stay at the lab. Thanks Göran! Also, I want to thank Mikael Fogelström, Tomas Löfwander, Margareta Wallquist, Jonas Sköldberg, Lars Tornberg, Jens Michelsen and Cecilia Holmquist for providing a very friendly environment. I had a pleasure to share a room with Cissi and Jens. I enjoyed it very much and it was fun. Additionally, I thank Lars for taking me out bouldering on the wonderful island Hönö.

I wish to thank Štěpán Holub from the Charles University in Prague who introduced me to the field of quantum computing, Antoni Lozano from Universitat Politècnica de Catalunya in Barcelona who supervised me during my Erasmus stay at Barcelona and become a friend, Čestmír Burdík, Miloslav Znojil and Vladimir Gerdt who I had stimulating discussions with during my short stay at the Joint Institute for Nuclear Research in Dubna, Russia.

On the private side, I want to thank the friends I have met in Barcelona. Thanks Luisa and Juan for letting my live in your apartment and teaching me Spanish, thanks Fabien for being a great friend and bringing me back to computer hacking, thanks Roxane for your smile and designing a wedding invitation for me and my girlfriend Hanka. I am going to ask you for this favor soon :-)) Thanks Anna for being at Barcelona at the same time I was. It was fun.

Very special thanks go to the small Czech community in Göteborg. Mainly to Petr, Dominik, Sandro, Vojta, Jana B., Filip with Jana M. and their kids, Honza, Tomáš, Zdeněk and Lenka. Additionally, to my friends in Prague who I regretfully now meet quite rarely. I am greeting Jiríček with Karolínka, Robi with Milča, Radek, Zornička, Radim, Mates, Canibal, Slunce, Helča and Polárnici. You all were and still are an important part of my life. Thank you for being my friends.

Finally, my greatest thanks to the people who are my family, my parents, my brother and sister and my girlfriend Hanka. I am very much indebted for your love and great support. Life is fine as you all are here.

Contents

| | | |
|----------|---|-----------|
| 1 | Introduction | 1 |
| 1.1 | Computation and information processing | 1 |
| 1.2 | Motivation and contribution of the thesis | 7 |
| 1.3 | Organization of the thesis | 8 |
| 2 | Basic definitions and postulates | 9 |
| 2.1 | Turing machines | 9 |
| 2.1.1 | Probabilistic Turing machines | 11 |
| 2.1.2 | Nondeterministic Turing machines | 13 |
| 2.2 | Complexity classes | 14 |
| 2.3 | Linear algebra | 15 |
| 2.3.1 | Hilbert space | 16 |
| 2.3.2 | Linear operators | 18 |
| 2.3.3 | Operator functions | 19 |
| 2.4 | The postulates of quantum mechanics | 20 |
| 2.5 | Density operator formalism | 22 |
| 3 | Quantum computing | 27 |
| 3.1 | Quantum computing | 27 |
| 3.1.1 | Quantum states | 27 |
| 3.1.2 | Measurement | 29 |
| 3.1.3 | Evolution of quantum states | 31 |
| 3.2 | Quantum Turing machines | 36 |
| 3.3 | Quantum circuit model | 40 |
| 3.3.1 | Single qubit gates | 41 |
| 3.3.2 | Two-qubit gates | 42 |
| 3.3.3 | Evolutionary universal set of gates | 46 |

| | | |
|----------|--|-----------|
| 4 | Phase estimation and applications | 51 |
| 4.1 | Fast quantum algorithms | 51 |
| 4.1.1 | Algorithm solving the abelian hidden subgroup problem | 52 |
| 4.1.2 | New quantum algorithms | 53 |
| 4.2 | Quantum Fourier transform based phase estimation | 54 |
| 4.2.1 | Quantum phase estimation | 55 |
| 4.2.2 | Advanced scheme for phase estimation | 56 |
| 4.2.3 | Efficient circuit for the quantum Fourier transform | 60 |
| 4.2.4 | Linear property of the phase estimation algorithm | 62 |
| 4.2.5 | Representative instances of the phase estimation problem | 64 |
| 4.2.5.1 | Factoring casted as a phase estimation problem | 64 |
| 4.2.5.2 | Finding the energy eigenvalues of a local Hamiltonian | 66 |
| 4.3 | Iterative phase estimation algorithm - IPEA | 69 |
| 4.3.1 | Motivation | 69 |
| 4.3.2 | IPEA derived from the Kitaev phase estimation | 72 |
| 4.3.3 | QFT-based PEA versus IPEA | 75 |
| 4.4 | IPEA applications | 77 |
| 4.4.1 | Iterative phase estimation: A two-qubit test-bed application | 77 |
| 4.4.2 | Multiround protocols for reference frame alignment | 85 |
| 4.5 | Quantum phase estimation conclusions | 89 |
| 5 | Quantum cryptography | 91 |
| 5.1 | Qubit authentication with a quantum key of minimum length | 94 |
| 6 | Conclusions | 99 |

List of Figures

| | | |
|------|---|----|
| 1.1 | Key points in the structure of decision problems | 5 |
| 2.1 | Turing machine action table | 10 |
| 2.2 | Stochastic transition matrix | 12 |
| 2.3 | Restricted set of probabilities | 13 |
| 3.1 | Graphical scheme of a quantum algorithm | 31 |
| 3.2 | Hypothetical cloning matrix | 33 |
| 3.3 | Amplitude amplification by inversion about mean | 34 |
| 3.4 | The power of destructive interference | 37 |
| 3.5 | The circuit model notation | 40 |
| 3.6 | The Bloch sphere | 42 |
| 3.7 | Universal gates for quantum and classical reversible computing | 43 |
| 3.8 | Controlled-U operation | 44 |
| 3.9 | Interpreting who is controlling who | 44 |
| 3.10 | Implementation of the SWAP gate | 46 |
| 4.1 | Eigenvalue kick-back circuit | 55 |
| 4.2 | Schematic diagram of the eigenvalue kick-back circuit | 56 |
| 4.3 | Advanced setup for the phase estimation | 56 |
| 4.4 | The PEA output probability distribution. | 58 |
| 4.5 | Two closest estimates of the phase | 59 |
| 4.6 | Success probability of observing a good estimator | 59 |
| 4.7 | Efficient circuit for the quantum Fourier transform over \mathbb{Z}_{2^m} | 62 |
| 4.8 | Detailed scheme for advanced phase estimation | 63 |
| 4.9 | Modular exponentiation by modular squaring | 66 |
| 4.10 | The phase estimation scheme of Abrams and Lloyd | 67 |
| 4.11 | The Aspuru-Guzik phase estimation scheme | 69 |

| | | |
|------|---|----|
| 4.12 | Kitaev's phase estimation circuit | 70 |
| 4.13 | A plot of outcome probabilities in the Kitaev PEA | 71 |
| 4.14 | Weighted interval intersection method | 72 |
| 4.15 | Additional iteration in the Kitaev PEA | 73 |
| 4.16 | Iterative phase estimation scheme (IPEA) | 74 |
| 4.17 | Bitwise error probabilities in the IPEA | 75 |
| 4.18 | The IPEA and an input eigenstate | 76 |
| 4.19 | Controlled powers of the simple Z-rotation operator | 79 |
| 4.20 | Controlled-NOT implementation using the XX coupling | 79 |
| 4.21 | Benchmark circuit I | 80 |
| 4.22 | Results of simulations I | 81 |
| 4.23 | Benchmark circuit II | 82 |
| 4.24 | Results of simulations II | 82 |
| 4.25 | Benchmark circuit III | 83 |
| 4.26 | Results of simulations III | 84 |
| 4.27 | The problem of phase reference alignment | 86 |
| 4.28 | The Grover amplification viewed as a mirroring | 87 |
| 4.29 | The Burgh and Bartlett protocol for phase reference alignment | 87 |
| 4.30 | Improved protocol for phase reference alignment | 87 |
| 5.1 | A simple scheme for error correction | 92 |
| 5.2 | Quantum steganography on the error correction level | 92 |
| 5.3 | A protocol for qubit authentication | 95 |

Chapter 1

Introduction

1.1 Computation and information processing

Computation and information processing in general always played an important role in human society. We can find mechanical devices performing computation in ancient astronomy, the Greek Antikythera mechanism (150 BCE) is the oldest known mechanical analog computer, we can see basic information processing such as storing data in a written form on the walls of Egyptian pyramids. Somewhat more advanced information processing in a modern sense is historically connected to solving equations. Arabic mathematician Muhammad ben Musa al-Khwarizmi (825 CE) seems to be the first who used a concept of repeatedly applying a set of reduction rules in order to reduce the complexity of the equation at hand. The reduction (al-jabr) later gave name to the algebra and the concept itself is now called an algorithm in honor of al-Khwarizmi.

Later on, advances in math and technology gave birth to the question of which problems/functions in principle can be calculated, and to which extent can a mechanical device (a machine) support an automatic problem reduction according to a given algorithm. These questions date back to Gottfried Leibnitz (1646 - 1716) who after having successfully constructed a mechanical calculating device dreamt of a formal language for a machine that could manipulate symbols and determine the truth values of mathematical statements.

In 1936, Alan Turing developed an abstract notation of a programmable universal computer, a Turing machine, in order to address Hilbert's Entscheidungsproblem (German for 'decision problem'). Using this abstraction, he was able to prove that there is no mechanical procedure which can be used to decide arbitrary statements in mathematics. In particular, Turing reduced the halting problem to the Entscheidungsproblem. The term 'mechanical procedure' does not directly refer to digital or analog implementation, however, classical mechanics was assumed implicitly.

Also Turing's result directly spawns a short proof of Gödel Incompleteness Theorem, which is otherwise very long. In modern terms, the original Gödel's proof (1931) consists mainly of defining a sort of computational model (a Gödel numbering) and the work of Turing is heavily influenced by that. A few months before Turing's result, Alonzo Church (adviser of Turing) finished his work on lambda calculus and proved that any function of positive integers can be mechanically computed only if it is recursive. Soon it turned out that lambda calculus and

Turing machine abstraction are equivalent (so are all new non-hypercomputation proposals) and the authors formulated the Church-Turing hypothesis:

'Any function "naturally to be regarded as computable" is computable by a Turing machine'.

The Church-Turing hypothesis cannot be mathematically proven and therefore it is sometimes proposed as a definition or as a physical law using a modern reformulation by David Deutsch (1985):

'Every finitely realizable physical system can be perfectly simulated by a universal model computing machine operating by finite means'.

A few years later, after Church and Turing's theoretical breakthrough, the use of digital electronics largely invented by Claude Shannon and the stored program architecture proposed by John von Neumann gave birth to the first real-life computing devices with all the capabilities of a Turing machine. The discovery of the transistor led to a rapid progress in technology and the digital computer era started.

It may be surprising that while previous computing devices were analog ones (based on mechanical or hydraulic principles), electronic analog computers were beaten by electronic digital computers. The reasons are purely technical. Practical computing with analog electronics is limited by the range over which the variables might vary, and, moreover, there is no theorem saying that noise can ever be efficiently suppressed. Today, electronic analog computing is still studied in the field of artificial neural networks with limited precision real weights. The main point is that some problems map more naturally to an analog computer than to a digital computer and therefore the computation might be faster due to reduced mapping costs.

It is worth to mention here, that an analog computer does not implicitly assume a scale-independent continuous phenomenon for a practical computation. In fact, a hypothetical device using real numbers with infinite precision which are harnessable for computation would allow for non-Turing (non-recursive) functions to be computed. This is called a hypercomputation. A hypercomputer would not only violate the Church-Turing hypothesis, but it would need to exploit a real continuous phenomenon, which we know is largely ruled out by quantum physics. The difference between digital computing and real analog computing might also be seen as the difference between computable numbers and algebraic systems vs. real numbers and differential systems.

On the theoretical ground, it was found that it is valuable to classify algorithms and their complexity according to time and space demands with respect to the size of the input. Several prominent complexity classes refer to problems which are solvable by a Turing machine in polynomial time, P , problems solvable in polynomial space, $PSPACE$, problems solvable in exponential time, EXP , and problems for which, if the answer is 'yes', then there is a polynomial time proof, NP . It follows that $P \subseteq NP \subseteq PSPACE \subseteq EXP$.

The empirical practice reveals that the vast majority of important problems with polynomial time/space demands have their complexity bounded by polynomials of low degree, while

problems with exponential time demands have the base not close to one. It seems there is a sharp difference among problems contained in NP , fast efficiently solvable problems on one hand and practically intractable problems on the other one. The contrast of $O(n^2)$ and $O(2^n)$ for $n > 100$ is overwhelming and therefore the structure of the NP class is of utmost importance.

In the early 1970's, Cook, Karp and Levin proved an existence of NP -complete problems, a subset of problems in NP , that are at least as difficult as any other problems in NP , and they actually showed that the decision version of many practical problems such as the Boolean satisfiability problem, the Traveling salesperson problem or the Clique problem are NP -complete. So far, we have no indication that NP -complete problems can be solved in polynomial time instead of exponential time on a Turing machine, and yet we are not able to prove or disprove the relation $P \neq NP$. Currently, the $P \neq NP$ relation is one of the six Millennium Prize Problems established by the Clay Mathematics Institute.

Clearly, we know problems contained in NP for which we did not find a polynomial runtime algorithm nor we proved them to be NP -complete. This is partly due to our lack of knowledge, and, moreover, if $P \neq NP$, then the Ladner's Theorem (1975) says that there must be intermediate problems between P and NP -complete. Next to several artificial problems developed just for the purpose of being NP -intermediate, natural problems such as primality testing, factoring of large integers and graph isomorphism were on the list together until recently.

Probabilistic computing

It was always considered that relaxing from deterministic computing to probabilistic computing might have a positive impact on problems solving. This expectation is connected to relatively fast converging iterative processes which when supplied with a 'right' guess do produce the correct answer most of the time. The right guess is given by empirical sampling over similar cases. In this view, the art of designing a probabilistic algorithm consists of replacing a right guess by a random guess while keeping the convergence of the process. Probabilistic algorithms that always return the correct answer within an expected (finite) runtime is called a Las Vegas simulation. If the expected runtime is bounded by a polynomial we talk about a zero-error probabilistic polynomial-time problems, class ZPP . With a modification to stop after a certain number of steps and output an approximate (possibly false) result we get a Monte Carlo simulation. If the error probability can be bounded and polynomial-runtime is kept the corresponding class is BPP . The most general class of probabilistic polynomial time problems is PP . It follows that $P \subseteq ZPP \subseteq BPP \subseteq PP \subseteq PSPACE$.

Advances in complexity theory and probabilistic algorithms led to so called Extended or Strong Church-Turing hypothesis due to Bernstein and Vazirani (1993):

'Any "reasonable" model of computation can be efficiently simulated on a probabilistic Turing machine'

The message of this hypothesis is three-fold. First, it is practically oriented. It refers to reasonable models of computation in order to stand aside from hypothetical hypercomputers

and classical extremely massive parallel systems where the space and/or energy requirements grow exponentially with the size of the input. An example of such model is the original Adleman DNA computer proposal (1994). Secondly, it allows a Turing machine to access a random number generator. This expresses the assumption that probabilistic algorithms might be stronger than deterministic ones. This assumption was supported mainly by a very fast probabilistic primality testing while deterministic testing needed exponential time. Third, the word 'efficiently' means that all reasonable computational devices regardless of their physical implementation are equivalent up to polynomial-time reductions. In other words, a problem classification is independent of machines while it is true that an architecture reflecting the inner structure of the problem may yield a polynomial speed-up compared to other architectures.

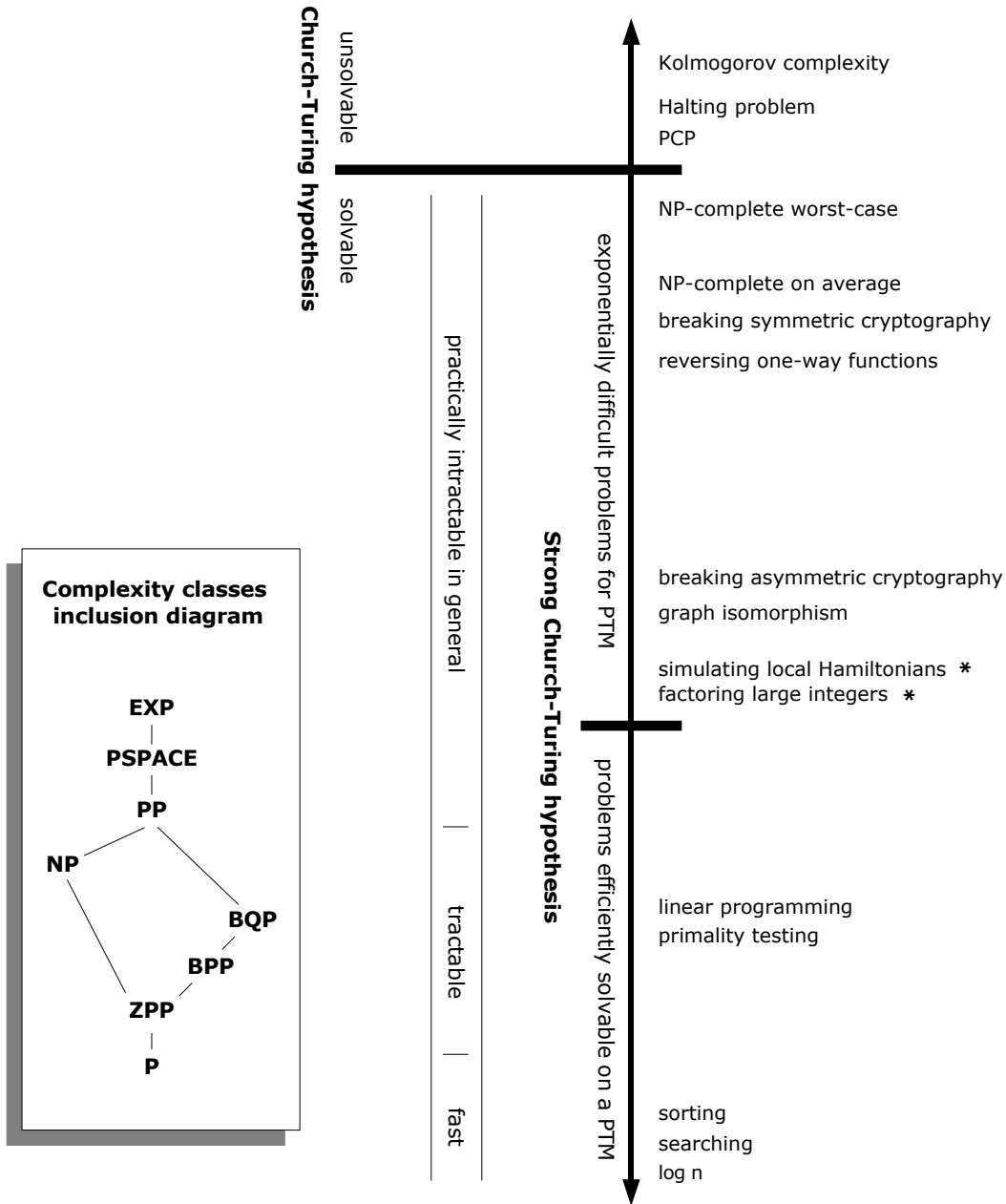
One puzzling problem is the question of where the power of probabilistic algorithms comes from. Practice reveals that pseudo-random number generators are sufficient for most probabilistic algorithms in order to be successful. Additionally, classical mechanics which is still the main framework for computational theory does not allow for true random number generators. In fact, only practically unpredictable numbers can be generated. For this reason it is now believed that P might be equal to BPP .

The assumption $P = BPP$ started a study branch called 'derandomization'. The random essence is at first reduced to a one call of a random number generator, and then replaced by a deterministic process. In 2002, one remarkable success of derandomization was a deterministic polynomial-time primality testing algorithm (the AKS test) developed by Agrawal, Kayal and Saxena. Primality testing nicely reflects advances in theoretical computer science. In 1975, the test was shown to be in NP , later in 1992, it fell to ZPP , and in 2002, it went to P . The AKS primality test is however much slower than the probabilistic one and therefore it is used only as a part of mathematical proofs.

This approach reflects problem solving in practice. Problem classification alone is sometime not the right figure of merit for its practical tractability. It only considers worst-case times, i.e. it gives us information that the problem never grows faster than a certain upper bound. In fact, relaxation to average cases (application specific) can reveal nice inner structure which allows for practically fast heuristics, random sampling, evolution techniques and so on. These methods are often of advantage in general.

Some P and NP -intermediate problems let alone NP -complete ones, do not show enough structure which can be exploited for practical speed-up on average. On the other hand they seem to naturally map on unconventional computing devices such as biological, chemical and quantum systems. Actually, these problems are usually problems from their respective fields. This mirrors the fact that information processing is always a physical process and some processes in their natural form are faster compared to others. In the future, we might end up with a computer equipped with several co-processors which operate by different means than digital electronics. Another tantalizing prospect related to unconventional computers is that they might exhibit even faster than polynomial speed-up for specific problems and thus be a threat to the Strong Church-Turing hypothesis. The most serious candidate that appears not to obey the Strong Church-Turing hypothesis is a yet to be built quantum computer.

Figure 1.1 summarizes currently known key points in the structure of decision problems and indicates the assumed power of quantum computers. PTM stands for a probabilistic Turing machine.



* efficiently solvable on a hypothetical quantum computer

Figure 1.1: Currently known key points in the structure of decision problems.

Quantum computing

Classical information processing is developed in the framework of classical mechanics. Since classical physics was superseded by quantum physics in many ways, it is natural to explore the potential of information processing in the framework of quantum mechanics. In 1982, Richard P. Feynman conjectured that quantum physics in general can be simulated on a probabilistic Turing machine only with an exponential slowdown and speculated about a universal quantum mechanical system capable of an efficient simulation. At about the same time, it was proved by Bennett, Fredkin and Toffoli that a universal computation can be done reversibly. This was an important step since processes in quantum physics are reversible. An abstract model for quantum computing, a quantum Turing machine, was created by David Deutsch (1985) as a tool for constructing a test of the many-world interpretation of quantum physics. Quaintly enough, the test now hinges on the question whether quantum information processing allows for strong artificial intelligence. The resemblance to the work of Turing is striking. They both developed an abstract toolkit with a great theoretical value on its own for further reasoning, and the reasoning at the end concerns artificial intelligence. Theoretical quantum information processing seems to have good prospects for deeper understanding of connections between computation, nature and life.

States in a classical computer are represented as strings from a finite alphabet, typically binary strings, while states of a quantum computer are represented by unit vectors in a finite dimensional complex Hilbert space. Complex amplitudes of these vectors are allowed to change continuously, but they cannot be directly observed. Therefore quantum computing can be seen as a special intermediate case between digital and real analog computing. Importantly, there is a threshold theorem for quantum error correction, as opposed to the pure analog case. Alternatively, quantum computing can be seen as generalized probabilistic computing, where non-negative real probabilities are replaced with complex amplitudes. The main new resources are quantum mechanical phenomena such as state superposition, interference and entanglement. Superposition together with interference provide a special kind of parallelism, while entanglement, especially when spatially shared, supports unique means of communication.

The first evidence (promise problems) of the assumed quantum computer power was given by D. Deutsch (1985), Deutsch and Jozsa (1992), Bernstein and Vazirani (1993), and D. Simon (1994). Although these promise problems concerning global properties of functions were somewhat artificial, the corresponding quantum circuits composed from a set of universal gates gave an exponential speed-up over the classical deterministic solutions. In 1994, Peter W. Shor published a bounded-error quantum polynomial time algorithm for large integer factoring and discrete logarithm calculations. The corresponding complexity class is called *BQP*. The factoring problem is believed to be classically intractable and currently the RSA public-key cryptosystem is based on its intractability. Later in 1999, Abrams and Lloyd described a *BQP* algorithm for finding energy eigenvalues and eigenvectors of quantum physical systems governed by a local Hamiltonian. Again, this is exponentially faster than many *ab initio* methods now used in physics and chemistry problems where local Hamiltonian takes place. It turned out that all of the above mentioned problems are instances of a hidden subgroup problem (HSP) over Abelian groups, and the problem-solution approach can be explained in terms of the quantum phase estimation algorithm (PEA). A key ingredient in this algorithm is a quantum Fourier transform (QFT) for which we actually know a quadratic size circuit

in the case of Abelian groups. Efficient solution for the HSP over non-Abelian groups is an open problem. It is known that an efficient circuit for the HSP over the symmetric group would lead to an efficient algorithm for the graph isomorphism problem. The HSP over the dihedral group is in the same manner connected to the shortest vector in a lattice problem. Recently, there has been progress in solving certain non-Abelian HSP utilizing transforms beyond QFT. A discovery of an efficient circuit for the Clebsch-Gordan transform over the Heisenberg group led to an efficient solution for the HSP over the Heisenberg group. The Clebsch-Gordan transform also led to a discovery of an efficient circuit for the Schur transform which might possibly open up for new quantum algorithms.

The question whether a quantum computer can solve efficiently NP -complete problems remains to be answered as well as for a classical computer. In 1996, L. Grover presented a sort of no-go for quantum exponential speed-up for completely unstructured problems. He developed an algorithm for unsorted database search which is provably optimal and yields only a square root speed-up. In particular, a desired item from an unsorted database of size N can be searched in time $O(\sqrt{N})$.

From the practical realization point of view, there is no known viable technology for a scalable quantum computer yet. Experiments have been performed with up to ten quantum bits. Problems are caused mainly due to environmental noise. Other problems arise from entanglement which makes the border line between a central processing unit and memory not that sharp as it is in a classical computer. One possible approach toward solving this problem is a distributed quantum computer. Here, relatively easy to stabilize small cores are interconnected with quantum and classical channels, and by exploiting quantum state teleportation and non-local control they cooperate on the problem solving.

Besides quantum computing, there is also a discipline called quantum cryptography. The protocol BB84 developed by Bennett and Brassard in 1984 solves the problem of secret-key distribution in an unconditionally secure way. This is assured for free by the quantum physics laws themselves. The first commercial products with the BB84 are already available. They work as point-to-point systems over a fiber optic line, up to 100km of length, at rates approximately 100kbits/sec. Another commercially available product in the form of a standard PCI card or an USB device is a quantum random number generator. The bit rate is over 16Mbits/sec. In general, special properties of a quantum physical system such as the no-cloning property of an unknown quantum state and the randomness in a wave function collapse allow for many quantum enhanced cryptographic primitives. These primitives range from authentication to covered communication (steganography) and more.

1.2 Motivation and contribution of the thesis

The Abrams-Lloyd algorithm for finding energy eigenvalues is potentially the most useful quantum algorithm known so far, and if a quantum computer is ever built, it will revolutionize quantum chemical calculations. Thus there is a growing consensus regarding investments into the experimental quantum computing.

In this thesis, attention is paid to small experimental testbed applications with respect to the quantum phase estimation algorithm, the core approach for finding energy eigenvalues. An iterative scheme for quantum phase estimation (IPEA) is derived from the Kitaev phase

estimation, a study of robustness of the IPEA utilized as a few-qubit testbed application is performed, and an improved protocol for phase reference alignment is presented. The derivation of the IPEA involves incorporating a classical postprocessing algorithm into a quantum circuit. The resulting quantum circuit then becomes more efficient. Such a result may be quite inspiring in the further development of quantum algorithms as the usual approach is to perform as much as possible of necessary calculations classically, since quantum resources are considered 'expensive'.

Additionally, a short overview of a quantum cryptography is given, with a particular focus on quantum steganography and authentication.

1.3 Organization of the thesis

The outline of this thesis is as follows. Chapter 2 defines the terms used in the thesis. A rather large part of this chapter pertains to Turing machines and portions of linear algebra. Chapter 3 provides a summary on quantum computing and introduces the quantum Turing machine and quantum circuit model. A special attention is paid to the thin borderline between quantum computing and classical computing. Chapter 4 considers fast quantum algorithms, quantum phase estimation and its iterative variants, and a few testbed circuits. This is the main chapter of the thesis. Chapter 5 contains the overview on quantum cryptography. Finally, Chapter 6 consists of conclusions and the direction for future work.

Chapter 2

Basic definitions and postulates

2.1 Turing machines

A Turing machine is a symbol-manipulating device equipped with a read/write head and possibly infinite cell-divided tape. The head reads the current symbol on the tape, changes its internal state accordingly, writes an output symbol if any, and then moves to the neighboring left or right cell. A computation consists of periodical repeating of these steps. At any moment, the computation is completely described by a Turing machine configuration.

Definition 2.1.1 (Turing machine configuration). A configuration of a Turing machine is an ordered triple:

- the contents of the tape,
- the current state,
- the position of the head. □

Despite the simplicity, the Church-Turing hypothesis says that a Turing machine can simulate the logic of any computer that could be possibly constructed. A Turing machine is said to be universal if it is able to simulate any other Turing machine. The smallest universal Turing machine has three symbols and two internal states [1]. No smaller universal Turing machine is possible. Other commonly used computational abstractions are a cellular automata and a gate model. Gates sets such as {NAND, fanout}, {NOR, fanout}, {AND, NOT} constitute universal sets. The Toffoli gate alone is a universal gate for reversible computing. Formally, a Turing machine is defined as follows.

Definition 2.1.2 (Turing machine). A Turing machine is an ordered six-tuple $M = (Q, \Sigma, b, q_0, F, \delta)$, where

1. Q is a *finite* set of states,
2. Σ is a finite set of tape symbols (alphabet),
3. $b \in \Sigma$ is the blank symbol,

4. $q_0 \in Q$ is the initial state,
5. $F \subseteq Q$ is the set of final or accepting states,
6. $\delta : \Sigma \times Q \rightarrow \Sigma \times Q \times \{L, R\}$ is a transition function, where L is shift left, R is shift right.

The transition function

$$\delta : (\text{current symbol, current state}) \longrightarrow (\text{output symbol, new state, shift left/right})$$

is usually given as an action table or a diagram. Sometimes it is written as

$$\delta : (\text{current symbol, current state, output symbol, new state, shift left/right}) \longrightarrow \{0, 1\}$$

in order to denote allowed transitions among configurations.

Example 1 (Turing machine action table). A simple action table for a machine

$$M = (\{q_0, q_1, q_2, q_f\}, \{'0', '1', b\}, b, q_0, \{q_f\}, \delta)$$

calculating the XOR function over input bits a and b is shown in the Figure 2.1. The result is stored at the place of the second input bit.

| a | b | XOR |
|-----|-----|-----|
| 0 | 0 | 0 |
| 0 | 1 | 1 |
| 1 | 0 | 1 |
| 1 | 1 | 0 |

| δ | '0' | '1' |
|----------|------------------|------------------|
| q_0 | ('0', q_1, L) | ('0', q_2, L) |
| q_1 | ('0', q_f, L) | ('1', q_f, L) |
| q_2 | ('1', q_f, L) | ('0', q_f, L) |

Figure 2.1: The XOR function and corresponding Turing machine action table.

The transition function δ makes a given machine M to behave like a computer with a fixed program. In this sense, only one function can be calculated. However, since the action table itself can be encoded as a string, we can say that a part of the tape is considered to be a program description and the rest is the input:

$$\underbrace{'0101011 \dots 1101 \dots 010'}_{\text{program}} \underbrace{\xrightarrow{\delta}}_{\text{input data}} \underbrace{'0101011 \dots 0011 \dots 100'}_{\text{program}} \underbrace{\text{output data}}.$$

The Turing machine from the Example 1 can be interpreted as recognizing two programs:

program '0': perform identity on the input bit,
 program '1': negate the input bit.

Due to string manipulation nature, Turing machines (and various automata) are often studied in terms of languages they accept.

Definition 2.1.3 (Accepting a language). A Turing machine accepts a language L iff

- for all strings $x \in L$: the machine halts in a final state,
- for all strings $x \notin L$: the machine halts in a non-final state or the machine enters an infinite loop. \square

The Turing machine from the Example 1 halts in a non-final state for the blank symbol b (no program or input bit is given).

Definition 2.1.4 (Time complexity). Let M be a Turing machine and $f(n)$ a function $f : \mathbb{N} \rightarrow \mathbb{N}$. We say that M has a time complexity $f(n)$ if for each input of size n the machine halts after at most $f(n)$ steps.

Definition 2.1.5 (Space complexity). Let M be a Turing machine and $f(n)$ a function $f : \mathbb{N} \rightarrow \mathbb{N}$. We say that M has a space complexity $f(n)$ if for each input of size n the machine utilizes at most $f(n)$ cells before halting.

2.1.1 Probabilistic Turing machines

Definition 2.1.6 (Probabilistic Turing machine). A probabilistic Turing machine is a generalized Turing machine where the transition function does not output only a single triple (output symbol, new state, shift), but a probabilistic distribution of such triples. Formally, the transition function assigns to each possible transition a non-negative probability:

$$\delta : Q \times \Sigma \times Q \times \Sigma \times \{L, R\} \longrightarrow [0, 1] \quad (2.1)$$

in such a way that the set of transitions from one configuration c_0 to all its direct successor-configurations c_1, \dots, c_k must fulfill the local probability condition

$$\sum_{i=1}^k p_i = 1, \quad (\text{LPC}) \quad (2.2)$$

where p_i is the probability of transition from c_0 to c_i . \square

It may happen that several different paths ℓ of computation lead to the same configuration c_x . In such a case, this configuration has a probability equal to the sum of probabilities of corresponding paths

$$p(c_x) = \sum_{\ell} p_x^{(\ell)}. \quad (2.3)$$

At any moment, the local probability condition assures that probabilities of all distinct configurations sum to one (global probability condition).

$$\sum_{\forall c_i \neq c_j} p_i = 1. \quad (\text{GPC}) \quad (2.4)$$

Definition 2.1.7 (Transition matrix). Let us have a Turing machine M . The transition matrix consists of elements $T_{i,j}$ defined as

$$T_{i,j} = \begin{cases} p_i & \text{if the machine goes from } c_j \text{ to } c_i \text{ in one step with probability } p_i, \\ 0 & \text{otherwise.} \end{cases} \quad (2.5)$$

□

Transition matrix for a probabilistic Turing machine is a stochastic matrix where each column sums to one (LPC). We can say that transition matrix T maps one normalized (GPC) linear combination (superposition) of configurations to another normalized linear combination of configurations. Thus stochastic matrices preserve the 1-norm (Eq. 2.6, page 16). For every column vector of configurations v , $\|T \cdot v\|_1 = \|v\|_1$.

Transitions allowed by stochastic matrices can also be viewed in terms of constructive interference. The constructive interference is to be understood as the possibility for different computational paths to end up in the same configuration and thus increase the total probability of that configuration. Probabilities of other configurations are proportionally smaller. One has to program a probabilistic Turing machine in such a way that the probability of a desirable configuration (correct result) is gradually enlarged, and probabilities of undesirable configurations subsequently diminishes. Destructive interference might allow for more progressive separation of desirable and undesirable configurations, but with a probabilistic Turing machine this is not possible since we have only non-negative probabilities.

Example 2 (Transition matrix and configuration tree). An example of stochastic transition matrix and a configuration tree starting in a configuration c_1 is shown in the Figure 2.2. Local and global probability conditions are manifestly fulfilled at each level of the tree. The computation halts in a configuration c_4 with probability $p(c_4) = 0.1 + 0.4 \cdot 0.4 + 0.4 \cdot 0.6 \cdot 1 = 0.5$ and in a configuration c_5 with probability $p(c_5) = 0.5$.

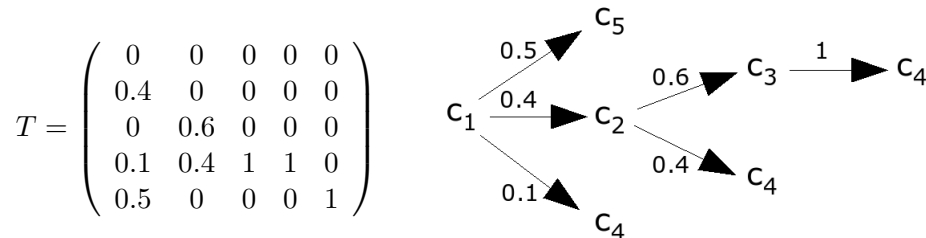


Figure 2.2: An example of stochastic transition matrix and configuration tree with root c_1 .

Definition 2.1.8 (Accepting a language on a probabilistic Turing machine). A probabilistic Turing machine PTM accepts a language L with completeness c and soundness s iff

- for all strings $x \in L$: PTM accepts x with probability $p > c$,
- for all strings $x \notin L$: PTM accepts x with probability $p \leq s$.

□

In order to study the power of a probabilistic Turing machine, we need to restrict real probabilities from the range $[0, 1]$ only to a certain 'universal' set of probabilities. Assumption that all real numbers from the range $[0, 1]$ with infinite precision can be used, would allow one to encode the hard problems right into them. C. H. Papadimitriou in [2] showed that for each probabilistic Turing machine M_1 there exists a probabilistic Turing machine M_2 with

probabilities restricted to the set $\{0, \frac{1}{2}, 1\}$ which simulates M_1 with at most polynomial slowdown. The probability $\frac{1}{2}$ shows a crucial dependence on the fair coin flipping if we assume a probabilistic Turing machine to be more powerful than a deterministic Turing machine.

Example 3 (Probabilistic Turing machine with restricted probabilities). Let M_1 be a probabilistic Turing machine using probabilities 0.4 and 0.6 at some part of the computation. Let M_2 be a probabilistic Turing machine with restricted probabilities $\{0, \frac{1}{2}, 1\}$. In the Figure 2.3, M_2 approximates M_1 with precision $\varepsilon \geq 1/2^3$.

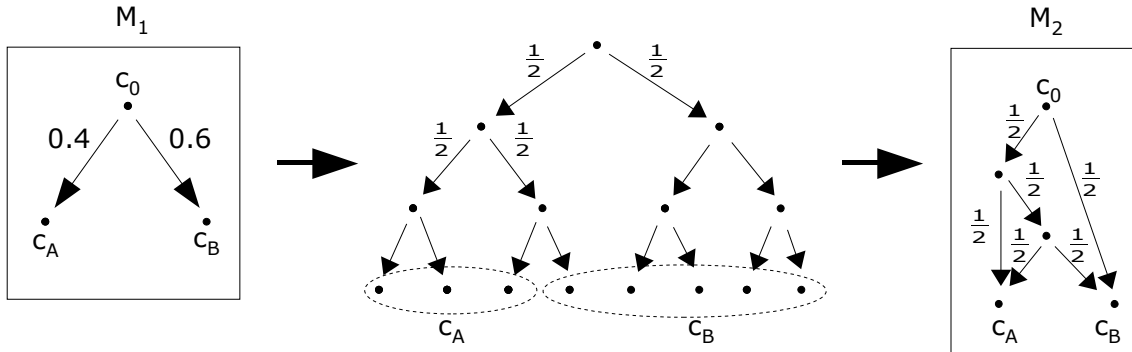


Figure 2.3: Probabilistic Turing machine with restricted probabilities.

2.1.2 Nondeterministic Turing machines

Definition 2.1.9 (Nondeterministic Turing machine). Nondeterministic Turing machine is a hypothetical kind of a probabilistic Turing machine. It explores all the possible computational paths at parallel, and if the solution exists, then with non-zero probability the right path is actually taken. Equivalently, we can say that the machine goes along the right computational path, and if there is a fork, it nondeterministically chooses with non-zero probability the right path again. Nondeterministic Turing Machine NTM accepts a language L iff

- for all $x \in L$: NTM accepts x with probability $p > 0$,
- for all $x \notin L$: NTM accepts x with probability $p = 0$.

□

Modern computers are usually not described as direct instances of a Turing machine. Their architecture is well described by abstract models as 'stored program' Random Access Machine (RAM) and synchronous/asynchronous Parallel Random Access Machine (PRAM). These models simplify algorithm analysis mainly due to grouping tape cells to fixed size 'words', which are uniquely indexed and can be thus directly accessed in a unit time as opposed to a Turing machine.

2.2 Complexity classes

Algorithms and problems are categorized into complexity classes.

Definition 2.2.1 (Polynomial-time, P). P is the class of decision problems deterministically solvable in polynomial time on a Turing machine. A language L is in a class P if

- for all $x \in L$: the machine accepts x with probability $p = 1$,
- for all $x \notin L$: the machine accepts x with probability $p = 0$,

and the time complexity is upper-bounded by some polynomial $poly(x)$.

Definition 2.2.2 (Nondeterministic polynomial-time, NP). NP is the class of decision problems nondeterministically solvable in polynomial time on a Turing machine. A language L is in a class NP if

- for all $x \in L$: the machine accepts x with probability $p > 0$,
- for all $x \notin L$: the machine accepts x with probability $p = 0$,

and the time complexity is upper-bounded by some polynomial $poly(x)$.

Equivalently, NP is the class of decision problems such that, if the answer is 'yes', then there is a deterministic polynomial time proof of this fact.

Definition 2.2.3 (Zero-error probabilistic polynomial-time, ZPP). ZPP is the class of decision problems probabilistically solvable in expected polynomial time on a Turing machine with zero error probability. A language L is in a class ZPP if

- for all $x \in L$: the machine accepts x with probability $p = 1$,
- for all $x \notin L$: the machine accepts x with probability $p = 0$,

and the time complexity is upper-bounded by some polynomial $poly(x)$.

Definition 2.2.4 (Bounded-error probabilistic polynomial-time, BPP). BPP is the class of decision problems probabilistically solvable in polynomial time on a Turing machine with bounded error probability. A language L is in a class BPP if

- for all $x \in L$: the machine accepts x with probability $p > \frac{2}{3}$,
- for all $x \notin L$ the machine accepts x with probability $p \leq \frac{1}{3}$,

and the time complexity is upper-bounded by some polynomial $poly(x)$.

The numbers $\frac{2}{3}$ and $\frac{1}{3}$ may be replaced to any a and b such that $a = \frac{1}{2} + poly^{-1}(x)$ and $b = \frac{1}{2} - poly^{-1}(x)$. The Chernoff bound [3] says that by repeating trials the error probability diminishes exponentially fast and the success probability is accordingly increased.

Definition 2.2.5 (Probabilistic polynomial-time, PP). PP is the class of decision problems probabilistically solvable in polynomial time on a Turing machine but the error probability need not to be bounded by a polynomial. Compared to the BPP class, the numbers a and b are allowed to scale as $\frac{1}{2} \pm \exp^{-1}(\text{poly}(x))$. A language L is in a class PP if

- for all $x \in L$: the machine accepts x with probability $p > 1/2$,
- for all $x \notin L$: the machine accepts x with probability $p \leq 1/2$,

and the time complexity is upper-bounded by some polynomial $\text{poly}(x)$.

Definition 2.2.6 (Polynomial-space, PSPACE). PSPACE is the class of decision problems solvable on a Turing machine whose space complexity is upper-bounded by some polynomial with respect to the size of the input.

Definition 2.2.7 (Exponential-time, EXP). EXP is the class of decision problems deterministically solvable a Turing machine whose time complexity is upper-bounded by an exponential function $\exp(\text{poly}(n))$, where $\text{poly}(n)$ represents a polynomial with respect to the size of the input n . \square

It follows that $P \subseteq ZPP \subseteq NP \subseteq PP \subseteq PSPACE \subseteq EXP$ and $P \subseteq ZPP \subseteq BPP \subseteq PP \subseteq PSPACE \subseteq EXP$. The relation of BPP and NP is unknown. However, it is believed that $P = BPP$ and $P \neq NP$. Detailed proofs and further information can be found in [2].

2.3 Linear algebra

Table 2.1 summarizes the Dirac notation for notions from linear algebra. This 'bra(c)ket' notation is widely used within quantum mechanics. Another formalism based on density matrices will be introduced later.

| Notation | Description |
|----------------------------------|---|
| $ v\rangle$ | Column vector. Known as 'ket'. |
| $\langle v $ | Row vector dual to $ v\rangle$. Known as 'bra'. |
| $\langle \psi v\rangle$ | Inner product of vectors $\langle \psi $ and $ v\rangle$. |
| $ \psi\rangle\langle v $ | Outer product of $ \psi\rangle$ and $\langle v $. |
| $ \psi\rangle \otimes v\rangle$ | Tensor product of $ \psi\rangle$ and $ v\rangle$. |
| $ \psi\rangle v\rangle$ | Abbreviated notation for tensor product of $ \psi\rangle$ and $ v\rangle$. |
| $ \psi, v\rangle$ | Abbreviated notation for tensor product of $ \psi\rangle$ and $ v\rangle$. |
| A^\dagger | Adjoint operator of the A matrix. $A^\dagger = (A^T)^* = (A^*)^T$. |
| $\langle \psi A v\rangle$ | Inner product of $\langle \psi $ and $A v\rangle$. |
| $\ v\ $ | Notation for a vector norm. |

Table 2.1: Dirac notation.

2.3.1 Hilbert space

Definition 2.3.1 (Vector space). A set $\mathcal{V} = (\mathcal{V}, +, \cdot)$ is called a vector space over a scalar field \mathcal{F} iff the operations $+$: $\mathcal{V} \times \mathcal{V} \rightarrow \mathcal{V}$ (vector addition) and \cdot : $\mathcal{F} \times \mathcal{V} \rightarrow \mathcal{V}$ (scalar multiplication) are defined and

1. $(\mathcal{V}, +)$ is a commutative group with a neutral element $\vec{0}$,
and for all $v, \psi \in \mathcal{V}$, $\alpha, \beta \in \mathcal{F}$:
2. $\alpha(\beta|v\rangle) = (\alpha\beta)|v\rangle$,
3. $(\alpha + \beta)|v\rangle = \alpha|v\rangle + \beta|v\rangle$,
4. $\alpha(|v\rangle + |\psi\rangle) = \alpha|v\rangle + \alpha|\psi\rangle$.

Definition 2.3.2 (Vector norm). Let \mathcal{V} be a vector space over a scalar field \mathcal{F} . A vector norm is a function $\|\cdot\| : \mathcal{V} \rightarrow \mathbb{R}^+$ satisfying for all $v, \psi \in \mathcal{V}$, $\eta \in \mathcal{F}$:

1. $\|v\| = 0 \Leftrightarrow v = \vec{0}$,
2. $\|\eta v\| = |\eta| \|v\|$,
3. $\| |\psi\rangle + |v\rangle \| \leq \|\psi\| + \|v\|$. (triangle inequality)

Definition 2.3.3 (p -norm). Let $p \geq 0$ be a real number and $v = (v_1, v_2, \dots, v_n)$ an n -dimensional vector. The p -norm is a vector norm defined as

$$\|v\|_p = \left(\sum_{i=1}^n |v_i|^p \right)^{\frac{1}{p}}. \quad (2.6)$$

For $p = 1$ we get the taxicab norm and for $p = 2$ we get the Euclidean norm.

Definition 2.3.4 (Inner product). Let \mathcal{V} be a vector space over the field of complex numbers. An inner product over the vector space \mathcal{V} is a function $\langle \cdot | \cdot \rangle : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{C}$ satisfying for all $v, \psi, \lambda \in \mathcal{V}$:

1. $\langle v|v\rangle \in \mathcal{R}$, $\langle v|v\rangle \geq 0$, $\langle v|v\rangle = 0 \Leftrightarrow |v\rangle = \vec{0}$,
2. $\langle v|\psi\rangle = \langle \psi|v\rangle^*$,
3. $\langle v|(|\psi\rangle + |\lambda\rangle) = \langle v|\psi\rangle + \langle v|\lambda\rangle$.

The inner product induces the 2-norm $\|v\|_2 = \sqrt{\langle v|v\rangle}$. A vector $|v\rangle$ is normalized or unit-vector under the 2-norm iff $\|v\|_2 = 1$.

Note. For a real or complex vector space \mathcal{V} with an inner product $\langle \cdot | \cdot \rangle$ and norm $\|\cdot\|$ the following inequality holds, $v, \psi \in \mathcal{V}$:

$$|\langle \psi|v\rangle| \leq \|\psi\| \cdot \|v\|. \quad (\text{Schwarz inequality})$$

Definition 2.3.5 (Completeness). Let \mathcal{V} be a vector space with the 2-norm $\|\cdot\|_2$ and $|v_i\rangle \in \mathcal{V}$ a sequence of vectors.

- $|v_i\rangle$ is a Cauchy sequence iff

$$\forall \epsilon > 0 \exists N > 0 \text{ such that } \forall n, m > N : \| |v_n\rangle - |v_m\rangle \|_2 < \epsilon.$$

- $|v_i\rangle$ is convergent iff

$$\exists |v\rangle \in \mathcal{V} \text{ such that } \forall \epsilon > 0 \exists N > 0 \forall n > N : \| |v_n\rangle - |v\rangle \|_2 < \epsilon.$$

Space \mathcal{V} is complete iff every Cauchy sequence converges.

Definition 2.3.6 (Hilbert space). A Hilbert space \mathcal{H} is a vector space with an inner product $\langle \cdot | \cdot \rangle$ that is complete under the induced norm $\|v\|_2 = \sqrt{\langle v | v \rangle}$.

Note. Standard quantum computing is defined only for finite dimensional complex Hilbert spaces, $\mathcal{H} = \mathbb{C}^n$, and normalized (unit) vectors.

Definition 2.3.7 (Kronecker delta function). The Kronecker delta function is a function of two variables (usually integers) defined as

$$\delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{otherwise.} \end{cases} \quad (2.7)$$

Definition 2.3.8 (Basis vectors). An enumerable set of normalized vectors $|e_i\rangle$ forms an orthonormal basis of a Hilbert space \mathcal{H} iff

1. $\forall i, j : \langle e_i | e_j \rangle = \delta_{ij}$,
2. any vector $|v\rangle \in \mathcal{H}$ can be written as $|v\rangle = \sum_i \alpha_i |e_i\rangle$.

Definition 2.3.9 (Dimension of a Hilbert space). The dimension of a Hilbert space \mathcal{H} is the number of vectors of a basis of \mathcal{H} . We write

$$\dim(\mathcal{H}) = |\{|e_i\rangle\}|, \quad (2.8)$$

where $\{|e_i\rangle\}$ denotes an orthonormal basis of \mathcal{H} .

Definition 2.3.10 (Tensor product of two Hilbert spaces). Let \mathcal{H}_A and \mathcal{H}_B be Hilbert spaces with basis $|e_i\rangle$ and $|f_j\rangle$ respectively. A tensor product

$$\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B = \{|v\rangle \otimes |\psi\rangle : |v\rangle \in \mathcal{H}_A, |\psi\rangle \in \mathcal{H}_B\} \quad (2.9)$$

is also a Hilbert space

- with base $|g_k\rangle = \{|e\rangle \otimes |f\rangle : |e\rangle \in |e_i\rangle, |f\rangle \in |f_j\rangle\}$ and
- inner product defined as $\langle a \otimes b | c \otimes d \rangle = \langle a | c \rangle \langle b | d \rangle$, where $|a\rangle, |c\rangle \in \mathcal{H}_A$ and $|b\rangle, |d\rangle \in \mathcal{H}_B$.
- $\dim(\mathcal{H}_{AB}) = \dim(\mathcal{H}_A) \cdot \dim(\mathcal{H}_B)$.

2.3.2 Linear operators

Definition 2.3.11 (Linear operator). Let \mathcal{V} be a vector space. Function $A : \mathcal{V} \rightarrow \mathcal{V}$ is a linear operator iff it is linear in its inputs, i.e.

$$A\left(\sum_i \alpha_i |e_i\rangle\right) = \sum_i \alpha_i (A|e_i\rangle), \quad |e_i\rangle \in \mathcal{V}. \quad (2.10)$$

In \mathbb{C}^n a linear operator A can be expressed as an n -by- n square matrix.

$$A = \begin{pmatrix} a_{0,0} & \cdots & a_{0,n-1} \\ \vdots & \ddots & \vdots \\ a_{n-1,0} & \cdots & a_{n-1,n-1} \end{pmatrix} = \sum_{i,j} a_{ij} |i\rangle\langle j|, \quad a_{ij} = \langle i|A|j\rangle. \quad (2.11)$$

Definition 2.3.12 (Adjoint Operator). Let A be a linear operator on a Hilbert space \mathcal{H} . A unique linear operator A^\dagger on \mathcal{H} satisfying

$$\langle v|A\psi\rangle = \langle A^\dagger v|\psi\rangle, \quad (2.12)$$

for all vectors $v, \psi \in \mathcal{H}$, is called an adjoint operator or Hermitian conjugate of the operator A . Additionally, we define

$$|v\rangle^\dagger \equiv \langle v|. \quad (2.13)$$

Note. Let A and B be linear operators on a Hilbert space \mathcal{H} , $|v\rangle \in \mathcal{H}$. Then

$$(AB)^\dagger = B^\dagger A^\dagger, \quad (2.14)$$

$$(A|v\rangle)^\dagger = \langle v|A^\dagger. \quad (2.15)$$

Definition 2.3.13. A linear operator A defined on a vector space \mathcal{V} is called

1. identity operator \mathbb{I} iff $A|v\rangle = |v\rangle$ for all vectors $|v\rangle \in \mathcal{V}$,
2. zero operator $\mathbb{0}$ iff $A|v\rangle = \vec{0}$ for all vectors $|v\rangle \in \mathcal{V}$,
3. normal iff $A^\dagger A = A A^\dagger$,
4. self-adjoint or Hermitian iff $A^\dagger = A$,
5. unitary iff $A^\dagger A = A A^\dagger = \mathbb{I}$,
6. idempotent iff $A^2 = A$,
7. projection iff A is self-adjoint and idempotent.

Theorem 2.3.1 (Spectral decomposition). *Any normal operator A on a vector space \mathcal{V} is diagonal to some orthonormal basis for \mathcal{V}*

$$A = \sum_i \lambda_i |i\rangle\langle i|, \quad (2.16)$$

where λ_i are the eigenvalues of A , $\{|i\rangle\}$ is an orthonormal basis for \mathcal{V} , and each $|i\rangle$ is an eigenvector of A corresponding to eigenvalue λ_i . For proof see [4] on page 72.

Definition 2.3.14 (Linear operator on a composed Hilbert space). Let A, B be linear operators on $\mathcal{H}_A, \mathcal{H}_B$ respectively. The tensor product

$$A \otimes B = \begin{pmatrix} a_{1,1}B & \dots & a_{1,m}B \\ \vdots & \ddots & \vdots \\ a_{m,1}B & \dots & a_{m,m}B \end{pmatrix} \quad (2.17)$$

is a linear operator on $\mathcal{H}_A \otimes \mathcal{H}_B$.

$$(A \otimes B) \left(\sum_{i,j} \alpha_i |e_i\rangle \otimes \beta_j |f_j\rangle \right) = \sum_{i,j} \alpha_i \beta_j (A|e_i\rangle \otimes B|f_j\rangle). \quad (2.18)$$

2.3.3 Operator functions

Definition 2.3.15 (Operator function). Let $A = \sum_i \lambda_i |i\rangle\langle i|$ be a spectral decomposition for a normal operator A . An operator function f on A is defined as

$$f(A) = \sum_i f(\lambda_i) |i\rangle\langle i|. \quad (2.19)$$

This allows us to define functions like square root, logarithm or exponential for operators.

Definition 2.3.16 (Trace of a matrix). The trace of a square matrix A is defined as

$$\text{tr}(A) = \sum_i a_{ii}. \quad (2.20)$$

Note. Let A and B be linear operators and $\eta \in \mathbb{C}$. Then

$$\bullet \quad \text{tr}(AB) = \text{tr}(BA), \quad (\text{cyclic property}), \quad (2.21)$$

$$\bullet \quad \begin{aligned} \text{tr}(A+B) &= \text{tr}(A) + \text{tr}(B), \\ \text{tr}(\eta A) &= \eta \text{tr}(A), \end{aligned} \quad (\text{linear property}). \quad (2.22)$$

From the cyclic property it follows that the trace of a square matrix A is invariant under the unitary similarity transformation $A \rightarrow UAU^\dagger$.

$$\text{tr}(UAU^\dagger) = \text{tr}(UU^\dagger A) = \text{tr}(A) \quad (2.23)$$

This property is important for the density operator formalism. See Section 2.5 on page 22.

2.4 The postulates of quantum mechanics

Postulate 1 (State space). *Associated to any isolated physical system is a complex vector space with inner product (Hilbert space) known as the state space of the system. The state of the system is completely described by its state vector, which is a unit vector in the system's state space.*

Postulate 2 (Evolution). *The time evolution of the state of a closed quantum system is described by the Schrödinger equation*

$$i\hbar \frac{\partial}{\partial t} |v\rangle = H|v\rangle, \quad (2.24)$$

where \hbar is the reduced Planck constant $\hbar \approx 1.05457 \cdot 10^{-34} \text{Js}$ and H is a fixed self-adjoint operator (hermitian) known as the Hamiltonian of the closed system.

In practice, it is convenient to absorb \hbar into H , effectively setting $\hbar = 1$. The solution of the equation (2.24) constitutes a time evolution operator

$$U(\Delta t) = e^{-iH\Delta t}. \quad (2.25)$$

The time evolution operator $U(\Delta t)$ is a unitary operator since the Hamiltonian H is a hermitian matrix, $H = H^\dagger$, and therefore the unitarity condition is fulfilled,

$$U(\Delta t)U(\Delta t)^\dagger = e^{-iH\Delta t} e^{iH\Delta t} = I.$$

We can reformulate the second postulate using unitary evolution:

The time evolution of a closed quantum system from the state $|v_1\rangle$ at time t_1 to the state $|v_2\rangle$ at time t_2 is described by a unitary operator $U = U(t_2 - t_1)$,

$$|v_2\rangle = U|v_1\rangle. \quad (2.26)$$

The correspondence between the discrete-time description of dynamics using unitary operators and continuous time description using Hamiltonians is one-to-one.

Postulate 3 (Quantum measurement). *A quantum measurement is described by a collection $\{M_m\}$ of measurement operators. These are operators acting on the state space of the system being measured. The index m refers to the measurement outcome r_m that may occur in the experiment. Measuring the system state $|v\rangle$ will give the result r_m with probability*

$$p(r_m) = \langle v|M_m^\dagger M_m|v\rangle, \quad (2.27)$$

and the state of the system reduces to the post-measurement state

$$|v'\rangle = \frac{M_m|v\rangle}{\sqrt{p(r_m)}}. \quad (2.28)$$

The operators $\{M_m\}$ satisfy the completeness equation

$$\sum_m M_m^\dagger M_m = I. \quad (2.29)$$

Using the completeness equation together with the normalization condition, we can see that probabilities sum to one:

$$1 = \sum_m p(r_m) = \sum_m \langle v | M_m^\dagger M_m | v \rangle. \quad (2.30)$$

Note. We say that states $|v\rangle$ and $|\psi\rangle$ are equivalent, $|v\rangle \cong |\psi\rangle$, up to the global phase factor, iff

$$|v\rangle = e^{i\Phi} |\psi\rangle, \quad \Phi \in \mathbb{R}.$$

The statistics of measurement predicted for these two states are the same!

$$\langle v | M_m^\dagger M_m | v \rangle = \langle \psi | e^{-i\Phi} M_m^\dagger M_m e^{i\Phi} | \psi \rangle = \langle \psi | M_m^\dagger M_m | \psi \rangle$$

The global phase must not be confused with the relative phase. Let $|0\rangle, |1\rangle$ form a basis of a Hilbert space \mathcal{H} . States $|v\rangle = \alpha|0\rangle + \beta|1\rangle$ and $|\psi\rangle = \alpha|0\rangle - \beta|1\rangle$ defined on \mathcal{H} do differ by a relative phase.

Projective measurement

For many applications a special class of measurements known as projective measurement is of importance. A projective measurement is described by a self-adjoint operator M , called an **observable**, with a spectral decomposition

$$M = \sum_m r_m P_m, \quad (2.31a)$$

where

$$\sum_m P_m = I \quad \text{and} \quad P_m P_n = \delta_{mn} P_m. \quad (2.31b)$$

P_m is an orthogonal projector onto the eigenspace of M with eigenvalue r_m . The eigenvalues r_m correspond to possible outcomes of the measurement. The probability of getting result r_m and afterward state's collapse are given by

$$p(r_m) = \langle v | P_m | v \rangle, \quad (2.32)$$

$$|v'\rangle = \frac{P_m |v\rangle}{\sqrt{p(r_m)}}. \quad (2.33)$$

A projective measurement allows to express the measurement

in an orthonormal basis $|m\rangle$

simply by defining

$$P_m = |m\rangle\langle m|.$$

The corresponding observable M is then given by a list of projectors P_m . Actually, projective measurement is equivalent to the general measurement postulate when augmented with the ability to perform unitary transformations. See [4] for details. Unless stated otherwise, projective measurement is the default measurement used in quantum computing.

Note. The average value of a projective measurement is

$$E(M) = \sum_m r_m p(r_m) = \sum_m r_m \langle v | P_m | v \rangle = \langle v | \left(\sum_m r_m P_m \right) | v \rangle = \langle v | M | v \rangle. \quad (2.34)$$

The average value $E(M)$ is often denoted by $\langle M \rangle$.

Postulate 4 (Composite systems). *The state space \mathcal{H} of a composite physical system is the tensor product of the state spaces \mathcal{H}_i of its components,*

$$\mathcal{H} = \bigotimes_i \mathcal{H}_i.$$

Moreover, if the subsystems are in the states $|v_i\rangle \in \mathcal{H}_i$, then the joint state $|v\rangle \in \mathcal{H}$ of the total system is

$$|v\rangle = \bigotimes_i |v_i\rangle.$$

Note (Entanglement). Let $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ be a Hilbert space. A joint state $|v\rangle \in \mathcal{H}$ that cannot be written as tensor product of some vectors $|v_A\rangle \in \mathcal{H}_A$, $|v_B\rangle \in \mathcal{H}_B$ is said to be entangled, otherwise we call this joint state a product state. The volume of entangled states is much higher than the volume of product states [Aubrun, Szarek].

In entangled states, unitary operators and measurement performed on one system affect the state of the second system. In product states, these operations affect only the state of the target component.

2.5 Density operator formalism

Next to the state vectors formalism, there exists an alternative density operator (density matrix) formalism. The postulates of quantum mechanics can be equivalently written using density operators. Density operator formalism is of advantage when describing individual subsystems of a composite quantum system or dealing with quantum systems whose state is not completely known.

Definition 2.5.1 (Density operator). Let a quantum system \mathcal{S} with associated Hilbert space \mathcal{H} be at a state $|v_i\rangle \in \mathcal{H}$ with probability p_i . The density operator for the system described by an ensemble $\{p_i, |v_i\rangle\}$ is defined as

$$\rho = \sum_i p_i |v_i\rangle\langle v_i|. \quad (2.35)$$

A density operator satisfies the conditions:

1. $\text{tr}(\rho) = 1$,
2. ρ is a normal operator and its spectrum consists of positive real numbers.

□

A quantum state represented by a density operator ρ is said to be a **pure state** iff

$$\text{tr}(\rho^2) = 1. \quad (2.36)$$

Otherwise the state is said to be **mixed**. For a pure state described by a state vector $|v\rangle$ the equation (2.35) reduces to

$$\rho = |v\rangle\langle v|. \quad (2.37)$$

Using the density operator formalism, equation (2.26) for temporal unitary evolution of a closed quantum system has the form

$$\rho' = U\rho U^\dagger. \quad (2.38)$$

This can be easily seen from a transformation

$$\sum_i p_i |v_i\rangle\langle v_i| \xrightarrow{U} \sum_i p_i U|v_i\rangle\langle v_i|U^\dagger. \quad (2.39)$$

Equations (2.27), (2.28) for the 3rd postulate of quantum mechanics have the form

$$p(r_m) = \text{tr} \left(M_m^\dagger M_m \rho \right), \quad (2.40)$$

$$\rho' = \frac{M_m \rho M_m^\dagger}{p(r_m)}. \quad (2.41)$$

For composite systems described by the 4th postulate where the individual components are in the states ρ_i , the joint state of the total system is

$$\rho = \bigotimes_i \rho_i. \quad (2.42)$$

The reduced density operator

When we deal with a subsystem of a larger system \mathcal{S} whose state is described by a density operator ρ , we need to find a function which will provide the correct measurement statistics for this subsystem. Such a function is a partial trace function and the provided statistics is called a reduced density operator. It can be shown that a partial trace is the unique function with the above written property.

Definition 2.5.2 (Partial trace). Let ρ_A and ρ_B be a density operator of a system A and B , respectively. The partial trace over system B is defined by

$$\text{tr}_B (\rho_A \otimes \rho_B) = \rho_A \text{tr} (\rho_B). \quad (2.43)$$

Definition 2.5.3 (Reduced density operator). Let ρ_{AB} be a density operator describing a composed state of physical systems A and B . The reduced density operator ρ_A for system A is defined by

$$\rho_A = \text{tr}_B(\rho_{AB}). \quad (2.44)$$

Example 4 (Tracing out). Let ρ_{AB} be a four-by-four density operator describing a composed state of two two-dimensional systems A and B . According to the definition, the reduced density operator ρ_A and ρ_B are calculated by partial trace over system B and A , respectively. There are two practical approaches how to perform this tracing out.

One approach follows directly from the definition of the partial trace. To be notationally clear, with respect to the state labelling, let us write

$$\begin{aligned} \rho_A &= \sum_{\phi, \phi' \in \{0,1\}} \alpha_{\phi, \phi'} |\phi\rangle\langle\phi'|, \\ \rho_B &= \sum_{\psi, \psi' \in \{0,1\}} \beta_{\psi, \psi'} |\psi\rangle\langle\psi'|, \end{aligned}$$

then

$$\begin{aligned} \rho_A \otimes \rho_B &= \sum_{\phi, \phi' \in \{0,1\}} \alpha_{\phi, \phi'} |\phi\rangle\langle\phi'| \otimes \rho_B \\ &= \begin{pmatrix} \alpha_{0,0} \rho_B & \alpha_{0,1} \rho_B \\ \alpha_{1,0} \rho_B & \alpha_{1,1} \rho_B \end{pmatrix} \\ &= \sum_{\phi, \phi' \in \{0,1\}} \alpha_{\phi, \phi'} |\phi\rangle\langle\phi'| \otimes \sum_{\psi, \psi' \in \{0,1\}} \beta_{\psi, \psi'} |\psi\rangle\langle\psi'| \\ &= \sum_{\phi, \phi', \psi, \psi' \in \{0,1\}} \alpha_{\phi, \phi'} \beta_{\psi, \psi'} |\phi\rangle|\psi\rangle\langle\phi'| \langle\psi'| \\ &= \begin{pmatrix} \alpha_{0,0} \beta_{0,0} & \alpha_{0,0} \beta_{0,1} & \alpha_{0,1} \beta_{0,0} & \alpha_{0,1} \beta_{0,1} \\ \alpha_{0,0} \beta_{1,0} & \alpha_{0,0} \beta_{1,1} & \alpha_{0,1} \beta_{1,0} & \alpha_{0,1} \beta_{1,1} \\ \alpha_{1,0} \beta_{0,0} & \alpha_{1,0} \beta_{0,1} & \alpha_{1,1} \beta_{0,0} & \alpha_{1,1} \beta_{0,1} \\ \alpha_{1,0} \beta_{1,0} & \alpha_{1,0} \beta_{1,1} & \alpha_{1,1} \beta_{1,0} & \alpha_{1,1} \beta_{1,1} \end{pmatrix} \\ &= \sum_{\phi, \psi, \phi', \psi' \in \{0,1\}} x_{\phi, \psi, \phi', \psi'} |\phi, \psi\rangle\langle\phi', \psi'| \\ &= \sum_{\gamma, \gamma' = 0}^3 x_{\gamma, \gamma'} |\gamma\rangle\langle\gamma'| \\ &= \begin{pmatrix} x_{0,0} & x_{0,1} & x_{0,2} & x_{0,3} \\ x_{1,0} & x_{1,1} & x_{1,2} & x_{1,3} \\ x_{2,0} & x_{2,1} & x_{2,2} & x_{2,3} \\ x_{3,0} & x_{3,1} & x_{3,2} & x_{3,3} \end{pmatrix}. \end{aligned}$$

Now we can see that the trace of the upper-left two-by-two submatrix gives us

$$\text{tr} \begin{pmatrix} x_{0,0} & x_{0,1} \\ x_{1,0} & x_{1,1} \end{pmatrix} = \text{tr}(\alpha_{0,0} \rho_B) = \alpha_{0,0} = x_{0,0} + x_{1,1},$$

since $\text{tr}(\rho_B)$ must be equal to one. System B is traced out. Accordingly, we can take a submatrix consisting of items on positions related to $\beta_{0,0}$ and trace out system A ,

$$\text{tr} \begin{pmatrix} x_{0,0} & x_{0,2} \\ x_{2,0} & x_{2,2} \end{pmatrix} = \text{tr}(\beta_{0,0} \rho_A) = \beta_{0,0} = x_{0,0} + x_{2,2}.$$

Thus the reduced density operator for the system A is

$$\begin{aligned} \rho_A = \text{tr}_B(\rho_{AB}) &= \text{tr}_B \begin{pmatrix} x_{0,0} & x_{0,1} & x_{0,2} & x_{0,3} \\ x_{1,0} & x_{1,1} & x_{1,2} & x_{1,3} \\ x_{2,0} & x_{2,1} & x_{2,2} & x_{2,3} \\ x_{3,0} & x_{3,1} & x_{3,2} & x_{3,3} \end{pmatrix} = \begin{pmatrix} \text{tr} \begin{pmatrix} x_{0,0} & x_{0,1} \\ x_{1,0} & x_{1,1} \end{pmatrix} & \text{tr} \begin{pmatrix} x_{0,2} & x_{0,3} \\ x_{1,2} & x_{1,3} \end{pmatrix} \\ \text{tr} \begin{pmatrix} x_{2,0} & x_{2,1} \\ x_{3,0} & x_{3,1} \end{pmatrix} & \text{tr} \begin{pmatrix} x_{2,2} & x_{2,3} \\ x_{3,2} & x_{3,3} \end{pmatrix} \end{pmatrix} \\ &= \begin{pmatrix} x_{0,0} + x_{1,1} & x_{0,2} + x_{1,3} \\ x_{2,0} + x_{3,1} & x_{2,2} + x_{3,3} \end{pmatrix}. \end{aligned}$$

The reduced density operator for the system B is

$$\begin{aligned} \rho_B = \text{tr}_A(\rho_{AB}) &= \text{tr}_A \begin{pmatrix} x_{0,0} & x_{0,1} & x_{0,2} & x_{0,3} \\ x_{1,0} & x_{1,1} & x_{1,2} & x_{1,3} \\ x_{2,0} & x_{2,1} & x_{2,2} & x_{2,3} \\ x_{3,0} & x_{3,1} & x_{3,2} & x_{3,3} \end{pmatrix} = \begin{pmatrix} \text{tr} \begin{pmatrix} x_{0,0} & x_{0,2} \\ x_{2,0} & x_{2,2} \end{pmatrix} & \text{tr} \begin{pmatrix} x_{0,1} & x_{0,3} \\ x_{2,1} & x_{2,3} \end{pmatrix} \\ \text{tr} \begin{pmatrix} x_{1,0} & x_{1,2} \\ x_{3,0} & x_{3,2} \end{pmatrix} & \text{tr} \begin{pmatrix} x_{1,1} & x_{1,3} \\ x_{3,1} & x_{3,3} \end{pmatrix} \end{pmatrix} \\ &= \begin{pmatrix} x_{0,0} + x_{2,2} & x_{0,1} + x_{2,3} \\ x_{1,0} + x_{3,2} & x_{1,1} + x_{3,3} \end{pmatrix}. \end{aligned}$$

A different approach involves direct formulas for tracing out. These formulas are derived from the trace definition (2.20) and the partial trace definition (2.43),

$$\rho_A = \text{tr}_B(\rho_{AB}) = \sum_{\phi, \phi' \in \{0,1\}} |\phi\rangle \left(\sum_{\psi \in \{0,1\}} \langle \phi, \psi | \rho_{AB} | \phi', \psi \rangle \right) \langle \phi' |, \quad (2.45)$$

$$\rho_B = \text{tr}_A(\rho_{AB}) = \sum_{\psi, \psi' \in \{0,1\}} |\psi\rangle \left(\sum_{\phi \in \{0,1\}} \langle \phi, \psi | \rho_{AB} | \phi, \psi' \rangle \right) \langle \psi' |. \quad (2.46)$$

An insight into these formulas gives a simple rule how to calculate a partial trace over a given system. Let ρ_{AB} be written in the form of

$$\rho_{AB} = \sum x_{\phi, \psi, \phi, \psi'} |\phi, \psi\rangle \langle \phi', \psi'|,$$

then the partial trace over the i -th system (qubit) is evaluated by replacing an outer product term $|x\rangle\langle y|$ for this system with a corresponding inner product $\langle x|y\rangle$.

For example, let $\rho_{AB} = x_{0,1,0,0} |01\rangle\langle 00| + x_{0,1,1,1} |01\rangle\langle 11| + \dots$. Then,

$$\text{tr}_B(\rho_{AB}) = (x_{0,1,0,0} |0\rangle\langle 0| \underbrace{\langle 1|0\rangle}_{=0} + x_{0,1,1,1} |0\rangle\langle 1| \underbrace{\langle 1|1\rangle}_{=1} + \dots) = x_{0,1,1,1} |0\rangle\langle 1| + \dots,$$

$$\text{tr}_A(\rho_{AB}) = (x_{0,1,0,0} |1\rangle\langle 0| \underbrace{\langle 0|0\rangle}_{=1} + x_{0,1,1,1} |1\rangle\langle 1| \underbrace{\langle 0|1\rangle}_{=0} + \dots) = x_{0,1,0,0} |1\rangle\langle 0| + \dots \quad \square$$

Chapter 3

Quantum computing

3.1 Quantum computing

Quantum computing is an abstract concept of computation defined within the framework of quantum mechanics. The postulates of quantum mechanics relate the set of allowed states to unit vectors in a complex Hilbert space, allowed transitions are described by unitary matrices and the measurement reveals complex amplitudes only indirectly via the corresponding probabilities. For the concept of quantum computing, it is convenient to adopt these rules in the following way.

3.1.1 Quantum states

The smallest nontrivial quantum mechanical system is a two-state system. In the context of quantum computing, it is called a quantum bit, or a qubit for short. Examples are two distinct polarization states of a photon, or the two spin directions of an electron in a magnetic field. The associated state space is a two-dimensional complex Hilbert space \mathcal{H}_2 which is spanned by an orthonormal basis $\mathcal{B} = \{|0\rangle, |1\rangle\}$ labeled in some canonical way. Labeling $|0\rangle$ and $|1\rangle$ refers to the standard (computational) basis defined as:

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (\text{standard basis}) \quad (3.1)$$

An arbitrary qubit state $|v\rangle$ can be written as a normalized linear combination (superposition) of basis states

$$|v\rangle = \alpha|0\rangle + \beta|1\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \quad \alpha, \beta \in \mathbb{C}, \text{ with } |\alpha|^2 + |\beta|^2 = 1. \quad (\text{qubit}) \quad (3.2)$$

The complex numbers α, β are so-called quantum mechanical amplitudes. The condition

$$|\alpha|^2 + |\beta|^2 = 1$$

follows from the 2-norm unity condition $\|v\|_2 = 1$. We have

$$\|v\|_2 = \sqrt{\langle v|v\rangle} = 1 \quad \Rightarrow \quad \langle v|v\rangle = 1,$$

and for the standard basis we directly get

$$\langle v|v\rangle = (\alpha^* \ \beta^*) \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \alpha\alpha^* + \beta\beta^* = |\alpha|^2 + |\beta|^2 = 1.$$

For any other orthonormal basis, e.g.

$$|0'\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad |1'\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad (\text{dual basis}) \quad (3.3)$$

the same is true.

In principle, it is also possible to work with three-state systems (qutrits) or general d -state systems (qudits), but there is no actual advantage of doing so, as well as for classical trits/dits. For this reason, we can restrict ourselves to qubits and larger systems composed out of qubits. The joint state of a composed system is given by the tensor product of its individual subsystem states. Let us have two qubits $|v\rangle, |\psi\rangle \in \mathcal{H}_2$, $|v\rangle = \alpha|0\rangle + \beta|1\rangle$, $|\psi\rangle = \gamma|0\rangle + \delta|1\rangle$. Their tensor product is equal to

$$|v\rangle|\psi\rangle = \alpha\gamma|0\rangle|0\rangle + \alpha\delta|0\rangle|1\rangle + \beta\gamma|1\rangle|0\rangle + \beta\delta|1\rangle|1\rangle,$$

and we can rewrite this equation in the form

$$|v\rangle|\psi\rangle = \alpha_{00}|00\rangle + \alpha_{01}|01\rangle + \alpha_{10}|10\rangle + \alpha_{11}|11\rangle = \sum_{i \in \{0,1\}^2} \alpha_i |i\rangle.$$

States $|00\rangle, |01\rangle, |10\rangle, |11\rangle$ are the standard basis vectors of the Hilbert space \mathcal{H}_{2^2} .

$$\begin{aligned} |00\rangle &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, & |01\rangle &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \\ |10\rangle &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, & |11\rangle &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \end{aligned}$$

Labels 00, 01, 10, 11 may be renamed to 0, 1, 2, 3 because they can be easily seen as a binary representation of these integers. Using this representation, we write

$$|v\rangle|\psi\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle + \alpha_2|2\rangle + \alpha_3|3\rangle = \sum_{i=0}^3 \alpha_i |i\rangle.$$

A general state of an n -qubit system (often called a quantum register), $|\psi\rangle \in \mathcal{H}_{2^n}$, is described as

$$|\psi\rangle = \sum_{i \in \{0,1\}^n} \alpha_i |i\rangle = \sum_{i=0}^{2^n-1} \alpha_i |i\rangle, \quad (\text{quantum register}) \quad (3.4)$$

where $\sum_i |\alpha_i|^2 = 1$ and $\{|i\rangle\}$ is the standard basis of an 2^n -dimensional Hilbert space \mathcal{H}_{2^n} . Equation (3.4) shows how a state space of quantum system grows exponentially with its physical size - that is the number of the qubits. It is due to the property of tensor product of Hilbert spaces (def. 2.3.10) that

$$\dim(\mathcal{H}_A \otimes \mathcal{H}_B) = \dim(\mathcal{H}_A) \cdot \dim(\mathcal{H}_B).$$

3.1.2 Measurement

It may seem that an n -qubit register is capable of storing exponentially more information than an n -bit register. However, the Holevo theorem [5] states that one can retrieve faithfully only n **bits** from an n -qubit register. Using other words, retrieving of classic information from a quantum register requires measurement and by the 3^{rd} postulate, we know that measurement destroys the superposition. In particular, let \mathcal{H} be a state space spanned by basis \mathcal{B} , then the projective measurement associated with some observable

$$M = \sum_m r_m P_m,$$

where P_m are orthogonal projectors

$$P_m = |m\rangle\langle m|, \quad |m\rangle \in \mathcal{B},$$

projects the state $|\psi\rangle \in \mathcal{H}$,

$$|\psi\rangle = \sum_i \alpha_i |i\rangle, \quad |i\rangle \in \mathcal{B},$$

to the post-measurement state

$$\frac{P_m |\psi\rangle}{\sqrt{p(r_m)}} = \frac{1}{|\alpha_m|} \sum_i \alpha_i |m\rangle\langle m|i\rangle = \frac{1}{|\alpha_m|} \sum_i \alpha_m |m\rangle \delta_{mi} = \frac{\alpha_m}{|\alpha_m|} |m\rangle \cong |m\rangle,$$

with probability

$$p(r_m) = \langle \psi | P_m | \psi \rangle = |\alpha_m|^2,$$

where r_m is the outcome of the measurement and $\alpha_m (|\alpha_m|)^{-1}$ is an unobservable global phase. It is also possible to measure only a part of the register at a time. For product states, the rest of the register is kept untouched, for entangled states, the whole register is affected.

Example 5 (Projective measurement).

Product state. Let $|\psi\rangle = \alpha|000\rangle + \beta|001\rangle + \gamma|010\rangle + \delta|011\rangle \in \mathcal{H}_{2^3}$, $\alpha, \beta, \gamma, \delta \neq 0$. It is a product state since $|\psi\rangle = |0\rangle (a|0\rangle + b|1\rangle) (c|0\rangle + d|1\rangle)$ for some a, b, c, d satisfying $\alpha = ac$, $\beta = ad$, $\gamma = bc$, and $\delta = bd$. Measuring the middle qubit with respect to the standard observable

$$M = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \Big|_{P_0=|0\rangle\langle 0|} + (-1) \cdot \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \Big|_{P_1=|1\rangle\langle 1|}$$

yields the post-measurement state

$$|\psi'\rangle = \begin{cases} c|000\rangle + d|001\rangle, & \text{with probability } |a|^2, \\ & \text{and the result of the measurement is } +1, \\ c|010\rangle + d|011\rangle, & \text{with probability } |b|^2, \\ & \text{and the result of the measurement is } -1. \end{cases}$$

The average value of this measurement is $\langle M \rangle = |a|^2 - |b|^2$. Results $+1$ and -1 are to be understood as observing 0 and 1, respectively, in the measured register.

Entangled state. Let $|\psi\rangle = \alpha|00\rangle + \beta|11\rangle \in \mathcal{H}_{2^2}$, $\alpha, \beta \neq 0$. It is an entangled state since $|\psi\rangle$ cannot be written as a tensor product of two qubits

$$(a|0\rangle + b|1\rangle)(c|0\rangle + d|1\rangle) = ac|00\rangle + ad|10\rangle + bc|10\rangle + bd|11\rangle,$$

because

$$\left(\begin{array}{l} ac = \alpha \Rightarrow ac \neq 0 \\ bd = \beta \Rightarrow bd \neq 0 \end{array} \right) \Rightarrow ad \neq 0, \quad bc \neq 0.$$

Measuring the left qubit yields the post-measurement state

$$|\psi'\rangle = \begin{cases} |00\rangle, & \text{with probability } |\alpha|^2, \text{ and the result of the measurement is } +1, \\ |11\rangle, & \text{with probability } |\beta|^2, \text{ and the result of the measurement is } -1. \end{cases}$$

It can be seen that the measurement of the left qubit uniquely determined the state of the right qubit.

□

Classical vs. probabilistic vs. quantum register

We can now compare a classical, probabilistic and quantum register composed out of n two-state systems.

| | |
|-------------------------|--|
| Classical register: | $x, \quad x \in \{0, 1\}^n.$ |
| Probabilistic register: | $\sum_x p_x x, \quad x \in \{0, 1\}^n, \quad p_x \in \mathbb{R}^+, \quad \sum_x p_x = 1.$ |
| Quantum register: | $\sum_x \alpha_x x\rangle, \quad x \in \{0, 1\}^n, \quad \alpha_x \in \mathbb{C}, \quad \sum_x \alpha_x ^2 = 1.$ |

A classical n -bit register can store any string from the set $\{0, 1\}^n$. Depending on the assigned semantic, these strings can be interpreted as nonnegative integers $\{0, \dots, 2^n - 1\}$, restricted set of real numbers, characters, and so on. The important point is that the register contains only one such string at a time. This is in contrast with a probabilistic register which contains a probabilistic distribution of all possible n -bit strings. These are exponentially many with respect to the length n . However, only one string is observed upon measurement. If we want to observe a different string, we need to repeat the whole computation as the former probabilistic distribution is no longer available nor it can be cloned right before measurement for obvious reasons. A quantum register can also contain a special distribution of all possible n -bit strings. To each possible string is assigned a complex amplitude and we talk about a quantum superposition. After a measurement of all qubits, the superposition is no longer available and we observe a single string only. In this sense, a quantum register is very similar to a probabilistic register.

3.1.3 Evolution of quantum states

However, a probabilistic and a quantum register differ in the transitions they can undergo. An evolution of a quantum register $|\psi\rangle$ is described by a unitary matrix U , $|\psi'\rangle = U|\psi\rangle$. Since for every unitary U there exists its inversion $U^{-1} = U^\dagger$, such an evolution is always reversible, $U^{-1}|\psi'\rangle = U^\dagger U|\psi\rangle = |\psi\rangle$. Models of classical computation are usually not reversible, nevertheless, Bennett in [6] showed the existence of reversible universal Turing machines, and Fredkin and Toffoli [7] showed a set of universal classical reversible gates for the circuit model. Therefore the restriction to reversible computation should not be considered as an obstacle. In principle, for any computable function $f : \{0, 1\}^n \rightarrow \{0, 1\}^m$, there is a unitary matrix U_f corresponding to the reversible version of f . A quantum computer is at least as powerful as a classical computer [8]. The operator U_f for a nonreversible classical function f is typically implemented to perform a bijective mapping

$$U_f : |x, y\rangle \longrightarrow |x, y \oplus f(x)\rangle, \quad (3.5)$$

where $x \in \{0, 1\}^n$, $y \in \{0, 1\}^m$ and \oplus denotes the bitwise addition modulo 2 (equivalently, bitwise exclusive disjunction). By setting the second register to zero, $|y\rangle = |0\rangle = |0^{(m)}\rangle = |0, \dots, 0\rangle$, we get a computation

$$U_f |x, 0\rangle = |x, f(x)\rangle. \quad (3.6)$$

Classical reversible functions and quantum specific transforms (inherently reversible) perform mapping

$$U_f : |x\rangle \longrightarrow |f(x)\rangle. \quad (3.7)$$

A quantum algorithm typically consists of several different transforms applied to various quantum registers or individual qubits. This can be graphically depicted as shown in the Figure 3.1. The figure shows a procedure consisting from three evolutionary steps U_1, U_2, U_3 applied over eight qubits to an initial state $|0\rangle = |0^{(8)}\rangle$; $|\psi\rangle = (U_3 \cdot (U_2 \cdot (U_1|0\rangle))) = U|0\rangle$.

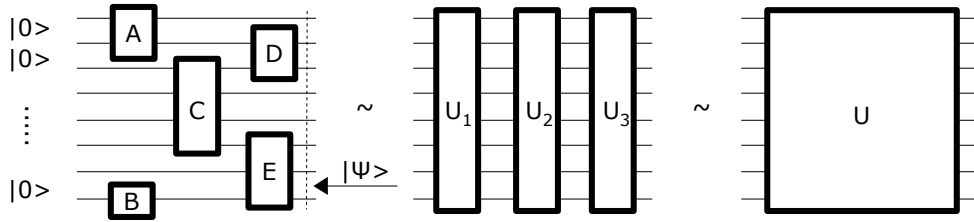


Figure 3.1: Graphical scheme of a quantum algorithm.

The overall unitary operation at the first step has the form $U_1 = A^{(2)} \otimes \mathbb{I}^{(5)} \otimes B^{(1)}$. Analogically, $U_2 = \mathbb{I}^{(2)} \otimes C^{(4)} \otimes \mathbb{I}^{(2)}$, and $U_3 = \mathbb{I}^{(1)} \otimes D^{(2)} \otimes \mathbb{I}^{(2)} \otimes E^{(3)}$. The upper index denotes the number of qubits over which the transform is applied.

Examples of important unitary transforms over n qubits are the Hadamard transform and the discrete Fourier transform. Note that $H^{(1)} = QFT^{(1)}$.

$$H^{(n)} : \frac{1}{\sqrt{2^n}} \sum_{x,y=0}^{2^n-1} (-1)^{x \cdot y} |y\rangle \langle x| \quad (\text{Hadamard transform}) \quad (3.8)$$

$$QFT^{(n)} : \frac{1}{\sqrt{2^n}} \sum_{x,y=0}^{2^n-1} e^{2\pi i xy/2^n} |y\rangle\langle x| \quad (\text{Fourier transform}) \quad (3.9)$$

The exponent $x \cdot y$ in (3.8) refers to the dot product of binary representations of x and y . This dot product is defined as $x \cdot y = \bigoplus_{i=1}^n x_i y_i$. It is easy to see that transforms (3.8) and (3.9) are indeed unitary, that is

$$H^{(n)} H^{(n)\dagger} = QFT^{(n)} QFT^{(n)\dagger} = \mathbb{I}^{(n)}.$$

The Hadamard transform is often used to create an equally weighted superposition of standard basis states

$$H^{(n)}|x\rangle = \frac{1}{\sqrt{2^n}} \sum_{y=0}^{2^n-1} (-1)^{x \cdot y} |y\rangle, \quad |x\rangle, |y\rangle \in \mathcal{B}, \quad (3.10)$$

$$H^{(n)}|0\rangle = \frac{1}{\sqrt{2^n}} \sum_{y=0}^{2^n-1} |y\rangle = \frac{1}{\sqrt{2^n}} (|0\rangle + |1\rangle + |2\rangle + \dots + |2^n - 1\rangle). \quad (3.11)$$

Quantum parallelism Since unitary operators are linear operators, they are linear in its inputs. Let us prepare two quantum registers in the following state

$$|\psi\rangle = (H^{(n)} \otimes \mathbb{I}^{(m)}) \underbrace{|0, \dots, 0\rangle}_n \underbrace{|0, \dots, 0\rangle}_m = \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x, 0\rangle,$$

then a single application of an operator U_f results in a quantum parallel calculation of values $f(x)$ for all $x \in \{0, \dots, 2^n - 1\}$.

$$|\psi'\rangle = U_f |\psi\rangle = U_f \left(\frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x, 0\rangle \right) = \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} U_f |x, 0\rangle = \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x, f(x)\rangle \quad (3.12)$$

However, this quantum parallelism alone is not very useful. A subsequent measurement would reveal exactly one pair $(x, f(x))$ chosen at random. Such a process can be perfectly simulated by a probabilistic computation. One idea might be to clone the final superposition $|\psi'\rangle$ so that we can perform additional measurements, but the 'No cloning theorem' prohibits unknown state cloning.

Theorem 3.1.1 (No cloning theorem). *An unknown quantum state cannot be cloned. Namely, there is no unitary transformation performing evolution*

$$|\chi\rangle|0\rangle \xrightarrow{U} |\chi\rangle|\chi\rangle \quad (3.13)$$

for any state $|\chi\rangle$.

Proof. Let us have two different orthonormal states $|\psi\rangle$ and $|\nu\rangle$, and $|\chi\rangle = \frac{1}{\sqrt{2}}(|\psi\rangle + |\nu\rangle)$. Assuming a 'cloning' U exists, we have $U(|\psi\rangle|0\rangle) = |\psi\rangle|\psi\rangle$, $U(|\nu\rangle|0\rangle) = |\nu\rangle|\nu\rangle$.

Additionally,

$$\begin{aligned} U(|\chi\rangle|0\rangle) &= U\left(\frac{1}{\sqrt{2}}(|\psi\rangle + |\nu\rangle)|0\rangle\right) = \frac{1}{\sqrt{2}}(U(|\psi\rangle|0\rangle) + U(|\nu\rangle|0\rangle)) = \frac{1}{\sqrt{2}}(|\psi\rangle|\psi\rangle + |\nu\rangle|\nu\rangle) \\ &\neq |\chi\rangle|\chi\rangle = \frac{1}{2}(|\psi\rangle|\psi\rangle + |\psi\rangle|\nu\rangle + |\nu\rangle|\psi\rangle + |\nu\rangle|\nu\rangle). \end{aligned}$$

□

Note that the 'No cloning theorem' trivially holds for a probabilistic register too. Figure 3.2 shows a hypothetical cloning matrix for a probabilistic one bit register. The matrix is not a valid stochastic matrix since its columns do not sum to one.

$$\begin{pmatrix} 1 & 0 & . & . \\ 0 & 1 & . & . \\ 1 & 0 & . & . \\ 0 & 1 & . & . \end{pmatrix} \cdot \begin{pmatrix} p_1 \\ p_2 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} p_1 \\ p_2 \\ p_1 \\ p_2 \end{pmatrix}$$

Figure 3.2: A hypothetical cloning matrix for a probabilistic one bit register.

The only possible way to make use of quantum parallelism is to employ quantum interference. Using interference phenomena, we can modify the superposition, for example, so that only a pair $(x, f(x))$ satisfying $f(x) = 1$ is measured with high probability. So far, we have no indication that interference as present in quantum mechanics can modify the superposition quickly enough so that NP -complete problems can be solved in polynomial time. Surprisingly, Abrams and Lloyd in [9] showed that even small non-linearities, if allowed by quantum physics, would enable for efficient NP -complete problems solving.

Example 6 (Grover amplitude amplification algorithm). Let us have a computable function $f : \{0, 1\}^n \rightarrow \{0, 1\}$ and a corresponding unitary operator U_f . We are looking for a string x such that $f(x) = 1$. An example of such problem is a Boolean satisfiability problem or a database search. An unsorted database search of size N is especially interesting problem since there is a complete lack of inner structure to be exploited. Classically, we need to check all N items, one by one, in order to find the right one in the worst case. On a quantum computer, we can use a different approach leading to a square root speed-up.

Let $n = \lceil \log_2 N \rceil$. Using a weighted superposition of basis states over n qubits, we can search through all N item with a single query to U_f

$$U_f\left(\frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle \otimes |0\rangle\right) = \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle \otimes |0 \oplus f(x)\rangle.$$

A small modification to the initial state of the second helps to change the computation in such a way that the values $f(x)$ are conveniently transferred to the amplitudes.

$$\begin{aligned} U_f\left(\frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\right) &= \frac{1}{\sqrt{2^n}\sqrt{2}} \left(\sum_{x=0}^{2^n-1} |x\rangle |0 \oplus f(x)\rangle - \sum_{x=0}^{2^n-1} |x\rangle |1 \oplus f(x)\rangle \right) \\ &= \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} (-1)^{f(x)} |x\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle). \end{aligned}$$

Now, amplitude $-1/\sqrt{2^n}$ marks the item(s) with $f(x) = 1$ and amplitude $1/\sqrt{2^n}$ marks the rest of items. The amplitude of the desired item can now be amplified using a transform known as inversion about mean.

$$H^{(n)} \left(2|0\rangle\langle 0| - \mathbb{I}^{(n)} \right) H^{(n)} \quad (\text{inversion about mean}) \quad (3.14)$$

Inversion about mean applied to a general superposition produces transform

$$\begin{aligned} & \left(H^{(n)} (2|0\rangle\langle 0| - \mathbb{I}^{(n)}) H^{(n)} \right) \sum_{x=0}^{2^n-1} \alpha_x |x\rangle = \left(\frac{2}{\sqrt{2^n}\sqrt{2^n}} \sum_{j,k=0}^{2^n-1} |j\rangle\langle k| - \mathbb{I}^{(n)} \right) \sum_{x=0}^{2^n-1} \alpha_x |x\rangle \\ &= \frac{2}{2^n} \sum_{j,k=0}^{2^n-1} |j\rangle\langle k| \sum_{x=0}^{2^n-1} \alpha_x |x\rangle - \mathbb{I}^{(n)} \sum_{x=0}^{2^n-1} \alpha_x |x\rangle = 2 \sum_{j=0}^{2^n-1} |j\rangle \cdot \underbrace{\left(\sum_{x=0}^{2^n-1} \alpha_x \right) \cdot \frac{1}{2^n}}_{\text{mean } \langle \alpha \rangle} - \sum_{x=0}^{2^n-1} \alpha_x |x\rangle \\ &= \underbrace{\sum_{j=0}^{2^n-1} 2\langle \alpha | j \rangle}_{\text{may be relabelled}} - \sum_{x=0}^{2^n-1} \alpha_x |x\rangle = \sum_{x=0}^{2^n-1} \left(-\alpha_x + 2\langle \alpha \rangle \right) |x\rangle . \end{aligned}$$

Repeated sequential action of the operator U_f and inversion about mean is shown in the Figure 3.3.

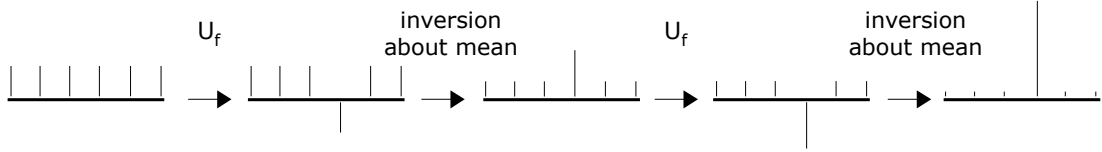


Figure 3.3: Amplitude amplification by inversion about mean.

It is clear that in order to stop the amplification at the right point one has to know at least approximately the number M of items satisfying $f(x) = 1$. The number of steps of amplification then grows like $O(\sqrt{N/M})$, see [10]. The number M can be exactly determined by the counting algorithm [4] if it is not roughly known by other means.

Recently, a similar type of amplification has been reported in energy transfer in plants during photosynthesis [11]. \square

A different example of a quantum algorithm is the period finding algorithm. This algorithm is the only quantum part of the Shor factoring algorithm [12]. With respect to the factoring, the task is to find the period r of a function $f_{k,N}(x) = k^x \pmod{N}$ for given coprimes k and N . The number N corresponds to the number to be factored and k is chosen at random. Accidentally, if k is not a coprime to N , then k is a factor of N . Once the period r is determined, a classical post-processing algorithm checks whether r leads to a non-trivial factor of N . This happens with high probability.

The period finding algorithm consists of three steps and we need two m -qubit registers, where m is chosen in the order of $\Theta(\log(N^2))$. First, we create a state with the period r

which we wish to determine in the first register. We do it by exploiting quantum parallelism and measurement on the second register. The circuit implementing U_f is essentially a reversible version of a classical circuit for modular exponentiation utilizing repeated squaring and requiring $O(m^3)$ universal gates.

$$\begin{aligned} \sum_x |x\rangle|0\rangle &\xrightarrow[\text{parallelism}]{U_f} \sum_x |x\rangle|k^x \bmod N\rangle \xrightarrow[\text{measurement}]{\text{second register}} \sum_{\{x' : y \equiv k^{x'} \bmod N\}} |x'\rangle|y\rangle \\ &= (|l\rangle + |l+r\rangle + |l+2r\rangle + \dots)|y\rangle \\ &= \sum_j |l+jr\rangle|y\rangle \end{aligned}$$

The result of the measurement is some non-negative integer y . Now, the superposition in the first register starts with a label $|l\rangle$, where l is equal to the smallest x such that $y \equiv k^x \pmod{N}$. We call l an offset, $l \leq r$. Note that in general 2^m is not an exact multiple of the unknown r and therefore the amount of terms in the sum $\sum_j |l+jr\rangle$ is not $2^m/r$ exactly. However, if 2^m is sufficiently large, then the small decay of the periodicity at x close to 2^m will have a negligible effect. Thus

$$|\{j\}| = \{0, 1, 2, \dots\} \approx \frac{2^m}{r}.$$

Second, we apply the quantum Fourier transform to the first register. Quantum Fourier transform over \mathbb{Z}_{2^m} requires $O(m^2)$ gates. The Fourier transform maps functions with period r to functions having non-zero values only at the amplitudes of the frequency $1/r$.

Thus we get rid of the offset and neglecting states with negligible amplitudes the number of terms in the superposition is strikingly decreased. It is now proportional to r . There are exactly r states with amplitude peaks and other states with smaller amplitudes are clustered around these peaks. Therefore, a subsequent measurement on the first register produces with high probability an m -bit non-negative integer

$$a \approx j_a \frac{2^m}{r},$$

where $j_a \in \{0, 1, 2, \dots, r-1\}$. Third, we take a rational number $a/2^m$ and using a continued fraction expansion we find a convergent which resembles j_a/r . Thus we can extract r . The continued fraction expansion can be computed using $O(m^3)$ gates for elementary arithmetic.

$$\frac{a}{2^m} = 0.\underbrace{a_{m-1}a_m \dots a_0}_{m\text{-bits}} = \frac{1}{c_1 + \frac{1}{c_2 + \frac{1}{c_3 + \dots}}} \Rightarrow \frac{j_a}{r}$$

Choosing m in the order of $\Theta(\log(N^2))$ ensures that r is revealed with high probability. A continued fraction expansion is a very important part of this algorithm.

3.2 Quantum Turing machines

The first formal definition of a quantum Turing machine, viewed as a quantum physical analogue of a probabilistic Turing machine, was given by D. Deutsch [13] in 1985. The existence of a fully universal quantum Turing machine which can simulate any other quantum Turing machine with at most polynomial slowdown was proven by Bernstein and Vazirani [14] in 1993. Additionally, they proved that there is an oracle relative to which there is a language that can be efficiently accepted by a quantum Turing machine but cannot be efficiently accepted by a bounded-error probabilistic Turing machine. The problem which was considered is called Recursive Fourier Sampling and the proposed quantum algorithm gives a quasipolynomial speedup, $O(n)$ versus $O(n^{\log n})$. This provides the first *formal* evidence that a quantum Turing machine is more powerful than a probabilistic Turing machine.

Informal evidence is present in promise problems such as the Deutsch-Jozsa problem [15], the Simon problem [16] or Shor's factoring algorithm [12]. However, these problems are only assumed to be hard, it is not proven. For example, the Deutsch-Jozsa problem was later shown to be in the complexity class BPP [14] and thus efficiently solvable on a probabilistic Turing machine. In this case, the quantum algorithm for the Deutsch-Jozsa problem only gives an advantage of deterministic approach instead of a probabilistic one. Regarding the Simon promise problem, its complexity was elaborated from the former expected quantum polynomial time to the exact quantum polynomial time [17]. The Simon promise problem was the first 'solid' problem where a quantum algorithm yields an exponential speedup.

Definition 3.2.1 (Quantum Turing machine). A quantum Turing machine is a generalized Turing machine. The machine has a *finite* control (finite set of states Q), possibly infinite tape, and a read/write head that recognizes a finite alphabet Σ . A transition function assigns to each transition a complex amplitude

$$\delta : Q \times \Sigma \times Q \times \Sigma \times \{L, R\} \rightarrow \mathbb{C} \quad (3.15)$$

and the corresponding finite transition matrix is unitary. A particular configuration is observed with a probability equal to the squared absolute value of the amplitude associated with that configuration. \square

The restriction to unitary matrices can be derived from the need of preserving the inner product as only unit vectors represent valid quantum states. A unitary matrix is the only finite dimensional square matrix that preserves the inner product. The property of a unitary matrix $U^\dagger U = \mathbb{I}$ is equivalent to the assertion that column vectors e_i of U are normalized and orthogonal;

$$\text{for all } i, j : \langle e_i | e_j \rangle = \delta_{ij}.$$

The normalization under the 2-norm, $\|e_i\|_2 = 1$, is to be compared with the local probability condition (Eq. 2.2) for a probabilistic Turing machine. Columns of a stochastic matrix are normalized under the 1-norm and for a probabilistic Turing machine this ensures that a normalized linear combination of configurations is mapped to another normalized linear combination of configurations (Eq 2.4). However, for a quantum Turing machine having normalized columns of transition matrix is not sufficient to ensure that normalized states are mapped again to normalized states. This is due to the interference phenomena.

It may happen that several different paths ℓ of computation lead to the same configuration c_x . In such a case, this configuration has an amplitude equal to the sum of amplitudes of corresponding paths. The configuration c_x is then observed with probability

$$p(c_x) = \left| \sum_{\ell} \alpha_x^{(\ell)} \right|^2, \quad \alpha_x^{(\ell)} \in \mathbb{C}. \quad (3.16)$$

We talk about a **constructive interference** if

$$\left| \sum_{\ell} \alpha_x^{(\ell)} \right|^2 \geq \sum_{\ell} \left| \alpha_x^{(\ell)} \right|^2, \quad (3.17)$$

and about a **destructive interference** if

$$\left| \sum_{\ell} \alpha_x^{(\ell)} \right|^2 < \sum_{\ell} \left| \alpha_x^{(\ell)} \right|^2. \quad (3.18)$$

For a probabilistic Turing machine, $\alpha_x^{(\ell)} \in \mathbb{R}^+$, we have only constructive interference in the form of

$$\left| \sum_{\ell} \alpha_x^{(\ell)} \right| = \sum_{\ell} \left| \alpha_x^{(\ell)} \right|, \quad (3.19)$$

which can be seen as a special case of the equation (3.17), up to the used p -norm.

Example 7 (Destructive interference). Let us have a quantum Turing machine whose transitions are described by a unitary matrix T and the computation starts from the configuration c_0 . A three-step evolution is shown in the Figure 3.4. Observe that at the second level we have non-zero probability to observe configuration c_1 . At the third level, the probability of observing c_1 is zero due to the destructive interference.

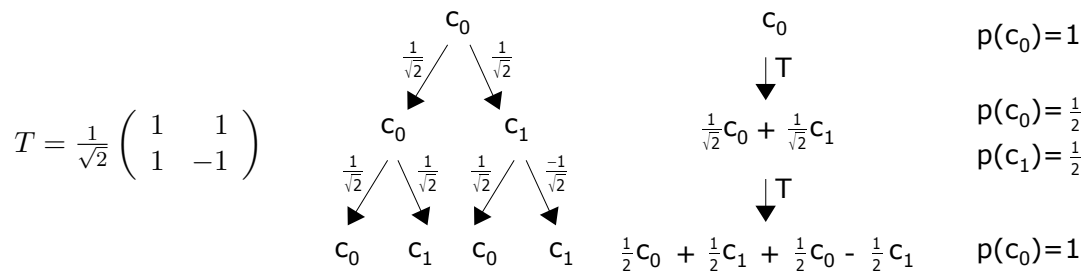


Figure 3.4: The power of destructive interference.

Definition 3.2.2 (Accepting a language on a quantum Turing machine). A quantum Turing machine QTM accepts a language L with completeness c and soundness s iff there is a configuration $c(x) = (x', q_f, m)$, where $q_f \in F$, $m \in \{L, R\}$, and the machine halts in a superposition of configurations $\sum_i \alpha_i |c_i\rangle$ such that

- for all strings $x \in L$: the probability $p = \sum_i \alpha_i |\langle c(x)|c_i\rangle|^2 > c$,
- for all strings $x \notin L$: the probability $p = \sum_i \alpha_i |\langle c(x)|c_i\rangle|^2 \leq s$.

Definition 3.2.3 (Exact or error-free quantum polynomial time, EQP). EQP is the class of decision problems efficiently solvable on a quantum Turing machine QTM with zero-error probability. A language L is in the class EQP if there exists a QTM such that

- for all $x \in L$: the machine accepts x with probability $p = 1$,
- for all $x \notin L$: the machine accepts x with probability $p = 0$,

and the time complexity is upper-bounded by some polynomial $poly(x)$. \square

The EQP class was defined in [14] and, unfortunately, there is no theory of universal quantum Turing machines for exact quantum computing. This is due to the formulation "if there exists a QTM such that ..." in the definition of the class. Therefore showing that a problem is in the EQP is of limited theoretical value. In the circuit model for quantum computing (the model is described in Section 3.3), Deutsch-Jozsa and Simon promise problems are known to have error-free efficient circuits while using only gates from a universal set of gates.

Definition 3.2.4 (Bounded-error quantum polynomial time, BQP). BQP is the class of decision problems solvable in a polynomial time on a quantum Turing machine with bounded error probability. A language L is in the class BQP if

- for all $x \in L$: the machine accepts x with probability $p > 2/3$,
- for all $x \notin L$: the machine accepts x with probability $p \leq 1/3$,

and the time complexity is upper-bounded by some polynomial $poly(x)$. \square

The class BQP is considered to be the class of feasible problems for a quantum computer. It is a quantum analogue of the class BPP. It contains problems such as factoring and the discrete logarithm problem [12], Pell's equation [18] and simulation of quantum systems having a local Hamiltonian [19]. The authors of [14, 20] proved that BQP contains BPP and is contained in PP; $BPP \subseteq BQP \subseteq PP$. The relation of BQP and NP is unknown.

Concluding remarks for quantum Turing machines

Lemma 3.2.1. Any quantum Turing machine M_1 can be simulated with constant slowdown by a quantum Turing machine M_2 which uses only real amplitudes.

Proof. Let us have a state $|\psi\rangle = \sum_k \alpha_k |k\rangle$. A complex number $\alpha = a + ib$ can be written as a two dimensional real vector:

$$\alpha = \begin{pmatrix} a \\ b \end{pmatrix} = a|0\rangle + b|1\rangle, \quad a, b \in \mathbb{R}.$$

Then for a system composed out of qubits and amplitudes $r_k \in \mathbb{R}$, we have

$$\begin{aligned} M_1 : \sum_{k \in \{0,1\}^n} \alpha_k |k\rangle &\longrightarrow M_2 : \sum_{k \in \{0,1\}^n} \begin{pmatrix} a_k \\ b_k \end{pmatrix} |k\rangle \\ &= \sum_{k \in \{0,1\}^n} a_k |0, k\rangle + b_k |1, k\rangle \\ &= \sum_{k \in \{0,1\}^{n+1}} r_k |k\rangle. \end{aligned}$$

The transition unitary matrix is changed accordingly:

$$U_{M_1} : \begin{pmatrix} \boxed{a + ib} & \cdots \\ \vdots & \ddots \end{pmatrix} \longrightarrow U_{M_2} : \begin{pmatrix} \boxed{a} & \boxed{-b} & \cdots \\ \boxed{b} & \boxed{a} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}.$$

Mutual simulation of quantum Turing machines with only real numbers is in greater details discussed in [14, 20]. Additionally, Adleman, DeMarrais and Huang in [20] proved that problems contained in BQP can be efficiently solved using only amplitudes from the set $\{0, \pm\frac{3}{5}, \pm\frac{4}{5}, \pm 1\}$. Kitaev [21] proved even a smaller set $\{0, \pm\frac{1}{\sqrt{2}}, \pm 1\}$ to be sufficient. This is to be compared with a sufficient set for a probabilistic Turing machine $\{0, \frac{1}{2}, 1\}$. In [20], the authors raised an interesting question whether $\text{BQP} = \text{BPP}_{\mathbb{R}}$, where $\text{BPP}_{\mathbb{R}}$ is appropriately defined BPP class for a probabilistic Turing machine with restricted set of real amplitudes (probabilities). An affirmative answer would imply that the assumed power of quantum Turing machine comes only from possibly negative amplitudes (destructive interference) and it is independent of the use of the 2-norm.

Proposition 3.2.2. *A quantum Turing machine can be simulated by a deterministic Turing machine.*

Proof. The evolution of a quantum system is described by an equation $|\psi_2\rangle = U|\psi_1\rangle$, where $|\psi_1\rangle, |\psi_2\rangle$ are complex column vectors and U is a unitary matrix. Thus a deterministic Turing machine can simulate the evolution by a matrix-vector multiplication.

Proposition 3.2.3. *Naive simulation of a quantum Turing machine on a deterministic Turing machine by a matrix-vector multiplication is not efficient.*

Proof. An n -qubit quantum system is described by an 2^n -dimensional complex vector. Corresponding unitary evolution operator is a 2^n -by- 2^n matrix. Thus the matrix-vector multiplication takes $O(2^{2n})$ steps. Simulation of a quantum physical system, described by a Hamiltonian H during time Δt , is even more complicated since it involves large matrix exponentiation

$$U(\Delta t) = e^{-iH\Delta t} = \sum_{k=0}^{\infty} \frac{(-iH\Delta t)^k}{k!}. \quad (3.20)$$

Lemma 3.2.4. *A quantum Turing machine with transitions restricted only to product states (no entanglement) can be simulated by a deterministic Turing machine in linear time.*

Proof. Since there is no entanglement, the qubits evolve independently of each other. A simulation of a single qubit is efficient - it involves only multiplication of two-by-two matrix with a two-dimensional vector. An n -qubit system is thus simulated in time $O(n)$.

Lemma 3.2.5. *A deterministic Turing machine can efficiently simulate a quantum Turing machine with transitions restricted only to states such that entanglement is localized in chunks. A chunk consists of at most $\log(\text{poly}(n))$ qubits and there are at most polynomially many chunks.*

Proof. Chunks can be simulated independently of each other. A simulation of one chunk is efficient. Chunks are at most polynomially many.

3.3 Quantum circuit model

The quantum circuit model is a computational model analogous to the standard Boolean circuit model. The model was introduced by D. Deutsch [22] in 1989 and called quantum computational networks. Deutsch in his work also identified a three-qubit gate, the Deutsch gate, which is universal for quantum computing. The Deutsch gate is a quantum analogue of the Toffoli gate.

A quantum circuit presents quantum gates in a sequence as they are applied to a quantum register. A quantum register comprises of individually addressable qubits which are depicted by horizontal lines (quantum wires) in a circuit diagram. The quantum register is initially prepared in the computational basis, typically in the state $|0, 0, \dots, 0\rangle$. Vertical lines denote joint operations over two or more qubits. This model allows us to express computational complexity of a given algorithm in the number of required elementary gates and/or a circuit depth. The notation for quantum circuitry is summarized in the Figure 3.5.



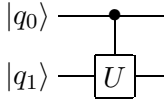





| | |
|---|--|
| Wire carrying a qubit | $ q_0\rangle$ ——— |
| Wire carrying a bit | c ===== |
| Projection onto $ 0\rangle$ and $ 1\rangle$ | $ q_0\rangle$ —  = |
| Unitary operation U | $ q_0\rangle$ —  — |
| Controlled-U operation | $ q_0\rangle$ —  — $ q_1\rangle$ —  — |
| Controlled-NOT operation | $ q_0\rangle$ —  — $ q_1\rangle$ —  — |
| Swap | $ q_0\rangle$ —  — $ q_1\rangle$ —  — |

Figure 3.5: The circuit model notation.

In 1993, Andrew Yao [23] elaborated the model and showed that any function computable in polynomial time on a quantum Turing machine has a polynomial-size quantum circuit. Two years later, several papers [24, 25, 26, 27, 28] addressed questions related to gates construction and decomposition, proofs of universality of almost all two-qubit gates, and in particular, universality of a two-qubit gate called Controlled-NOT accompanied with single qubit gates.

Since single qubit gates can be seen as continuous operators, they can be at best only approximated to some precision ε by a circuit consisting of gates from a fixed set of single qubit gates. The Solovay-Kitaev theorem [21, 29] proves that this approximation is indeed possible and induces only a polylogarithmic overhead. Importantly, Bernstein and Vazirani [14] showed that the total error caused by a sequence of imperfect (approximated) gates is at most the sum of errors of individual gates.

3.3.1 Single qubit gates

Operations on a single qubit are described by 2-by-2 unitary matrices. Some of the most important unitary matrices are Pauli matrices X, Y, Z and identity matrix I .

$$X \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \quad Y \equiv \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}; \quad Z \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}; \quad I \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (3.21)$$

Pauli matrices give rise to three useful classes of unitary matrices - the rotation operators

$$R_x(\phi) \equiv e^{-i\phi X/2} = \cos \frac{\phi}{2} I - i \sin \frac{\phi}{2} X = \begin{pmatrix} \cos \frac{\phi}{2} & -i \sin \frac{\phi}{2} \\ -i \sin \frac{\phi}{2} & \cos \frac{\phi}{2} \end{pmatrix}, \quad (3.22)$$

$$R_y(\phi) \equiv e^{-i\phi Y/2} = \cos \frac{\phi}{2} I - i \sin \frac{\phi}{2} Y = \begin{pmatrix} \cos \frac{\phi}{2} & -\sin \frac{\phi}{2} \\ \sin \frac{\phi}{2} & \cos \frac{\phi}{2} \end{pmatrix}, \quad (3.23)$$

$$R_z(\phi) \equiv e^{-i\phi Z/2} = \cos \frac{\phi}{2} I - i \sin \frac{\phi}{2} Z = \begin{pmatrix} e^{-i\phi/2} & 0 \\ 0 & e^{i\phi/2} \end{pmatrix}. \quad (3.24)$$

One reason why $R_{\{x,y,z\}}(\phi)$ matrices are referred to as rotation operators is the Bloch sphere interpretation of their actions.

Bloch sphere. A Bloch sphere is a unit sphere in Euclidean space \mathbb{R}^3 . A qubit state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ parameterized by two complex numbers $\alpha, \beta : |\alpha|^2 + |\beta|^2 = 1$, can be rewritten as

$$|\psi\rangle = e^{i\delta} \left(\cos \frac{\theta}{2} |0\rangle + e^{i\varphi} \sin \frac{\theta}{2} |1\rangle \right), \quad (3.25)$$

where $\delta, \theta, \varphi \in \mathbb{R}$. The global phase factor $e^{i\delta}$ can be ignored because it has no observable effect. The numbers θ, φ , interpreted as polar coordinates, define a point on a Bloch sphere, see Figure 3.6. We write the (x, y, z) -coordinates as a unit Bloch vector

$$\hat{r} = (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta). \quad (3.26)$$

In the Bloch sphere picture, rotation operators $R_{\{x,y,z\}}(\phi)$ corresponds to rotating a qubit state about \hat{x}, \hat{y} and \hat{z} axes by ϕ radians. Moreover, since rows and columns of a unitary matrix are orthonormal, i.e. $UU^\dagger = U^\dagger U$, arbitrary single qubit U can be written as

$$U = e^{i\delta} \begin{pmatrix} e^{i(-\alpha/2-\gamma/2)} \cos \frac{\beta}{2} & -e^{i(-\alpha/2+\gamma/2)} \sin \frac{\beta}{2} \\ e^{i(\alpha/2-\gamma/2)} \sin \frac{\beta}{2} & e^{i(\alpha/2+\gamma/2)} \cos \frac{\beta}{2} \end{pmatrix}, \quad (3.27)$$

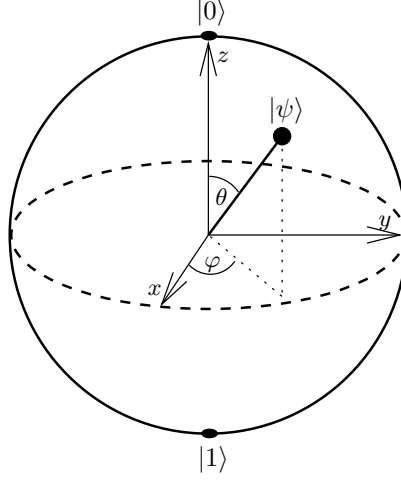


Figure 3.6: The Bloch sphere.

which actually corresponds to rotations about only two axes

$$U = e^{i\delta} R_z(\alpha) R_y(\beta) R_z(\gamma). \quad (3.28)$$

Equation (3.28) is called the Z-Y decomposition of a single qubit gate. There is also a Z-X decomposition following from the substitution of R_y rotation by R_x rotation: $R_y(\beta) = R_z(\pi/2) R_x(\beta) R_z(-\pi/2)$. In fact, arbitrary single qubit gate can be written as

$$U = e^{i\delta} R_{\hat{n}}(\alpha) R_{\hat{m}}(\beta) R_{\hat{n}}(\gamma), \quad (3.29)$$

where \hat{m} and \hat{n} are non-parallel real unit vectors in three dimensions. Therefore all single qubit gates can be seen as operators performing rotations of a qubit state on a Bloch sphere.

Important single qubit gates which appear very often in quantum circuits are the Hadamard gate (denoted H), $\pi/8$ gate (denoted T), and phase gate (denoted S):

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}; \quad T = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{pmatrix}; \quad S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}. \quad (3.30)$$

It follows that $H = e^{i\pi/2} R_z(\pi/2) R_x(\pi/2) R_z(\pi/2)$, $T = e^{i\pi/8} R_z(\pi/4)$, and $S = T^2$.

Several useful gate identities are:

- $X^2 = Y^2 = Z^2 = I$,
- $HZH = X$, $HXH = Z$, $HYH = -Y$, $HTH = R_x(\pi/4)$,
- $XR_z(\theta)X = R_z(-\theta)$, $XR_y(\theta)X = R_y(-\theta)$.

3.3.2 Two-qubit gates

The set of all gates over two qubits consists of separable and unseparable gates. Separable gates can be decomposed to single qubit gates and in this sense they are not considered to be

two-qubit gates. Separable gates can be always simulated by single qubit gates. On the other hand, results of unseparable gates related to entanglement cannot be simulated by single qubit gates. For this reason, unseparable gates are referred to as entangling gates.

DiVincenzo [24] has shown that certain two-qubit gates are already adequate in quantum computing. Barenco in [26] improved on results of DiVincenzo and original work of Deutsch [22] by identifying a large class of two-qubit gates of form

$$A(\phi, \alpha, \theta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & e^{i\alpha} \cos \theta & -ie^{i(\alpha-\phi)} \sin \theta \\ 0 & 0 & -ie^{i(\alpha+\phi)} \sin \theta & e^{i\alpha} \cos \theta \end{pmatrix}, \quad (3.31)$$

which is sufficient for construction of the Deutsch gate, $D(\theta)$, see Figure 3.7a. Parameters ϕ, α , and θ in $A(\phi, \alpha, \theta)$ and $D(\theta)$ are fixed irrational multiples of π and of each other.

Without depending on an exact proof [22] of quantum universality of the Deutsch gate, we can say the following. Due to irrationality properties that are required, a sequence of $D(\theta_0)$ gates with a chosen parameter θ_0 can generate the whole family of gates, $D(\theta)$, asymptotically. The same applies to the $A(\phi, \alpha, \theta)$ gate, where, in particular

$$A^n(\phi_0, \alpha_0, \theta_0) = A(\phi_0, n\alpha_0 \pmod{2\pi}, n\theta_0 \pmod{2\pi}).$$

The family of gates $D(\theta)$ comprises the Toffoli gate, $\text{Toffoli} = D(\pi)$, and since $A(\phi_0, \alpha_0, \theta_0)$ gate is sufficient for construction of the $D(\theta_0)$ gate, this implies that the class of gates $A(\phi, \alpha, \theta)$ is capable of classical universal reversible computation, at least. The Toffoli gate is depicted in Figure 3.7b.

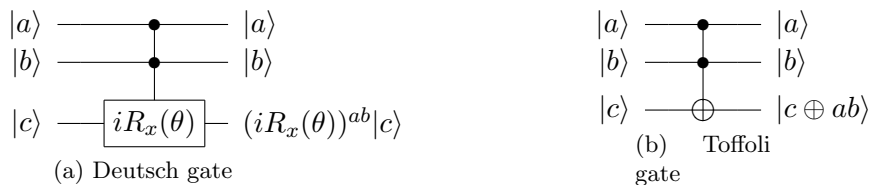


Figure 3.7: Universal gates for quantum and classical reversible computing.

To achieve the power of a quantum computer as assumed by the BQP class, the Toffoli gate must be accompanied with a gate which can manipulate the interference phenomena. The Hadamard gate or $A(\phi, \alpha, \theta)$ gates are examples of such gates. Therefore sets $\{\text{Toffoli}, H\}$ and $\{A(\phi, \alpha, \theta)\}$ constitute quantum universal sets in computational sense.

Definition 3.3.1 (Computationally universal set). A finite set of unitary gates G is called quantum computationally universal if every quantum Turing machine can be efficiently simulated to arbitrary fixed accuracy by a circuit composed out of gates from the set G . \square

Quantum computational universality of the set $\{\text{Toffoli}, H\}$ was proven by Y. Shi [30] in 2002. This result is in line with the minimal set of real amplitudes sufficient for universal quantum Turing machine; see Lemma 3.2.1 on page 38. Moreover, since the Hadamard transform is the Fourier transform over the group \mathbb{Z}_2 , quantum computing can be roughly viewed as

the Fourier transform plus a classical reversible computing. Indeed, the quantum part of the Shor factoring algorithm consists exactly of a reversible circuit calculating the function $f_{k,N}(x) = k^x \pmod{N}$ for given coprimes k and N , and the Fourier transform over the Abelian group \mathbb{Z}_{2^n} .

Universality of the $A(\phi, \alpha, \theta)$ gate as proven by Barenco [26] is also remarkable. No classical reversible universal gate with only two input and output bits is known. The Toffoli gate has three input/output wires and so has the primordial Deutsch gate. Furthermore, Deutsch, Barenco and Ekert in [27] and Lloyd [25] showed that almost any two-qubit gate is quantum universal. Thus the $A(\phi, \alpha, \theta)$ gate is not a special gate in this sense. It seems that universal computation independently of the classical or quantum framework is deeply embodied in nature.

However, since $A(\phi, \alpha, \theta)$ already contains the essence of quantum universal computing, it is useful to study its structure. A matrix representation of the gate shows a layout of form

$$\begin{pmatrix} I & 0 \\ 0 & U \end{pmatrix},$$

where U is a single qubit gate. Such a two-qubit gate applied to two qubits effectively acts like a Controlled-U gate, see Figure 3.8. The two input qubits are known as the control qubit and the target qubit. If the control qubit is set, i.e. $|c\rangle = |1\rangle$, then U is applied to the target qubit, otherwise identity gate is applied, $U^0 = I$, and thus both qubits are kept untouched.

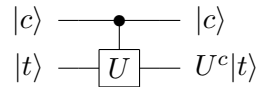


Figure 3.8: Controlled-U operation.

Due to the 'controlling' interpretation of some entangling gates they are also referred to as controlling gates. But, this term might be misleading, since it is not always clear who is controlling who. See Figure 3.9, where Hadamard gates lead to the exchange of the control.



Figure 3.9: Difficulties with interpreting who is controlling who.

Out of controlling gates, the Controlled-NOT (CNOT) is a prototypical operation. Its action is described as

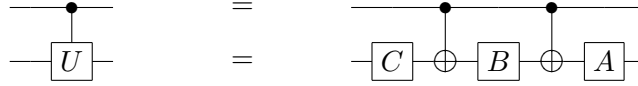
$$CNOT : |c\rangle|t\rangle \longrightarrow |c\rangle|t \oplus c\rangle, \quad (3.32)$$

and the matrix representation is

$$CNOT = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \quad (3.33)$$

The importance of the CNOT gate comes from the fact that it is relatively easy to understand its action, and moreover, there is a constructive proof that any controlled single qubit gate U , controlled- U , can be constructed using at most two CNOTs and three single qubit gates. This proof is due to Barenco *et. al.* [28]. Since the $A(\phi, \alpha, \theta)$ gate is sufficient for universal quantum computing this implies that CNOT and all single qubit gates are sufficient too.

Theorem 3.3.1 (Controlled- U simulation). *Let U be a single qubit gate. Then there exist single qubit gates A, B, C such that $ABC = I$ and $U = AXBXC$. Controlled- U is then implemented by the following circuit.*



The control qubit decides whether operation ABC or $AXBXC$ will be applied to the target qubit.

Proof. By Z-Y decomposition, $U = R_z(\alpha) R_y(\beta) R_z(\gamma)$ for some $\alpha, \beta, \gamma \in \mathbb{R}$. Set $A = R_z(\alpha) R_y(\beta/2)$, $B = R_y(-\beta/2) R_z(-(\alpha + \gamma)/2)$ and $C = R_z((-\alpha + \gamma)/2)$. Then

$$ABC = R_z(\alpha) \underbrace{R_y(\beta/2) R_y(-\beta/2)}_I \underbrace{R_z(-(\alpha + \gamma)/2) R_z((-\alpha + \gamma)/2)}_{R_z(-\alpha)} = I$$

and using the identity $I = XX$, we have

$$\begin{aligned} AXBXC &= R_z(\alpha) R_y(\beta/2) X R_y(-\beta/2) R_z(-(\alpha + \gamma)/2) X R_z((-\alpha + \gamma)/2) \\ &= R_z(\alpha) R_y(\beta/2) \underbrace{X R_y(-\beta/2) (X X)}_{R_y(\beta/2)} \underbrace{R_z(-(\alpha + \gamma)/2) X R_z((-\alpha + \gamma)/2)}_{R_z((\alpha + \gamma)/2)} \\ &= R_z(\alpha) R_y(\beta) R_z(\gamma). \end{aligned}$$

□

Barenco *et. al.* [28] in fact showed that arbitrary gate over n qubits can be constructed using only CNOTs and single qubit gates. With an assumption that arbitrary single qubit gates are available or can be efficiently approximated, the number of CNOTs basically describes the length of a circuit since at most one single qubit gate makes sense between two successive CNOTs acting on the same qubit.

The authors of [28] gave the upper bound $O(n^3 4^n)$ and the lower bound $\Omega(4^n)$ on the number of needed CNOTs for n -qubit entangling gates. Since the vast majority of all gates over n qubits are indeed entangling gates this implies that computationally interesting unitary transforms requiring only polynomially many CNOTs are rare. Clearly, those rare ones can be hardly found by automatized brute-force decomposition. The upper bound $O(n^3 4^n)$ was derived using the QR-decomposition [31]. The lower bound $\Omega(4^n)$ is a consequence of $2^n \times 2^n$ degrees of freedom of an n -qubit gate.

In 2003, Vartiainen *et. al.* [32] improved the upper bound to $O(4^n)$. Vidal and Dawson [33] studied the lower bound for the limit case, $n = 2$, and showed that arbitrary unitary transformation over two qubits can be decomposed to a circuit containing at most three

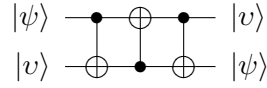


Figure 3.10: The SWAP gate implementation using three CNOT gates.

CNOTs and eight single qubit gates. An example of a gate requiring three CNOTs is the SWAP gate, see Figure 3.10.

Additional reason for choosing CNOT as an exemplary two-qubit gate is an easy entangled pair creation from a trivial product state $|0, 0\rangle$.

Example 8 (Entangled pair creation).

$$\begin{aligned}
 |0, 0\rangle &\xrightarrow{H \otimes I} \frac{1}{\sqrt{2}}(|0, 0\rangle + |1, 0\rangle) \\
 &\xrightarrow{CNOT} \frac{1}{\sqrt{2}}(|0, 0\rangle + |1, 1\rangle)
 \end{aligned}$$

3.3.3 Evolutionary universal set of gates

In order to capture most of the abstraction offered by a circuit model, it is useful to step from quantum computational universality to more strict evolutionary universality where every unitary evolution may be arbitrarily well approximated by a circuit composed out of gates from a fixed finite set. For example, the set $\{\text{Toffoli}, \text{H}\}$ is evolutionary universal for gates with real entries but it clearly cannot be evolutionary universal for gates with complex entries.

To formally define evolutionary universality we need a notion of approximation. For that purpose we define a distance function between two unitary operators U and U' as

$$E(U, U') = \|U - U'\| = \max_{\|\psi\|=1} \|(U - U')|\psi\rangle\|, \quad (3.34)$$

where the maximum is over all normalized quantum states $|\psi\rangle$, and we say that operator U' approximates operator U to within ε if

$$E(U, U') \leq \varepsilon.$$

The justification of this distance function comes from the fact that measurement outcomes are the only things that matter regardless whether U or U' was actually used. Let $M = \sum_m r_m P_m$, where $\sum_m P_m = I$, be an observable with orthogonal projectors P_m . Then $p_U(r_m)$ and $p_{U'}(r_m)$ are probabilities of obtaining the outcome r_m if U and U' were performed, respectively. The distance function $E(U, U')$ as defined by (3.34) obeys the rule that if $E(U, U')$ is small then the difference between outcome probabilities is small. In particular,

$$|p_U(r_m) - p_{U'}(r_m)| \leq 2E(U, U'). \quad (3.35)$$

Proof. Let $|\psi\rangle$ be an initial state.

$$\begin{aligned}
|p_U(r_m) - p_{U'}(r_m)| &= \left| \langle \psi | U^\dagger P_m U | \psi \rangle - \langle \psi | U'^\dagger P_m U' | \psi \rangle \right| \quad (\text{projective measurement}) \\
&= \left| \langle \psi | U^\dagger P_m U | \psi \rangle - \langle \psi | U^\dagger P_m U' | \psi \rangle + \langle \psi | U^\dagger P_m U' | \psi \rangle - \langle \psi | U'^\dagger P_m U' | \psi \rangle \right| \\
&= \left| \langle \psi | U^\dagger P_m (U - U') | \psi \rangle + \langle \psi | (U^\dagger - U'^\dagger) P_m U' | \psi \rangle \right| \\
&= \left| \langle \psi | U^\dagger P_m | \Delta \rangle + \langle \Delta | P_m U' | \psi \rangle \right| \quad (\text{substitution } |\Delta\rangle = (U - U')|\psi\rangle) \\
&\leq \left| \langle \psi | U^\dagger P_m | \Delta \rangle \right| + \left| \langle \Delta | P_m U' | \psi \rangle \right| \quad (\text{triangle inequality}) \\
&\leq \sqrt{\langle \psi | U^\dagger U | \psi \rangle \langle \Delta | P_m | \Delta \rangle} + \sqrt{\langle \psi | U'^\dagger U' | \psi \rangle \langle \Delta | P_m | \Delta \rangle} \quad (\text{Schwarz ineql.}) \\
&= 2 \left(\langle \Delta | P_m | \Delta \rangle \right)^{1/2} \quad (|\psi\rangle \text{ is normalized, and } U, U' \text{ preserve the norm}) \\
&\leq 2 \left(\sum_m \langle \Delta | P_m | \Delta \rangle \right)^{1/2} = 2 \left(\langle \Delta | \left(\sum_m P_m \right) | \Delta \rangle \right)^{1/2} = 2 \left(\langle \Delta | \Delta \rangle \right)^{1/2} \\
&= 2 \|\Delta\| = 2 \|(U - U')|\psi\rangle\| \\
&\leq 2E(U, U').
\end{aligned}$$

Definition 3.3.2 (Evolutionary universal set). A finite set of unitary gates G is called evolutionary universal if for every unitary gate U and $\varepsilon > 0$ there exists a sequence of gates $u'_1, u'_2, \dots, u'_k \in G$ and ancilla state $|a\rangle$ such that

$$\max_{\|\psi\|=1} \left\| U(|\psi\rangle) \otimes |a\rangle - U'_1 U'_2 \cdots U'_k (|\psi\rangle \otimes |a\rangle) \right\| \leq \varepsilon, \quad (3.36)$$

where U'_i gate corresponds to the gate u'_i surrounded with identity gates for an appropriate number of qubits and ancilla $|a\rangle$ is an efficiently preparable state by means of gates from G . A set G typically consists of gates acting on 1, 2 and/or 3 qubits. \square

The definition of evolutionary universal set utilizes a special form of the distance function where part of an input state is regarded as an ancilla state. The ancilla state allows to sets which are missing gates of small dimension still to be said to be evolutionary universal in all dimensions.

Well-known evolutionary universal sets include:

- $\{D(\theta)\}$, Deutsch [22],
- $\{A(\phi, \alpha, \theta)\}$, Barenco [28],
- $\{\text{Toffoli, S, H}\}$, Shor [34],
- $\{\text{Controlled-S, H}\}$, Kitaev [21],
- $\{\text{Controlled-NOT, H, T}\}$, Boykin *et. al.* [35].

Gates in these sets create only a tiny fraction of gates which can be implemented in a more or less straightforward way using current technologies. In light of evolutionary universality of almost any two-qubit gate [27, 25] it should be easy to find different evolutionary universal sets using only the most feasible gates for a given physical system. Surprisingly, a borderline

between evolutionary universal sets and sets which are not even quantum computationally universal can be very thin.

D. Gottesman and E. Knill [36, 37] identified an important class of quantum circuits which can be efficiently simulated classically. This is the class of stabilizer circuits introduced in study of quantum error-correcting codes [38]. Consequently, the set of gates which generates this class of quantum circuits cannot be quantum computationally universal.

Theorem 3.3.2 (Gottesman-Knill).

Every circuit involving only the following elements:

- *state preparation in the computational basis,*
- *gates Controlled-NOT, H, S, and the Pauli gates,*
- *and measurements of observables in the Pauli group,*

can be efficiently simulated on a classical computer.

Stabilizer circuits embrace not only all linear quantum error-correcting codes but schemes for entanglement purification [39, 40], quantum state teleportation [41], superdense coding [42], and Greenberger-Horne-Zeilinger paradox [43] too. Anders and Briegel [44] presented recently an algorithm based on graph-state formalism which simulates stabilizer circuits consisting of n qubits in time and space $O(n \log(n))$. This result makes a numerical study of the above mentioned problems easily handled and shows how subtle is the power of quantum computation.

An important figure of merit of a particular evolutionary universal set is whether its gates can operate in a fault-tolerant manner. Foundations of fault-tolerant quantum circuits were laid by P. W. Shor [34] in 1996. He suggested to perform the computation and error-correction on encoded logical qubits without a need to ever decode them, and in such a way that an error on a single physical qubit cannot propagate to other physical qubits in the same logical qubit. This imposes constraints on both the type of unitary operations that can be performed and error-correcting codes used to encode logical qubits. For example, gates that perform rotation by an irrational multiple of π cannot be realized in a fault-tolerant way [35]. Shor showed that the set {Toffoli, H, S} and CSS codes [45, 46] conform to the restrictions and that a quantum circuit with t gates can tolerate $O(1/\log^c t)$ amounts of inaccuracy and decoherence per gate for some small constant c at a cost of polylogarithmic increase of the circuit. This result was later improved to hold for all codes encompassed by the stabilizer formalism [47]. Further utilization of concatenation of codes led to the threshold theorem [38, 48, 49], which says that if the basic error rate is below some threshold, we can do arbitrary long computations. Currently the threshold is above 10^{-3} . Next to the set {Toffoli, H, S}, the Kitaev's set {Controlled-S, H} and the set {Controlled-NOT, H, T} are fault-tolerant too.

Note that general methods for classical fault-tolerant computation with noisy gates also require logarithmic increase in the size of circuits [50].

The set {Controlled-NOT, H, T} is the most commonly used evolutionary universal set. Hadamard and $\pi/8$ gates can efficiently approximate arbitrary single qubit gate U and hereby together with the Controlled-NOT are sufficient to approximate the set {Controlled-NOT, all single qubit gates}. The possibility of approximation follows from the fact that a sequence

$$HTHT = R_x(\pi/4) R_z(\pi/4) = e^{-i(\pi/8)X} e^{-i(\pi/8)Z} = R_{\hat{n}}(\theta_c) \quad (3.37)$$

produces a rotation along the axis

$$\hat{n} = (\cos \pi/8, \sin \pi/8, \cos \pi/8)$$

by a constant angle θ_c which is incommensurable with π [35]; accordingly,

$$HR_{\hat{n}}(\theta_c)H = THTH = R_{\hat{m}}(\theta_c) \quad (3.38)$$

with

$$\hat{m} = (\cos \pi/8, -\sin \pi/8, \cos \pi/8).$$

Using Equation (3.29) we know that a single qubit gate U can be written as

$$U = R_{\hat{n}}(\alpha) R_{\hat{m}}(\beta) R_{\hat{n}}(\gamma),$$

and, additionally, thanks to θ_c/π being irrational, a sequence of gates $R_{\{\hat{m}, \hat{n}\}}(\theta_c)$ can approximate any $R_{\{\hat{m}, \hat{n}\}}(\theta)$ gate:

$$R_{\{\hat{m}, \hat{n}\}}^k(\theta_c) = R_{\{\hat{m}, \hat{n}\}}(k\theta_c \pmod{2\pi}). \quad (3.39)$$

Thus for given $\xi > 0$ there exists a positive integer k such that

$$E(R_{\{\hat{m}, \hat{n}\}}(\theta), R_{\{\hat{m}, \hat{n}\}}^k(\theta_c)) \leq \xi. \quad (3.40)$$

The question now is how the error accumulates while using approximate gates. Suppose that gates in a sequence $U_1 U_2$ are replaced with their approximate version U'_1 and U'_2 . Then

$$\begin{aligned} E(U_1 U_2, U'_1 U'_2) &= \|U_1 U_2 - U'_1 U'_2\| \\ &= \|U_1 U_2 - U_1 U'_2 + U_1 U'_2 - U'_1 U'_2\| \\ &\leq \|U_1(U_2 - U'_2)\| + \|(U_1 - U'_1)U'_2\| \quad (\text{triangle inequality}) \\ &\leq \|U_1\| \|U_2 - U'_2\| + \|U_1 - U'_1\| \|U'_2\| \quad (\text{submultiplicative prop. of Eq. (3.34)}) \\ &= E(U_2, U'_2) + E(U_1, U'_1), \quad (\|U_1\| = \|U'_2\| = 1) \end{aligned}$$

and, by induction, we can conclude that the error incurred by approximate gates adds at most linearly:

$$E(U_1 U_2 \dots U_m, U'_1 U'_2 \dots U'_m) \leq \sum_{j=1}^m E(U_j, U'_j). \quad (3.41)$$

Thus if each gate in a sequence with t gates is accurate to $\xi = \varepsilon/t$ then the sequence is accurate to ε . In conclusion, there exist positive integers a, b, c such that

$$E(U, U') = E(R_{\hat{n}}(\alpha) R_{\hat{m}}(\beta) R_{\hat{n}}(\gamma), R_{\hat{n}}^a(\theta_c) R_{\hat{m}}^b(\theta_c) R_{\hat{n}}^c(\theta_c)) \leq \varepsilon. \quad (3.42)$$

Regarding the rate of convergence, a sequence of angles θ_c fills in the interval $[0, 2\pi)$ roughly uniformly and therefore it takes at most $O(1/\varepsilon)$ Hadamard and $\pi/8$ gates to approximate a single qubit gate to accuracy ε . Surprisingly, the Solovay-Kitaev theorem discovered by Solovay [29] and Kitaev [21] proves much faster convergence.

Theorem 3.3.3 (Solovay-Kitaev). *Let $SU(d)$ denote the group of d -dimensional unitary matrices with unit determinant. Let G be a finite set of gates in $SU(d)$ containing its own inverse such that the group generated by G is dense in $SU(d)$. Let $\varepsilon > 0$ be given. There is a constant $c \approx 4$ such that for any $U \in SU(d)$ there exists a finite sequence $U' = U'_1 U'_2 \dots U'_l$ of gates from G of length $l = O(\log^c(1/\varepsilon))$ such that $E(U, U') \leq \varepsilon$.*

The Solovay-Kitaev theorem for $SU(2)$, the set of all single qubit gates up to the overall phase, manifestly applies to Hadamard and $\pi/8$ gates: $\{H, H^\dagger = H, T, T^\dagger = T^7\} = \{H, T\} \in SU(2)$, up to the overall phase, and according to Equation (3.42) the set $\{H, T\}$ generates a dense subgroup in $SU(2)$.

The constant c is different for different approaches to the proof of the theorem. The algorithm described in [51] produces a sequence of length $O(\log^{3+\delta}(1/\varepsilon))$, where δ is a chosen positive real number. With additional restrictions imposed on G and usage of ancilla qubits there is an algorithm described in [52] which produces a sequence of length $O(\log^2(1/\varepsilon) \log(\log(1/\varepsilon)))$. The running time of both algorithms is polylogarithmic since they effectively produce a gate of the output sequence per step. Moreover, the authors of [53] gave a non-constructive proof that for a suitable G a sequence of length $O(\log(1/\varepsilon))$ can be achieved. They also showed this result to be the best possible. However, it is not known which sets G are suitable and whether the compiling algorithm will be efficient.

In conclusion, the Solovay-Kitaev theorem implies that an approximate circuit simulating a circuit with n qubits, $f(n)$ CNOTs and consequently $O(f(n))$ single qubit gates to an accuracy ε requires $O(f(n) \log^c(f(n)/\varepsilon))$ gates from a universal set.

Chapter 4

Phase estimation and applications

4.1 Fast quantum algorithms

The discovery of the fast quantum factoring algorithm [12] by Shor in 1994 was immediately followed by a more general work of A. Yu. Kitaev [54], who using group-theoretic approach gave an efficient algorithm for the abelian stabilizer problem. This problem subsumes the Deutsch-Jozsa problem [15], the Simon problem [16], factoring & discrete logarithms and many other problems as particular instances. A study of techniques used by Shor and Kitaev showed that the discovery of an efficient circuit for quantum Fourier transform over finite abelian groups enables efficient solutions to the abelian Hidden Subgroup Problem (HSP) and the approach can be described as quantum Fourier sampling. For example, an efficient quantum algorithm for solving Pell's equation [18] has been found using this framework.

Definition 4.1.1 (Abelian hidden subgroup problem). Let G be a finite abelian group, X a finite set, and $f : G \rightarrow X$ a function such that there exists a subgroup $H \subset G$ for which f separates cosets of H , i.e. $f(g_1) = f(g_2)$ if and only if g_1 and g_2 are members of the same coset gH . By querying a quantum oracle for performing the unitary transform $U_f|g\rangle|x\rangle = |g\rangle|x \oplus f(g)\rangle$, for $g \in G, x \in X$, and \oplus an appropriately chosen binary operation on X , determine a generating set for H .

A central challenge is to determine the hidden subgroup in time $O(\text{poly}(\log |G|))$ including encoding of elements of G and X in terms of bitstrings, oracle calls, quantum circuit for (optimal) measurement and any needed classical postprocessing time. In the case of abelian groups, the key point is that an efficient circuit for the Fourier transform is known and the transform performs mapping from the group to its dual while respecting subgroups and cosets.

Let G be a finite abelian group. The characters of G are homomorphisms $\varkappa : G \rightarrow \mathbb{C} \setminus \{0\}$, that is mapping such that $\varkappa(g_1 + g_2) = \varkappa(g_1)\varkappa(g_2)$ assuming additive structure of G and multiplicative structure of $\mathbb{C} \setminus \{0\}$. The characters form a group, called the dual group \hat{G} ,

$|\hat{G}| = |G|$, and using the characters the Fourier transform over the group G is given by:

$$FT_G = \frac{1}{\sqrt{|G|}} \sum_{i,j} \chi_j(i) |j\rangle \langle i|, \quad \text{i.e.} \quad |g\rangle \xrightarrow{FT_G} \frac{1}{\sqrt{|G|}} \sum_j \chi_j(g) |j\rangle. \quad (4.1)$$

For the case $G = \mathbb{Z}_N$, the characters are described as $\chi_j(k) = e^{-2\pi i j k / N} = \omega^{jk}$.

Corresponding to any subgroup $H \subseteq G$, there is a set of elements in G orthogonal to H :

$$H^\perp = \{g \in G \mid \chi_g(h) = 1 \text{ for all } h \in H\}. \quad (4.2)$$

The set of elements H^\perp forms the orthogonal subgroup $H^\perp \subseteq \hat{G}$ to the subgroup $H \subseteq G$, where $|H^\perp| = |G|/|H|$. The Fourier transform over G , FT_G , maps a subgroup H to its orthogonal subgroup H^\perp . In particular, an equally weighted superposition on H is mapped to equally weighted superposition over H^\perp :

$$\frac{1}{\sqrt{|H|}} \sum_{h \in H} |h\rangle \xrightarrow{FT_G} \frac{\sqrt{|H|}}{\sqrt{|G|}} \sum_{k \in H^\perp} |k\rangle = \frac{1}{\sqrt{|H^\perp|}} \sum_{k \in H^\perp} |k\rangle, \quad (4.3)$$

and for a particular coset of H :

$$|gH\rangle = \frac{1}{\sqrt{|H|}} \sum_{h \in H} |gh\rangle \xrightarrow{FT_G} \frac{1}{\sqrt{|H^\perp|}} \sum_{k \in H^\perp} \chi_g(k) |k\rangle. \quad (4.4)$$

The standard approach for solving the abelian HSP is to create a random coset state, perform a quantum Fourier transform and sample the resulting state from H^\perp . Since $(H^\perp)^\perp = H$, determining a generating set for H^\perp determines H uniquely. For example, in the factoring algorithm we have $f_{k,N} : x \rightarrow k^x \pmod{N}$, underlying group is \mathbb{Z}_M , where $M \gg N$, the hidden subgroup is $H = \langle r \rangle = \{0, r, 2r, \dots\}$, and $H^\perp = \langle M/r \rangle = \{0, M/r, 2M/r, \dots\}$.

4.1.1 Algorithm for solving the abelian HSP

1. Create a random coset state

Create a superposition over all elements of G in the first register and call the oracle:

$$\begin{aligned} |0, 0\rangle &\xrightarrow{FT_G} \frac{1}{\sqrt{|G|}} \sum_{g \in G} |g, 0\rangle \\ &\xrightarrow{U_f} \frac{1}{\sqrt{|G|}} \sum_{g \in G} |g, f(g)\rangle. \end{aligned}$$

Next, measure the second register. This results in a uniformly chosen coset $g_i H$ of H in the first register:

$$\frac{1}{\sqrt{|H|}} \sum_{h \in H} |g_i h, f(g_i)\rangle.$$

The second register is no longer in use now.

2. Fourier sampling

Perform the Fourier transform on the first register again:

$$\frac{1}{\sqrt{|H|}} \sum_{h \in H} |g_i h\rangle \xrightarrow{FT_G} \frac{1}{\sqrt{|H^\perp|}} \sum_{k \in H^\perp} \chi_{g_i}(k) |k\rangle.$$

The phase $\chi_{g_i}(k)$ has no effect on a subsequent measurement and uniformly random $k \in H^\perp$ is obtained.

3. H reconstruction

Reconstruct H from polynomially many samples.

Note. Since the measured value $f(g_i)$ from the second register is actually never used, this measurement might in fact be skipped. In the first step of the algorithm, we prepare the state

$$|0, 0\rangle \xrightarrow{FT_G} \frac{1}{\sqrt{|G|}} \sum_{g \in G} |g, 0\rangle \xrightarrow{U_f} \frac{1}{\sqrt{|G|}} \sum_{g \in G} |g, f(g)\rangle,$$

and after discarding (tracing out) the second register, due to the promise that f separates cosets, the first register is then in a mixed state

$$\rho_H = \frac{|H|}{|G|} \sum_{\tilde{g}} |\tilde{g}H\rangle \langle \tilde{g}H| \quad \text{with} \quad |\tilde{g}H\rangle = \frac{1}{\sqrt{|H|}} \sum_{h \in H} |\tilde{g}h\rangle,$$

where \tilde{g} is a coset representative. The state ρ_H is called the hidden subgroup state. Implicit measurement, however, simplifies the analysis.

4.1.2 New quantum algorithms

Attempts to find new fast quantum algorithms which might make some of known hard problems tractable naturally focus on various generalizations of the abelian HSP and underlying techniques. Efficient algorithms for non-abelian HSP are of particular importance because the HSP over the symmetric group encompasses the graph isomorphism problem and the HSP over the dihedral group subsumes the shortest vector in a lattice problem.

Efficient algorithms for the HSP over non-abelian groups such as 'almost abelian' groups [55] or 'near Hamiltonian' groups [56] have been successfully derived from the abelian approach, but in general straightforward quantum Fourier sampling fails since there is no dual group in the non-abelian case. An important result was the result of Ettinger *et. al.* [57] who showed that the query complexity of non-abelian HSP is polynomial as well as for abelian HSP. This means that there exists a measurement (a unitary transform followed by a projective measurement) over polynomially many samples of the hidden subgroup state ρ_H which reveals information sufficient to reconstruct H . It was shown that for certain groups (including the dihedral group [58]) the measurement may be performed by repeated measurement on a single register containing the hidden subgroup state ρ_H while other groups (including the symmetric group [59, 60]) require measurement across multiple copies of ρ_H . The problem is that the measurement even if correctly identified does not always have to have an efficient implementation. Moreover, classical postprocessing is not guaranteed to be efficient as well.

Measurement across multiple copies of the hidden subgroup state was for the first time used by Kuperberg [61] who gave a subexponential but not polynomial algorithm for the dihedral HSP. Utilizing measurement across multiple registers likewise, Bacon *et. al.* [62] gave recently an efficient algorithm for solving the HSP over certain semidirect product groups of the form $\mathbb{Z}_p^r \rtimes \mathbb{Z}$, for a fixed r (including the Heisenberg group, $r = 2$) and prime p . Bacon *et. al.* achieved this result in the framework of so-called 'pretty good measurement' (PGM) [63] and later reinterpreted the result using the Clebsch-Gordan transform [64]. PGM turned out to be an optimal measurement for the dihedral HSP, however, no efficient implementation is known in this case. The Clebsch-Gordan transform is the key component in the recently discovered efficient circuit for the Schur transform [65]. The Schur transform is known to be useful in tasks such as an estimate of the spectrum of a density operator [66] or communication without a shared reference frame [67]. Whether the Schur transform opens up for advances in non-abelian HSP is an open question.

An alternative effort in finding new quantum algorithms focuses on what kind of other structures may be efficiently revealed on a quantum computer. Partial results along this lines are discussed for the hidden shift problem [68, 69] and hidden nonlinear structures over finite fields [70].

4.2 Quantum Fourier transform based phase estimation

While theory based on a quantum Fourier sampling provides a rather unifying view of the core of many quantum algorithms, the actual problem-solution approach is often quite different. One of the most fruitful approaches up to date is to cast the problem as an estimation of an eigenvalue of a certain unitary operator. Additional insight is that since unitary operators preserve the 2-norm the eigenvalues are of the form $\lambda = e^{2\pi i\phi}$ for some phase $\phi \in [0, 1)$. Hence the goal is to estimate the phase.

In general, designing efficient and numerically stable algorithms for finding the eigenvalues of an operator is one of the most important problems in linear algebra. Eigenvalues of small operators can be found by finding roots of characteristic polynomial, but this method cannot be used for large operators, see Abel-Ruffini theorem [71]. One implication of the theorem is an expected iterative nature for any general eigenvalue algorithm. A stable classical method for finding eigenvalues is the QR-algorithm [72] which iteratively executes the QR-decomposition [31]. For an $N \times N$ operator U , each QR iteration requires $O(N^3)$ operations. Other methods utilize calculation of U, U^2, U^3, \dots (power iteration) and the overall complexity of these algorithms scale as $O(N^3)$ too. Thus eigenvalues (or even a single eigenvalue) of an operator of size $2^n \times 2^n$ are found inefficiently with respect to n . A quantum computer can often perform better given certain constraints on U . These constraints relate to the existence of efficient quantum circuits (efficient with respect to the size of the input n) for performing controlled powers U^{2^k} . Such constraints are not known to be helpful for classical algorithms.

4.2.1 Quantum phase estimation

The main idea in quantum phase estimation is to consider a register in a basis containing eigenvectors of the unitary operator related to the problem being solved. From the definition of eigenvalues and eigenvectors, when an operator U acts on one of its eigenvectors the result is a scaling of this vector by a corresponding eigenvalue:

$$U|\psi\rangle = \lambda|\psi\rangle.$$

In another words, the eigenvalue is kicked in front of the register. The trick behind quantum phase estimation is to exploit the eigenvalue kick-back in combination with the ability to efficiently perform controlled powers U^{2^k} and transfer the phase $2\pi(2^k\phi)$ into the relative phase of qubit in another register. Since U is a linear operator, the same concept applies to an arbitrary input superposition of eigenvectors of U .

The very basic scheme which utilizes only a single controlled- U gate is shown in the Figure 4.1. The upper line is the ancillary qubit which is measured, and the lower line represents space of n qubits on which U operates. Initially the ancillary qubit is set to $|0\rangle$ and the lower line register to an eigenstate $|\psi\rangle$ of the operator U with eigenvalue $e^{2\pi i\phi}$.

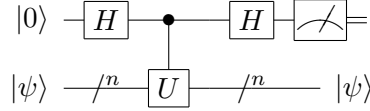


Figure 4.1: Eigenvalue kick-back circuit.

The process of estimation can be divided into two stages. At the first stage, the eigenvalue $e^{2\pi i\phi}$ is kicked in front of the $|1\rangle$ component of the first qubit:

$$\begin{aligned} |0\rangle|\psi\rangle &\xrightarrow{H\otimes I} \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle|\psi\rangle + |1\rangle|\psi\rangle) \\ &\xrightarrow{c-U} \frac{1}{\sqrt{2}}(|0\rangle|\psi\rangle + |1\rangle U|\psi\rangle) = \frac{1}{\sqrt{2}}(|0\rangle|\psi\rangle + e^{2\pi i\phi}|1\rangle|\psi\rangle) \\ &= \frac{1}{\sqrt{2}}(|0\rangle + e^{2\pi i\phi}|1\rangle)|\psi\rangle. \end{aligned}$$

The second stage is to apply the Hadamard transform to the first qubit in order to make the phase an observable quantity in the standard basis:

$$\frac{1}{\sqrt{2}}(|0\rangle + e^{2\pi i\phi}|1\rangle)|\psi\rangle \xrightarrow{H\otimes I} \frac{1}{2} \left((1 + e^{2\pi i\phi})|0\rangle + (1 - e^{2\pi i\phi})|1\rangle \right) |\psi\rangle.$$

The outcome '0' is observed with probability

$$P_0 = \left| \frac{1 + e^{2\pi i\phi}}{2} \right|^2 = \frac{1}{4} |(1 + \cos 2\pi\phi + i \sin 2\pi\phi)|^2 = \cos^2 \pi\phi. \quad (4.5)$$

Accordingly, the outcome '1' is observed with probability $P_1 = 1 - P_0 = \sin^2 \pi\phi$. The whole procedure is often modelled by a diagram shown in the Figure 4.2. The gate u_ϕ represents a

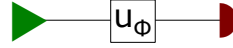


Figure 4.2: Schematic diagram of the eigenvalue kick-back circuit.

phase shift caused by some physical process and maps a prescribed single qubit state $|0\rangle + |1\rangle$ to $|0\rangle + e^{2\pi i\phi}|1\rangle$.

The question is now how to determine the phase in the most efficient way. By repeating the simple eigenvalue kick-back procedure M -times, P_0 and consequently ϕ can be determined to an accuracy of $1/\sqrt{M}$. This result follows from the central limit theorem: for large M the error on average decreases as Δ/\sqrt{M} , where Δ^2 is the variance of the measurement results associated to the physical apparatus. Thus one needs go to through at least $M \sim 2^{2m}$ independent rounds to obtain m accurate binary digits of ϕ . The scaling $1/\sqrt{M}$ is known as the standard quantum limit (SQL). However, using an advanced setup, the scaling may be improved up to $1/M$ (the Heisenberg limit). Giovannetti *et. al.* [73] showed that after employing the phase shift u_ϕ at total M -times the scaling $1/M$ is the general lower bound to the estimation error. A setup which achieves the Heisenberg limit is shown in the Figure 4.3. This setup with a register of ancillary qubits exploits namely a sequential applying of the phase shift u_ϕ to the same qubit. Dam *et. al.* [74] proved that this scheme is indeed optimal.

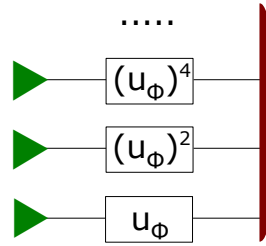


Figure 4.3: Advanced setup for the phase estimation.

4.2.2 Advanced scheme for phase estimation

The algorithm based on the powers of the phase shift is called the phase estimation algorithm (PEA). The PEA was for the first time used by Kitaev [54] in his work on the abelian stabilizer problem and the modern formulation of the PEA based on the quantum Fourier transform is due to Cleve *et. al.* [75]. The phase estimation problem itself is defined as follows.

Definition 4.2.1 (Phase estimation problem). Let U be a $2^n \times 2^n$ unitary transform. Let $|\psi_1\rangle, |\psi_2\rangle, \dots, |\psi_{2^n}\rangle$ denote eigenvectors of U with corresponding eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_{2^n}$, where λ_j is of the form $e^{2\pi i\phi_j}$, i.e.,

$$U = \sum_j e^{2\pi i\phi_j} |\psi_j\rangle\langle\psi_j|.$$

Given a set of black-boxes (oracles) capable of preparing a particular eigenstate $|\psi\rangle$ and performing the controlled- U^{2^k} operation for $k = 0, 1, 2, \dots, m-1$, efficiently determine phase $\phi \in [0, 1)$ to the precision of order $1/2^m$ with reasonable high probability. \square

A proof that the advanced scheme (Figure 4.3) for phase estimation reaches the Heisenberg limit follows. Right before the measurement the state of m ancillary qubits can be described as:

$$\begin{aligned}
& \frac{1}{\sqrt{2^m}} \bigotimes_{l=m-1}^2 \left(|0\rangle + e^{2\pi i(2^l \phi)} |1\rangle \right) \otimes \left(|0\rangle + e^{2\pi i(2^1 \phi)} |1\rangle \right) \otimes \left(|0\rangle + e^{2\pi i(2^0 \phi)} |1\rangle \right) = \\
& \frac{1}{\sqrt{2^m}} \bigotimes_{l=m-1}^2 \left(|0\rangle + e^{2\pi i(2^l \phi)} |1\rangle \right) \otimes \left(|00\rangle + e^{2\pi i\phi} |01\rangle + e^{2\pi i2\phi} |10\rangle + e^{2\pi i3\phi} |11\rangle \right) = \\
& \frac{1}{\sqrt{2^m}} \bigotimes_{l=m-1}^2 \left(|0\rangle + e^{2\pi i(2^l \phi)} |1\rangle \right) \otimes \sum_{k=0}^3 e^{2\pi i k \phi} |k\rangle = \\
& \frac{1}{\sqrt{2^m}} \sum_{k=0}^{2^m-1} e^{2\pi i k \phi} |k\rangle.
\end{aligned} \tag{4.6}$$

In order to prepare this superposition, the phase shift u_ϕ has been employed $\sum_{l=0}^{m-1} 2^l = 2^m - 1 \cong 2^m$ -times. Now, the essential role of the (inverse) quantum Fourier transform is the ability to perform the transformation

$$\frac{1}{\sqrt{2^m}} \sum_{k=0}^{2^m-1} e^{2\pi i k \phi} |k\rangle \longrightarrow |j\rangle, \tag{4.7}$$

where j is a non-negative integer and $j/2^m$ is a good estimator for ϕ with high probability.

Let us use the following notation:

$$|\nu\rangle = \frac{1}{\sqrt{2^m}} \sum_{k=0}^{2^m-1} e^{2\pi i k \phi} |k\rangle, \quad \text{and} \tag{4.8}$$

$$QFT^\dagger : |k\rangle \longrightarrow \frac{1}{\sqrt{2^m}} \sum_{j=0}^{2^m-1} e^{-2\pi i k j / 2^m} |j\rangle. \tag{4.9}$$

Then,

$$QFT^\dagger |\nu\rangle = \frac{1}{\sqrt{2^m}} \sum_{k=0}^{2^m-1} e^{2\pi i k \phi} \left(\frac{1}{\sqrt{2^m}} \sum_{j=0}^{2^m-1} e^{-2\pi i k j / 2^m} |j\rangle \right) \tag{4.10}$$

$$= \frac{1}{2^m} \sum_{k=0}^{2^m-1} \sum_{j=0}^{2^m-1} e^{2\pi i k(\phi - j/2^m)} |j\rangle \quad (\text{change the order of sums}) \tag{4.11}$$

$$= \sum_{j=0}^{2^m-1} \left(\frac{1}{2^m} \sum_{k=0}^{2^m-1} e^{2\pi i k(\phi - j/2^m)} \right) |j\rangle. \tag{4.12}$$

Therefore the probability to measure a particular outcome j is

$$p_j = \left| \frac{1}{2^m} \sum_{k=0}^{2^m-1} e^{2\pi i k(\phi - j/2^m)} \right|^2. \quad (4.13)$$

Figure 4.4 shows the output probability distribution for $m = 3$ and various ϕ .

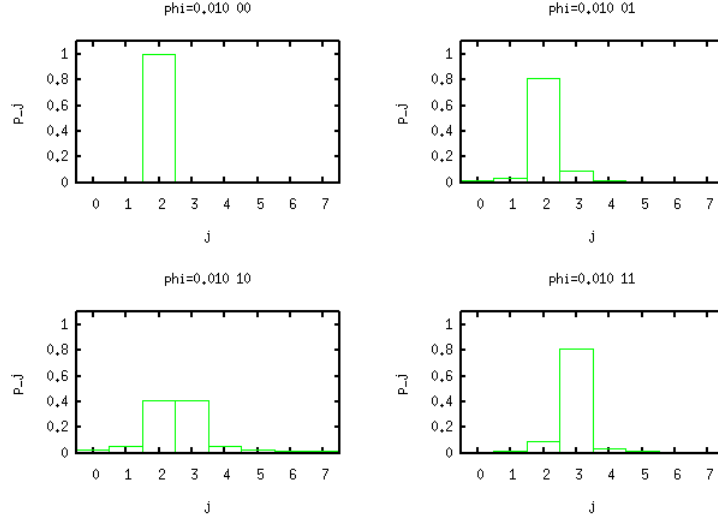


Figure 4.4: Output probability distribution for $m = 3$ and various ϕ .

In order to analyze (Eq. 4.13) and to understand the Figure 4.4 let us write ϕ as $\phi = \tilde{\phi} + \delta 2^{-m}$, where $\tilde{\phi} = 0.\phi_1\phi_2\dots\phi_m$ denotes the first m bits of the binary expansion and $0 \leq \delta < 1$ is a remainder. A special case where $\delta = 0$ means that ϕ has a binary expansion with no more than m bits, i.e. $\phi = \tilde{\phi} = j/2^m$ for some $j \in \{0, 1, \dots, 2^m - 1\}$. Given $\delta = 0$, the probability of observing j such that $j/2^m = \phi$ is

$$p_j(\phi = j/2^m) = \left| \frac{1}{2^m} \sum_{k=0}^{2^m-1} e^{2\pi i k(\phi - j/2^m)} \right|^2 = \left| \frac{1}{2^m} \sum_{k=0}^{2^m-1} 1 \right|^2 = 1. \quad (4.14)$$

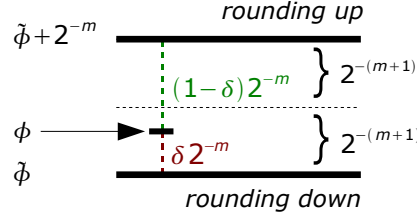
Thus the Fourier transform deterministically reveals the exact phase in this case. In other cases, $0 < \delta < 1$, the best we can hope for is to observe

$$j \begin{cases} \text{such that } j/2^m = \tilde{\phi} = 0.\phi_1\phi_2\dots\phi_m & \text{(rounding down),} \\ \text{such that } j/2^m = \tilde{\phi} + 2^{-m} & \text{(rounding up),} \end{cases}$$

with sufficiently large probability. Both estimates $\tilde{\phi}$ and $\tilde{\phi} + 2^{-m}$ guarantee error smaller than $1/2^m$. Figure 4.5 details the situation.

To determine the lower bound of the probability to observe either $\tilde{\phi}$ or $\tilde{\phi} + 2^{-m}$, we simplify (Eq. 4.13) using the formula for the sum of a geometric series and trigonometric identities:

$$p_j = \left| \frac{1}{2^m} \sum_{k=0}^{2^m-1} e^{2\pi i k(\phi - j/2^m)} \right|^2 = \frac{1}{2^{2m}} \left| \frac{1 - e^{2\pi i(2^m\phi - j)}}{1 - e^{2\pi i(\phi - j/2^m)}} \right|^2 = \frac{1}{2^{2m}} \frac{\sin^2(\pi(2^m\phi - j))}{\sin^2(\pi(\phi - j/2^m))}. \quad (4.15)$$

Figure 4.5: Two closest estimators of the phase ϕ .

Using the substitution $\phi = \tilde{\phi} + \delta 2^{-m}$, we get

$$P_{down} = p_j(j/2^m = \tilde{\phi}) = \frac{1}{2^{2m}} \frac{\sin^2(\pi\delta)}{\sin^2(\pi\delta 2^{-m})}, \quad (4.16)$$

and

$$P_{up} = p_j(j/2^m = \tilde{\phi} + 2^{-m}) = \frac{1}{2^{2m}} \frac{\sin^2(\pi(1-\delta))}{\sin^2(\pi(1-\delta)2^{-m})}. \quad (4.17)$$

The total probability is then $P = P_{down} + P_{up}$. Corresponding function plots for $0 < \delta < 1$ and $m = 10$ are shown in the Figure 4.6.

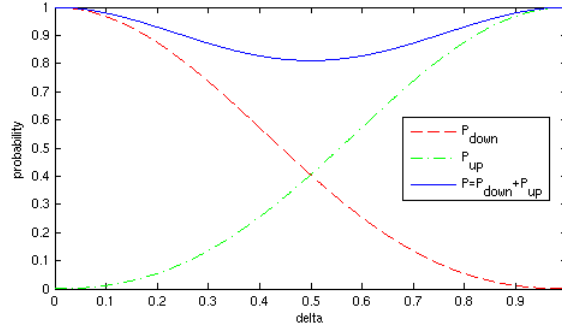


Figure 4.6: Success probability of observing a good estimator.

The success probability P decreases monotonically for increasing m and in the limit $m \rightarrow \infty$, the lower bound is

$$P(\delta = 1/2) = P_{down}(\delta = 1/2) + P_{up}(\delta = 1/2) = \frac{4}{\pi^2} + \frac{4}{\pi^2} > 0.81. \quad (4.18)$$

Here, I provide sketch of a simple proof which is conceptually different from the conventional proof based on the ratio of the minor arc length to the chord length [4, 75].

$$\delta > \sin(\delta), \quad \text{for } 0 < \delta, \quad (\text{follows from the Taylor series of } \sin \delta) \quad (4.19a)$$

$$\pi\delta > \sin(\pi\delta), \quad (4.19b)$$

$$\frac{\pi\delta}{2^m} > \sin\left(\frac{\pi\delta}{2^m}\right), \quad \text{for } m = 0, 1, 2, \dots \quad (4.19c)$$

$$\pi\delta > 2^m \sin\left(\frac{\pi\delta}{2^m}\right), \quad (4.19d)$$

$$P_{down}(0 < \delta \leq 1/2) = \left(\frac{\sin(\pi\delta)}{2^m \sin\left(\frac{\pi\delta}{2^m}\right)}\right)^2 \geq \frac{\sin^2(\pi\delta)}{(\pi\delta)^2} \geq \frac{\sin^2\left(\frac{\pi}{2}\right)}{\left(\frac{\pi}{2}\right)^2} = \frac{4}{\pi^2}, \quad (4.20)$$

$$P_{up}(1/2 \leq \delta < 1) = \left(\frac{\sin(\pi\delta)}{2^m \sin\left(\frac{\pi\delta}{2^m}\right)}\right)^2 \geq \frac{\sin^2(\pi\delta)}{(\pi\delta)^2} \geq \frac{\sin^2\left(\frac{\pi}{2}\right)}{\left(\frac{\pi}{2}\right)^2} = \frac{4}{\pi^2}. \quad (4.21)$$

□

In conclusion, the advanced scheme (Figure 4.3) can be used to determine the phase with precision of order $1/2^m$ after the phase shift u_ϕ has been used at total 2^m times. Thus the advanced scheme reaches the Heisenberg limit. In particular, the accuracy of order $1/2^m$ is reached with a success probability > 0.81 independently of m .

In contexts where a higher probability of success is desirable, the success probability can be amplified to $1 - \varepsilon$ for any $\varepsilon > 0$ by inflating m to $m' = m + O(\log(\frac{1}{\varepsilon}))$ and keeping only m most significant bits. For a proof see Cleve *et. al.* [75]. A different approach, which avoids implementing U^{2^k} for $k > m$, is to repeat the whole procedure a number of times, choosing the most frequent result.

4.2.3 Efficient circuit for the quantum Fourier transform

The next step in the search for efficient algorithms solving instances of the phase estimation problem is to find an efficient circuit for the quantum Fourier transform over \mathbb{Z}_{2^m} . An approach which is essentially the inverse of the derivation used in (Eq. 4.6) leads us to

$$QFT : |j\rangle \rightarrow \frac{1}{\sqrt{2^m}} \sum_{k=0}^{2^m-1} e^{2\pi i j k / 2^m} |k\rangle \quad \left| \begin{array}{l} k = \sum_{l=1}^m k_l 2^{m-l}, \quad k_l \in \{0, 1\}, \end{array} \right. \quad (4.22)$$

$$= \frac{1}{\sqrt{2^m}} \sum_{k=0}^{2^m-1} e^{2\pi i j \left(\sum_{l=1}^m k_l 2^{m-l}\right) / 2^m} |k\rangle \quad \left| \begin{array}{l} M = 2^m, \quad |k\rangle = \bigotimes_{l=1}^m |k_l\rangle, \end{array} \right. \quad (4.23)$$

$$= \frac{1}{\sqrt{M}} \sum_{k_1=0}^1 \dots \sum_{k_m=0}^1 e^{2\pi i j \sum_{l=1}^m k_l 2^{-l}} |k_1, k_2, \dots, k_m\rangle \quad (4.24)$$

$$= \frac{1}{\sqrt{M}} \sum_{k_1=0}^1 \dots \sum_{k_m=0}^1 e^{2\pi i j (k_1 2^{-1} + k_2 2^{-2} + \dots + k_m 2^{-m})} |k_1, k_2, \dots, k_m\rangle \quad (4.25)$$

$$= \frac{1}{\sqrt{M}} \sum_{k_1=0}^1 \dots \sum_{k_m=0}^1 e^{2\pi i j k_1 2^{-1}} e^{2\pi i j k_2 2^{-2}} \dots e^{2\pi i j k_m 2^{-m}} |k_1, k_2, \dots, k_m\rangle \quad (4.26)$$

$$= \frac{1}{\sqrt{M}} \sum_{k_1=0}^1 \dots \sum_{k_m=0}^1 \bigotimes_{l=1}^m e^{2\pi i j k_l 2^{-l}} |k_l\rangle \quad (4.27)$$

$$= \frac{1}{\sqrt{M}} \sum_{k_1=0}^1 \dots \sum_{k_m=0}^1 \bigotimes_{l=1}^{m-1} e^{2\pi i j k_l 2^{-l}} |k_l\rangle \otimes e^{2\pi i j k_m 2^{-m}} |k_m\rangle \quad (4.28)$$

$$= \frac{1}{\sqrt{M}} \sum_{k_1=0}^1 \dots \sum_{k_{m-1}=0}^1 \bigotimes_{l=1}^{m-1} e^{2\pi i j k_l 2^{-l}} |k_l\rangle \otimes \sum_{k_m=0}^1 e^{2\pi i j k_m 2^{-m}} |k_m\rangle \quad (4.29)$$

$$= \frac{1}{\sqrt{M}} \bigotimes_{l=1}^m \sum_{k_l=0}^1 e^{2\pi i j k_l 2^{-l}} |k_l\rangle \quad (4.30)$$

$$= \frac{1}{\sqrt{M}} \bigotimes_{l=1}^m \left(e^{2\pi i j (0) 2^{-l}} |0\rangle + e^{2\pi i j (1) 2^{-l}} |1\rangle \right) \quad (4.31)$$

$$= \frac{1}{\sqrt{M}} \bigotimes_{l=1}^m \left(|0\rangle + e^{2\pi i j 2^{-l}} |1\rangle \right) \quad (4.32)$$

$$= \frac{1}{\sqrt{M}} \left(|0\rangle + e^{2\pi i 0 \cdot j_m} |1\rangle \right) \left(|0\rangle + e^{2\pi i 0 \cdot j_{m-1} j_m} |1\rangle \right) \dots \left(|0\rangle + e^{2\pi i 0 \cdot j_1 j_2 \dots j_m} |1\rangle \right) \quad (4.33)$$

Let us write the input state $|j\rangle$ in the tensor product form $|j\rangle = \bigotimes_{l=1}^m |j_l\rangle$, $j_l \in \{0, 1\}$. Then the first factor in (Eq. 4.33) can be prepared via the Hadamard transform applied to the state $|j_m\rangle$.

$$H|j_m\rangle = \frac{1}{\sqrt{2}}(|0\rangle + (-1)^{j_m}|1\rangle) = \frac{1}{\sqrt{2}}(|0\rangle + e^{2\pi i 0 \cdot j_m}|1\rangle) \quad (4.34)$$

To prepare the second factor in (Eq. 4.33) we need a gate capable of adjusting the relative phase. A suitable gate is a z-rotation gate

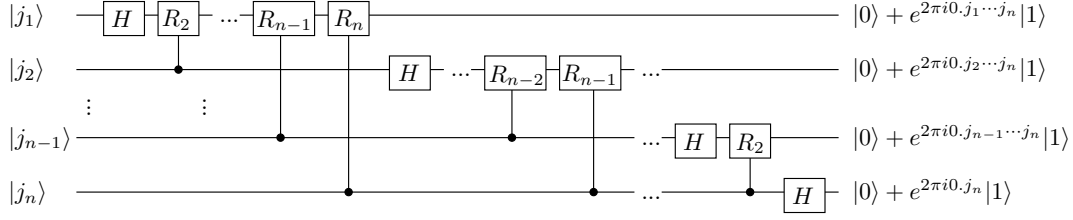
$$R_k = \begin{pmatrix} 1 & 0 \\ 0 & e^{2\pi i / 2^k} \end{pmatrix}. \quad (4.35)$$

By applying the Hadamard transform and the gate R_2 conditioned on the bit j_m to the state $|j_{m-1}\rangle$ we get

$$\begin{aligned} R_2^{j_m} H|j_{m-1}\rangle &= R_2^{j_m} \frac{1}{\sqrt{2}}(|0\rangle + e^{2\pi i 0 \cdot j_{m-1}}|1\rangle) = \frac{1}{\sqrt{2}}(|0\rangle + e^{2\pi i 0 \cdot j_{m-1}} e^{2\pi i j_m / 2^2}|1\rangle) \\ &= \frac{1}{\sqrt{2}}(|0\rangle + e^{2\pi i 0 \cdot j_{m-1} j_m}|1\rangle) \end{aligned} \quad (4.36)$$

The other factors can be prepared accordingly. The corresponding circuit, derived from the product representation (Eq. 4.33), for the quantum Fourier transform over \mathbb{Z}_{2^m} is shown in the Figure 4.7. Note that the order of qubits becomes reversed and swap gates should be used if needed.

The circuit utilizes $(m^2 + m)/2 = O(m^2)$ gates. An obvious bottleneck is that some of these gate require an exponential precision up to $\xi = 1/2^m$. By the Solovay-Kitaev theorem, each

Figure 4.7: Efficient circuit for the quantum Fourier transform over \mathbb{Z}_{2^m} .

such gate will require $O(\log^c(1/\xi)) = O(m^c)$, where $c \approx 4$, gates from a fixed universal set. This implies an efficient circuit, however, the corresponding polynomial is of unacceptable degree. Barenco *et. al.* [76] studied approximated quantum Fourier transform (AQFT) where gates R_k for $k \geq \log_2 m + 2$ (i.e. a logarithmic-depth AQFT) are neglected and showed that the accuracy of the QFT can be achieved after $O((m/\log m)^3)$ iterations of the AQFT. D. Cheung [77] improved significantly on this result and showed that the output probability distribution of a single logarithmic-depth AQFT iteration is in fact very close to the QFT output distribution. In particular, Cheung gave a lower bound

$$P_{down,up}(\delta = 1/2) \geq \frac{4}{\pi^2} - \frac{1}{4m},$$

for $m \geq 4$. Therefore the logarithmic-depth AQFT can be considered a direct replacement for the QFT. The amount of R_k gates is now $O(m \log m)$ and the desired gate accuracy is $\xi = 1/2^{\log m}$. Using the Solovay-Kitaev theorem, only $O(m \log m \log^c m) = O(m^2)$ gates from a fixed universal set are required in order to implement that transform.

The history of efficient circuits for a quantum Fourier transform goes back to D. Deutsch. In 1989, Deutsch [22] described an efficient circuit for QFT over \mathbb{Z}_2^m , that is the Hadamard transform. Shor [12] inspired by the Simon promise problem [16], where the Hadamard transform was used, realized how to efficiently implement QFT over the group \mathbb{Z}_q for certain special values of q . In 1994, D. Coppersmith improved this result to hold for \mathbb{Z}_{2^m} and one year later Kitaev [54] gave an efficient circuit for a quantum Fourier transform over an arbitrary finite abelian group. Somewhat later, human-competitive circuits for QFT (and other transforms) were also discovered using genetic programming [78, 79].

4.2.4 Linear property of the advanced scheme

Details of the advanced scheme for phase estimation including the efficient measurement in the Fourier basis (implemented by the Fourier transform + projection onto the computational basis) is shown in the Figure 4.8.

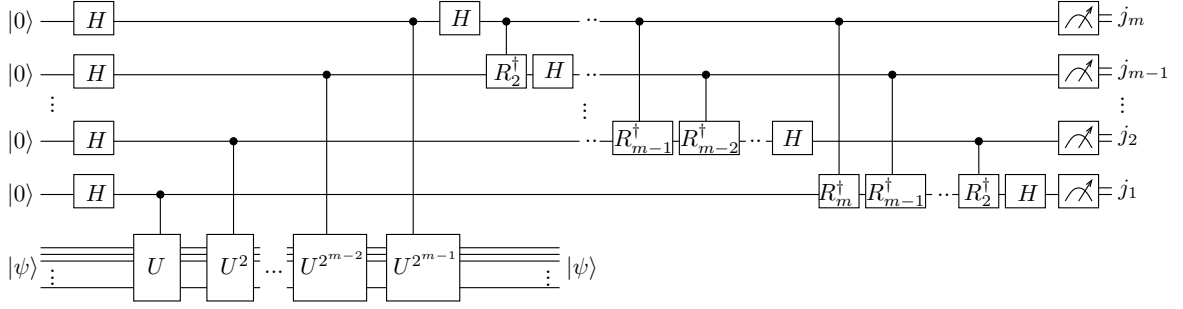


Figure 4.8: Detailed scheme for advanced phase estimation. The state $|\psi\rangle$ is promised to be an eigenstate of U with an eigenvalue $e^{2\pi i\phi}$. The output given by $\sum_{k=1}^m j_k 2^{-k} = j/2^m$ is a good estimate of ϕ , $|\phi - j/2^m| < 1/2^m$, with high probability.

Essentially the scheme transforms the input state $|0, \psi\rangle$ in the following way:

$$|0, \psi\rangle \quad (4.37a)$$

$$\xrightarrow{H \otimes I} \frac{1}{\sqrt{2^m}} \sum_j |j, \psi\rangle \quad (4.37b)$$

$$\xrightarrow{\text{controlled-}U^j} \frac{1}{\sqrt{2^m}} \sum_j |j, U^j \psi\rangle = \frac{1}{\sqrt{2^m}} \sum_j e^{2\pi i j \phi} |j, \psi\rangle \quad (4.37c)$$

$$\xrightarrow{QFT^\dagger \otimes I} \sum_j \alpha_j |j, \psi\rangle, \quad (4.37d)$$

where states $|j, \psi\rangle$ such that $|\phi - j/2^m| < 1/2^m$ have dominating probabilities $|\alpha_j|^2$ and thus one of them is observed with high probability. Therefore the overall transform can be described as

$$|0, \psi\rangle \xrightarrow{\text{phase estimation}} |\blacktriangleleft \phi \blacktriangleright, \psi\rangle, \quad (4.38)$$

where $\blacktriangleleft \phi \blacktriangleright$ denotes a good estimator. Importantly, since U is a linear operator, the same concept applies to an input superposition of eigenvectors of U . Thus we can avoid preparing a (possibly unknown) eigenstate, at the cost of introducing some additional randomness into the scheme.

Let $|\varsigma\rangle = \sum_s \gamma_s |\psi_s\rangle$ denote an arbitrary quantum state expanded in terms of eigenstates $|\psi_s\rangle$ of U . Then

$$|0, \varsigma\rangle = \sum_s \gamma_s |0, \psi_s\rangle \xrightarrow{\text{phase estimation}} \sum_s \gamma_s |\blacktriangleleft \phi_s \blacktriangleright, \psi_s\rangle, \quad (4.39)$$

where $\blacktriangleleft \phi_s \blacktriangleright$ denotes a good estimator related to the eigenstate $|\psi_s\rangle$. A subsequent measurement on the first register yields one estimator $\blacktriangleleft \phi_s \blacktriangleright$ chosen at random with probability $|\gamma_s|^2$. An important side effect of the measurement is a collapse of the second register to the corresponding eigenvector (or a superposition of eigenvectors related to a particular degenerative eigenvalue). Therefore the phase estimation scheme can be used to prepare eigenstate(s). Note that the eigenstate remains unknown to us.

4.2.5 Representative instances of the phase estimation problem

4.2.5.1 Factoring casted as a phase estimation problem

Kitaev [54, 21] showed that the order finding algorithm (the only quantum part of the factoring algorithm) is equivalent to the scheme where we estimate an eigenvalue of the unitary operator U given by 'modular multiplication by k ':

$$U : |y\rangle \longrightarrow |ky \pmod{N}\rangle. \quad (4.40)$$

Let r denote the unknown order, i.e. the least positive integer such that $k^r = 1 \pmod{N}$ for given coprimes k and N . The main observation done by Kitaev is that states defined as

$$|\psi_s\rangle \equiv \frac{1}{\sqrt{r}} \sum_{j=0}^{r-1} e^{-2\pi ijs/r} |k^j \pmod{N}\rangle, \quad (4.41)$$

for integer $0 \leq s \leq r-1$ are eigenstates of U . Additionally, since $U^r|y\rangle = |k^r y \pmod{N}\rangle = |y\rangle$ and thus $U^r = I$, the eigenvalues of U must be of the form $\lambda_s = e^{2\pi is/r}$. We can verify that

$$U|\psi_s\rangle = \frac{1}{\sqrt{r}} \sum_{j=0}^{r-1} e^{-2\pi ijs/r} |k^{j+1} \pmod{N}\rangle = e^{2\pi is/r} |\psi_s\rangle = \lambda_s |\psi_s\rangle. \quad (4.42)$$

It is also useful to rewrite (Eq. 4.41) in the form

$$\frac{1}{\sqrt{r}} \sum_{s=0}^{r-1} e^{2\pi ijs/r} |\psi_s\rangle = |k^j \pmod{N}\rangle, \quad (4.43)$$

and by setting $j = 0$ to derive

$$\frac{1}{\sqrt{r}} \sum_{s=0}^{r-1} |\psi_s\rangle = |1\rangle. \quad (4.44)$$

Thus the state $|1\rangle$ trivially constitutes an equally weighted superposition of eigenstates $|\psi_s\rangle$.

The process of a phase estimation applied to the 'modular multiplication by k ' operator U can be described as follows.

$$|0\rangle|1\rangle \quad (4.45a)$$

$$\xrightarrow{H \otimes I} \frac{1}{\sqrt{2^m}} \sum_{j=0}^{2^m-1} |j\rangle|1\rangle \quad (4.45b)$$

$$\xrightarrow{\text{controlled-}U^j} \frac{1}{\sqrt{2^m}} \sum_{j=0}^{2^m-1} |j\rangle U^j |1\rangle \quad \Bigg| \Rightarrow \frac{1}{\sqrt{2^m}} \sum_{j=0}^{2^m-1} |j\rangle |k^j \pmod{N}\rangle \quad (4.45c)$$

$$= \frac{1}{\sqrt{2^m}} \sum_{j=0}^{2^m-1} |j\rangle U^j \left(\frac{1}{\sqrt{r}} \sum_{s=0}^{r-1} |\psi_s\rangle \right) \quad (4.45d)$$

$$= \frac{1}{\sqrt{2^m}} \sum_{j=0}^{2^m-1} |j\rangle \left(\frac{1}{\sqrt{r}} \sum_{s=0}^{r-1} e^{2\pi i j s / r} |\psi_s\rangle \right) \quad (4.45e)$$

$$= \frac{1}{\sqrt{2^m}} \frac{1}{\sqrt{r}} \sum_{j=0}^{2^m-1} \sum_{s=0}^{r-1} e^{2\pi i j s / r} |j\rangle |\psi_s\rangle \quad (4.45f)$$

$$\xrightarrow{QFT^\dagger \otimes I} \sim \frac{1}{\sqrt{r}} \sum_{s=0}^{r-1} |\blacktriangleleft s/r \blacktriangleright\rangle |\psi_s\rangle \quad (4.45g)$$

Given the outcome $a = \blacktriangleleft s/r \blacktriangleright$ for some s , the continued fractions algorithm efficiently produces numbers s' and r' with no common factor, such that $s'/r' = s/r$. For $m = \Theta(\log(N^2))$ the number r' is equal to the order r with high probability.

Efficient implementation of controlled- U^{2^l} gates

In order for the period finding algorithm to be practically useful, we need to make sure that all its steps are performed efficiently with respect to the size of the input $m = \Theta(\log(N^2))$. So far we know that initialization of the input state to $|0\rangle|1\rangle$ is efficient, measurement in the Fourier basis is efficient and the classical postprocessing is efficient. The implementation of the controlled- U^{2^l} gates, where $0 \leq l \leq m - 1$, based on the repeated modular squaring can be performed efficiently as well. Naive implementation consisting only of repeated application of the controlled- U gate is obviously not efficient with respect to m .

Let us have gates for modular multiplication and modular squaring (these are manifestly efficient):

$$U_{mul} : |a\rangle|b\rangle \longrightarrow |a\rangle|ab \pmod{N}\rangle, \quad (\text{modular multiplication}) \quad (4.46)$$

$$U_2 : |a\rangle \longrightarrow |a^2 \pmod{N}\rangle. \quad (\text{modular squaring}) \quad (4.47)$$

The implementation of the controlled- U^{2^l} gate based on the modular squaring consists of four steps:

1. Prepare two registers. The first is a scratch register initially set to $|k\rangle$. The second register contains an arbitrary state $|y\rangle$.
2. Apply the squaring gate l -times to the scratch register.

$$|k, y\rangle \rightarrow |k^2 \pmod{N}, y\rangle \rightarrow |k^4 \pmod{N}, y\rangle \rightarrow \dots \rightarrow |k^{2^l} \pmod{N}, y\rangle \quad (4.48)$$

3. Conditioned on some bit j_l perform the modular multiplication.

$$|k^{2^l} \pmod{N}, y\rangle \longrightarrow |k^{2^l} \pmod{N}, k^{2^l} y \pmod{N}\rangle \quad (4.49)$$

4. Set the scratch register to $|k\rangle$ again, so it can be reused. This can be performed either by 'uncomputing' or a measurement based reset operation.

The procedure is shown in the Figure 4.9.

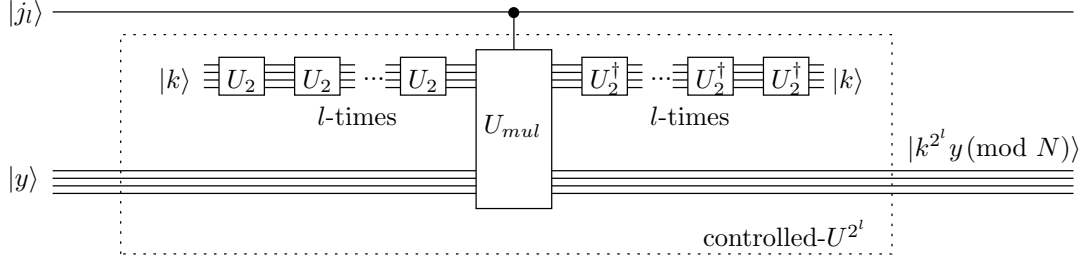


Figure 4.9: Efficient implementation of the controlled- U^{2^l} gate, where U performs a modular multiplication by k .

4.2.5.2 Finding the energy eigenvalues of a local Hamiltonian

In 1982, R. P. Feynman suggested that a quantum computer might be capable of an efficient simulation of other quantum physical systems [80]. S. Lloyd [19] showed that an efficient simulation is indeed possible at least for quantum systems with a local Hamiltonian. Systems with a local Hamiltonian capture a large part of relevant physics. The question then arises, what kind of properties can be efficiently extracted from the simulation. Abrams and Lloyd [81] focused on static properties and showed that the phase estimation procedure can be used to find efficiently the energy eigenvalues if approximate eigenvectors are available.

Let \mathcal{H} be a (local) Hamiltonian with energy eigenvalues λ_s and eigenstates $|\psi_s\rangle$. This Hamiltonian generates a unitary evolution $U = e^{-i\mathcal{H}t}$ during time t . Energy eigenvalues are mapped to the phase of eigenvalues of U .

$$\mathcal{H} = \sum_s \lambda_s |\psi_s\rangle\langle\psi_s|, \quad U = e^{-i\mathcal{H}t} = \sum_s e^{-i\lambda_s t} |\psi_s\rangle\langle\psi_s| = \sum_s e^{2\pi i\phi_s} |\psi_s\rangle\langle\psi_s|. \quad (4.50)$$

Let $|\zeta\rangle$ denote a promised approximate eigenvector $|\psi\rangle$. In general, accessing an approximate eigenvector is a very hard problem [82, 83]. Therefore the Abrams-Lloyd algorithm is in practice limited to cases where a rough approximate eigenvector can be efficiently obtained by classical 'ab initio' methods which result in a known wave function [81]. Another possibility is to exploit the adiabatic state preparation procedure [84, 85]. However, the power of adiabatic state preparation is yet not fully explored and its further study will determine the usefulness of the Abrams-Lloyd algorithm.

The phase estimation scheme applied to the operator U and an approximate eigenvector $|\zeta\rangle$ reveals a good estimator of the phase ϕ with a probability $|\langle\psi|\zeta\rangle|^2$. The corresponding energy eigenvalue is then calculated as

$$\lambda = 2\pi(1 - \phi)/t. \quad (4.51)$$

Clearly, time t should be chosen such that $\phi \approx 1/2$ in order to avoid a long sequence of initial zeroes in the binary expansion of ϕ . A satisfying value for time t can be calculated using classical approximation methods for rough estimates of λ . The dependence on a rough estimate of λ is not that limiting as the dependence on an approximate eigenvector. Potential initial zeroes contribute to the overall complexity of the algorithm as a (large) hidden constant.

The scheme for phase estimation as proposed by Abrams and Lloyd [81] is shown in the Figure 4.10. This scheme is functionally identical with the scheme in the Figure 4.8. The

only difference is that an additional quantum logic is used to determine how many times should be a controlled- U gate applied to a particular ancilla qubit j_i in order to implement a controlled- U^{2^i} gate. This is intentional in order to stress that the controlled- U^{2^i} gates can be implemented only by a long sequences of applications of a plain controlled- U gate. There is no shortcut similar to the one used in the period finding algorithm.

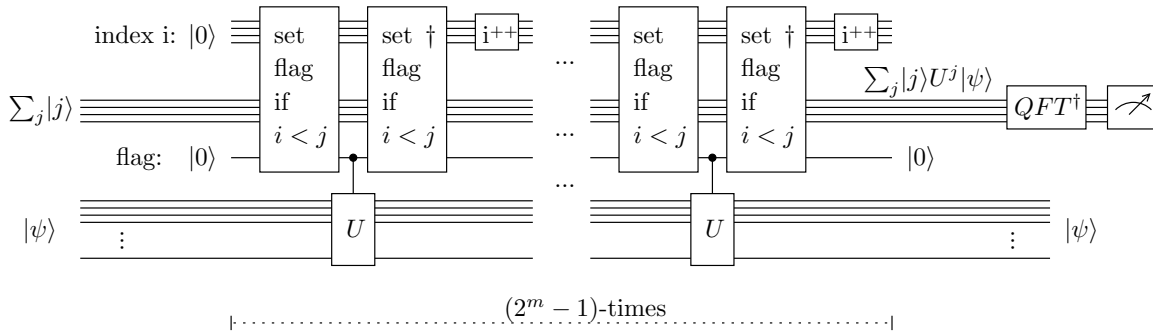


Figure 4.10: The phase estimation scheme of Abrams and Lloyd.

Assuming access to an approximate eigenvector and an efficient implementation of U (and controlled- U consequently) with respect to the size of the problem n (size of the system being simulated), the algorithm for finding energy eigenvalues and preparing eigenstates performs as follows. The scheme operates at the Heisenberg limit and therefore the accuracy of order $1/2^m$ implies at most 2^m employments of the controlled- U gate. The number m is the amount of ancillary qubits which are measured. In practice, the precision is a fixed parameter and amounts to the overall complexity as a hidden constant. For example, quantum chemistry [85] requires $m \approx 20$. This is in sharp contrast with the order finding algorithm where the size of the ancilla register scales with the size of the problem (order of few thousands of qubits for breaking the RSA public-key cryptography). Considering m a fixed parameter, the algorithm is efficient by virtue of an efficient circuit performing the controlled- U gate. A quantum simulation with $n > 30$ outperforms classical computers dramatically.

Efficient simulation of quantum systems with a local Hamiltonian

A local Hamiltonian describing dynamics of a quantum system with n particles can be written as a sum over many local interactions

$$\mathcal{H} = \sum_{k=1}^L \mathcal{H}_k, \quad (4.52)$$

where \mathcal{H}_k acts on at most a constant number of particles, and L is a polynomial in n . The key to quantum simulation algorithms is the convergence of the Trotter formula which allows to approximate the evolution

$$U = e^{-i\mathcal{H}t}. \quad (4.53)$$

An exact simulation is not possible since the different \mathcal{H}_k do not have to commute with each other. Therefore, the direct approach

$$e^{-i\mathcal{H}t} = e^{-i\mathcal{H}_1t} e^{-i\mathcal{H}_2t} \dots e^{-i\mathcal{H}_Lt} \quad (4.54)$$

is not correct.

Theorem 4.2.1 (Trotter formula). *Let A and B be hermitian operators, then for any real t ,*

$$\lim_{q \rightarrow \infty} \left(e^{iAt/q} e^{iBt/q} \right)^q = e^{i(A+B)t}. \quad (4.55)$$

□

By the definition of a matrix exponentiation and for some $\delta < 1$, we have

$$e^{-i\mathcal{H}_A\delta} e^{-i\mathcal{H}_B\delta} = \left(I - i\mathcal{H}_A\delta + O(\delta^2) \right) \left(I - i\mathcal{H}_B\delta + O(\delta^2) \right) \quad (4.56)$$

$$= I - i(\mathcal{H}_A + \mathcal{H}_B)\delta + O(\delta^2), \quad (4.57)$$

and

$$e^{-i(\mathcal{H}_A + \mathcal{H}_B)\delta} = I - i(\mathcal{H}_A + \mathcal{H}_B)\delta + O(\delta^2). \quad (4.58)$$

The asymptotic equations (4.57) and (4.58) permit us to write

$$e^{-i(\mathcal{H}_A + \mathcal{H}_B)\delta} = e^{-i\mathcal{H}_A\delta} e^{-i\mathcal{H}_B\delta} + O(\delta^2). \quad (\text{for } \delta < 1) \quad (4.59)$$

Now, the equation (4.59) and the convergence of the Trotter formula allow us to perform the whole simulation by simulating the local evolution operators over short discrete time slices and then repeat the sequence of these local simulations a couple of times. In particular, for $q > t$,

$$e^{-i(\mathcal{H}_1 + \mathcal{H}_2 + \dots + \mathcal{H}_L)t} = \left(e^{-i\mathcal{H}_1(t/q)} e^{-i\mathcal{H}_2(t/q)} \dots e^{-i\mathcal{H}_L(t/q)} \right)^q + O\left(q(t/q)^2\right) \quad (4.60)$$

$$= \left(e^{-i\mathcal{H}_1(t/q)} e^{-i\mathcal{H}_2(t/q)} \dots e^{-i\mathcal{H}_L(t/q)} \right)^q + O\left(t^2/q\right). \quad (4.61)$$

Therefore by choosing $q \geq (t^2/\varepsilon)$, we can implement the evolution $U = e^{-i\mathcal{H}t}$, where \mathcal{H} is a sum over local interactions, $\mathcal{H} = \sum_{k=1}^L \mathcal{H}_k$, within error tolerance of $O(\varepsilon)$ using $O(L(t^2/\varepsilon))$ 'basic' gates of the form $e^{-i\mathcal{H}_k(t/q)}$. Since L is polynomial in the size of the problem n and \mathcal{H}_k acts on at most a constant number of particles, the whole simulation requires an amount of gates of order

$$O\left(\text{poly}(n) \frac{t^2}{\varepsilon}\right), \quad (4.62)$$

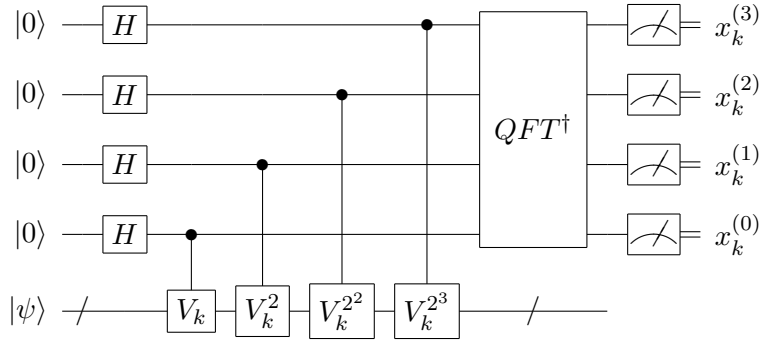
and thus is efficient.

A similar result is known for Hamiltonians of the form $\mathcal{H} = \bigotimes_{k=1}^L \sigma_{p(k)}$, where $\sigma_{p(k)}$ are Pauli matrices.

4.3 Iterative phase estimation algorithm - IPEA

4.3.1 Motivation

The Abrams-Lloyd algorithm is potentially the most important quantum algorithm known so far. Even a small quantum computer with as many as 100 qubits at total is expected to outperform any classical computer in certain scopes of quantum chemistry calculations. The issue of a relevance of a small quantum computer for real chemistry problems has been recently studied by Aspuru-Guzik *et. al.* [85]. In the work, the authors carried out a modified phase estimation scheme where the size of the ancilla register is fixed and a high accuracy of order $1/2^m$ is achieved via m iterations of the scheme. Using this scheme and an adiabatic state preparation, simulated calculations of the water and lithium hydride molecular ground-state energies were performed. This required only 8 and 11 qubits for storing an approximate eigenvector (the molecular ground-state wave function) and four qubits in the fixed-size ancilla register. These simulations strengthened the conjecture of Abrams and Lloyd that quantum computers of tens to hundreds of qubits can match and exceed classical calculations. The k -th iteration of the Aspuru-Guzik scheme for phase estimation (recursively defined iterative PEA) is shown in the Figure 4.11.



$$V_k = \left(e^{-2\pi i \varphi_{k-1}} V_{k-1} \right)^2 \quad \text{for } k = 1, 2, \dots, m,$$

where $V_1 = U$ and $\varphi_k = 0.x_k^{(0)} x_k^{(1)} x_k^{(2)} x_k^{(3)}$.

Figure 4.11: The Aspuru-Guzik phase estimation scheme.

In the first iteration the operator V_1 is set to U , where U simulates the evolution of some physical system. The input state $|\psi\rangle$ is promised to be an eigenstate of U with an eigenvalue $e^{2\pi i \phi}$. Let us denote the binary expansion of ϕ as $\phi = 0.\phi_1\phi_2\phi_3\dots$. The result of the first iteration is a four bit estimate φ_1 of the unknown phase ϕ . With high probability this estimate is accurate at least in order of $1/2$. The estimate φ_1 enters the next iteration with the operator $V_2 = e^{-2\pi i(2\varphi_1)}U^2$. The phase shift in the operator V_2 effectively enables to measure a four bit estimate φ_2 of a phase $(2(\phi - \varphi_1) \bmod 1) = (0.\phi_2\phi_3\phi_4\phi_5\dots - 0.x_1^{(1)}x_1^{(2)}x_1^{(3)})$. Now, the number $(\varphi_1 + \varphi_2/2)$ constitutes an estimate of ϕ which is accurate at least in order of $1/4$ with high probability. Analogically, each additional iteration yields one more bit of ϕ . In the

k -th iteration, we work with a phase

$$\begin{array}{cccccccccccccccc}
 & \phi_1 & \phi_2 & \phi_3 & \dots & \dots & \phi_{k-2} & \phi_{k-1} & \cdot & \phi_k & \phi_{k+1} & \phi_{k+2} & \phi_{k+3} & \phi_{k+4} & \dots \\
 - & x_1^{(0)} & x_1^{(1)} & x_1^{(2)} & x_1^{(3)} & & & & & & & & & & & \\
 \vdots & & & & & & & & & & & & & & & \\
 - & & & & & \dots & x_{k-2}^{(0)} & x_{k-2}^{(1)} & \cdot & x_{k-2}^{(2)} & x_{k-2}^{(3)} & & & & \\
 - & & & & & \dots & x_{k-1}^{(0)} & \cdot & x_{k-1}^{(1)} & x_{k-1}^{(2)} & x_{k-1}^{(3)} & & & & ,
 \end{array} \tag{4.63}$$

measure a corresponding estimate φ_k , and the final estimate of ϕ accurate to at least m binary digits is given by

$$\blacktriangleleft \phi \blacktriangleright = \sum_{k=1}^m \varphi_k / 2^{k-1}. \tag{4.64}$$

In conclusion, the Aspuru-Guzik scheme requires $O(m)$ iterations, $O(4m) = O(m)$ measurements, and the plain controlled-U gate is employed $O(2^{m+3}) = O(2^m)$ times in order to create required operators controlled- V_k^c for $c = 1, 2, 4, 8$.

Example 9 (Aspuru-Guzik phase estimation).

Let $\phi = 0.011101101\dots$. The progress of obtaining gradually more accurate estimates of ϕ is captured in the following table.

| iteration k | working phase | result φ_k | final estimate |
|---------------|----------------|--------------------|------------------------------|
| 1 | 0.011101101... | 0.0111 | 0.0111 \rightarrow 0.1 |
| 2 | 0.00001101... | 0.0001 | 0.01111 \rightarrow 0.10 |
| 3 | 0.1111101... | 0.0000 | 0.011110 \rightarrow 0.100 |

Table 4.1: Sample run of the Aspuru-Guzik phase estimation. □

Another scheme for phase estimation which can be performed in an iterative manner with a fixed-sized ancilla register is the original procedure used by Kitaev in [54]. Let ϕ denote the unknown phase of an eigenvalue $e^{2\pi i\phi}$ related to some unitary operator U and its eigenstate $|\psi\rangle$. Kitaev observed that if one knows estimates of the numbers $2^{k-1}\phi \pmod{1}$, for $k = 1, 2, \dots, m$, then there exists a polynomial classical algorithm which can reconstruct ϕ to at least m binary digits. The circuit which determines an estimate of $2^{k-1}\phi \pmod{1}$ is shown in the Figure 4.12.

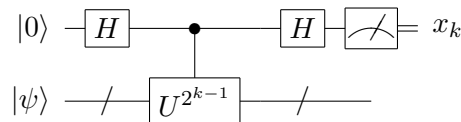


Figure 4.12: Kitaev's phase estimation circuit.

The outcome $x_k = 0$ is observed with probability $P_0 = \cos^2(\pi 2^{k-1}\phi)$. Accordingly, the probability that $x_k = 1$ is $P_1 = 1 - P_0 = \sin^2(\pi 2^{k-1}\phi)$. See Figure 4.13.

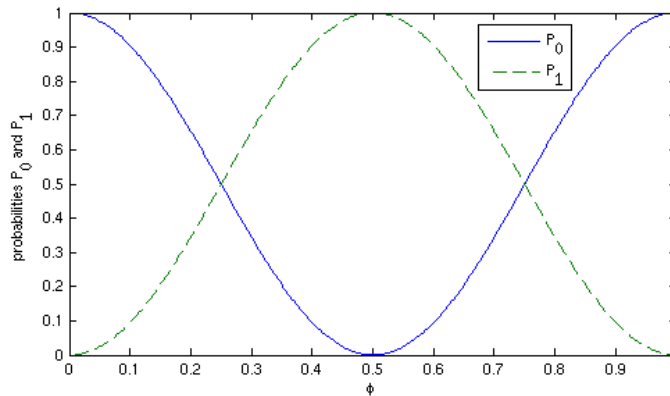


Figure 4.13: A plot of outcome probabilities in the Kitaev PEA.

Exploiting independent trials of the circuit and using the Chernoff bound, it is possible to determine whether $2^k \phi \pmod{1}$ is closer to zero ("0.0"), i.e. $P_0 \geq P_1$, or to one half ("0.1") with error probability $\leq \varepsilon_k$ after $O(\log(1/\varepsilon_k))$ trials. By setting $\varepsilon_k \leq \varepsilon/m$, all the values x_k , for $k = 0, 1, \dots, m-1$, can be correctly determined with error probability $\leq \varepsilon$. Therefore this scheme requires at total $O(m \log(m/\varepsilon))$ iterations (measurements) for estimating ϕ to m binary digits. The plain controlled-U gate is employed $O(2^m \log(m/\varepsilon))$ times at total in order to implement required operators controlled- U^{2^k} .

Compared to the Aspuru-Guzik scheme, the Kitaev phase estimation is asymptotically slower by a small logarithmic factor. However, in the context of estimating energy eigenvalues where m is a fixed constant between 20 and 30, the performance of the Kitaev algorithm is actually better. Besides, it requires only a single ancillary qubit and the circuit is easier to implement.

The question arises, whether there exists an iterative scheme for phase estimation (IPEA) which does not contain the logarithmic factor (reaches the true Heisenberg limit) and utilizes only a single ancilla qubit. Such a scheme would be of utmost importance for experimental implementation with a very limited number of qubits. The answer is yes and such a scheme can be derived from the textbook (QFT-based) phase estimation (Figure 4.8). The derivation requires knowledge of how the entanglement is localized in the circuit and understanding of the information flow in the measurement in the Fourier basis. In particular, the key point is a discovery of an adaptive implementation of the measurement in the Fourier basis (so called semiclassical QFT) by Griffiths and Niu [86]. The relevance of incorporating the semiclassical QFT into the Shor's circuit for factoring was first noticed by C. Zalka [87], then by Mosca and Ekert [88] in the context of solving the hidden subgroup problem by means of the phase estimation approach, and by Childs et. al. [89] for distinguishing Hamiltonians in some cases.

Regarding the problem of finding energy eigenvalues, Abrams and Lloyd [81] and Aspuru-Guzik et. al. [85] do not employ IPEA although parts of their respective works focus on a relevance of the algorithm for a quantum computer with a very limited amount of qubits. Properties of the IPEA in the context of finding energy eigenvalues have been recently studied by Knill et. al. [90] and Dobšíček et. al. [A.1].

4.3.2 IPEA derived from the Kitaev phase estimation

In contrast to the QFT-based PEA (Figure 4.8), the Kitaev PEA (Figure 4.12) does not explicitly use the quantum Fourier transform as such. Therefore, in some sense, the Kitaev PEA allows to solve abelian HSP-like problems without using the relatively extensive knowledge behind quantum Fourier sampling. Moreover, it is possible to derive the IPEA scheme from the Kitaev PEA by incorporating a classical postprocessing algorithm used to construct the estimate $\langle \phi \rangle$ into the quantum circuit. Such a result may be quite inspiring in the further development of quantum algorithms as the usual approach is to perform as much as possible of necessary calculations classically, since quantum resources are considered 'expensive'.

There are many possible efficient classical algorithms suitable for postprocessing in the Kitaev PEA (no specific algorithm was given by Kitaev in [54]). I propose to use an algorithm which can be described as a weighted interval intersection method. The core of this method is outlined in the Figure 4.14. To each of the results x_k is assigned an interval with a weight $1/2^k$.

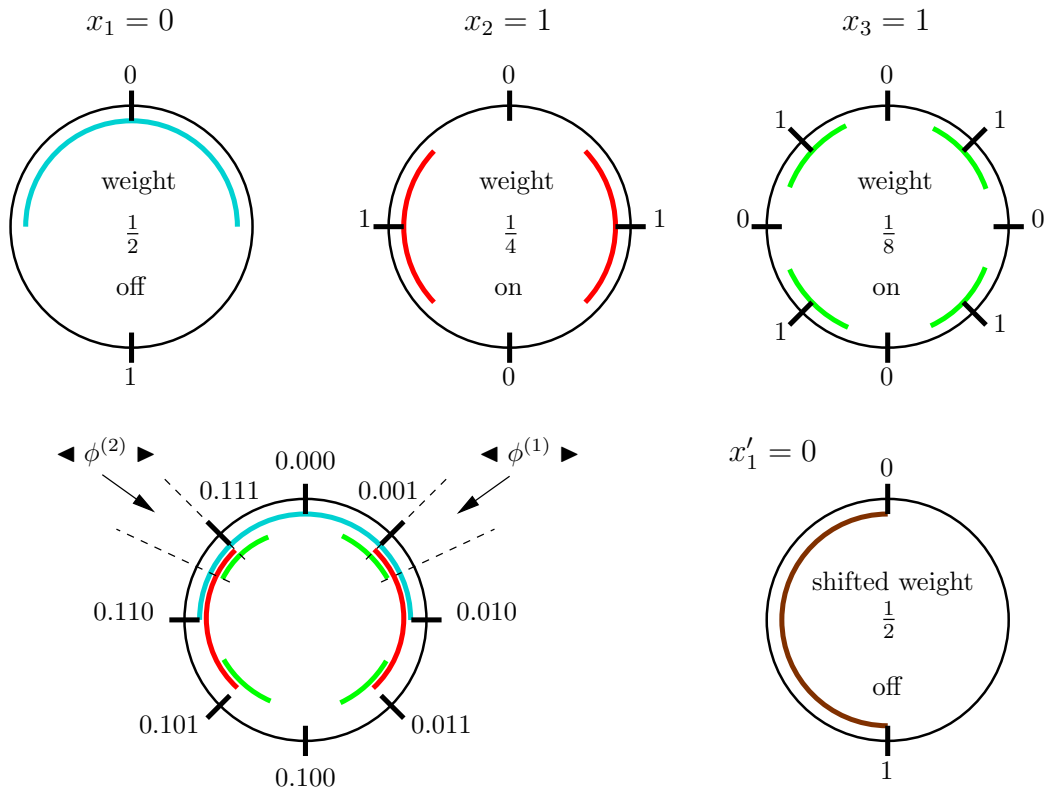


Figure 4.14: A weighted interval intersection method as a postprocessing algorithm for the Kitaev PEA. In general, there are two candidate estimates and an intersection with additional shifted weight- $1/2$ interval determines the final estimate. The figure shows a sample run for $\phi = 0.110110\dots$ which gives the following estimates x_k for values of $2^{k-1}\phi \pmod{1}$: $x_1 = 0$, $x_2 = 1$, and $x_3 = 1$.

The weight also determines the spanning range of the interval. By making intersections of the two largest intervals per step, we end up with two small intervals of size $1/2^{m+1}$ after $m - 1$

steps. Let us associate a candidate estimate $\blacktriangleleft \phi^{(1)} \blacktriangleright$ to the value in the center of one of the two small intervals. This ensures that the candidate contains information sufficient to reveal ϕ to $m+2$ binary digits. Accordingly, there is a second candidate $\blacktriangleleft \phi^{(2)} \blacktriangleright = 1 - \blacktriangleleft \phi^{(1)} \blacktriangleright$. The last step is to decide which candidate is the right one. We perform one more intersection with an interval assigned to the outcome x'_1 of an additional iteration of the Kitaev scheme where to the phase $2\pi\phi$ (contained in the ancilla qubit) is added a phase $2\pi(0.01)$. A circuit corresponding to this additional iteration is shown in the Figure 4.15.

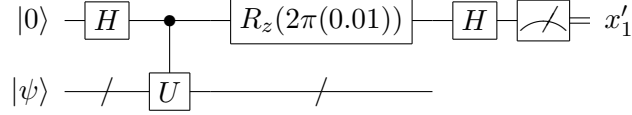


Figure 4.15: Additional iteration in the Kitaev PEA.

An important observation from the Figure 4.14 is that an interval with the smallest weight imposes the largest constraints to the spanning range of the other intervals. In other words, an interval with a small weight, if determined at first, can help to keep the spanning range of other intervals small. While this is not a good classical strategy (there are too many small ranges to keep track of), nothing prevent us to obtain the least significant bit x_m at first and use this information to improve the search for the bit x_{m-1} in the quantum circuit. This is a sort of feedback improvement similar to the one used in the Aspuru-Guzik scheme.

Let us write ϕ as $\phi = \tilde{\phi} + \delta 2^{-m}$, where $\tilde{\phi} = 0.\phi_1\phi_2\dots\phi_m$ denotes the first m bits of the binary expansion and $0 \leq \delta < 1$ is a reminder. In order to determine the bit ϕ_m , we perform an iteration of the Kitaev PEA for $k = m$, i.e. the gate controlled- $U^{2^{m-1}}$ is used. A single trial gives us the probability $P_0 = \cos^2(\pi(0.\phi_m + \delta/2))$ to measure $x_m = 0$ and the probability $P_1 = \sin^2(\pi(0.\phi_m + \delta/2))$ to measure $x_m = 1$. Thus for

$$\phi_m = \begin{cases} 0 & \Rightarrow P_0 = \cos^2(\pi\delta/2), \\ 1 & \Rightarrow P_1 = \sin^2(\pi/2 + \pi\delta/2) = \cos^2(\pi\delta/2), \end{cases} \quad (4.65)$$

and therefore

$$P_m = P(x_m = \phi_m) = \cos^2(\pi\delta/2). \quad (4.66)$$

In the next iteration of the Kitaev PEA, $k = m-1$, the phase of the ancilla qubit is $2\pi(0.\phi_{m-1}\phi_m + \delta/4)$. Assuming we know the bit ϕ_m , we can perform a Z rotation with angle $\omega_{m-1} = -2\pi(0.0x_m)$. Then the conditional measurement probability becomes $P_0 = \cos^2(\pi(0.\phi_{m-1}0 + \delta/4))$ and

$$P_{m-1} = P(x_{m-1} = \phi_{m-1} \mid x_m = \phi_m) = \cos^2(\pi\delta/4). \quad (4.67)$$

Note that P_{m-1} is significantly larger than P_m . Accordingly, we use bits x_m and x_{m-1} to improve the probability that x_{m-2} will be equal to ϕ_{m-2} . The k -th iteration of this IPEA is shown in the Figure 4.16. Generally, using the feedback angle $\omega_k = -2\pi(0.0x_{k+1}x_{k+2}\dots x_m)$ the conditional probability P_k for a bit x_k to be equal to ϕ_k is

$$P_k = \cos^2(\pi 2^{k-1} (\delta 2^{-m})). \quad (4.68)$$

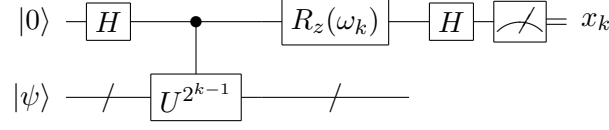


Figure 4.16: The k -th iteration of the iterative phase estimation (IPEA). The feedback angle depends on the previously measured bits through $\omega_k = -2\pi(0.0x_{k+1}x_{k+2}\dots x_m)$, and $\omega_m = 0$. Note that k is iterated backwards from m to 1.

Note that the equation (4.68) shows an exponential increase of P_k with decreasing k . The overall probability for the IPEA to extract $\tilde{\phi} = 0.\phi_1\phi_2\dots\phi_m$, i.e. an m -bit rounded down estimate of ϕ , is then

$$P_{\text{down}}(0 \leq \delta < 1) = P_m P_{m-1} \dots P_1 = \prod_{k=1}^m \cos^2(\pi 2^{k-1} (\delta 2^{-m})) \quad (4.69)$$

$$= \begin{cases} 1 & \text{for } \delta = 0, \\ \frac{\sin^2(\pi\delta)}{2^{2m} \sin^2(\pi\delta 2^{-m})}, & \text{for } 0 < \delta < 1, \end{cases} \quad (4.70)$$

using the trigonometric identity

$$\prod_{k=0}^{m-1} \cos(2^k \alpha) = \frac{\sin(2^m \alpha)}{2^m \sin(\alpha)} \quad (4.71)$$

for $\alpha \neq 0$. Using a similar derivation, the probability to extract $\tilde{\phi} + 2^{-m}$ (rounding up) is

$$P_{\text{up}}(0 \leq \delta < 1) = \frac{\sin^2(\pi(1-\delta))}{2^{2m} \sin^2(\pi(1-\delta)2^{-m})}. \quad (4.72)$$

Equations (4.70) and (4.72) are identical to the QFT-based PEA equations (4.16) and (4.17), respectively, see page 59. Therefore, the IPEA shares with the QFT-based PEA the same lower bound on the probability to observe an estimate of ϕ accurate to at least m binary digits with error probability $\varepsilon < 1 - 8/\pi^2$. In conclusion, the IPEA requires $O(m)$ iterations (measurements) in order to achieve the precision of order $1/2^m$ and the plain controlled- U gate is employed $O(2^m)$ times. Thus the true Heisenberg limit is achieved utilizing only a single ancilla qubit.

Similarly to the QFT-based PEA, success probability can be amplified up to $1 - \varepsilon$ for $\varepsilon > 0$ by extracting $m' = m + O(\log(1/\varepsilon))$ bits and keeping only m most significant ones. However, this approach is not favorable with respect to the costs of implementing gates controlled- U^{2^k} for $k > m$. The plain controlled- U gate would be needed

$$O\left(2^m 2^{\log(1/\varepsilon)}\right) \quad (4.73)$$

times. Per bit repetitions and majority voting is a better option. First, we observe that the bare bitwise error probability

$$1 - P_k = \sin^2(\pi 2^{k-1} (\delta 2^{-m})) \quad (4.74)$$

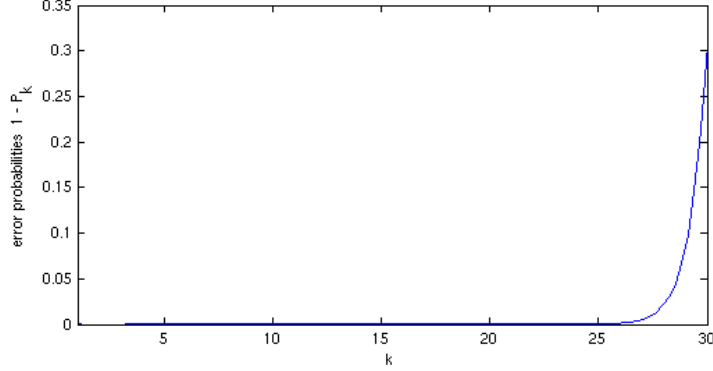


Figure 4.17: Bitwise error probabilities $\sin^2(\pi 2^{k-1} (\delta 2^{-m}))$ for $m = 30$ and $\delta = 3/8$.

decreases exponentially with decreasing k . This implies that only logarithmically many least significant bits have non-negligible error probabilities $1 - P_k$; see Figure 4.17.

Additionally, the probability of a bit x_k to be correctly determined after N_k independent trials is

$$P'_k = \sum_{s=0}^{\lfloor \frac{N_k-1}{2} \rfloor} \binom{N_k}{s} (P_k)^{N_k-s} (1 - P_k)^s, \quad (4.75)$$

given that x_k is chosen using majority voting from the results. Thus the effective error probability decreases exponentially with the number of repetitions, according to the binomial distribution. Therefore repeating the iterations for the $O(\log(1/\varepsilon))$ least significant bits an $O(\log(1/\varepsilon))$ number of times gives an error probability smaller than ε , independently of m . The total amount of iterations (measurements) then goes like $O(m + \log^2(1/\varepsilon))$ and the plain controlled- U gate is utilized

$$\sum_{k=1}^m 2^{k-1} + \lceil \log(1/\varepsilon) \rceil \sum_{t=1}^{\lceil \log(1/\varepsilon) \rceil} 2^{m-t} = O(2^m \log(1/\varepsilon)) \quad (4.76)$$

times. The scaling $O(2^m \log(1/\varepsilon))$ is a consequence of the fact that the least significant bits are the most expensive bits in terms of constructing the corresponding controlled gates. However, the factor $\log(1/\varepsilon)$ is more pleasing than $2^{\log(1/\varepsilon)}$.

4.3.3 QFT-based PEA versus IPEA

The IPEA can be considered as a direct viable alternative to the QFT-based PEA, especially for experimental quantum devices with a very limited amount of qubits. Next to a smaller number of qubits the IPEA requires, an additional advantage for experiments is the possibility to exchange a long coherence time of an input eigenstate for an input state preparation per iteration (and/or repetition). See Figure 4.18. The possibility of having a choice is valid for problems where we work with a particular (approximate) input eigenstate since we can prepare this eigenstate for each iteration.

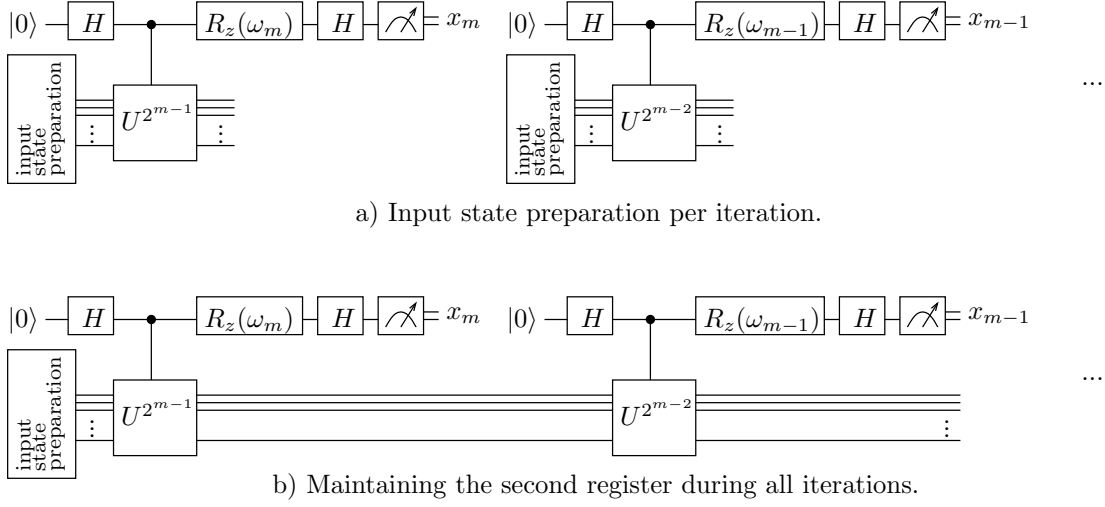


Figure 4.18: The IPEA and an input eigenstate.

Thanks to the linear property of any PEA scheme (Eq. 4.39, Section 4.2.4), it is not a problem to work with an approximate eigenstate at the cost of some additional randomness, however, this randomness now enters (complicates) each iteration. Depending on the quality of the approximation of the eigenstate, repetitions of each iteration are needed to make sure that the measured bit is related to the eigenvalue corresponding to the desired eigenstate. Otherwise, we may end up in a situation where we work with a different eigenvalue in each iteration.

In problems where an eigenstate is completely unknown and a state with special properties is used instead (e.g. a weighted superposition of eigenstates in the factoring problem), the corresponding register must be maintained during all iterations and repetitions. This is due to the fact that each measurement on the ancilla qubit results in a collapse in the second register. The resulting 'collapsed' state is unknown to us.

Optimal quantum circuits for general phase estimation

The question of finding optimal quantum circuits for general phase estimation has been addressed by Dam *et. al.* [74]. It was shown that the QFT-based PEA is indeed an optimal circuit (so are derived iterative schemes) and further optimization with respect to some cost function can be done only by optimizing the input state in the ancilla register. A common cost function is the window cost function that allows any error smaller than ξ :

$$C_W^\xi(\phi, \blacktriangleleft \phi \blacktriangleright) = \begin{cases} 0 & \text{if } |\phi - \blacktriangleleft \phi \blacktriangleright| < \xi, \\ 1 & \text{otherwise.} \end{cases} \quad (4.77)$$

Minimization of this cost function leads to the ancilla input state

$$|ancilla_{optW}\rangle = \frac{1}{\sqrt{2^m}} \sum_{j=0}^{2^m-1} |j\rangle. \quad (4.78)$$

This state is created by the initial Hadamard gates in the default QFT-based PEA setting. Since the state is a rather simple product state, the initial Hadamard gate in each IPEA iteration serves as well. Thus the QFT-based PEA and the IPEA are capable of producing optimal estimates with respect to the window cost function. Another commonly used cost function is the (in)fidelity cost function

$$C_F(\phi, \blacktriangleleft \phi \blacktriangleright) = \sin^2\left(\frac{\phi - \blacktriangleleft \phi \blacktriangleright}{2}\right), \quad (4.79)$$

and the minimum cost is achieved with the initial state

$$|ancilla_{opt_F}\rangle = \sum_{j=0}^{2^m-1} \sqrt{\frac{2}{2^m+1}} \sin\left(\frac{(j+1)\pi}{2^m+1}\right) |j\rangle. \quad (4.80)$$

The state $|ancilla_{opt_F}\rangle$ is derived in [91]. An important observation is that this state is an entangled state. In this case, an analytical calculation shows that the information flow is such that it cannot be simulated by the IPEA. This implies that there are cost functions for which the IPEA is not able to yield optimal estimates. To which degree entangled ancilla input states can be approximated in the IPEA is an open question.

4.4 IPEA applications

4.4.1 Iterative phase estimation: A two-qubit test-bed application

The ability to perform single qubit gates and two-qubit (entangling) gates has been demonstrated for all major quantum computation technology candidates such as superconducting quantum computation, cavity quantum electro-dynamic computation and ion trap quantum computation. Having a very limited amount of not yet reliable qubits, the question is what kind of test-bed applications can be performed. For example, five to seven qubits are already sufficient to run the smallest instances of Shor's or Grover's algorithms, but experiments with only two qubits have been so far limited only to testing Bell's inequality [92] or doing quantum state tomography [93]. Phase estimation in its iterative form (IPEA) shows up as a new test-bed application with only two qubits.

Here, I deal with three different simulations of the IPEA applied to a simple Z-rotation operator

$$U = \begin{pmatrix} e^{-i\alpha} & 0 \\ 0 & e^{i\alpha} \end{pmatrix} = \begin{pmatrix} e^{-2\pi i\phi} & 0 \\ 0 & e^{2\pi i\phi} \end{pmatrix} \quad (4.81)$$

on a two qubit system. In the simulations, the underlying Hamiltonians

$$\mathcal{H}_{xx} = B_{x_1} X^{(1)} + B_{z_1} Z^{(1)} + B_{x_2} X^{(2)} + B_{z_2} Z^{(2)} + \gamma(X^{(1)} \otimes X^{(2)}), \quad (\text{XX coupling}) \quad (4.82)$$

$$\mathcal{H}_{zz} = B_{x_1} X^{(1)} + B_{z_1} Z^{(1)} + B_{x_2} X^{(2)} + B_{z_2} Z^{(2)} + \gamma(Z^{(1)} \otimes Z^{(2)}), \quad (\text{ZZ coupling}) \quad (4.83)$$

where $B_{z_1, z_2}, B_{x_1, x_2}, \gamma$ are tunable parameters (absorbing the reduced Planck constant \hbar) and Z, X are Pauli matrices, and noise models were chosen with respect to [94]. The Hamiltonians are related to superconducting quantum computation [95].

The operator (4.81) has been identified during analysis as especially suitable since: (1) it is diagonal in the qubit eigenbasis, thus the initial preparation of its eigenstate is straightforward, (2) the phase to be measured can be controlled directly, and (3) controlled powers of this gate can be done with a very short circuit. These properties allow to keep the complexity of a real experiment low and isolate well the performance of single qubit rotations and coupling terms.

Simulations with focus on single qubit rotations performance and overall setup stability

In general, any single qubit gate $U \in SU(2)$ can be decomposed to at most three successive rotations about two non-parallel axis. Given the form of Hamiltonians (4.82) and (4.83), the natural arising single qubit rotations are R_z and R_x for both of them, and we can write

$$U = R_z(\alpha) R_x(\theta) R_z(\beta). \quad (4.84)$$

Using the Z-X decomposition (4.84), a decomposition of the gate U^2 can be obtained as

$$\begin{aligned} U^2 &= \left(R_z(\alpha) R_x(\theta) R_z(\beta) \right) \left(R_z(\alpha) R_x(\theta) R_z(\beta) \right) \\ &= R_z(\alpha) R_x(\theta) R_z(\alpha + \beta) R_x(\theta) R_z(\beta) \\ &= R_z(\alpha + \nu_1) R_x(\nu_2) R_z(\beta + \nu_3) \end{aligned} \quad (4.85)$$

for ν_1, ν_2 and ν_3 such that

$$R_x(\theta) R_z(\alpha + \beta) R_x(\theta) = R_z(\nu_1) R_x(\nu_2) R_z(\nu_3). \quad (4.86)$$

Solving the equation (4.86) gives us

$$\nu_2 = \begin{cases} 2 \arcsin \left(\sin \theta \cos \frac{\alpha + \beta}{2} \right), & \text{if (*) or (**),} \\ 2 \left(\pi - \arcsin \left(\sin \theta \cos \frac{\alpha + \beta}{2} \right) \right), & \text{otherwise,} \end{cases} \quad (4.87)$$

(*) $\sin(\alpha + \beta) = 0$ and $\cos \theta > 0$,
(**) $\cos \theta \sin(\alpha + \beta) > 0$,

$$\nu_1 = \arcsin \left(\sin \frac{\alpha + \beta}{2} \cos \frac{\nu_2}{2} \right), \quad (4.88)$$

$$\nu_3 = \nu_1. \quad (4.89)$$

Analogically, a decomposition of the gate U^{2^k} is easily calculated in the following loop:

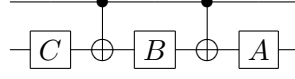
```

repeat  $k$  times:
    calculate  $\nu_2$ 
    calculate  $\nu_1$ 
     $\alpha \leftarrow \alpha + \nu_1$ 
     $\theta \leftarrow \nu_2$ 
     $\beta \leftarrow \beta + \nu_1$ 
return  $U^{2^k} \leftarrow R_z(\alpha) R_x(\theta) R_z(\beta)$ .

```

(4.90)

Knowing the angles α, θ, β corresponding to the single qubit gate U^{2^k} , we can implement (Theorem 3.3.1) its controlled version with a circuit



where

$$\begin{aligned} A &= R_z(\alpha) R_x(\theta/2) R_z(-\pi/2), \\ B &= R_z(\pi/2) R_x(-\theta/2) R_z(-(\alpha + \beta + \pi)/2), \\ C &= R_z((\beta - \alpha + \pi)/2). \end{aligned} \tag{4.91}$$

This construction can be pretty much simplified for the operator (4.81). See the circuit in the Figure 4.19. Essentially, the circuit exploits the identity $XR_z(\alpha)X = R_z(-\alpha)$.

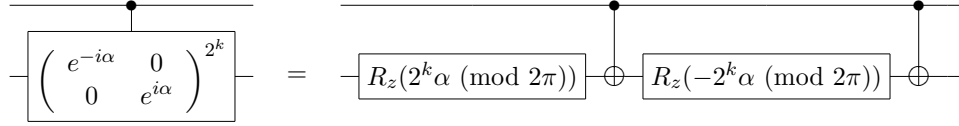


Figure 4.19: Controlled powers of the simple Z-rotation operator.

Now, the question is how to implement the Controlled-NOT gate. The XX coupling term in the Hamiltonian (4.82) produces during time t an unitary evolution described by a matrix

$$e^{-i\gamma(X \otimes X)t} = \begin{pmatrix} \iota_1 & 0 & 0 & \iota_2 \\ 0 & \iota_1 & \iota_2 & 0 \\ 0 & \iota_2 & \iota_1 & 0 \\ \iota_2 & 0 & 0 & \iota_1 \end{pmatrix}, \quad \text{where } \iota_1 = \cos \gamma t, \quad \iota_2 = -i \sin \gamma t. \tag{4.92}$$

Hadamard rotations modify (4.92) to the diagonal form

$$(H \otimes H) e^{-i\gamma(X \otimes X)t} (H \otimes H) = \text{diag}(e^{-i\gamma t}, e^{i\gamma t}, e^{i\gamma t}, e^{-i\gamma t}), \tag{4.93}$$

and additional single qubit gates lead (after simplification) to the Controlled-NOT implementation shown in the Figure 4.20.

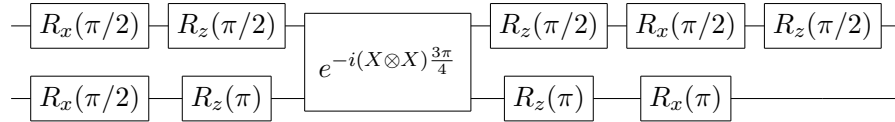


Figure 4.20: Controlled-NOT implementation using the XX coupling; $\gamma t = 3\pi/4$.

A complete minimized gate sequence implementing the k -th step of the iterative phase estimation (Figure 4.16) applied to the operator (4.81) is shown in the Figure 4.21. After m iterations (measurements), we obtain an m -bit estimate of a phase $\phi = \alpha/2\pi$ with high probability.

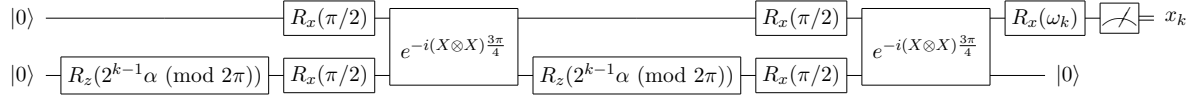


Figure 4.21: The k -th step of the IPEA applied to the operator (4.81), implemented on a two-qubit system with XX coupling. Note the block structure.

For simulations of a real experiment, we used a noise model with accurate initialization and measurement, but imperfect gates. The unitary operation of the i -th gate can be parameterized as a rotation of a certain angle φ_i around some axis (X, Z or XX). The angle is a product of the tunable strength of the corresponding control or interaction term in the Hamiltonian λ_i and time t_i the interaction is switched on, $\varphi_i = \lambda_i t_i$. Assuming a precise timing, a fluctuating interaction/control strength

$$\lambda_i \longrightarrow \lambda_i + \delta\lambda_i \quad (4.94)$$

leads to fluctuations in the rotation angle

$$\varphi_i \longrightarrow \varphi_i \left(1 + \frac{\delta\lambda_i}{\lambda_i}\right). \quad (4.95)$$

The fluctuations are assumed to be evenly distributed in the interval

$$-\frac{\Delta}{2} < \frac{\delta\lambda_i}{\lambda_i} < \frac{\Delta}{2}, \quad (4.96)$$

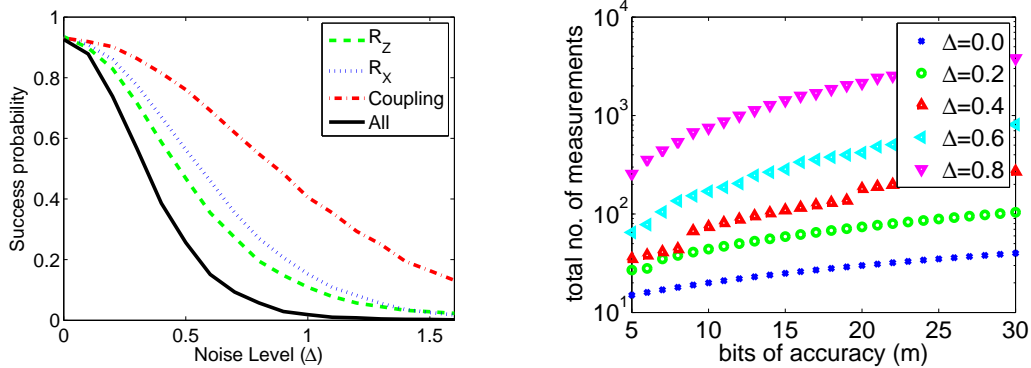
where Δ is a dimensionless parameter, indicating the strength of the noise. As an example, for $\Delta = 1$ a $\pi/2$ X-rotation will be replaced with an X-rotation with a random angle, evenly distributed between $\pi/4$ and $3\pi/4$.

The success probability of the circuit as a function of the noise level Δ is shown in the Figure 4.22a.

As can be seen the setup is quite robust towards errors up to $\Delta \sim 0.3$ and the R_z gates are the most sensitive gates. When the noise level is larger, repeated measurements on individual bits are necessary. To achieve an overall success probability of $1 - \varepsilon$, we need to increase the per bit success probability to $(1 - \varepsilon/m)$, using N_k trials (repeated measurements). The number N_k can be determined in the following way. We write down the estimates obtained from few sample runs of the circuit. From these results it is possible to calculate per bit success probabilities P_k . Next, we set the desired per bit success probability to $P'_k = 1 - \varepsilon/m$ and N_k can be calculated using the equation (4.75) (see Page 75). Essentially, N_k can be read out from the plot of the function (4.75). A convenient way is to express this cumulative distribution function in terms of the regularized incomplete beta function

$$P'_k = \sum_{s=0}^{\frac{N_k-1}{2}} \binom{N_k}{s} (P_k)^{N_k-s} (1 - P_k)^s = \text{betainc}\left(P_k, \frac{N_k+1}{2}, \frac{N_k+1}{2}\right) \quad \text{for } N_k = 1, 3, 5, \dots \quad (4.97)$$

in order to avoid numerical problems with a huge binomial coefficient, and calculate N_k using the following Matlab[®] 6 compatible code:



(a) The success probability of the IPEA correctly determining $m = 10$ binary digits of the phase $\phi = \alpha/2\pi$, as a function of the noise level Δ . The curves shows errors in all Z, X, XX-coupling gates separately, as well as in all gates simultaneously.

(b) The total number of measurements as a function of the number of bits m and noise level Δ . The noise level Δ is the same for all gates.

Figure 4.22: Results of simulations of the circuit shown in the Figure 4.21.

```
function nreps = NRepetitions (p0, perr)
% nreps is the number of trials needed to determine a bit with
% perr probability of error if
% p0 is the single shot success rate

nreps = ceil(fzero(@(x) myfun(x,p0,perr), [1 1000000]));

function f = myfun (x, p0, perr)
f = betainc (1-p0, (x+1)/2, (x+1)/2) - perr;
```

Figure 4.22b shows the total number of measurements needed to obtain the phase $\phi = \alpha/2\pi$ with m accurate bits under different noise levels and an error probability $\varepsilon < 0.05$. As one can see, repeated bitwise measurements works well up to rather high noise levels. The total number of measurements scales as $O(m \log(1/\varepsilon))$. However, above a certain noise level, when the external noise dominates over the intrinsic errors arising from the remainder δ (see Page 73), the error probability per bit ε_k becomes almost independent of the bit position k . Besides other problems, the feedback improvement does not work any more. Then N_k on average is proportional to $\log(1/\varepsilon_k) = \log(m/\varepsilon)$ and total number of measurements scales as $O(m \log(m/\varepsilon))$ with a relatively large hidden constant.

The circuit shown in the Figure 4.21 can be also rearranged using the ZZ-coupling. The results of the simulations are very similar. However, the rearranged circuit does not have the nice block structure which might be useful in a real experimental setup.

Simulations with focus on the coupling term

The construction used to implement controlled powers of the operator (4.81) shown in the Figure 4.19 is not the only possible construction given the underlying Hamiltonians (4.82)

and (4.83). Unitary gates

$$ZZ(\gamma t) = e^{-i\gamma(Z\otimes Z)t} = \text{diag}(e^{-i\gamma t}, e^{i\gamma t}, e^{i\gamma t}, e^{-i\gamma t}) \quad (4.98)$$

and

$$XX(\gamma t) = e^{-i\gamma(X\otimes X)t} \quad (4.99)$$

arising from the ZZ and XX coupling terms, respectively, can be used straightforwardly to achieve the same goal without intermediate construction of the two controlled-NOT gates. Figure 4.23a shows a complete minimized circuit implementing the k -th step of the IPEA applied to the operator (4.81), exploiting the ZZ-coupling directly. A circuit making use of the XX-coupling is shown in the Figure 4.23b.

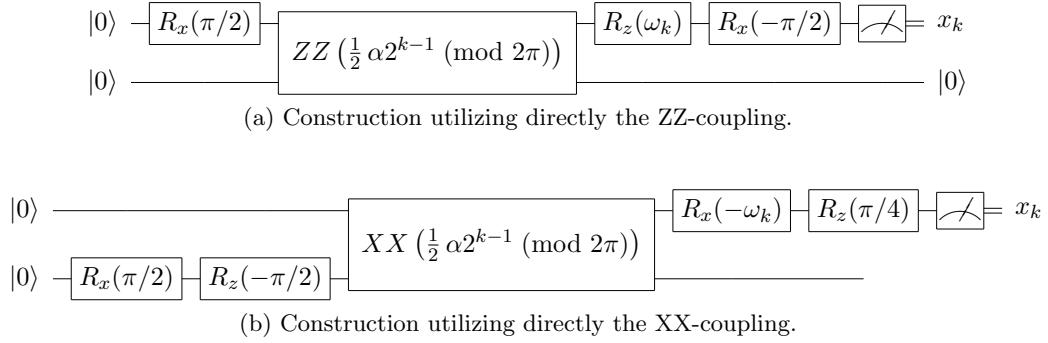


Figure 4.23: The k -th step of the IPEA applied to the operator (4.81), implemented on a two-qubit system. Implementation makes direct use of the coupling terms in the Hamiltonians (4.82) and (4.83).

In the simulations of a real experiment under the noise model described in the previous section, both circuits performed almost identically. Figure 4.24a shows the success probability as a function of the noise level Δ for a circuit with the ZZ-coupling.

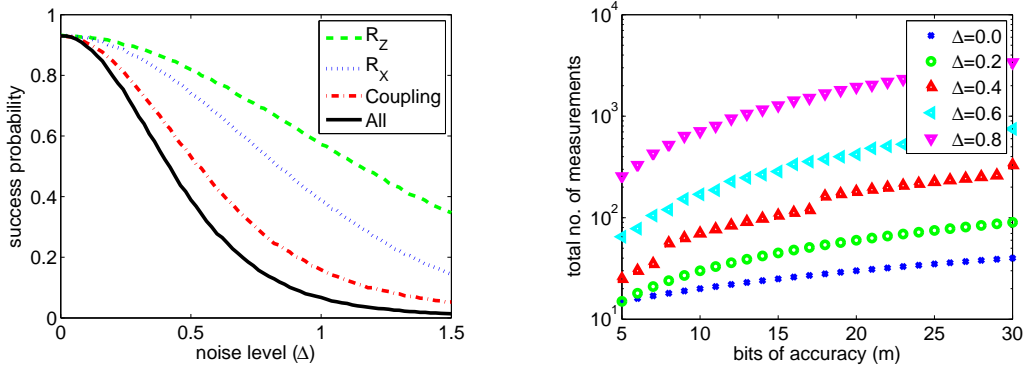


Figure 4.24: Results of simulations of the circuit shown in the Figure 4.23a.

The success probability is somewhat better compared to the result shown in the Figure 4.22a and as expected the coupling gate is the most sensitive gate. Importantly, the feedback gate

$R_z(\omega_k)$ is quite robust compared to the other gates. The total number of measurements needed to obtain the phase with m accurate bits under different noise levels and an error probability $\varepsilon < 0.05$ is then shown in the Figure 4.24b. These results are very much the same as results shown in the Figure 4.22b.

Simulations of a setup close to the Abrams-Lloyd algorithm

The two previous setups represented simple test-bed applications (benchmarks) where an input parameter α is plugged in as $\sim (2^k \alpha \pmod{2\pi})$ in each iteration of the IPEA and the obtained outcome $\blacktriangleleft \phi \blacktriangleright$ is expected to exhibit strong correlations with α , $\blacktriangleleft \phi \blacktriangleright \doteq \alpha/2\pi$. A different (stronger) benchmark can be performed within a scenario where the coupling strength γ is supposed to be unknown and the goal of the IPEA is to estimate γ .

Let us consider a circuit shown in the Figure 4.25. The ZZ gate is parameterized similarly as in the Figure 4.23a, but since the parameter α is not known now, the modulo operation is not present any more.

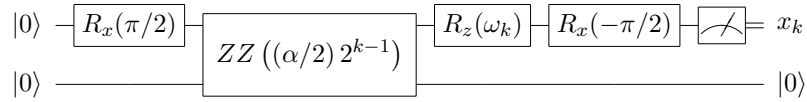


Figure 4.25: The k -th step of the IPEA applied to the operator (4.81), implemented on a two-qubit system using the entangling ZZ gate. The parameter α is considered to be unknown. The circuit estimates the coupling strength γ of the ZZ gate.

The elementary entangling gate in this circuit is the $ZZ(\alpha/2)$ gate and it describes an unitary evolution $e^{-i\gamma(Z\otimes Z)t}$ produced during a time interval t , typically order of nanoseconds. As follows from the definition of the ZZ gate (Eq. 4.98), $\alpha = 2\gamma t \pmod{2\pi}$. Since the outcome of the IPEA is an estimate of the value $\phi = \alpha/2\pi$, an estimate of γ is then equal to

$$\blacktriangleleft \gamma \blacktriangleright = \frac{\pi \blacktriangleleft \phi \blacktriangleright}{t}. \quad (4.100)$$

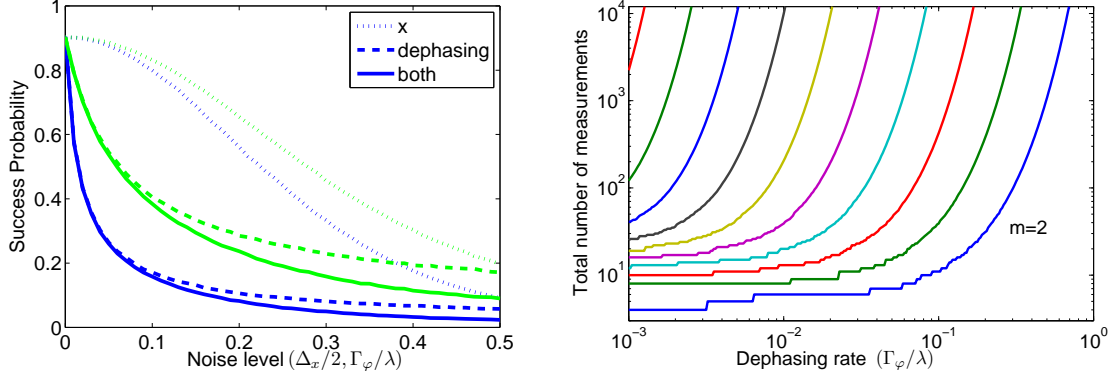
In the view of the fact that γ is unknown, the only possibility how to implement the powers of the $ZZ(\alpha/2)$ gate, as required by the steps of the IPEA, is to create a long sequences of applications of the plain $ZZ(\alpha/2)$ gate. This is equivalent to an evolution during a time interval $(2^{k-1} t)$:

$$(ZZ(\alpha/2))^{2^{k-1}} = ZZ((\alpha/2) 2^{k-1}) = e^{-i\gamma(Z\otimes Z)2^{k-1} t}. \quad (4.101)$$

In the simulations of this setup while considering superconducting qubits, it is necessary to take into account the effect of pure dephasing in the computational basis with rate Γ_φ as a result of the environmental noise and long time intervals $(2^{k-1} t)$ [94, 95]. Dephasing is a process in which a physical qubit without any error corrections is losing its phase information. In our setup, the accuracy of the $ZZ((\alpha/2)2^{k-1})$ gate will be affected with a dephasing rate proportional to Γ_φ/γ . In addition, it is considered imperfect X-rotations of the form $R_x(\pm\pi/2 + \delta_x)$, where δ_x is a normally distributed random angle with variance Δ_x . Errors in

the $R_z(\omega_k)$ gate are omitted, since as compared to errors in other gates, they do not contribute to the decrease of the success probability too significantly, see Figure 4.23a.

The success probability of the circuit as a function of dephasing rate Γ_φ/γ and variance Δ_x is shown for $m = 5$ and $m = 7$ in the Figure 4.26a. The circuit is rather robust against X-rotation errors, while being much more sensitive to dephasing.



(a) The success probability of the IPEA to correctly determine the phase $\phi = \alpha/2\pi$, with precision better than 2^{-5} (upper/green line) and 2^{-7} (lower/blue line) as a function of the noise level. The three cases of pure X-gate errors, pure dephasing, and both types of noise acting simultaneously are considered. The simulation was averaged over evenly distributed α .

(b) The total number of measurements needed to obtain the phase $\phi = \alpha/2\pi$ with precision better than 2^{-m} ($2 \leq m \leq 11$), with error probability $\varepsilon < 0.05$.

Figure 4.26: Results of simulations of the circuit shown in the Figure 4.25.

In order understand this sensitivity, an analytical expression was carried out. The per bit success probability without errors as derived in the Section 4.3.2 is

$$P_k = \cos^2(\pi 2^{k-1} (\delta 2^{-m})) = \frac{1}{2} \left(1 + \cos(\pi 2^k (\delta 2^{-m})) \right). \quad (4.102)$$

The errors modify (4.102) into

$$P_k^{err} = \frac{1}{2} \left(1 + e^{-\Delta_x^2 - |\alpha| 2^k (\Gamma_\varphi/\gamma)} \cos(\pi 2^k (\delta 2^{-m})) \right). \quad (4.103)$$

It is evident that the factor 2^k brought by dephasing into the exponent of (4.103) is responsible for the circuit's increased sensitivity to the environmental noise. As in the previous simulations, the success probability can be improved by repeated measurements of each bit. The number of per bit repetitions N_k is proportional to $\exp(2^k (\Gamma_\varphi/\gamma))$, to neutralize the effect of the 'noise term' $\exp(-\Delta_x^2 - |\alpha| 2^k (\Gamma_\varphi/\gamma))$. Figure 4.26b shows the total number of measurements needed to obtain the phase ϕ with m accurate bits under different noise levels and an error probability $\varepsilon < 0.05$. For a realistic dephasing rate $0.01 < \Gamma_\varphi/\gamma < 0.1$ [94], between 5 and 8 binary digits of ϕ can be extracted with less than 10^4 measurements.

4.4.2 Multiround protocols for reference frame alignment

The IPEA has been so far presented as a viable alternative to the QFT-based PEA in situations where the number of available qubits is a crucial limiting factor. Besides, the IPEA and other iterative variants for phase estimation can also be useful in communication tasks. In communication between two parties, a common frame of reference is usually needed. This is not less true for quantum communication where e.g. to encode information into the horizontal or vertical polarization of a photon, we have to agree on the spatial direction of these axes. Think of ground to satellite quantum cryptography [96]. While some quantum information processing tasks can be achieved completely without a shared reference frame by using entangled states of multiple systems [67, 91], this is not true in general, and a shared reference frame is always a valuable resource. The question is how to align (share) a reference frame in the most efficient way, optimally at the Heisenberg limit.

Protocols for reference frame alignment which do not require transmission of large entangled states have been studied by Rudolph and Grover [97], and de Burgh and Bartlett [98]. This is of importance since large entangled states are difficult to be communicated faithfully. Essentially, the protocols which have been discovered are instances of the Kitaev phase estimation procedure. Due to its iterative (multiround) nature, these protocols are referred to as multiround protocols. Protocols utilizing the QFT-based PEA are called single round protocols.

Phase reference alignment

Phase reference alignment corresponds to the problem of synchronizing distant clocks. There are two classical approaches to distant clock synchronization. The Einstein light-pulse synchronization protocol and the Eddington adiabatic clock transfer scheme. Einstein synchronization uses a light pulse sent between the clocks, in order to estimate the transmission time. In the Eddington scheme, a wristwatch synchronized to the first clock is sent to the place of the second clock.

Entanglement enhanced versions of both protocols have been suggested [99, 100], as well as a new protocol for clock synchronization based on shared prior entanglement [101]. In the quantum version of the Eddington protocol [100], the crucial idea is to encode the time difference into the relative phase of physical systems (qubits) having a well-defined energy difference, so called 'ticking' qubits. Then the QFT-based PEA is used to estimate this phase.

Burgh and Bartlett [98] pointed out that the protocol [100] is, however, using quantum operations in such a way that the two parties would need an a priori shared phase reference (e.g. synchronized clocks), just to be able to agree on the operations. The multiround protocol of Burgh and Bartlett (BB) solves this problem and additionally does not require entanglement. The framework for the BB protocol is as follows.

Let us assume that Alice and Bob share an inertial reference frame (relativistic effect can be neglected) and agree on a free Hamiltonian of two level atomic ticking qubit

$$\mathcal{H}_{tick} = \frac{\omega}{2}Z = \frac{\omega}{2}(|0\rangle\langle 0| - |1\rangle\langle 1|). \quad (4.104)$$

This means they can agree on the upper and lower energy eigenstates and the frequency standard ω . States $|0\rangle$ and $|1\rangle$ can also be understood as defining a common shared Z axis

of the Bloch sphere that describes the two dimensional qubit (atomic) Hilbert space. The shared Z axis viewed as a resource is enough for Alice and Bob to agree on a qubit state

$$|\psi\rangle = \cos\theta|0\rangle + e^{i\xi}\sin\theta|1\rangle, \quad (4.105)$$

up to the phase ξ . They would need a shared phase reference to agree on a qubit state completely. Let us assume that both Alice and Bob define their local phase reference relatively to their clocks that are ticking at the same frequency ω . The time difference of their clocks is given by $\Delta t = t|_B - t|_A$ (Bob's clock is ahead). Then for example, a state described by Alice as $|\psi\rangle_A = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ corresponds to the state $|\psi\rangle_B = (|0\rangle + e^{i(\omega\Delta t)}|1\rangle)$ in Bob's frame.

The state of the ticking qubit evolving under the free Hamiltonian \mathcal{H}_{tick} can be pictured as the Bloch vector rotating anticlockwise about the Z axis, however, it is more convenient to work in the rotating frame in which the qubit no longer ticks. In this rotating frame, it is assumed that Alice and Bob use the Rabi pulses to interact with the qubit. The Rabi Hamiltonian is defined as

$$\mathcal{H}_{Rabi}(\alpha) = \frac{\Omega}{2}(X \cos\alpha + Y \sin\alpha), \quad (4.106)$$

where α is the phase angle of the laser, defined relatively to the local clock. The Rabi pulse applied for time $t = \beta/\Omega$ corresponds to a unitary operator

$$U(\alpha, \beta) = e^{-i\alpha Z/2} e^{-i\beta X/2} e^{i\alpha Z/2} = R_z(\alpha) R_x(\beta) R_z(-\alpha). \quad (4.107)$$

Using the operator $U(\alpha, \beta)$ both parties, Alice and Bob, can relatively to their clock implement arbitrary single qubit gates. For example, the NOT gate $X = U(0, \pi)$, the Hadamard gate $H = XU(\pi/2, \pi/2)$, and the $\pi/8$ gate $T = HU(0, \pi/4)H$. It follows that an operation $U(\alpha|_B, \beta) = U_B(\alpha|_B)$ performed by Bob corresponds to the operation $U_A(\alpha|_A - \phi)$ performed by Alice, where $\phi = \omega\Delta t$;

$$U_B(\alpha|_B) = U_A(\alpha|_A - \phi) = R_z(-\phi) U_A(\alpha|_A) R_z(\phi). \quad (4.108)$$

The relation of operations U_A and U_B as described by the equation (4.108) can be translated into the language of the Bloch sphere as shown in the Figure 4.27. Alice and Bob share the Z axis, but their respective X -axes X_A and X_B do differ by an unknown angle $\phi = \omega\Delta t$. The problem of clock synchronization is hereby reduced to estimating the angle ϕ .

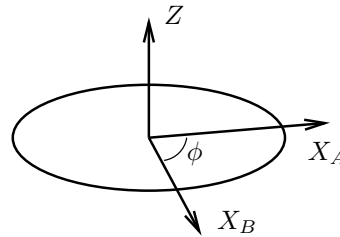


Figure 4.27: Alice and Bob share the Z axis. However, their X -axes differ by an unknown angle ϕ . Working in the rotating frame simplifies this picture by being static and not rotating anticlockwise about the Z axis with the angular frequency ω .

The question is now how to encode an angle $\sim (2^k\phi)$ into the relative phase of a qubit to enable the phase estimation procedure. The BB protocol uses a clever phase amplifying

technique known from the Grover search algorithm [10, 4] which can be graphically viewed as mirroring a qubit state about the X_A and X_B axes. Let $|\psi\rangle_A = \cos\theta|0\rangle + e^{i\xi}\sin\theta|1\rangle$ be a qubit state described in the Alice's phase reference frame. The result of mirroring the state $|\psi\rangle_A$ about the X_B axis and then about the X_A axis is shown in the Figure 4.28.

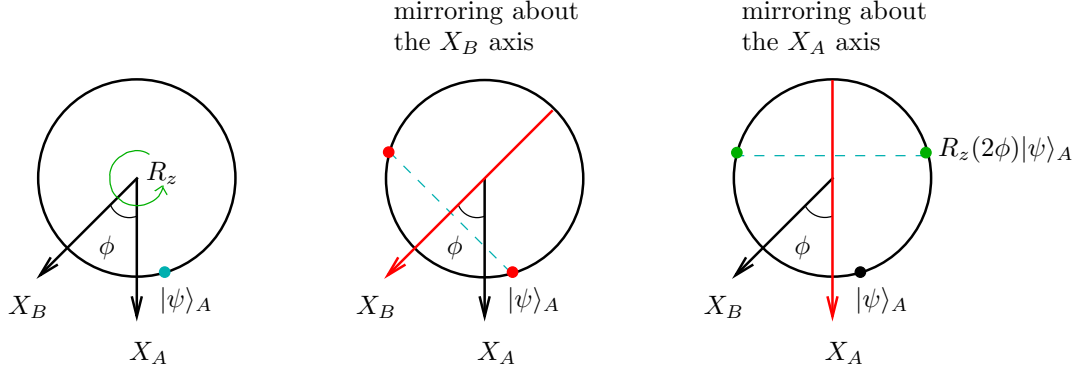


Figure 4.28: Grover phase amplification method graphically viewed as a mirroring technique.

In practice, this mirroring effect can be achieved by a qubit exchange between Alice and Bob. Alice prepares a state $|\psi\rangle_A$ and sends the qubit to Bob, Bob applies his NOT gate X_B , sends the qubit back, and Alice applies her NOT gate X_A . The net gain of this exchange expressed in Alice's phase reference frame is that a relative phase of the qubit is increased by an angle 2ϕ ,

$$X_A X_B |\psi\rangle_A = X_A \underbrace{\left(R_z(-\phi) X_A R_z(\phi) \right)}_{R_z(\phi)} |\psi\rangle_A = R_z(2\phi) |\psi\rangle_A. \quad (4.109)$$

It is now easy to see that $X_A X_B H_A |0\rangle_A = \frac{1}{\sqrt{2}}(|0\rangle + e^{i(2\phi)}|1\rangle)_A$ and the phase ϕ can be estimated using the Kitaev phase estimation applied to the operator $(X_A X_B)$, see Figure 4.29. This is the BB protocol for phase reference alignment.

$$|0\rangle_A \text{ --- } \boxed{H_A} \text{ --- } \boxed{(X_A X_B)^{2^{k-1}}} \text{ --- } \boxed{H_A} \text{ --- } \boxed{\text{meter}} = x_k$$

Figure 4.29: The Burgh and Bartlett protocol for phase reference alignment.

Utilizing the IPEA derived in the Section 4.3.2, the scaling of the BB protocol can be improved. In order to do so, Alice only needs to modify the scheme by inserting the $R_z(\omega_k)$ feedback. Given the Rabi Hamiltonian (4.106), the $R_z(\omega_k)$ gate can be implemented by a sequence of pulses producing $(H_A U_A(0, \omega_k) H_A)$. The IPEA applied to the operator $(X_A X_B)$ is shown in the Figure 4.30.

$$|0\rangle_A \text{ --- } \boxed{H_A} \text{ --- } \boxed{(X_A X_B)^{2^{k-1}}} \text{ --- } \boxed{H_A} \text{ --- } \boxed{U_A(0, \omega_k)} \text{ --- } \boxed{\text{meter}} = x_k$$

Figure 4.30: Improved protocol for phase reference alignment.

Spatial axes alignment

The problem of establishing either a single direction in space or an orthogonal trihedron (xyz axes) using quantum enhanced methods was studied extensively in the past, see [97] for a list of references. A particular multiround protocol making no use of entanglement is discussed by Rudolph and Grover in [97]. The framework for the RG protocol is as follows.

Let us assume that Alice and Bob define their spatial axes relatively to their single qubit rotations axes. Additionally, let the rotation matrix describing the change from Alice's to Bob's frame of reference be given by

$$R = R_z(\xi) R_y(\theta) R_z(\phi) = e^{-i(\xi+\phi)/2} \begin{pmatrix} \cos \frac{\theta}{2} & -e^{i\phi} \sin \frac{\theta}{2} \\ -e^{i\xi} \sin \frac{\theta}{2} & e^{\phi+\xi} \cos \frac{\theta}{2} \end{pmatrix}. \quad (4.110)$$

The goal is to estimate the Euler angles ϕ, θ, ξ .

Rudolph and Grover showed explicitly that the k -th bit describing the angle θ (Kitaev-like approach) can be determined by applying the transform

$$U(k) = (Z_B Z_A)^{2^{k-1}} = (Z_B R^\dagger Z_B R)^{2^{k-1}} = \begin{pmatrix} \cos(2^{k-1}\theta) & -e^{i\phi} \sin(2^{k-1}\theta) \\ e^{-i\phi} \sin(2^{k-1}\theta) & \cos(2^{k-1}\theta) \end{pmatrix} \quad (4.111)$$

to a qubit prepared in the state $|\psi\rangle_B = |0\rangle$. The state $|\psi\rangle_B$ is expressed in the Bob's reference frame. The protocol goes as follows. Bob prepares the state $|\psi\rangle_B$ and sends the qubit to Alice, Alice applies her Z_A gate, sends the qubit back, and Bob applies his Z_B gate. The qubit exchange is repeated 2^{k-1} times and then Bob performs the measurement to estimate the bit x_k . The outcome '0' is observed with probability

$$P_0 = \left| \cos(2^{k-1}\theta) \right|^2 = \cos^2(2^{k-1}\theta) \quad (4.112)$$

and the outcome '1' with probability

$$P_1 = \left| e^{-i\phi} \sin(2^{k-1}\theta) \right|^2 = \sin^2(2^{k-1}\theta). \quad (4.113)$$

The other angle ξ or ϕ needed to align the spatial axes can be estimated by changing the transform that Alice and Bob perform and/or the initial state Bob prepares. The important point is that it is possible to make the outcome probabilities dependent only on a specific angle at time. Intriguingly, in the light of the IPEA, it is not clear how to improve the scaling of the RG protocol by using a feedback. That is to find a unitary transform $V(\omega_k)$ such that the outcome probabilities remain dependent only on θ and ω_k , for example. In fact, I was not able to find such transform, nor to rigorously prove that it does not exist. One open direction I keep working on is to focus on the order of angles to be estimated.

4.5 Quantum phase estimation conclusions

In this chapter, I discussed the broad issue of quantum phase estimation algorithm as a practical approach for problem solving on a quantum computer. Alternatively, it is possible to use the framework of quantum Fourier sampling. An example of a problem which is more naturally understood using the quantum Fourier sampling is the Shor factoring algorithm [12]. On the other side of the problem spectrum is the Abrams-Lloyd algorithm [81] which is rather well described using the phase estimation.

In contrast to the modern quantum phase estimation based on the Fourier transform, the original Kitaev PEA [54] did not explicitly use the quantum Fourier transform and Kitaev's approach looked completely different from Shor's. However, soon it was recognized that the Kitaev's approach can be explained in terms of the Fourier transform over \mathbb{Z}_2^n and that a transform over \mathbb{Z}_{2^n} would be in fact more appropriate. Hence Kitaev's approach was unified with Shor's QFT approach for factoring and the QFT-based PEA was born [4, 75]. While this is perfectly fine on the big scale of understanding the power of a quantum computer, a few interesting aspects of the original Kitaev work faded out.

Guzik *et. al.* [85] building on the work of Abrams and Lloyd even carried out their own iterative phase estimation procedure in order to reduce the number of required qubits at time. Paradoxically, the Kitaev's circuit would require less resources under given circumstances and it is more robust to the environmental noise. When I started working on the phase estimation algorithm, my task was to investigate how to implement an instance of the Abrams-Lloyd PEA on three or less qubits. The stepping stones were the QFT-based PEA and the work of Guzik *et. al.*. As a first step, I rediscovered the Kitaev's circuit by reducing gates in the Guzik PEA to absolute minimum. Hereupon, I focused on the exigent classical postprocessing and incorporated it to the quantum circuit. Later, the validity of the obtained IPEA was verified using the adaptive measurement in the Fourier basis [86] (semiclassical QFT).

The QFT-based PEA was implemented experimentally a few times [102, 103] on three qubits and two-bit phase estimates were obtained. Using a benchmark circuit of similar complexity, I showed that a much higher precision can be achieved using the IPEA instead.

In the Section 4.3.3, differences between the QFT-based PEA and the IPEA were discussed. Using the IPEA, repetitions needed due to the environmental noise and/or a high success probability require less resources compared to the QFT-based PEA. Next, in the context of the Abrams-Lloyd algorithm, the IPEA allows to exchange a long coherence time of an input eigenstate for a per iteration input state preparation. On the other hand, the QFT-based PEA enables to employ entangled ancilla input states which can be optimized with respect to a given cost function.

The question of yielding optimal estimates with respect to the fidelity cost function without using entanglement was discussed a lot in multi-round protocols for reference frame alignment [97, 98]. It was concluded that presented protocols do not produce such optimal estimates. An existence of better protocols was left as an open problem. It should be clear from this thesis that the proposed protocols could not produce desired optimal estimates, since they use the Kitaev circuit which has a small logarithmic penalty compared to the IPEA. Moreover, an exact simulation of information flow in the QFT-based PEA cannot be achieved using small

fixed-size ancilla registers present in iterative schemes. Now, an open problem is to which degree entangled ancilla input states can be approximated in the IPEA.

Chapter 5

Quantum cryptography

Special properties of quantum physical systems such as the no-cloning property of an unknown quantum state and the randomness of a wave function collapse allow for many quantum enhanced cryptographic primitives. These primitives range from authentication, secret key generation/distribution to covered communication (steganography) and more. Up to date, the BB84 protocol [104] and its variants [105, 106] for quantum key generation are the most celebrated quantum cryptography primitives.

For myself, I am particularly interested in quantum steganography and its applications related to watermarking protocols for authentication. Similarly to the classical steganography, hidden communication is achieved via embedding a message into a redundant part of a (innocent looking) cover medium. The main difficulties for a steganographer are to identify the redundant part of the cover medium and insert a message in an unobtrusive way. For watermarking, it is additionally desirable for the message to be tied with the cover medium so well that the message cannot be removed by a third party without damaging the cover medium substantially.

Embedding methods may differ to a large extent in the medium access level. The main levels are (1) physical carrier level, (2) error correction code level, and (3) data formats and protocol level. Steganography at the physical carrier level is somewhat specific in the sense that the theory of information coding does not play a big role, but instead it is a clear race in the technology available to the steganographer and steganalyst. To put it more to the context, let us take a look at the first commercially available products for quantum key generation [107, 108]. For these devices to work as securely as proposed by the theory, single photon emitters and detectors are crucial. Nevertheless, single photon sources are often substituted with an industry-standard weak coherent pulse approach. Whilst practical, this approach suffers from non-zero probability of multiple-photons events. Apart from the introduced security loophole [109, 110], these extra photons can be, in principle, used for steganographic purposes while legitimate quantum key generation seems to happen.

Regarding quantum steganography on the error correction level, first steps along this line were accomplished by J. Gea-Banacloche in [111]. It was proposed to encode the message to be hidden into the error syndrome. Compared to the classical counterpart, quantum error

correction codes (QECC) are very suitable for steganographic purposes. This is due to the fact that quantum errors are continuous and quantum gates are very faulty, and therefore QECC schemes require a rather big bunch of extra physical qubits to create and preserve logical qubits. Especially, at higher levels of concatenated QECC [112], it is plausible to assume that errors occur not too often, and the capacity of the QECC scheme can serve as a steganographic channel.

An example of simple QECC scheme is shown in the Figure 5.1. Gea-Banacloche used the same scheme to present his ideas.

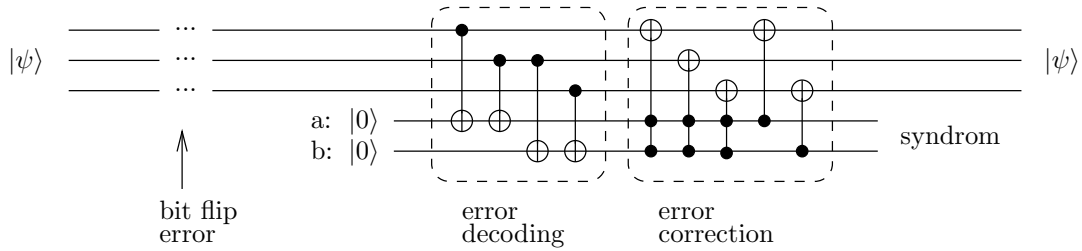


Figure 5.1: A circuit to decode the error syndrome of the 3-qubit repetition code and correct the error accordingly. Note that the correction part does not require measurement on the ancilla qubits a and b .

The state

$$|\psi\rangle = \alpha|000\rangle + \beta|111\rangle = \alpha|0_L\rangle + \beta|1_L\rangle$$

represents a logical qubit protected against a single bit flip error on any of the three physical qubits, using the 3-qubit repetition code [113, 4]. After the decoding & correction stage, the state of the ancilla qubits a and b holds the information on what had happened

- $|00\rangle_{ab}$: no error occurred,
- $|10\rangle_{ab}$: bit flip error $X^{(1)}$ on the first physical qubit,
- $|11\rangle_{ab}$: bit flip error $X^{(2)}$ on the second physical qubit,
- $|01\rangle_{ab}$: bit flip error $X^{(3)}$ on the third physical qubit.

Now, assuming that no accidental error will happen, it may be the steganographer who deliberately introduces an error and herewith transfers two bits of information. A convenient method for introducing an error (inserting a message) is to use the decoding & correction block in a reversed mode, see Figure 5.2.

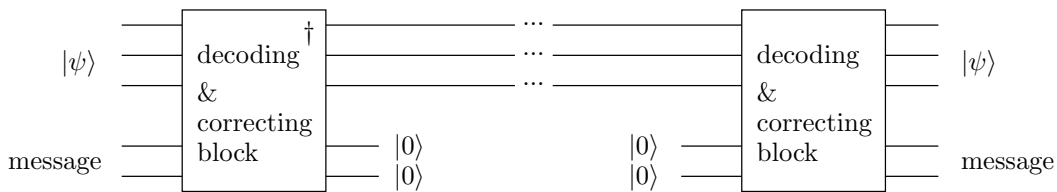


Figure 5.2: Quantum steganography on the error correction level.

This steganographic channel can be also used for a simple authentication. Alice and Bob a priori agree on a special message (a tag). Then later, when Alice sends a state $|\psi\rangle$ to Bob,

Bob considers the received state as authentic only if the embedded message corresponds to the agreed tag. Additionally, it is required that a third party (with an access to a transmission channel) should not be able to change the state $|\psi\rangle$ without affecting the tag as well. It was shown that this requirement can be fulfilled only if the tag is a quantum superposition of the form

$$\iota|00\rangle_{ab} + \kappa|01\rangle_{ab} + \mu|10\rangle_{ab} + \nu|11\rangle_{ab}$$

and not a classical digital tag, e.g. '10'. Effectively, the operation (decoding & correction block)[†] encodes the coefficients ι, κ, μ, ν together with the coefficients α, β in such a way that the transmitted state $|\psi\rangle$ becomes encrypted and it is impossible for a third party to modify α, β without affecting the tag. In fact, the tag becomes a watermark.

These results indicate that quantum authentication implies encryption and that it is impossible to use digital signing to authenticate quantum states. Indeed, this was proven by Barnum *et. al.* [114] in a more general study on authentication of quantum states. A limit case study addressing authentication of a single qubit using a shared quantum key of minimum length has been conducted by Curty *et. al.* [115]. Particular details of this study are discussed in the Section 5.1.

Finally, we discuss quantum steganography at the level of data formats and protocols. In digital data formats, tens of methods are known that exploit redundancy in JPEG images [116], MP3 music files [117], binary executable files [118] and so on. However, there are no such formats in the quantum domain by this time. It is too early to predict how and when will the technology reach this stage. Nevertheless, some general methods such as modifying the least significant bits can be ported to the quantum domain quite easily [119].

Let us assume that a classical information, e.g. bits representing a picture, is going to be transmitted using qubits (a quantum channel). Each bit is mapped in some standard way to a qubit state. Additionally, since the carried information is a classical information and the mapping is publicly known, anybody with an access to the quantum channel can make copies of the qubits and check what is being transmitted.

Under this scenario, Alice and Bob want to establish a steganographic channel. It is enough for them to merely agree on a non-standard base encoding and measurement on the qubits which corresponds to the least significant bits. Hereby they set up a secret channel. Other parties who are not aware of this deal treat all qubits in a standard way. As a consequence, these parties may obtain wrong bit values upon measurement, but due to the low significance of corresponding bits it can hardly be detected. Additionally, once the qubit carrying hidden information is collapsed by a measurement, no later leakage of the stego key (the deal of Alice and Bob) enables to obtain the past hidden information. There is no classical analogue of this pretty striking property. Of course, the difficult part for the steganographer is to properly identify which bits are the least significant ones. A variant of the above described method has been successfully used by K. Martin [120] to establish a steganographic channel within the BB84 protocol. Arbitrary third party sees only that a legitimate quantum key generation takes place.

Regarding quantum steganography with entanglement, it is possible to build on schemes which are based on the superdense coding [42]. Basically, Alice and Bob start with a shared entangled pair of qubits. Then, (1) Alice applies a local unitary operation to her qubit, (2) Alice sends the qubit to Bob, (3) Bob performs joint measurement on both qubits, and (4)

the result represents a delivered two-bit message. A third party having an access to the sent off qubit while in traffic cannot get even a single bit of that message.

5.1 Qubit authentication with a quantum key of minimum length

Curty *et. al.* [115] studied the question whether a single qubit can be authenticated with a quantum key of minimum length. It was shown that, unlike in the case of quantum-authenticated classical bit value [121], this is not possible. Nevertheless, it is interesting to see how the protocol works since all what is ultimately needed to make the protocol operating securely is to increase the length of the key [122].

The protocol [115] is as follows. Alice wants to send a general quantum state (possibly a mixed state) described by the density operator ρ_M acting on a two-dimensional message space M . As in a classical authentication, a tag must be attached to the message, in order to allow to Bob to convince himself about the authenticity of the message. Let the tag be given by a density operator ρ_T acting on a two dimensional tag space T . In principle, the tag can be larger, however, this is advantageous only in a presence of noise. No extra security is added. The tag space T has to be divided into two orthogonal subspaces. One subspace represents a valid tag, while the other represents an invalid tag. Without loss of generality, the state $\rho_T = |0\rangle\langle 0|$ can be fixed as a valid tag.

The space of tagged messages is then defined as $\epsilon = M \otimes T$ and the tagged message is given by $\rho_\epsilon = \rho_M \otimes \rho_T$. On the space ϵ a unitary coding set $\{I_\epsilon, U_\epsilon\}$ is defined, where I_ϵ is the identity operation and U_ϵ a unitary transformation. A shared quantum key has the form of a maximally entangled EPR pair, e.g. $|\psi\rangle_{ab} = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$. Alice and Bob each own one qubit of that EPR pair. The key can be publicly known as well as the coding set. The state of the global system (key + tagged message) is given by

$$\rho_{abe} = |\psi\rangle\langle\psi|_{ab} \otimes \rho_\epsilon = |\psi\rangle\langle\psi|_{ab} \otimes \rho_M \otimes \rho_T. \quad (5.1)$$

Encoding stage. Alice wants to send an authenticated message to Bob. In order to do so, Alice performs an encoding operation $E_{a\epsilon}$ which acts on her part of the key and the tagged message. With respect to the coding set $\{I_\epsilon, U_\epsilon\}$, the encoding operation is defined as

$$E_{a\epsilon} = |0\rangle\langle 0|_a \otimes I_b \otimes I_\epsilon + |1\rangle\langle 1|_a \otimes I_b \otimes U_\epsilon. \quad (5.2)$$

Once the encoding operation is applied, Alice sends the encoded tagged message to Bob. The state of the global system after the encoding operation is given by

$$\begin{aligned} \rho_{abe}^{enc} &= E_{a\epsilon} \rho_{abe} E_{a\epsilon}^\dagger \\ &= \frac{1}{2} \left(|01\rangle\langle 01| \otimes \rho_\epsilon - |01\rangle\langle 10| \otimes \rho_\epsilon U_\epsilon^\dagger - |10\rangle\langle 01| \otimes U_\epsilon \rho_\epsilon + |10\rangle\langle 10| \otimes U_\epsilon \rho_\epsilon U_\epsilon^\dagger \right). \end{aligned} \quad (5.3)$$

Decoding stage. Bob receives the encoded tagged message and decodes it using a decoding operation

$$D_{b\epsilon} = I_a \otimes |0\rangle\langle 0|_b \otimes U_\epsilon^\dagger + I_a \otimes |1\rangle\langle 1|_b \otimes I_\epsilon. \quad (5.4)$$

5.1. QUBIT AUTHENTICATION WITH A QUANTUM KEY OF MINIMUM LENGTH 95

The state of the global system after the decoding is described as

$$\begin{aligned}
 \rho_{ab\epsilon}^{dec} &= D_{b\epsilon} \rho_{ab\epsilon}^{enc} D_{b\epsilon}^\dagger \\
 &= \frac{1}{2} \left(|01\rangle\langle 01| \otimes \rho_\epsilon - |01\rangle\langle 10| \otimes (\rho_\epsilon U_\epsilon^\dagger) U_\epsilon - |10\rangle\langle 01| \otimes U_\epsilon^\dagger (U_\epsilon \rho_\epsilon) + |10\rangle\langle 10| \otimes U_\epsilon^\dagger (U_\epsilon \rho_\epsilon U_\epsilon^\dagger) U_\epsilon \right) \\
 &= |\psi\rangle\langle \psi|_{ab} \otimes \rho_\epsilon.
 \end{aligned} \tag{5.5}$$

Verification. Bob takes decoded tagged message, which is formally given by $\rho_\epsilon^{dec} = \text{tr}_{ab}(\rho_{ab\epsilon}^{dec})$, and measures the tag portion. If the outcome belongs to a valid tag subspace of space \mathbb{T} , then the extracted message is considered to be authentic. Using the gate model, the protocol can be depicted as shown in the Figure 5.3.

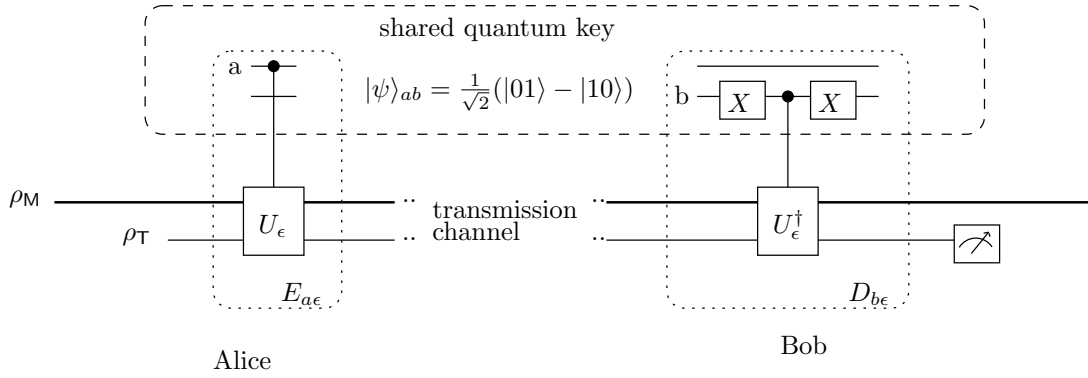


Figure 5.3: A protocol for qubit authentication with a quantum key of minimum length.

Message attack

Let us now consider a 'message attack' performed by an adversary party called Eve. This party with a full access to a public transmission channel sees a mixed state

$$\begin{aligned}
 \rho_\epsilon^E &= \text{tr}_{ab}(\rho_{ab\epsilon}^{enc}) \\
 &= \frac{1}{2} \left(\rho_M \otimes |0\rangle\langle 0|_{\mathbb{T}} + U_\epsilon (\rho_M \otimes |0\rangle\langle 0|_{\mathbb{T}}) U_\epsilon^\dagger \right).
 \end{aligned} \tag{5.6}$$

Her task is to find a transformation Q_E which, applied to ρ_ϵ^E , will modify the ρ_M keeping the tag portion intact. The authors of [115] gave an existential proof that such a transformation always exists regardless of the choice of U_ϵ , thus the protocol is not secure. However, the form of Q_E with respect to U_ϵ and the range possible damages was not discussed. In a later paper [122], it was shown that if the key is longer, then there exists a larger coding set such that Eve's desired Q_E does not exist.

An interesting point is what exactly happens when the operation U_ϵ is a separable gate. Classically, the tag must be a function of both the key and the message. Now, since the mixed state (5.6) is achievable even if Alice and Bob had shared only a classical one bit key, this indicates that U_ϵ should be an entangling gate to limit Eve's actions, at least. However, the shared key is a quantum one and the question is whether entanglement present in the key can compensate for non-entangling operation U_ϵ .

Let U_ϵ be a separable gate of the form $U_\epsilon = U_M \otimes U_T$, then

$$\rho_\epsilon^E = \frac{1}{2} \left(\rho_M \otimes |0\rangle\langle 0|_T + (U_M \rho_M U_M^\dagger) \otimes (U_T |0\rangle\langle 0|_T U_T^\dagger) \right). \quad (5.7)$$

For $Q_E = R \otimes I$, where $Q_E \in \epsilon$, $R \in M$ and $I \in T$, we have

$$\begin{aligned} \rho_\epsilon^{Q_E} &= Q_E \rho_\epsilon^E Q_E^\dagger \\ &= \frac{1}{2} \left(R \rho_M R^\dagger \otimes |0\rangle\langle 0|_T + \left(R (U_M \rho_M U_M^\dagger) R^\dagger \right) \otimes \left(U_T |0\rangle\langle 0|_T U_T^\dagger \right) \right). \end{aligned} \quad (5.8)$$

And after Bob's decoding operation we get

$$\rho_\epsilon^{Q_E, dec} = \frac{1}{2} \left(R \rho_M R^\dagger \otimes |0\rangle\langle 0|_T + \left(U_M^\dagger R U_M \rho_M U_M^\dagger R^\dagger U_M \right) \otimes \left(U_T^\dagger U_T |0\rangle\langle 0|_T U_T^\dagger U_T \right) \right) \quad (5.9)$$

and $\text{tr}_M \left(\rho_\epsilon^{Q_E, dec} \right) = |0\rangle\langle 0|_T$.

Hence the adversary party is always able to change the message ρ_M keeping the tag portion intact, and there are **no constraints** on the operation R . Having some statistics of usually sent messages, R can be even optimized to cause maximal damage. This result also implies that there is no added security in the protocol by using quantum key instead of a classical key.

Secret-key discussion

Authenticating quantum data makes sense only in a scenario where a reliable technology for quantum information processing is available. This means that there are no extra costs in using quantum keys instead of classical ones. Moreover, quantum keys have better key-management properties thanks to the no-cloning theorem, and, strikingly, in the above described protocol the EPR pair does not collapse so it can be in principle reused.

But anyway, reliable storing of an EPR pair requires a periodic entanglement purification [40]. In between two successive refreshments the EPR pair should be inevitably considered as corrupted to some degree. A well designed protocol has to be insensitive to a small corruption (entanglement purification guarantees to output only almost pure states), and a non-negligible corruption should result in a high probability of rejecting received messages as authentic ones.

Let us correlate the corruption of the EPR pair with the quantity p of a maximally mixed state in the mixture. The state of the quantum key is then described by a density operator

$$\rho_\psi^p = (1-p) |\psi\rangle\langle\psi|_{ab} + p \frac{I_{ab}}{4}. \quad (5.10)$$

When the protocol is executed with this mixed key the resulting global state of the system is described as

$$\rho_{abc\epsilon}^{dec} = D_{b\epsilon} \left(E_{a\epsilon} \left(\rho_\psi^p \otimes \rho_M \otimes |0\rangle\langle 0|_T \right) E_{a\epsilon}^\dagger \right) D_{b\epsilon}^\dagger, \quad (5.11)$$

and the decoded message after tracing out the key is

$$\begin{aligned} \rho_\epsilon^{dec} &= \text{tr}_{\rho_\psi^p} \left(\rho_{abc\epsilon}^{dec} \right) \\ &= \frac{2-p}{p} (\rho_M \otimes |0\rangle\langle 0|_T) + \frac{p}{4} \left(U_\epsilon^\dagger (\rho_M \otimes |0\rangle\langle 0|_T) U_\epsilon + U_\epsilon (\rho_M \otimes |0\rangle\langle 0|_T) U_\epsilon^\dagger \right). \end{aligned} \quad (5.12)$$

5.1. QUBIT AUTHENTICATION WITH A QUANTUM KEY OF MINIMUM LENGTH 97

Here, we can see that the probability of ρ_ϵ^{dec} passing Bob's verification procedure is unpleasantly high even for apparently corrupted key. Therefore, in this protocol and its variants a classical key should be preferred.

Chapter 6

Conclusions

In this thesis, I have addressed the emerging field of quantum computing. Since at the time of starting my PhD studies I was a complete newcomer to the field, a relatively large part of the thesis deals with portions of linear algebra essential to understanding the framework of quantum mechanics, and with comparison of deterministic, probabilistic and quantum Turing machines in order to emphasize the fundamental differences. The quantum gate model and the thin borderline between quantum evolutionary universal sets of gates and not even classically universal sets of gates are discussed as well. At this point, the thesis hopefully provides a self contained description of the main issues related to quantum computing.

The major part of the thesis concerns the quantum phase estimation algorithm. This algorithm presents a fundamental approach for problem solving on a quantum computer and it can be used to well explain most of known quantum algorithms which yield an exponential speed-up compared to their classical counterparts. The two most prominent quantum algorithms are the Shor factoring algorithm and the Abrams-Lloyd algorithm for finding energy eigenvalues. In Section 4.2.5, it was shown how to cast both algorithms as a phase estimation problem and it was stressed why both algorithms are efficient with respect to the input size of the problem.

The novel work I did encompasses a new simple proof of the lower bound of the PEA success probability, a derivation of an iterative scheme for quantum phase estimation (IPEA) from the Kitaev phase estimation, a study of robustness of the IPEA utilized as a few-qubit testbed application and an improved protocol for phase reference alignment. There is also a chapter concerning quantum steganography and authentication which provides additional insights into quantum enhanced protocols, however, no substantial contribution in this field is presented.

The design and study of robustness of small testbed applications currently represents one of the chief goals in the quantum computing field. The famous theorem for fault-tolerant quantum computation uses a very general noise model and it is not clear at all whether the theorem actually holds in an experimental setting. Testbed applications provide an important way how to realize the potential of many-body quantum systems as quantum computers and give understanding to the effects of environmental noise. It might happen that our universe is not generous enough to allow for robust large-scale quantum computing, however, quantum

information processing plays an eminent role in forthcoming quantum sensors and similar nanoscale single purpose devices. Think of the energy transfer in plants during photosynthesis [11].

Regarding future work, it is planned to work out a complete example of a quantum chemistry calculation (the Abrams-Lloyd algorithm). That is to choose a small quantum system such as the Helium atom, map a corresponding truncated Hamiltonian to the states of two to three qubits, simulate the evolution, adiabatically prepare an approximate ground state and utilize the IPEA to estimate the ground state energy.

Bibliography

- [1] Alex Smith. The Wolfram 2,3 Turing machine is universal. The Wolfram research prize, 2007. <http://www.wolframscience.com/prizes/tm23/>.
- [2] C. H. Papadimitriou. *Computational complexity*. Addison-Wesley, 1994.
- [3] Herman Chernoff. The Chernoff bound. http://en.wikipedia.org/wiki/Chernoff_bound.
- [4] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.
- [5] Alexander Holevo. Information theoretical aspects of quantum measurements. *Problems of Information Transmission (USSR)*, 9 (2):31–42, 1973.
- [6] Charles H. Bennett. Logical reversibility of computation. *IBM Journal of Research and Development*, 17:525–532, 1973.
- [7] Eduard Fredkin and Tommaso Toffoli. Conservative logic. *International Journal of Theoretical Physics*, 21 (3–4):219–253, 1982.
- [8] Paul A. Benioff. Quantum mechanical models of Turing machines that dissipate no energy. *Physical Review Letters*, 48:1581–1585, 1982.
- [9] Daniel S. Abrams and Seth Lloyd. Nonlinear quantum mechanics implies polynomial-time solution for NP-complete and #P problems. *Physical Review Letters*, 81:3992–3995, 1998.
- [10] Lov K. Grover. Quantum mechanics helps in searching for a needle in a haystack. *Physical Review Letters*, 79:325–328, 1997.
- [11] Gregory S. Engel, Tessa R. Calhoun, Elizabeth L. Read, Tae-Kyu Ahn, Tomáš Mančal, Yuan-Chung Cheng, Robert E. Blankenship, and Graham R. Fleming. Evidence for wavelike energy transfer through quantum coherence in photosynthetic systems. *Nature*, 446:782–786, 2007.
- [12] Peter W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal of Computing*, 26:1484–1509, 1997.
- [13] David Deutsch. Quantum theory, the Church-Turing principle and the universal quantum computer. *Proceedings of the Royal Society of London. Series A*, 400:97–111, 1985.

- [14] Ethan Bernstein and Umesh Vazirani. Quantum complexity theory. *SIAM Journal on Computing*, 26 (5):1411–1473, 1997.
- [15] David Deutsch and Richard Jozsa. Rapid solutions of problems by quantum computation. *Proceedings of the Royal Society of London. Series A*, 439:553–558, 1992.
- [16] David R. Simon. On the power of quantum computation. *SIAM Journal of Computing*, 26:1474–1483, 1997.
- [17] Gilles Brassard and Peter Hoyer. An exact quantum polynomial-time algorithm for Simon’s problem. In *Israel Symposium on Theory of Computing Systems*, pages 12–23, 1997.
- [18] Sean Hallgren. Polynomial-time quantum algorithms for Pell’s equation and the principal ideal problem. *Journal of the ACM*, 54 (1):1–19, 2007.
- [19] Seth Lloyd. Universal quantum simulators. *Science*, 273:1073–1078, 1996.
- [20] Leonard M. Adleman, Jonathan DeMarrais, and Ming-Deh A. Huang. Quantum computability. *SIAM Journal on Computing*, 26 (5):1524–1540, 1997.
- [21] Alexei Yu. Kitaev. Quantum computations: Algorithms and error correction. *Russian Mathematical Surveys*, 52 (6):1191–1249, 1997.
- [22] David Deutsch. Quantum computational networks. *Proceedings of the Royal Society of London. Series A*, 425:73–90, 1989.
- [23] Andrew Chi-Chih Yao. Quantum circuit complexity. In *Proceedings of the 34th Annual Symposium on Foundations of Computer Science*, pages 352–361. Institute of Electrical and Electronic Engineers Computer Society Press, 1993.
- [24] David P. DiVincenzo. Two-qubit gates are universal for quantum computation. *Physical Review A*, 51:1015–1022, 1995.
- [25] Seth Lloyd. Almost any quantum logic gate is universal. *Physical Review Letters*, 75:346–349, 1995.
- [26] Adriano Barenco. A universal two-bit gate for quantum computation. *Proceedings of the Royal Society of London. Series A*, 449:679–683, 1995.
- [27] David Deutsch, Adriano Barenco, and Artur Ekert. Universality in quantum computation. *Proceedings of the Royal Society of London. Series A*, 449:669–677, 1995.
- [28] Adriano Barenco, Charles H. Bennett, Richard Cleve, David P. DiVincenzo, Norman Margolus, Peter Shor, Tycho Sleator, John A. Smolin, and Harald Weinfurter. Elementary gates for quantum computation. *Physical Review A*, 52:3457–3467, 1995.
- [29] Robert Solovay. Unpublished manuscript. 1995.
- [30] Yaoyun Shi. Both Toffoli and Controlled-NOT need little help to do universal quantum computation. *Quantum Information and Computation*, 3 (1):84–92, 2003.

- [31] James E. Gentle. *Numerical Linear Algebra for Applications in Statistics*. Springer, 1998.
- [32] Juha J. Vartiainen, Mikko Miöttönen, and Martti M. Salomaa. Efficient decomposition of quantum gates. *Physical Review Letters*, 92:177902, 2004.
- [33] Guifre Vidal and Christopher M. Dawson. Universal quantum circuit for two-qubit transformations with three controlled-not gates. *Physical Review A*, 69:010301(R), 2004.
- [34] Peter W. Shor. Fault-tolerant quantum computation. In *IEEE Symposium on Foundations of Computer Science*, pages 56–65, 1996.
- [35] P. Oscar Boykin, Tal Mor, Matthew Pulver, Vwani Roychowdhury, and Farrokh Vatan. A new universal and fault-tolerant quantum basis. *Information Processing Letters*, 75:101–107, 2000.
- [36] Daniel Gottesman. The Heisenberg representation of quantum computers. Technical report, arXiv:quant-ph/9807006, 1998.
- [37] Emanuel Knill. Knill - Gottesman private communication. 1998.
- [38] Daniel Gottesman. *Stabilizer Codes and Quantum Error Correction*. PhD thesis, California Institute of Technology Pasadena, California, 1997.
- [39] David Deutsch, Artur Ekert, Richard Jozsa, Chiara Macchiavello, Sandu Popescu, and Anna Sanpera. Quantum privacy amplification and the security of quantum cryptography over noisy channels. *Physical Review Letters*, 77:2818–2821, 1996.
- [40] Charles H. Bennett, Gilles Brassard, Sandu Popescu, Benjamin Schumacher, John A. Smolin, and William K. Wootters. Purification of noisy entanglement and faithful teleportation via noisy channels. *Physical Review Letters*, 76:722–725, 1996.
- [41] Charles H. Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wootters. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Physical Review Letters*, 70:1895–1899, 1993.
- [42] Charles H. Bennett and Stephen J. Wiesner. Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states. *Physical Review Letters*, 69:2881–2884, 1992.
- [43] Daniel M. Greenberger, Michael A. Horne, and Anton Zeilinger. *Bell's Theorem, Quantum Theory and Conceptions of the Universe*. Springer, 1989.
- [44] Simon Anders and Hans J. Briegel. Fast simulation of stabilizer circuits using a graph-state representation. *Physical Review A*, 73:022334, 2006.
- [45] A. R. Calderbank and Peter W. Shor. Good quantum error-correcting codes exist. *Physical Review A*, 54:1098–1105, 1996.
- [46] Andrew Steane. Multiple-particle interference and quantum error correction. *Proceedings of the Royal Society of London. Series A*, 452:2551–2577, 1996.

- [47] David P. DiVincenzo and Peter W. Shor. Fault-tolerant error correction with efficient quantum codes. *Physical Review Letters*, 77:3260–3263, 1996.
- [48] Andrew M. Steane. Overhead and noise threshold of fault-tolerant quantum error correction. *Physical Review A*, 68:042322, 2003.
- [49] Emanuel Knill. Quantum computing with realistically noisy devices. *Nature*, 434:39–44, 2005.
- [50] Nicholas Pippenger, George D. Stamoulis, and John N. Tsitsiklis. On a lower bound for the redundancy of reliable networks with noisy gates. *Information Theory, IEEE Transactions on*, 37 (3):639–643, 1991.
- [51] Alexei Yu. Kitaev, Alexander H. Shen, and Mikhail N. Vyalyi. *Classical and Quantum Computation*. American Mathematical Society, Volume 47 of Graduate Studies in Mathematics, 2002.
- [52] Alexei Yu. Kitaev. Quantum error correction with imperfect gates. In *Proceeding of the Third International Conference on Quantum Communication and Measurement*, pages 181–188, 1997.
- [53] Aram W. Harrow, Benjamin Recht, and Isaac L. Chuang. Efficient discrete approximations of quantum gates. *Journal of Mathematical Physics*, 43:4445, 2002.
- [54] Alexei Yu. Kitaev. Quantum measurements and the abelian stabilizer problem. *Electronic Colloquium on Computational Complexity*, 3 (3):1–22, 1996.
- [55] Michelangelo Grigni, Leonard J. Schulman, Monica Vazirani, and Umesh Vazirani. Quantum mechanical algorithms for the nonabelian hidden subgroup problem. *Combinatorica*, 24 (1):137–154, 2004.
- [56] Dmitry Gavinsky. Quantum solution to the hidden subgroup problem for Poly-Near Hamiltonian groups. *Quantum Information and Computation*, 4:229–235, 2004.
- [57] Mark Ettinger, Peter Hoyer, and Emanuel Knill. Hidden subgroup states are almost orthogonal. Technical report, arXiv:quant-ph/9901034, 1999.
- [58] Mark Ettinger and Peter Hoyer. On quantum algorithms for noncommutative hidden subgroups. *Advances in Applied Mathematics*, 25 (3):239–251, 2000.
- [59] Cristopher Moore, Alexander Russell, and Leonard J. Schulman. The symmetric group defies strong Fourier sampling. In *Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science*, pages 479–490, 2005.
- [60] Sean Hallgren, Cristopher Moore, Martin Rötteler, Alexander Russell, and Pranab Sen. Limitations of quantum coset states for graph isomorphism. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, 2006.
- [61] Greg Kuperberg. A subexponential-time quantum algorithm for the dihedral hidden subgroup problem. *SIAM Journal on Computing*, 35 (1):170–188, 2005.

- [62] Dave Bacon, Andrew M. Childs, and Wim van Dam. From optimal measurement to efficient quantum algorithms for the hidden subgroup problem over semidirect product groups. In *Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science*, 2005.
- [63] Dave Bacon, Andrew M. Childs, and Wim van Dam. Optimal measurements for the dihedral hidden subgroup problem. *Chicago Journal of Theoretical Computer Science*, 2006 (2):1–25, 2006.
- [64] Dave Bacon. How a Clebsch-Gordan transform helps to solve the Heisenberg hidden subgroup problem. Technical report, arXiv:quant-ph/0501044, 2005.
- [65] Dave Bacon, Isaac L. Chuang, and Aram W. Harrow. Efficient quantum circuits for Schur and Clebsch-Gordan transforms. *Physical Review Letters*, 97:170502, 2006.
- [66] Michael Keyl and Reinhard F. Werner. Estimating the spectrum of a density operator. *Physical Review A*, 64:052311, 2001.
- [67] Stephen D. Bartlett, Terry Rudolph, and Robert W. Spekkens. Classical and quantum communication without a shared reference frame. *Physical Review Letters*, 91:027901, 2003.
- [68] Wim van Dam, Sean Hallgren, and Lawrence Ip. Quantum algorithms for some hidden shift problems. *SIAM Journal on Computing*, 36 (3):763–778, 2006.
- [69] Andrew M. Childs and Wim van Dam. Quantum algorithm for a generalized hidden shift problem. In *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 1225–1232, 2007.
- [70] Andrew M. Childs, Leonard J. Schulman, and Umesh V. Vazirani. Quantum algorithms for hidden nonlinear structures. In *The 48th Annual Symposium on Foundations of Computer Science*, 2007.
- [71] Niels Abel and Paolo Ruffini. The Abel-Ruffini theorem. http://en.wikipedia.org/wiki/Abel-Ruffini_theorem.
- [72] Gene H. Golub and Charles F. van Loan. *Matrix Computations*. The John Hopkins University Press, 1989.
- [73] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. Quantum metrology. *Physical Review Letters*, 96:010401, 2006.
- [74] Wim van Dam, G. Mauro D’Ariano, Artur Ekert, Chiara Macchiavello, and Michele Mosca. Optimal quantum circuits for general phase estimation. *Physical Review Letters*, 98:090501, 2007.
- [75] Richard Cleve, Artur Ekert, Chiara Macchiavello, and Michele Mosca. Quantum algorithms revisited. *Proceedings of the Royal Society of London, Series A*, 454:339–354, 1998.
- [76] Adriano Barenco and Artur Ekert. Approximate quantum Fourier transform and decoherence. *Physical Review A*, 54:139–146, 1996.

- [77] Donny Cheung. Improved bounds for the approximate QFT. Technical report, arXiv:quant-ph/0403071, 2004.
- [78] Lee Spector. *Automatic Quantum Computer Programming: A Genetic Programming Approach*. Kluwer Academic Publishers, 2004.
- [79] Paul Massey, John A. Clark, and Susan Stepney. Human-competitive evolution of quantum computing artefacts by genetic programming. *Evolutionary Computation*, 14(1):21–40, 2005.
- [80] Richard P. Feynman. Simulating physics with computer. *International Journal of Theoretical Physics*, 21:467–488, 1982.
- [81] Daniel S. Abrams and Seth Lloyd. Quantum algorithm providing exponential speed increase for finding eigenvalues and eigenvectors. *Physical Review Letters*, 83:5162–5165, 1999.
- [82] Julia Kempe, Alexei Yu. Kitaev, and Oded Regev. The complexity of the local hamiltonian problem. *SIAM Journal on Computing*, 35(5):1070–1097, 2006.
- [83] Dorit Aharonov, Daniel Gottesman, and Julia Kempe. The power of quantum systems on the line. Technical report, arXiv:0705.4077, 2007.
- [84] Dorit Aharonov and Amnon Ta-Shma. Adiabatic quantum state generation and statistical zero knowledge. In *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, pages 20–29, 2003.
- [85] Alán Aspuru-Guzik, Anthony D. Dutoi, Peter J. Love, and Martin Head-Gordon. Simulated quantum computation of molecular energies. *Science*, 309:1704–1707, 2005.
- [86] Robert B. Griffiths and Chi-Sheng Niu. Semiclassical fourier transform for quantum computation. *Physical Review Letters*, 76:3228–3231, 1996.
- [87] Christof Zalka. Fast versions of Shor’s quantum factoring algorithm. Technical report, arXiv:quant-ph/9806084, 1998.
- [88] Michele Mosca and Artur Ekert. The hidden subgroup problem and eigenvalue estimation on a quantum computer. *Lecture Notes in Computer Science*, 1509:174–188, 1999.
- [89] Andrew M. Childs, John Preskill, and Joseph Renes. Quantum information and precision measurement. *Journal of Modern Optics*, 47:155–176, 2000.
- [90] Emanuel Knill, Gerardo Ortiz, and Rolando D. Somma. Optimal quantum measurements of expectation values of observables. *Physical Review A*, 75:012328, 2007.
- [91] Stephen D. Bartlett, Terry Rudolph, and Robert W. Spekkens. Reference frames, superselection rules, and quantum information. *Reviews of Modern Physics*, 79:555–609, 2007.
- [92] David L. Moehring, Martin J. Madsen, Boris B. Blinov, and Cris Monroe. Experimental Bell inequality violation with an atom and a photon. *Physical Review Letters*, 93:090410, 2004.

- [93] Matthias Steffen, M. Ansmann, R. McDermott, N. Katz, Radoslaw C. Bialczak, Erik Lucero, Matthew Neeley, E. M. Weig, A. N. Cleland, and John M. Martinis. State tomography of capacitively shunted phase qubits with high fidelity. *Physical Review Letters*, 97:050502, 2006.
- [94] Göran Johansson. Private communication, 2007.
- [95] Göran Wendin and Vitaly S. Shumeiko. *Handbook of Theoretical and Computational Nanotechnology*, volume 3. ASP, Los Angeles, 2006.
- [96] Gerald Gilbert and Michael Hamrick. Practical quantum cryptography: A comprehensive analysis (part one). Technical Report MTR00W0000052, The MITRE Corporation, 2000.
- [97] Terry Rudolph and Lov Grover. Quantum communication complexity of establishing a shared reference frame. *Physical Review Letters*, 91:217905, 2003.
- [98] Mark de Burgh and Stephen D. Bartlett. Quantum methods for clock synchronization: Beating the standard quantum limit without entanglement. *Physical Review A*, 72:042301, 2005.
- [99] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. Quantum-enhanced positioning and clock synchronization. *Nature*, 412:417–419, 2001.
- [100] Isaac L. Chuang. Quantum algorithm for distributed clock synchronization. *Physical Review Letters*, 85:2006–2009, 2000.
- [101] Richard Jozsa, Daniel S. Abrams, Jonathan P. Dowling, and Colin P. Williams. Quantum clock synchronization based on shared prior entanglement. *Physical Review Letters*, 85:2010–2013, 2000.
- [102] Jae-Seung Lee, Jaehyun Kim, Yongwook Cheong, and Soonchil Lee. Implementation of phase estimation and quantum counting algorithms on an NMR quantum-information processor. *Physical Review A*, 66:042316, 2002.
- [103] Jingfu Zhang, Gui Lu Long, Zhiwei Deng, Wenzhang Liu, and Zhiheng Lu. Nuclear magnetic resonance implementation of a quantum clock synchronization algorithm. *Physical Review A*, 70:062322, 2004.
- [104] Charles H. Bennett and Gilles Brassard. Quantum cryptography: Public key distribution and coin tossing. In *Proceedings of IEEE International Conference on Computer, Systems, and Signal Processing*, pages 175–179, 1984.
- [105] Charles H. Bennett. Quantum cryptography using any two nonorthogonal states. *Physical Review Letters*, 68:3121–3124, 1992.
- [106] Artur K. Ekert. Quantum cryptography based on Bell’s theorem. *Physical Review Letters*, 67:661–663, 1991.
- [107] id Quantique. Cerberis network security system. White paper, 2007. <http://idquantique.com>.

- [108] MagiQ Technologies. MagiQ's QPN. White paper, 2007. <http://magiqtech.com>.
- [109] Gilles Brassard, Norbert Lütkenhaus, Tal Mor, and Barry C. Sanders. Limitations on practical quantum cryptography. *Physical Review Letters*, 85:1330–1333, 2000.
- [110] Jan-Ake Larsson. A practical trojan horse for Bell-inequality-based quantum cryptography. *Quantum Information and Computation*, 2 (6):434–442, 2002.
- [111] Julio Gea-Banacloche. Hiding messages in quantum data. *Journal of Mathematical Physics*, 43 (9):4531–4536, 2002.
- [112] Emanuel Knill and Raymond Laflamme. Concatenated quantum codes. Technical report, arXiv:quant-ph/9608012, 1996.
- [113] Peter W. Shor. Scheme for reducing decoherence in quantum computer memory. *Physical Review A*, 52:R2493–R2496, 1995.
- [114] Howard Barnum, Claude Crépeau, Daniel Gottesman, Adam Smith, and Alain Tapp. Authentication of quantum messages. In *Proceedings of the 43rd Symposium on Foundations of Computer Science*, pages 449–458, 2002.
- [115] Marcos Curty, David J. Santos, and Esther Pérez. Qubit authentication. *Physical Review A*, 66:022301, 2002.
- [116] Hideki Noda, Michiharu Niimi, and Eiji Kawaguchi. High-performance JPEG steganography using quantization index modulation in DCT domain. *Pattern Recognition Letters*, 27 (5):455–461, 2006.
- [117] Fabien A. P. Petitcolas. Mp3stego. Open source code, <http://www.petitcolas.net/fabien/steganography/mp3stego/>, 2006.
- [118] Rakan El-Khalil and Angelos D. Keromytis. Hydan: Hiding information in program binaries. In *Proceedings of the sixth International Conference on Information and Communications Security*, pages 187–199, 2004.
- [119] Marcos Curty and David J. Santos. Protocols for quantum steganography. In *2nd Bielefeld Workshop on Quantum Information and Complexity*, pages 12–14, 2000.
- [120] Keye Martin. Steganographic communication with quantum information. In *9th Information Hiding*, 2007.
- [121] Marcos Curty and David J. Santos. Quantum authentication of classical messages. *Physical Review A*, 64:062309, 2001.
- [122] Esther Pérez, Marcos Curty, David J. Santos, and Priscila Garcia-Fernández. Quantum authentication with unitary coding sets. *Journal of Modern Optics*, 50 (6):1035–1047, 2003.