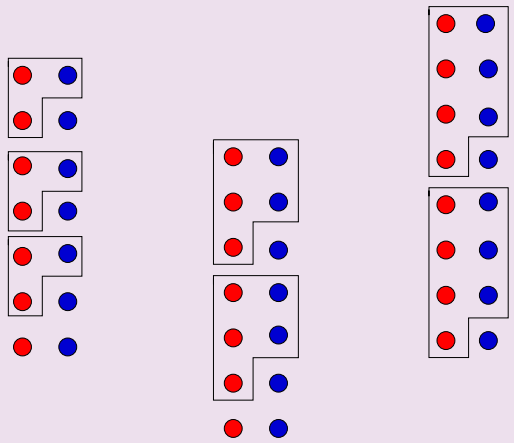
DOMINIK JANZING

Computer Science Approach to Quantum Control



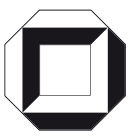
universitätsverlag karlsruhe

Dominik Janzing

Computer Science Approach to Quantum Control

Computer Science Approach to Quantum Control

von
Dominik Janzing



universitätsverlag karlsruhe

Habilitation, Universität Karlsruhe (TH)
Fakultät für Informatik, 2006

Impressum

Universitätsverlag Karlsruhe
c/o Universitätsbibliothek
Straße am Forum 2
D-76131 Karlsruhe
www.uvka.de

Universitätsverlag Karlsruhe 2006
Print on Demand

ISBN-13: 978-3-86644-083-8
ISBN-10: 3-86644-083-9

Contents

Preface: Algorithms - not only for Computational Problems

The term "algorithm" lies at the heart of computer science. Following a simple definition in the web [1], an algorithm is "a finite set of well-defined instructions for accomplishing some task which, given an initial state, will result in a corresponding recognizable end-state". This definition leaves the task itself unspecified and also means, as an example, that a set of instructions to prepare tomato salad is an algorithm. For obvious reasons, it is common in computer science to consider an algorithm as a set of instructions solving a *computational problem*. Following Horowitz and Sahni [2], one has further the following criteria:

1. *Input:* there are zero or more quantities which are externally supplied;
2. *Output:* at least one quantity is produced;
3. *Definiteness:* if we trace out the instructions of an algorithm, then for all cases the algorithm will terminate after a finite number of steps.
4. *Effectiveness:* every instruction must be sufficiently basic that it can in principle be carried out by a person using only pencil and paper. It is not enough that each operation be definite as in (3) but it must also be feasible.

During the 90s the idea of building a quantum computer attracted broad interdisciplinary interest in computer science, physics, and mathematics. Such a machine works with quantum superpositions of logically different states, and a completely new type of algorithms, so-called quantum algorithms, were invented. The elementary operations in quantum algorithms cannot be described as classical logical computation steps. These algorithms contain instructions which are not "sufficiently basic that they can in principle be carried out by a person using only pencil and paper" (see item 4.). Instead, they would require special 'quantum paper' on which one could write superpositions of different numbers. A quantum computation can only be performed with hardware which enables such superpositions between logically different states. This works only in physical systems which are sufficiently isolated from their environment to show typical quantum phenomena, and it is more likely feasible for small systems than for macroscopic systems. Therefore an important part of quantum computing research is the development of methods for controlling physical systems on the nanoscopic, microscopic, or mesoscopic scale. These systems can, for instance, be ions, atoms, molecules, or electromagnetic fields. Typical tools to manipulate their physical states are laser beams or other kind of electromagnetic waves like high-frequency radiation.

Apart from the goal of developing hardware for the quantum computer, it is clear that controlling tiny objects is a challenging task in its own right, having many potential applications. Focussing electron or light beams in a microscope or building smaller transistors which are triggered by single electrons for future electronic chips are only two examples of applications where states of quantum systems have to be controlled and measured. The idea of quantum computing encouraged theoreticians and experimentalists to further develop existing control technologies. On the other hand, it seems as if theoretical and experimental tools created for quantum computing could be also useful for non-computational problems. Among others, the following non-computational applications of quantum control techniques are closely related to quantum computing:

1. **NMR Spectroscopy:** Nuclear magnetic resonance (NMR) experiments are used in medicine, biology, chemistry, and physics to analyze organic and an-organic matter. Roughly speaking, the spins of the atoms in the analyzed material can be considered as little magnets. These 'magnets' are rotated by an electromagnetic pulse. While they turn back to their original position, they emit electro-magnetic pulses which give insight in the chemical structure. In many experiments the matter is subjected to rather complex sequences of electromagnetic pulses since the emitted radiation after this treatment is characteristic for the specific material and the interactions among its nuclear spins. The design of these pulse sequences is analogous to algorithmic problems in quantum computing (QC) even though these methods existed in NMR long before they were rediscovered for QC.

2. **Quantum Lithography and Microscopy:** Optical lithography is a primary tool in the microchip industry for transferring circuit images onto substrates. Similiarly as in microscopy, the resolution of the image is limited by the wavelength of the light beam, the so-called *diffraction limit*. Recent research suggests that the diffraction limit can in principle be beaten using non-classical light with quantum correlated photons. The proposals to prepare this type of light use a sequence of optical devices which is designed in strong analogy to the design of quantum circuits from elementary gates.

3. **Cooling Molecular Systems:** Modern cooling techniques which lower the temperature of atomic or molecular systems (in order to prepare it for further experiments) use rather sophisticated sequences of control operations. In the context of NMR experiments, there exists even a proposal for a *cooling algorithm* which can at the same time be considered as a data compression algorithm if the atomic states are interpreted as logical values of bits. Even though the original intention of this proposal was to initialize the system for quantum computing, it shows that the non-computing task of cooling a system can be treated algorithmically.

This work should elucidate what we can learn from quantum computing and communication research for non-computational problems:

- **Quantum Algorithmic Thinking and Complexity Theory:** The principle of concatenating elementary transformations that generate complex physical processes is not restricted to computational processes. Present research results indicate that

it makes sense to think about which physical processes require complex sequences of elementary operations and to develop a complexity theory for non-computational processes. Our own research has shown that one can even connect such a complexity theory with the conventional complexity theory of computer science.

- **Simplicity of Models:** The central concept of quantum information theory is a quantum-bit ('qubit'). It is, physically speaking, a two-level system which is certainly the simplest thinkable quantum system. In the 90s it was realized that there were still interesting questions unsolved dealing with just two qubits. After *quantum theory* existed almost for a century, *quantum information* research came up with models and unsolved questions which are surprisingly simple compared to the examples and problems of quantum systems in textbook quantum mechanics. This has already deepened the understanding of quantum laws.

- **Physical Relevance of Information:** The crucial role of the quantity *entropy* has already been generally accepted in thermodynamics, showing that, in principle, information is also an important concept in physics. However, in studying conventional thermodynamic machines like steam engines it seems rather academic to interpret heat transfer as *information* transfer. Likewise, it seems academic to interpret information flow among computing devices as heat transfer. In contrast, the information in a quantum computer is stored in simple physical degrees of freedom and the relation between heat and information becomes much more obvious on this elementary level. This allows thermodynamic insights to be obtained which have strong analogies to the mathematical issues of information theory like coding problems. We will show how these insights can be used to derive new bounds on the physical resource requirements of computation.

For the above reasons, the intention of this thesis is to advocate a quantum computer science approach to a control theory for nanoscopic objects. It is organized as follows.

In Chapter 1 we sketch the central ideas of quantum computation and information as a basis for this thesis.

Chapter 2 shows that these ideas help to rethink fundamental problems of different scientific disciplines: First, quantum computing research gave a more explicit operational meaning to the ingrediences of conventional quantum mechanics: quantum state preparation, dynamics and measurements can be seen as *algorithms* based on some elementary operations. Second, QC provides clear and simple toy models to understand thermodynamics which allows to derive thermodynamic constraints for molecular processes which do not exist in conventional thermodynamics. These constraints may be relevant to understand potential thermodynamic limits of low power computation. Third, quantum computing has made more obvious that the laws of physics determine which computation problems can be solved efficiently. In the same way they determine which non-computational tasks can be done efficiently - there appears to be no substantial difference. Fourth, we consider a quantum computing toy model to demonstrate that causal reasoning in every-day life has to take into account quantum effects. Even though this problem is not strongly related to the rest of this work it shows that simple models taken from quantum information research may lead to insights for problems which seem completely disconneted to usual computation problems.

In Chapter 3 we explain why an appropriately general definition of algorithms and their complexity is required to treat non-computational problems in an algorithmic way.

In Chapter 4 we present algorithms for non-computational problems. In contrast to Chapter 2 where problems of state preparation and measurements are considered for reasons of pure research, we consider here problems which stem from potential applications. We relate the complexity of some of those natural control problems to the complexity theory of computational problems. This can be seen as a first step for developing complexity classes for non-computational problems.

The intention of this thesis is to show connections between rather different issues in the field. For the details we refer to the literature.

Chapter 1

Quantum Information and Computation

The main subject of this thesis is not quantum computing itself; rather it should show that a broad variety of applications in future nano technology could use complex control algorithms which are in strong analogy to algorithms in quantum computation. In order to demonstrate this, we first present some of the basic concepts of quantum computing (for details we refer to [3]). They will provide a useful basis to discuss non-computational control problems. In particular, the simplicity of the standard model quantum computer is an attractive venue to discuss many of these issues.

1.1 Standard Model of a Quantum Computer

The power of quantum computing lies in the fact that the quantum register allows quantum superpositions between different inputs. This makes it, roughly speaking, to some extent possible to process many different inputs at once. The most important element of quantum computing and quantum information theory is the quantum bit, called "qubit".

In contrast to a classical bit, it can have not only the values 0 and 1, but also linear superpositions between them. According to quantum theory [4], the mathematics of superpositions can be described by Hilbert spaces. We will use *bra* and *ket* notation, i.e., vectors in a space $\mathcal{H}$ are denoted by symbols $|\psi\rangle$ and vectors in its dual by $\langle\phi|$. The inner product is written as $\langle\phi|\psi\rangle$, bra and ket in the reversed order denote rank-one operators: $|\psi\rangle\langle\phi|$ is defined by the intuitive equation

$$(|\psi\rangle\langle\phi|)|\alpha\rangle := |\psi\rangle\ \langle\phi|\alpha\rangle\,.$$

Now we define the unit of quantum information, i.e., a qubit:

Definition 1 (Quantum-Bit = Qubit)
A qubit is the quantum generalization of a bit. The states of a qubit are the one-dimensional subspaces of $\mathbb{C}^2$ where the span of the two canonical basis vectors, denoted by $|0\rangle$ and $|1\rangle$, correspond to the logical states 0 and 1, respectively. If a state is the one-dimensional space $\mathbb{C}|\psi\rangle$ where $|\psi\rangle$ is the unit vector $c_0|0\rangle + c_1|1\rangle$, the values $|c_0|^2$ and $|c_1|^2$ are the probabilities to find a qubit in the state $|0\rangle$ or $|1\rangle$, respectively, if its logical state is read out.

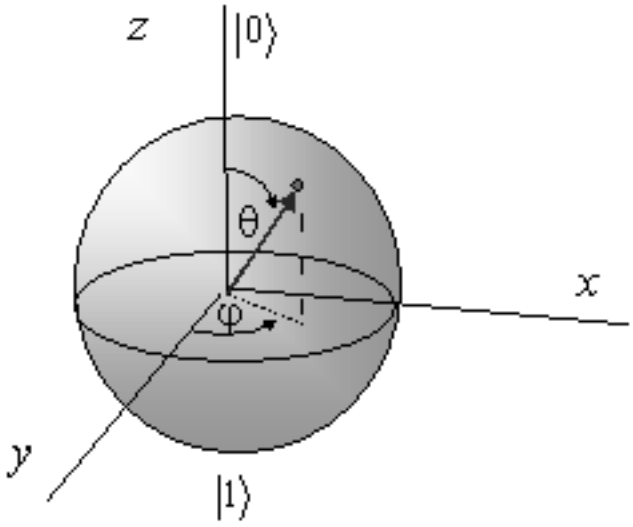


Figure 1.1: The state of a qubit can be represented by a vector in the Bloch sphere.

It is often useful to represent the one-dimensional subspaces of $\mathbb{C}^2$ by unit vectors in the Bloch sphere, a ball in $\mathbb{R}^3$. The correspondence is given as follows. The span of

$$\cos\frac{\theta}{2}|0\rangle + e^{i\varphi}\sin\frac{\theta}{2}|1\rangle$$

corresponds to the unit vector characterized by angles θ and φ as seen in Fig. 1.1. This is possibly the most natural visualization of the state of a qubit since the direction in 3-dimensional space given by the *Bloch vector* is for every spin-1/2 particle directly interpreted as the direction of its magnetic moment. Spin-1/2 particles with their two basis states 'spin up' and 'spin down' are a good example for physical systems representing qubits.

However, such an intuitive representation exists only for one qubit. An n-qubit register is a system consisting of n qubits. Note that this statement is not as 'harmless' as it seems since the state of an n-qubit register is not described by the states of its n components, i.e., its state cannot be represented by n Bloch vectors:

Definition 2 (State of a Quantum Register)
The states of an n-qubit register are given by the one-dimensional subspaces of $\mathcal{H} := (\mathbb{C}^2)^{\otimes n}$. The canonical basis states of $\mathcal{H}$ are in one-to one correspondence to binary words of length n.

One may interpret this fact by stating that quantum theory allows superpositions of binary words b since every unit vector

$$|\psi\rangle := \sum_b c_b|b\rangle \tag{1.1}$$

defines a state. This is the decisive difference to any classical computational devices which work with some kind of 'fuzzy' logical values between 0 and 1: The 'fuzzy-value' of the whole register would always be given by describing the fuzzy values of all its

components. The superposition principle is a general feature of quantum systems: If $|\psi_1\rangle$ and $|\psi_2\rangle$ are possible state vectors then $c_1|\psi_1\rangle + c_2|\psi_2\rangle$ is for every two complex numbers c_1, c_2 with $|c_1|^2 + |c_2|^2 = 1$ also a possible state vector.[1] A decisive phenomenon of the quantum world, namely *entanglement*, then follows: A generic superposition of binary words cannot be written as a product

$$|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle \otimes \cdots \otimes |\psi_n\rangle$$

where each $|\psi_j\rangle$ is a state vector in $\mathbb{C}^2$ (see [3] and references therein for details). In other words, the state of a *single* qubit is undefined if no such factorization exists.

It is interesting to note that the models of quantum computing implicitly contain an important statement on the connection between information and physics which holds for classical computation, too: the two states of a classical bit must be represented by physical states which are mutually orthogonal. This means that whenever the logical value of a bit is changed, this is based on a quantum dynamical evolution of a state moving between two orthogonal Hilbert space vectors. Even though this insight was probably not a new idea stemming from QC research, the language used in QC made it more obvious and the abstract concept of representing information by qubits can also be useful to explore device independent limits of classical computation. The essential advantage of a qubit compared to a bit for describing *classical* computation is that the continuous physical process changing the value of the bit cannot be described on bits, but it can be described on qubits [5]. Note that this difference can even be the basis for an axiomatic derivation of the laws of quantum mechanics [6, 7]. In other words, the concept of *classical information* is indeed consistent with quantum mechanics as long as one restricts the attention to a set of mutually orthogonal states. We emphasize that this consistency does not require any multi-particle or macroscopic limits. But we claim that the description of *dynamics of classical information* requires either the concept of quantum information or some macroscopic limits. This statement is the leading motivation in Subsection 2.2.7 where a quantum bound for copying timing information is derived and in Subsection 2.2.10 where we argue that quantum information transfer is useful for clock synchronization protocols.

1.1.1 Quantum Gates and Quantum Circuits

In classical computation it is well-known that every boolean circuit can be generated from circuits with two input bits and one output bit. One can even implement every boolean circuit using only NAND-gates or only NOR gates. To some extent as an analogy to this result, it is well-known in quantum computing that every unitary transformation on $\mathcal{H} = (\mathbb{C}^2)^{\otimes n}$ can be generated by so-called two-qubit gates, i.e., unitary transformations which act non-trivially on two components of the n-fold tensor product [8]. Note that in this framework a gate is a *physical process* in contrast to classical circuits where one thinks of AND, NAND, OR, NOR, and NOT gates usually as hardware devices in which input and output can exist *simultaneously*. In quantum computing, there is a fundamental law which forbids having input and output simultaneously: the no-cloning theorem [9]

[1] Note that the superposition refers to *state vectors* even though the latter is only defined up to a phase factor for a given state, therefore one should rather speak about superposition of *state vectors* than of *states*.

states that the content of a quantum register, the *quantum information* cannot be copied without disturbing it. We define:

Definition 3 (2-Qubit Quantum Gate)
A gate acting on the qubit pair $(i, i+1)$ *is a process which changes the state* $|\psi\rangle$ *of the quantum register into*

$$\mathbf{1}^{\otimes i-1} \otimes U \otimes \mathbf{1}^{\otimes n-i-1} |\psi\rangle \,,$$

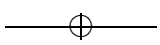
where U *is a unitary operator on* $\mathbb{C}^2 \otimes \mathbb{C}^2$*. In general,* U *may act on non-adjacent pairs* (i, j)*, but this is difficult to write symbolically.*

The following gates are widely used as elementary transformations.

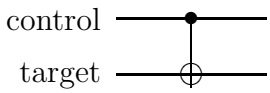
1. **NOT:** This gate is given by the linear extension of the permutation $|0\rangle \leftrightarrow |1\rangle$ and is described by the Pauli matrix

$$\sigma_x := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} .$$

   It is drawn as

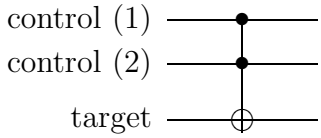


2. **C-NOT:** The controlled-not gate inverts the target qubit if the control qubit is in the $|1\rangle$ state. If the control qubit is in the $|0\rangle$ state the target is unchanged. Superpositions in the control qubit lead in general to entanglement between control and target qubit.



   The C-NOT gate is a good model for copy-operations in general. Since it copies the logical state 0 or 1 to the target, it destroys superpositions between $|0\rangle$ and $|1\rangle$.

3. **TOFFOLI:** a double-controlled not. It inverts the state of the target iff both control qubits are in the state $|1\rangle$.



Even though the standard model quantum computer is often formulated with one and two qubit gates only, we mention the Toffoli gate since it plays a central role in the theory of classical reversible computation. Compositions of Toffoli gates can compute arbitrary boolean functions. In contrast to AND, NAND, OR, NOR gates, the output of such a Toffoli gate allows the reconstruction of the input completely. This fact can be used [10] to prove that computation does not necessarily involve logically irreversible operations, which is an important feature from the thermodynamical point of view [11].

All these gates are classical in the sense that they do not create superpositions between binary words. Certainly quantum computing requires also gates which do not only permute basis states.

1. **HADAMARD:** Its matrix is

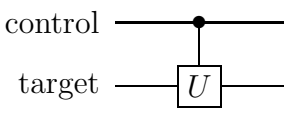
$$H := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} .$$

It is drawn as

2. **C-U:** The controlled-U-gate is for any unitary U on 1 or more qubits defined by

$$C - U := |0\rangle\langle 0| \otimes \mathbf{1} + |1\rangle\langle 1| \otimes U$$

where $|0\rangle\langle 0|$ and $|1\rangle\langle 1|$ denote orthogonal projections onto the spaces spanned by $|0\rangle$ and $|1\rangle$, respectively. The C-U gate implements U on the target if the control qubit is in the 1 state. It is drawn by



If U is diagonal in the computational basis one calls C-U a *controlled phase shift*.

The question of which set $\mathcal{U}$ of gates is universal in the sense that concatenation of elements in $\mathcal{U}$ can approximate every n-qubit unitary transformation has attracted broad

interest. A simple example of a universal set is a C-NOT together with arbitrary one-qubit rotations [12]. Another example [13] is the Hadamard gate together with the controlled phase shift $C - \sigma_z^{1/2}$, which implements the phase shift

$$\sigma_z^{1/2} := \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$$

whenever the control qubit is 1.

1.1.2 Readout of Quantum Registers

In the standard model it is usually assumed that the possible measurements are given by the readout of single qubits. If all qubits are measured, the values $|c_b|^2$ in eq. (1.1) are the probabilities for obtaining the binary word b as result. If one measures only the logical states of all qubits in the subset $j := (j_1, j_2, \ldots, j_k)$ the probability for the outcome $x_1, \ldots x_k$ is obtained by a sum over all $|c_b|^2$ in the cylinder set given by all binary words $b \in \{0,1\}^n$ with

$$b_{j_l} = x_l \text{ with } l = 1, \ldots, k\,.$$

If $P_{x,j}$ is the projector onto the space spanned by all $|b\rangle$ with b in this cylinder set the probability for the result x in the state $|\psi\rangle$ is given by

$$\langle\psi|P_{x,j}|\psi\rangle\,.$$

This suggests the more general quantum-mechanical principle of calculating probabilities for measurement outcomes from projections (see Subsection 1.2.1). It will be further generalized later, leading to an abundance of possible measurements even for a single qubit.

1.1.3 Quantum Circuits and Algorithms

A sequence of quantum gates is usually called a quantum circuit. A quantum algorithm is a rule for generating quantum circuits in such a way that the circuits solve a given problem [14]. This rule for creating the circuit is actually a classical algorithm. It may be surprising that such a classical "meta-algorithm" is needed; the problem is that a completely quantum Turing machine would have superpositions between inputs with different running time of the algorithm. This leads to paradoxical superpositions of terminal states and states which tells the machine to continue [15, 16]. We briefly mention that a similar problem will appear in Subsection 2.2.9 when we describe a computation process which is performed by a closed physical system. Such an autonomous dynamics can never stop if the free energy of the system is finite. As a solution to this problem we present a model which avoids this problem by encoding the computational result in the *time average* of the infinitely repeated dynamics.

In order to describe how to use quantum gates for computation we have to clarify what it means for a quantum circuit to *compute* a function f. Let $f : \{0,1\}^n \to \{0,1\}^k$ be arbitrary. In order to realize this as a unitary mapping we could choose U such that

$$U(|a\rangle \otimes |y\rangle) = |1\rangle \otimes |f(a) \oplus y\rangle$$

for any input $a \in \{0,1\}^n$ and $y \in \{0,1\}^k$. If $n = k$ and f is a bijective function, we can also have

$$U|a\rangle = |f(a)\rangle\,.$$

Then, by linearity, a superposition of different inputs a leads to a superposition of different outputs $f(a)$. However, since the state vector cannot be read out it is not clear why this kind of 'parallelism' should help. This is the same situation as in classical probability theory wherein the probability distribution of randomly chosen inputs is reflected in the distribution of the outputs, but it is impossible to determine the probability distribution from a single experiment. In quantum information, one is tempted[2] to believe in the possibility of such a readout if one considers the 'quantum state' as a physical object instead of a statistical description about its preparation (for the problems of interpretations of quantum mechanics we refer to [17]). Hence the essential challenge in finding quantum algorithms is to use the parallelism in a more sophisticated way. We briefly mention the two most famous such algorithms: Grover's search and Shor's factoring algorithm.

Grover's algorithm searches binary words which satisfy a given condition. Given a function

$$f : \{0,1\}^n \to \{0,1\}\,,$$

the output is a binary string b of length n such that $f(b) = 1$ whenever it exists. In contrast to the best known classical algorithm (with running time $O(2^n)$) it has only running time $O(2^{n/2})$. The essential ideas are: (1) Given a superposition of all binary words

$$\sum_{0\le b<N} c_b|b\rangle \quad N = 2^n\,,$$

quantum mechanics offers the possibility of inverting the coefficients of all binary words satisfying the condition $f(b) = 1$ and thus, to obtain the state

$$\sum_b c_b(-1)^{f(b)}|b\rangle\,.$$

(2) There is a transformation which 'reflects' all coefficients about the average

$$\sum_b c_b|b\rangle \mapsto \sum_b (2c_b - c)|b\rangle\,,$$

with $c := \sum_b c_b$. The concatenation of (1) and (2) amplifies the modulus of all coefficients c_b which satisfy $f(b) = 1$. Repeating the amplification procedure, one generates superpositions which consist primarily of binary words b with $f(b) = 1$.

Shor's algorithm efficiently factors composite numbers into primes. Prior to its invention it was already known [18] that this problem can be reduced to the (still difficult) question of order-finding: Given natural numbers x and n, determine the smallest $r \in \mathbb{N}$ such that $x^r = 1(mod\ n)$. The essential idea is here, that it is possible to efficiently generate the superposition state

$$\frac{1}{\sqrt{N}} \sum_b |b\rangle \otimes |x^b(mod\ n)\rangle\,.$$

[2]this is the error in some beginners-proposals for quantum algorithms.

Note that r is now the period of the map $b \mapsto x^b$. It can be found by the quantum discrete Fourier transform which is the linear extension of

$$|a\rangle \mapsto \frac{1}{\sqrt{N}} \sum_b e^{iab} |b\rangle \, .$$

In the context of this thesis we mention that the order-finding problem can be thought of as determining the period of a classical dynamical system, namely the dynamics on the set $\{0, \ldots, n-1\}$ given by $a \mapsto ax \ (mod\ n)$. It would be interesting to know whether this could be used for investigations of dynamical systems, i.e., to acquire information about recurrence time and orbit lengths of given initial points.

The statement that Grover's and Shor's algorithms outperform the best known classical algorithms refers to the standard model of quantum computing. Therefore it is based on the hope that one will succeed in controlling quantum systems in such a way that the basic operations sketched in Subsection 1.1.1 can be implemented. To measure the complexity of a quantum circuit one can either count the number of gates or the number of time steps and assume that gates acting on disjoint qubit pairs can be implemented simultaneously. For classical circuits an important complexity measure is their depth, i.e., the length of the number of gates that are passed from the input to the output when the longest path is considered [19]. Since the spatial propagation of information from the input to the output of one gate translates to one time step we define the depth of a quantum circuit as follows:

Definition 4 (Depth of a Quantum Circuit)
The number of time steps of a quantum circuit is the depth. Each time step may contain several gates acting on disjoint qubit pairs or qubits.

1.1.4 Realization

Here we give a (rather incomplete) overview of some important proposals for the realization of the standard model quantum computer. We select two of the famous proposals.

- **Ion Trap QC:** (Cirac-Zoller proposal [20]) A string of several ions is trapped in a dynamic magnetic field. The qubit basis states are represented by two different internal states of the ions. The ions can be addressed separately in order to drive transitions between their internal states. This allows the implementation of arbitrary single qubit gates. Due to the repulsive electrostatic forces between the ions they can perform collective oscillations. The different quantum states of this oscillator is used as 'data bus': By applying appropriate laser pulses to the ions one can cause an information transfer between them which is mediated by the oscillation mode.

- **NMR QC:** The nuclear spins of different atoms are used as qubits. The states $|0\rangle$ and $|1\rangle$ are given by the 'spin up' and 'spin down' states of the spin-1/2 nuclei. The reference frame is given by a static external magnetic field which causes an energy gap between $|0\rangle$ and $|1\rangle$. Single qubit gates can be performed by electro-magnetic radiation of appropriate frequency. The natural pair-interaction between them is used to implement two-qubit gates. Using so-called selective decoupling techniques

(where appropriate one-qubit rotations can cancel unwanted terms, see Subsection 3.1.3) it is possible to switch off all but one of the pair interactions. This can be used to implement a controlled-phase gate [21].

Meanwhile, there are many more proposals (see e.g. [3] as well as many other examples on the quant-ph archive). Our choice above contains no judgment on the proposals. We mentioned (1) because it is probably one of the most direct implementations of the *standard model* quantum computer: Its basic operations are really initialization of basis states, implementation of one and two qubit gates, and readout of single qubits. We mentioned (2) since it is an important example of a new understanding of what constitutes a quantum algorithm. Present day NMR is a nice example for a non-computing quantum algorithm.

The discussion whether quantum computing may be realizable in the future must focus on two questions:

1. Will it be possible to isolate larger quantum systems sufficiently to keep quantum information alive over a sufficiently long time period?

2. Will it be possible to implement the required control operations?

It seems as if both questions would be issues of experimental physics only. However, one should not neglect the mathematical and computer science aspect of both questions:

First it is not yet clear how noisy the system is allowed to be in order to still allow quantum computation. Whereas it is often argued that quantum computing on large registers would require large-scale entanglement which is extremely fragile, error correcting codes generate superpositions which are less fragile [22]. Second it is not clear how much the set of operations can be reduced without dropping the power of universal quantum computing. The 'one-way' quantum computer [23] shows, for instance, that the preparation of appropriate initial states together with single qubit readout is already sufficient for quantum computing.

1.2 Arbitrary Quantum Systems as Quantum Registers

Clearly quantum registers need not necessarily consist of qubits. One can also define *qudits*, i.e., systems having d basis states instead of 2. A qudit is already the most general abstract description of an arbitrary finite-dimensional quantum system. Whereas a physical example for a qubit is a spin-1/2 system, qudits are for instance given by spin-$(d-1)/2$ systems. Quantum computing on infinite dimensional systems has also been proposed [24, 25] and algorithms for transferring quantum information between continuous and discrete "registers" have been considered by [26]. Therefore one could interpret every quantum system as quantum register or a part of it and every physical process as a generalized logical operation. This point of view makes it more obvious that the questions what can be computed efficiently and which non-computational processes could be implemented efficiently are quite close together.

1.2.1 States, Dynamics and Measurements in Textbook Quantum Mechanics

To show in which way quantum computing research brought another point of view into quantum theory, we sketch how the latter is introduced in most textbooks. A crucial role in the formal description of a quantum system is the Hilbert space $\mathcal{H}$ which may be finite or infinite. An example for the former is $\mathbb{C}^2$, the qubit, which may e.g. be the spin configuration of a spin-1/2 particle, and for the latter the wave function of a Schrödinger particle (e.g. an electron). It is given by a square-integrable complex valued function ψ on $\mathbb{R}^3$ such that $|\psi(x)|^2$ is the probability density for finding the particle at the position $x \in \mathbb{R}^3$. A mathematically correct discussion of infinite dimensional quantum systems involves some tools from functional analysis, while the finite dimensional systems considered in quantum computing allow much simpler formulations of interesting quantum control problems. We will therefore mainly restrict to the finite case. The most important concepts of quantum mechanics are as follows [4, 27].

- **Preparations (States):** Pure states are the one-dimensional subspaces of $\mathcal{H}$, mixed states are the positive operators on $\mathcal{H}$ with trace 1. If only pure states are considered the corresponding density operator $|\psi\rangle\langle\psi|$ is usually represented by the state vector $|\psi\rangle$ with the additional remark that this representation is not unique. States represent the preparation or, more precisely, our *knowledge* about the preparation.

  Often one needs *mixed states*: If a source emits either of the quantum states $|\psi_j\rangle$ with probability p_j then our knowledge about the system is described by the positive operator ρ with trace one given by

  $$\rho := \sum_j p_j |\psi_j\rangle\langle\psi_j| \,.$$

  Even though every positive trace-one operator can be written as a mixture over pure states in this sense it is (from the philosophical point of view) doubtful to assume that the system "is in one of the states $|\psi_j\rangle\langle\psi_j|$ but we only do not know it". One reason is that the decomposition into pure states is not unique, in contrast to classical probability theory where each measure on a set has a unique representation as a (possible uncountable) mixture of point measures ("pure states"). The second reason is the following. Given any pure entangled state on a bipartite system (see "system composition" below), the reduced state on one component is necessarily mixed even though one knows the pure state of the composed system.

  Density operators $\rho, \tilde{\rho}$ are perfectly distinguishable whenever they are supported by mutually orthogonal subspaces, i.e., the kernel of ρ contains the complement of the kernel of $\tilde{\rho}$. This is equivalent to $tr(\rho\tilde{\rho}) = 0$ since the operators are non-negative. We will simply call ρ and $\tilde{\rho}$ *mutually orthogonal*.

- **Dynamics:** The dynamical evolution of a closed system is described by the one-parameter group $(U_t)_{t\in\mathbb{R}}$ mapping the state $|\psi\rangle$ onto $U_t|\psi\rangle$ after the time t. The evolution is determined by $U_t = \exp(-iHt)$, where H is the Hamiltonian of the system. This self-adjoint operator has to be determined by physical knowledge

about the system. It constitutes the physical content of quantum mechanics in contrast to the purely abstract setting. By manipulating classical control fields, one can change the Hamiltonian of the system. Apart from the dynamics, the Hamiltonian also determines the thermodynamics of the system (see below).

- **Measurements (Observables):** Every physical quantity is represented by a self-adjoint operator A acting on $\mathcal{H}$ with the following interpretation: If P_j is the spectral projection of A corresponding to the eigenvalue λ_j, then

$$tr(\rho P_j)$$

is the probability to obtain the result λ_j, where the density operator of the system is ρ. According to Lüder's postulate, the conditioned post-measurement state is $P_j\rho P_j/tr(P_j\rho)$ if the result was λ_j. If a d-dimensional quantum system is considered as a qudit, one assumes usually that there is a preferred measurement basis, its logical states $|0\rangle, \ldots, |d-1\rangle$.

- **System Composition:** If two systems are described by Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively, then the composed system has the Hilbert space

$$\mathcal{H}_A \otimes \mathcal{H}_B\,.$$

The fact that the Hilbert space of an n-qubit register is described by $(\mathbb{C}^2)^{\otimes n}$, is a special case of this postulate. A quantum system with n-dimensional space can be always considered as composition of two systems whenever n is not a prime number. It is a matter of taste whether one wants to consider only systems with prime dimension as elementary. However, it seems reasonable to define subsystems such that they agree either with physical particles or with systems that exist at different locations in space. This is because the question which operations are elementary and which ones are not should be somehow connected with the chosen tensor product structure. Since the fundamental physical interactions are pair-interactions between particles, operations acting on two particles are 'supported by nature'. This leads naturally to two-qudit gates. We will elaborate on this point in Subsection 3.1.2.

- **System Restriction:** Given a joint density operator ρ of a bipartite system with Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. The partial trace over B is the unique density operator ρ^A which coincides with ρ when subjected to any measurement on A. Formally this means:

$$tr(A\rho^A) = tr((A \otimes \mathbf{1})\rho)\,.$$

The partial trace generalizes the marginal distribution of a classical joint probability measure on a Cartesian product.

The operations which occur by restricting the effect of a unitary map to a subsystem are so-called completely positive (CP) trace-preserving maps [28, 3]. We will make use of them in what follows.

1.3 Tools of Quantum Communication Theory

The main application of quantum communication theory at present is quantum cryptography. Here we instead apply quantum communication tools to the thermodynamics of clock synchronization in Section 2.2 and to problems of causal reasoning from classical statistical data in everyday life in 2.4, refering the reader also interested in cryptography to [3] and references therein.

Therefore we will only briefly explain those tools which are necessary for our applications and mention only the context in which they are usually used. The selection of topics is by no means representative for the issues which are important in the usual context.

1.3.1 Entropy and Information

The natural generalization of the Shannon entropy

$$S(p) := -\sum_j p_j \log_2 p_j$$

of a discrete probability distribution (p_j) is the von-Neumann entropy of a density operator

$$S(\rho) := -tr(\rho \log_2 \rho)\,.$$

It coincides with the Shannon entropy of the eigenvalues of ρ.

An important problem in classical information theory [29] is of optimally distinguishing between a set of possible probability distributions $p^{(1)}, p^{(2)}, \ldots,$. This problem occurs, for instance, when an input j of an information channel leads to the probability distribution $p^{(j)}$ on the set of possible outputs. Then it is known that the receiver can obtain the information

$$I = S(\sum_j q_j\, p^{(j)}) - \sum_j q_j S(p^{(j)})$$

about the input, when the input j is chosen with probability q_j. The simplest quantum generalization of this quantity is the Holevo-information.

Definition 5 (Holevo-Information)
Given an ensemble $\rho^{(j)}$ of quantum states which are chosen with probability q_j, the Holevo Information of the ensemble is:

$$I_H := S(\sum_j q_j \rho^{(j)}) - \sum_j q_j S(\rho^{(j)})\,.$$

The Holevo-Information is, however, only an upper bound on the so-called *accessible information* about j which can be obtained by performing optimal measurements on the states. Note that the problem to optimally distinguish between a set of quantum states is not only relevant for issues of quantum computing and quantum communication. In Subsection 2.2.8 we will, for instance, consider a model for a classical computing device which interacts with its environment and therefore transfers some information about the computation to its surroundings. Statements about the released information have clearly to refer to *quantum* states in order to be sufficiently general.

1.3.2 Quantum and Classical Correlations

Bipartite quantum systems play an important role in quantum communication protocols. Each of two parties, say Alice and Bob, has a quantum system ('generalized quantum register'). The composite system Hilbert space is the tensor product of the Hilbert spaces of the subsystems:

$$\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \,.$$

An uncorrelated state is a state which is a tensor product

$$\rho = \rho^A \otimes \rho^B \,.$$

For an uncorrelated state the von-Neumann entropy $S(\rho)$ is equal to the sum of the entropies $S(\rho^A)$ and $S(\rho^B)$ of both subsystem. For a general state one defines:

Definition 6 (Quantum Mutual Information)
The difference

$$I(A:B) := S(\rho^A) + S(\rho^B) - S(\rho) \tag{1.2}$$

is the mutual information between A and B [3]. The symbols ρ^A and ρ^B denote here the restriction of ρ to A and B by partial trace, respectively.

This definition is a natural generalization of the classical mutual information between two random variables [29] defined in terms of Shannon entropies even though its meaning is less clear.

Furthermore the following concept is crucial in quantum information theory:

Definition 7 (Separable and Entangled States)
A separable state is a convex sum of uncorrelated states

$$\rho = \sum_j p_j \, \rho_j^A \otimes \rho_j^B \,.$$

with positive probabilities p, i.e., $\sum_j p_j = 1$. A state which is not separable is called entangled.

Entanglement is a form of non-classical correlations which can only occur in quantum systems. In contrast to classical stochastics where a joint probability distribution can also be decomposed into a convex sum or integral of probability distributions without correlations, entanglement is the same sort of correlations which also occurs in pure states. A pure state is entangled if and only if its state vector is not a tensor product.

Separable states are sometimes considered classically correlated. It is, however, not commonly agreed that this is appropriate since the correlations of separable states may have also non-classical aspects. This has most explicitly been stated by Ollivier and Zurek [30]. Here we explain their ideas in our own words. Assume Alice sends Bob either of n states ρ_j^B with $j = 1, \ldots, n$ and Bob should guess which one it was by implementing a measurement. If the density operators ρ_j^B do not mutually commute it is not clear which measurement basis he should choose. It is certainly commonly agreed that the choice of the optimal basis is a typical question of quantum estimation theory [31]. Consider a

third observer who does not share Alice's knowledge about which of the states was sent. He only knows the set of possible states and the probabilities p_j for the choice of each. According to his description the bipartite system is in a correlated state

$$\rho = \sum_j p_j \, \rho_j^A \otimes \rho_j^B \, ,$$

where ρ_j^A denote some states of Alice's mind or memory storing the information about which state was sent. If her mind or memory is perfect, they are perfectly distinguishable, i.e., orthogonal density operators. When Bob applies measurements on his system in order to obtain some information on j his intention could also be interpreted as trying to find out something about the state of Alice's memory. This is certainly also a question of *quantum* information since we have accepted that it is a question of quantum information theory to get information about j. As realized in [30], for non-commuting ρ_j^A there is a difference between the mutual information between Alice and Bob and the information that Bob can obtain about Alice's system by applying the measurement (P_l) to his own system. The latter quantity is given by

$$S(\rho^A) - \sum_l q_l S(\rho_l^A) \, , \tag{1.3}$$

where q_l is the probability for obtaining the measurement outcome l and

$$\rho_l^A := \frac{1}{q_l} tr_B((\mathbf{1} \otimes P_l)\rho(\mathbf{1} \otimes P_l))$$

is the conditional post-measurement state of Alice given that the outcome was l. The minimum of the difference between (1.2) and (1.3) over all measurements was called 'quantum discord' in [30]. Even though it was not explicitly stated there, it is clear that the definition could be modified by also allowing POVM measurements [32, 33]. It should only briefly be mentioned that in the special case above where Alice's states are perfectly distinguishable this modified discord is exactly the difference between the Holevo-information and the accessible information [3] of the state ensemble p_j, ρ_j^B. As we will see later, states with discord will appear in the context of clock synchronization.

1.3.3 Secret Correlations and Entanglement

We have already explained in Section 1.2 that mixed states appear in quantum theory in two different situations: First, if we prepare randomly some pure state without knowing which one was prepared and second, if we consider the restriction of a pure bipartite pure state to one of the subsystems. It is irrelevant for the statistics of all possible measurements performed on this system which possibility lead to the mixture. Likewise, if Alice and Bob share a mixed joint state it is irrelevant for the correlation between their measurement results whether their joint state is mixed because it is the restriction of an entangled state shared with a third party. However, there are two reasons why the difference matters nevertheless. For *cryptography* it is certainly important whether there is some third system which shares some information with Alice and Bob. Therefore it is useful for cryptography when Alice's and Bob's correlations are due to the *entanglement* of a pure bipartite state. Then there cannot be any correlation to a third party.

Secondly the difference is important from the thermodynamic point of view: If a communication protocol between Alice and Bob creates correlations with their environment (which is, by definition, a system where they do not have any access to) the process is irreversible and is accompanied by a loss of free energy (see Section 2.2). This will be important when we consider the problem to synchronize clocks such that no correlation with the environment occurs.

Chapter 2

Old Fundamental Problems in Light of Modern Quantum Information Research

This work advocates the use of tools from quantum information and computation to revisit fundamental questions in physics and computer science. The application to computer science requires perhaps less justification. Supposing someday quantum computers are built and outperform classical computers this would certainly greatly impact the paradigms of computer science. For physics, there are mainly two reasons to use QC tools also to revisit the fundamental issues: First the clarity and simplicity of models are attractive. Second the terminology of quantum information theory abstracts from specific hardware in constrast to the traditional language of physics which tends to refer rather to the theory of specific systems. Examples of this include the terms 'canonically conjugated observables' in quantum theory (which exist only for infinite systems) or 'pressure and volume' in thermodynamics (which are not appropriate to describe the thermodynamics of spin variables). Hence I prefer the concept of qudits (or their infinite dimensional generalization) as an abstract model of an unspecified quantum system. As we will see, the abstraction will help to shed light on device-independent limits of quantum control.

2.1 Revisiting Quantum Mechanics from an Algorithmic Point of View

In books on the foundations of quantum mechanics [27, 28], states are often introduced as equivalence classes of preparation procedures (for quantum objects) which lead to the same statistics in every potential measurement. Similarly, observables are equivalence classes of measurement procedures which lead to the same statistics given the same preparation procedure. This approach expresses already the demand to give an operational meaning to the ingrediences of quantum theory instead of speculating 'what a quantum state really is'. Quantum computing research can be considered as the most ambitious effort since the beginning of quantum theory to find such an operational meaning: It was stated in [34, 3] that every proposal for quantum computing should clarify (0) how the quantum information is represented in a sufficiently protected manner, (1) how

appropriate initial states are prepared, (2) how a sufficient set of unitary transformations is implemented, and (3) how the output is measured. It would be interesting to know which states are difficult to prepare and which observables are difficult to measure, and more ambitiously, one would like to find general principles which do not rely too much on features of a particular physical system. The word 'difficult' can have different meanings in different contexts. From a computer science point of view it seems to refer to some sort of complexity theory in the sense of counting elementary operations. But 'difficult' could also mean that a particular state preparation or measurement would require controlling the system very precisely and isolating it strongly to protect it against being disturbed by the environment. This kind of 'thermodynamic' difficulty may differ substantially from the complexity theoretical criterion as will be seen in Subsection 2.1.2.

2.1.1 Operational Meaning of States, Observables, and Unitary Maps

We introduce states as equivalence classes of preparation procedures. The most elementary preparation procedure is not to act on the quantum system at all. But what is the state of a system which has not been subjected to any external treatment? A central postulate of thermodynamics is that it is in a thermal equilibrium state:

- **Thermal Equilibrium:** When subjected to an environment with temperature T each system is after a while described by the Gibbs state

  $$\gamma_T := e^{-\frac{H}{kT}}/tr(e^{-\frac{H}{kT}}),$$

  where H is the Hamiltonian and k is Boltzmann's constant. Hence the eigenvalues p_j of γ_T are computed from the eigenvalues E_j of H according to the *Boltzmann* formula

  $$p_j = e^{-\frac{E_j}{kT}}/\sum_l e^{-\frac{E_l}{kT}}$$

This implies that the 'most natural initial state' of a quantum register is not necessarily the state $|0\ldots0\rangle$, it is rather the thermal equilibrium depending on the system Hamiltonian. In present day NMR experiments it is one of the main problems that the Gibbs state is highly mixed [35].

The abstract setting above raises the following questions: Is it possible to prepare an arbitrary state vector $|\psi\rangle \in \mathcal{H}$? Is it possible to implement an arbitrary unitary transformation U on $\mathcal{H}$? Is it be possible to measure every self-adjoint operator? If one assumes that the standard model of the quantum computer will be realized in the future one could answer the questions above. The answers could easily be generalized to qudits.

1. **Unitary Operations:** Every unitary operation on $(\mathbb{C}^2)^{\otimes n}$ can at least approximatively be generated from one- and two-qubit gates. A central issue for proposals of QC implementations is how they can be realized. Since the arguments below rely on the feasibility of sufficiently many basic operations we will add some remarks showing why one- and two-qubit gates are natural basic operations at the end of this subsection.

2. **State Preparation:** Initially, the system is described by its equilibrium state ρ_T according to the temperature T. Measuring the logical state of the register yields some basis state $|b\rangle$ with some binary word b. Applying an appropriate unitary operation U, one may obtain any desired pure state $|\psi\rangle$ with arbitrary accuracy. Mixed states can be obtained by randomizing this procedure using an external random generator. For large numbers of qubits or qudits the preparation of a generic state would certainly require a huge number of gates [36, 37]. One could also define a Kolmogoroff complexity for quantum states [38] based on the complexity of the description of preparation procedures. In [39] an efficient scheme was given to prepare a state of the form

$$|\psi\rangle := \sum_{j=0}^{N-1} \sqrt{p(j)}\, e^{i\phi(j)} |j\rangle \,,$$

whenever p is a probability distribution on the $N = 2^n$ basis states and ϕ specifies the corresponding phase under the assumption that p and ϕ can be computed efficiently on a classical computer and that all $p(j)$ are bounded from above by some constant multiple of $1/N$. In keeping with the general intended message of this thesis, it is interesting to note that Grover's algorithm is used to solve this non-computational problem.

3. **Von-Neumann Measurements:** Given arbitrary unitary operations, one can measure every self-adjoint operator A. Let U be a unitary that diagonalizes A, i.e., $UAU^\dagger$ is diagonal in the computation basis. Let $(|\psi_j\rangle)_{j=0,\ldots,N-1}$ be an orthonormal set of eigenvectors of A. If $U|\psi_j\rangle = |j\rangle$ we may measure A by applying U and measuring in the computational basis [40]. One obtains the result j with probability $\langle\psi_j|A\psi_j\rangle$. If one demands that the post-measurement state is an eigenstate of A one can implement $U^\dagger$ afterwards. For degenerate A (i.e. if not all eigenspaces are one-dimensional) this does not imply that the procedure *projects* onto the eigenspaces.

4. **Von-Neumann Measurements with Lüders' Postulate:** If one would like to have a measurement with minimal disturbance such that the state vector is *projected* onto the eigenspace [41], one has to get around the following problem: measurements of a subset of $k < n$ of the n qubits could only implement projections onto subspaces of dimension 2^{n-k}. However, observables having arbitrary dimensional eigenspaces can easily be measured using additional qubits, so called 'ancillas': after the ancillas are all set to 0, one implements the mapping

$$U(|j\rangle \otimes |l\rangle) = |j\rangle \otimes |l \oplus f(j)\rangle$$

where the numbers $f(0), \ldots, f(2^n - 1)$ enumerate the eigenvalues and $f(j)$ is the number of the eigenvalue of $|\psi_j\rangle$. Afterwards one measures the ancillas and implements $U^\dagger$.

5. **Generalized Measurements (POVMs):** Even though generalized measurements are usually not considered in quantum mechanics textbooks, they are older than quantum computing research [28]. This concept of advanced quantum theory describes every measurement with outcomes in a measurable set Σ as positive

operator-valued measure (POVM) on Σ, i.e., a map $S \mapsto M_S$ for every measurable set $S \subset \Sigma$ such that $M_S \geq 0$ and $M_\Sigma = \mathbf{1}$. For every state ρ

$$p(S) := tr(\rho M_S)$$

is the probability for obtaining an outcome in S. It is by no means clear how to implement a generic POVM. However, in principle one can use the Naimark extension [42, 43] in which the system is embedded into a larger system and the POVM measurement is reduced to a von-Neumann measurement on the larger system. Note that this is also called 'frame dilation' in other contexts. On the quantum computer one could reduce every POVM with finitely many outcomes to von-Neumann measurements.

6. **Reversibility of Quantum Dynamics:** Understanding the difference between the future and the past is considered a deep fundamental problem of science and philosophy [44, 45]. One of the most difficult questions in this context is to understand the second law of thermodynamics: Whereas the laws of micro-physics and particle physics seem to be symmetric with respect to time reflection there is no such symmetry on the level of thermodynamics. In my opinion, reversibility of microphysics could mean both of the following statements, my impression being that they are sometimes not clearly distinguished:

   - **Formal Time Inversion Symmetry:** Given a basis of energy eigenstates ($|E_j\rangle$), the anti-linear map

     $$R : \sum_j c_j |E_j\rangle \mapsto \sum_j \overline{c}_j |E_j\rangle$$

     can be considered as a time-reflection due to

     $$R \exp(-iHt) R = \exp(iHt)\,.$$

     When replacing all observables A with RAR one seems to observe the time-reversed process. For a Schrödinger wave function one replaces, for instance the momentum P (which is up to a factor the velocity of the particle) with $-P$.

   - **Implementing the Inverse Time-Evolution:** Note that the formal symmetry above does not answer the control-theoretic question of how to implement a transformation which transforms

     $$\exp(-Ht)|\psi\rangle$$

     back to $|\psi\rangle$. We may also consider a system with non-degenerate density operator ρ. The only possibility to convert

     $$\exp(-iHt)\rho \exp(iHt)$$

     back to ρ is to apply the unitary operation $\exp(iHt)$, i.e., to *implement* the inverse of the time-evolution. But it can be more difficult to implement $\exp(iHt)$

than $\exp(-iHt)$. The reason for this kind of "arrow of time" is simply that H is the real Hamiltonian and not $-H$. This argument does not necessarily apply when the Hamiltonian of the system stems from an external control field like the Hamiltonian of a spin system in a magnetic field, where H can be reversed by reversing the field. The situation we have in mind is that H is an interaction between particles like the electrostatic forces between two electrons which cannot be reversed. In [46] we have considered the complexity of *simulating* $-H$ when H is present, i.e., to intersperse the natural dynamics in such a way that the effective dynamics runs backwards. In Subsection 3.1.3 we consider the problem of inverting the dynamics of n interacting spin-1/2-particles by appropriate sequences of single-spin operations[1]. Depending on the specific type of the interaction this refocusing sequence involves necessarily $n-1$ time steps. We have argued that this holds also if the interaction between many pairs are small, e.g. due to large distances between them. This suggests that the inversion of n-particle dynamics requires in the generic case sophisticated control sequences with a complexity which increases with n.

In a system where not many control operations are available it is not clear how to invert the time evolution at all. However, whenever the set of operations are sufficient to allow a transfer of quantum information between the system and a register of a universal quantum computer, an efficient implementation of time inversion is probably possible. This is implied by a generalized understanding of the strong Church-Turing principle: given the assumption that the quantum computer can simulate the original dynamics $\exp(-iHt)$ efficiently, it can certainly also simulate the reversed evolution $\exp(iHt)$ by reversing the whole circuit. This implies that *irreversible processes* in the sense that no implementation of $\exp(iHt)$ is feasible could only take place in systems which allow no sufficient interface to a universal quantum computer.

The question of which control operations are the most elementary ones is certainly an essential question of algorithmic quantum control. For state preparation we have already mentioned that it is natural to consider the preparation of an equilibrium state as elementary as long as it is the real temperature of the environment surrounding the system. To achieve different temperatures would already be a non-trivial control problem. It is less clear which measurements and unitary control operations are basic. From the pragmatic point of view this is just an issue of experimental physics. Nevertheless, it would be interesting to understand in which way the mathematical structure of the fundamental interactions in physics support some operations and not others. First of all, the fact that these forces are pair-interactions is closely related to the special role of one- and two-qubit or two-qudit gates. This is discussed in Subsection 3.1.2 where the discrete quantum circuits are replaced by time-dependent pair-interactions and it is shown that this computation model differs little from the two-qubit gate model. Unfortunately, this analogy is not completely convincing since the fundamental forces between particles are not controllable, they are simply *always present*. At first sight it seems as if controllable pair-interactions would require three-particle interactions as the following model suggests. Let H_0 and H_1 acting on $\mathcal{H}_A \otimes \mathcal{H}_B$ be two interactions between two particles A and B.

[1] the so-called refocusing technique in NMR [47]

We would like H_0 to be present when our controlling device is in state $|0\rangle$ and H_1 when it is switched to $|1\rangle$. This is given by an interaction among controller and A and B which reads

$$|0\rangle\langle 0| \otimes H_0 + |1\rangle\langle 1| \otimes H_1 \,.$$

This seems to be at least a three-particle interaction which would be unphysical. However, there are several solutions to this problem. One possibility is that $\mathcal{H}_A$ and $\mathcal{H}_B$ are not the whole Hilbert spaces of the particles. The spaces $|0\rangle \otimes \mathcal{H}_A$ and $|1\rangle \otimes \mathcal{H}_A$ could, for instance, be different subspaces of the whole Hilbert space of particle A. Then the control operation consists of bringing A into either of the subspaces where the interaction looks like either of both Hamiltonians H_i.

A second possibility is that H_0 and H_1 are only *effective* interactions. They may describe, for instance, the situation when independent unitary operations on $\mathcal{H}_A$ and $\mathcal{H}_B$ of the form

$$U := U_A \otimes U_B$$

are applied to a system which intersperse the natural time evolution of a fixed Hamiltonian H (see Subsection 3.1.2). To control the implementation of the single-particle operations U_A and U_B requires only pair-interactions between the particles and the controlling device.

In [48] we tried to understand in which way the structure of the interactions between system and controlling device determines which control operations on the system are easy to implement. In doing so, we considered a model of quantum control where a system and its controller are both represented by quantum systems with a *fixed* interaction between them. The only access to the system is given by operations on the controller. This raises the question which operations on the controller are possible which would clearly lead to an infinite chain of controllers and its meta-controllers. The arbitrariness of the borderline between system and controller is well-known in discussions of the quantum measurement process. In this context, the borderline is called the *Heisenberg cut* [49]. Surprisingly, I do not know of any remark in the literature that the Heisenberg cut appears also in quantum *control* if one takes into account that interactions cannot be changed; they are simply given by the laws of particle physics. To understand the resource requirements of quantum control from this point of view, the state of the whole controlling device is the program determining the sequence of control operations and the dynamics of system and controller is the autonomous dynamics generated by a large Hamiltonian. An interesting result of the models in [48] is that even though the access to the controller is unrestricted there is a distinguished set of observables and unitaries which allow particularly simple implementations.

In [50] we have reduced the control possibilities further and constructed a 'programmable Hamiltonian'. which simulates universal gate operations if its initial state is chosen appropriately. It can be considered as a continuous time version of a cellular automaton. Even though our model requires 10-qubit interactions there is no apparent reason why this should not be possible with more realistic interactions. Meanwhile I found nearest-neighbor-interactions among qutrits on a 2D-lattice which implement an autonomous quantum computer [51]. The remarks above relate, at least in a vague sense, the physical resource requirements of quantum computing with space and time requirements of such a continuous-time quantum cellular automaton to implement a certain control sequence on a part of its register.

2.1.2 Micro- and Macro-Physics and Fault-Tolerance of Quantum Gates

Since the early days of quantum mechanics it has been an issue of many debates why physics in every-day life behaves classically and not according to the superposition principle. The famous 'Gedankenexperiment' with Schrödinger's cat was intended to show the counter-intuitive implications of quantum mechanics for every-day life: A cat being in a superposition between alive and dead. Textbook quantum mechanics does not restrict the superpositions principle such that those counter-intuitive superpositions are a priori excluded.

The hypothesis of quantum computing research is: Every state in the Hilbert space $(\mathbb{C}^2)^{\otimes n}$ of a quantum register can indeed be prepared. This means that one could even prepare the 'cat state'

$$\frac{1}{\sqrt{2}}(|0\ldots0\rangle + |1\ldots1\rangle)\,. \tag{2.1}$$

If every qubit is given by the spin of a particle and $|0\rangle, |1\rangle$ are the states with spin up and down, respectively, this would be a state with undefined magnetization, i.e., only if one measures the magnetization "the system decides" whether it is a magnet in positive or in negative z-direction.

The reason that superpositions like the cat state above are surprising is that some properties of a physical system have such an immediate effect on the surrounding of the system that it does not require a measurement instrument to investigate them. This is, for instance, the case for the mean magnetization of a many-spin system which determines the magnetic field around the system. Let us state these philosophical issues more explicitly. Consider an observable A and a state ρ that commutes with A, allowing the interpretation that A has at the moment a certain value, but we just do not know it. If ρ is, for instance, a superposition of eigenstates of A the statement "A has a definite value but we do not know it" contradicts the principles of quantum theory. However, the generic situation is between those two cases: If ρ, for instance, *almost* commutes with A, then the uncertainty of A-measurements is partly due to our missing knowledge and partly due to inherent quantum uncertainty. Given some definition of *macroscopic observables* we consider *macro-realism* as the statement that for all *macroscopic observables* A the trace norm of the commutator $[\rho, A]$ is small for all states ρ in nature. The consideration of the trace-norm is justified by the observation [52] that for any pure state ρ, the expression $\| [A, \rho] \|_{tr}$ coincides with the standard deviation of A up to the factor 2.

Even though we do not have a definition for macroscopic observables, we have argued above[2] that all observables of the form

$$\overline{a} := \frac{1}{n}\sum_{j=1}^{n} a_j$$

are macroscopic, where a_j is the 1-qubit operator a acting on qubit j. Accepting this definition, one would tend to find those states ρ more 'natural' and less paradoxical which almost commute with all observables of this type. It is easy to check that the cat state (2.1) satisfies

$$\| [|\psi\rangle\langle\psi|, \overline{\sigma_z}] \|_1 = 2\,,$$

[2] in agreement with Poulin [53], see also our arguments in [52]

which is already the maximum of a commutator with two norm-one observables.

It would be an attractive idea to assume that such counter-intuitive states which do not almost commute with macroscopic observables require preparation procedures with high complexity. Then macro-realism would, so to speak, be supported by complexity theory. But this is easy to disprove because the following scheme prepares the n-qubit cat state (2.1) using a circuit of depth $O(\log_2 n)$.

1. Start with the state $|\psi_0\rangle := |0\rangle^{\otimes n}$.

2. Perform a Hadamard gate on the first qubit and obtain

$$|\psi_1\rangle := 1/\sqrt{2}(|0\rangle + |1\rangle) \otimes |0\rangle^{\otimes n-1} .$$

3. Perform a C-NOT controlled by qubit 1 with qubit 2 as target. This prepares a cat state on 2 qubits.

4. Given a cat state on a 2^k-qubit cluster obtain a cat state on a 2^{k+1} qubit cluster by applying 2^k C-NOT gates. Each of them is controlled by one qubit in the 2^k cluster and has an arbitrary target qubit in the remaining set.

To show that $O(\log_2 n)$ is optimal we proved in [40] that every state prepared by a quantum circuit with depth d from a product state satisfies

$$\|[\rho, \overline{a}]\|_1 \leq \sqrt{\frac{2}{n}} 2^d .$$

We also considered the question whether one could create macroscopic superpositions by any *measurements* which can be implemented using circuits with depth less than $\Omega(\log_2 n)$ In our model the measurement is given by performing a unitary U, measuring a subset of $k \leq n$ qubits in the computational basis and performing $U^\dagger$ afterwards. We have shown that such a measurement (P_j) satisfies

$$\|[\overline{a}, P_j]\|_1 \leq \frac{2^d}{\sqrt{2n}} \tag{2.2}$$

if the circuit U has depth d. In other words, measuring observables which do not almost commute with $\overline{a}$ requires logarithmically increasing depth in n. Summarizing these results, the preparation complexity of states which are not consistent with macro-realism increases with n, even though the increase is only logarithmic. Therefore the arguments do certainly not show that there are complexity theoretic limitations for creating macroscopic superpositions in a quantum computer.

However, one can nevertheless maintain the claim that it is 'hard' to prepare states which contradict macro-realism on a hypothetical mega-qubit quantum computer since cat states require an *extremely closed* system as well as *reliable* gates. This is similar to the widely-accepted view that macrorealism is an effect of the environment [54]. But usually it is only argued that the environment destroys macroscopic superpositions very quickly. Here we want to argue that they are not even created when the reliability of the gates is not sufficient.

To give an intuition for the idea, we consider a quantum circuit which provides simultaneously a probabilistic algorithm for approximative $\overline{\sigma_z}$ measurements of n qubits and a model for the decoherence caused by the environment. Depending on the measurement accuracy demanded we use k additional qubits initialized to $|0\cdots 0\rangle$. Then we apply k C-NOT gates with the k ancillas as targets, where we choose randomly k of the n qubits as control inputs. After counting the number l of ancillas in the upper state $|1\rangle$ the value $l\,k/n$ gives an estimation for $\overline{\sigma_z}$. One can easily show that the output density operator ρ of the n qubit register has only a small commutator with $\overline{\sigma_z}$. This could be a good model for the effect of the environment if one decides independently and randomly for each qubit whether it is a control qubit for some C-NOT. But the environment will not always 'measure' in the σ_z basis. If the C-NOT is conjugated by some unitary U on the control qubit the environment 'measures'

$$\overline{U^\dagger \sigma_z U}\,.$$

Note that a measurement with respect to some randomly chosen basis in $\mathbb{C}^2$ (according to the unique $SU(2)$ invariant probability distribution) has the same effect on the measured system as the depolarizing channel

$$D(\rho) := (1-\epsilon)\rho + \epsilon\frac{1}{2}\mathbf{1}\,,$$

with some small $\epsilon > 0$. Hence a weakly polarizing channel on n qubits acts as if all observables $\overline{a}$ are approximatively measured and leads to states ρ where all commutators $[\overline{a}, \rho]$ are small. The detailed version of the intuition above can be found in [52]. We have proven that every state in the image of $D^{\otimes n}$ satisfies

$$\|[\rho, \overline{a}]\|_1 \leq 2\left(\frac{1}{\epsilon n(1-\alpha^2)} + \frac{1}{\sqrt{n\epsilon\alpha}}\right)$$

for every $1 > \alpha > 0$. In the asymptotics for large n we obtain a commutator with $\overline{a}$ in the order of $1/\sqrt{\epsilon n}$. For separable states one obtains $O(1/\sqrt{n})$. Furthermore we have shown that each observable B with $\|B\| \leq 1$ defines an entanglement witness: Given a partition of the n qubits into subsets ("clusters") of size $l_1, \ldots, l_k$ (with $\sum_j l_j = n$). Then every state which has no entanglement among the clusters satisfies

$$tr(\rho[\overline{a}, B]) \leq \frac{2}{n}\sqrt{\sum_j l_j^2}\,.$$

This shows that quantum uncertainty of macroscopic observables requires large scale entanglement.

The fact that even a weakly polarizing channel makes it impossible to obtain states with large quantum uncertainty of $\overline{a}$ implies that the generation of those states requires extremely reliable gates, since it is natural to assume that every real gate implementation depolarizes the qubit with some probability [52]. The possibility of quantum error correction is irrelevant for this result if interpreted correctly, since our result refers to the *physical level*, whereas error correction refers to the *logical level*. It is clear that error correction [22] allows the preparation of cat states on the logical qubits with arbitrarily

high accuracy if a certain reliability threshold for the gates can be guaranteed. Nevertheless, all observables $\overline{a}$ defined on the *original physical qubits* have essentially only classical uncertainty. The reason for this phenomenon is a separation of scales: Clearly the logical cat state will also have some quantum uncertainty of $\overline{a}$ but it will not be visible on the scale of $\|\overline{a}\|$. To explain the appearance of a classical world as an effect of the decoherence caused by the environment is meanwhile standard [54]. However, the approach above shows that macroscopic superpositions not only decohere quickly, they are even impossible to prepare with inaccurate operations. Taking into account that accurate gate implementations will typically require complex experimental setup this statement can also interpreted as complexity theoretic limitation for violating the type of macrorealism considered above.

The sensitivity of macroscopic superpositions to inaccuracies of the preparation procedure has also another aspect. The state of a many-particle system is typically *exchangeable*. An n qubit state is called exchangeable if (1) it is symmetric with respect to permutation of qubits (2) for every $m \in \mathbb{N}$ there is a permutation symmetric state on $n+m$ qubits such that its restriction to n qubits coincides with ρ. The notion of exchangeability formalizes a natural preparation procedure for macroscopically large ensembles of particles: The preparation procedure not only treats all particles in an equivalent way, the preparation procedure is not specifically adapted to the number n of qubits. In other words, there is a family of preparation procedures generating states ρ_n such that the restriction of ρ_{n+m} to an arbitrary selection of n qubits is ρ_n. This is, in some sense, also a fault-tolerance of the preparation procedure: It is not necessary to specify the number n of qubits *before* running the preparation procedure, one may run it for some $n' \geq n$ and select n qubits afterwards. Due to the quantum de Finetti theorem each exchangeable state has the form

$$\rho = \int \nu^{\otimes n} d\lambda(\nu)$$

with an appropriate measure λ on the set of 1-qubit states ν. Such states are almost undisturbed by $\overline{a}$-measurements if the latter are smoothed in an appropriate way (for details see [53]).

2.2 Revisiting Thermodynamics

The relation between thermodynamics and information theory is not really easy to explain as long as one considers *classical* physical systems. Its physical states are points in a continuous phase space which require infinite information to describe. As a generalization of Shannon entropy of discrete probability distributions (p_j) to probability densities $(p(x))$ one uses often the integral

$$S(p) = \int p(x) \log_2 p(x) dx + c$$

with an arbitrary additive constant c. The relation of such a quantity to bits (as the elementary unit of information) is less obvious than in the discrete case and the arbitrariness of c reflects its unclear meaning [55]. Certainly one can define classical discrete models by for instance discretizing the phase space or by neglecting superpositions in quantum

systems (as often done in Ising models in statistical physics). The problem is that classical models with finitely many states cannot have continuous reversible time-evolution [6]. Hence classical thermodynamics suffers from a lack of simple models.

In the following we will therefore often consider the thermodynamics of quantum mechanical two-level systems. The role of quantum superpositions will only be marginal. The decisive input from quantum information theory in these considerations is to consider unitary transformations on collections of two-level systems as simplest interesting examples for thermodynamically reversible operations on physical objects. This gives a first impression how closely related thermodynamic machines like heat engines and refrigerators are to computing devices. Some ideas for a complexity theory of thermodynamical machines will be presented in Section 3.2

2.2.1 Understanding Thermodynamics by Toy Models

Even though the laws of thermodynamics are directly relevant in every day life they are commonly not well understood and in particular the second law is sometimes considered mysterious even by physicists. The following questions are among the essential ones:

- Why is it not possible to use all the heat around us to generate other forms of energy even though heat is also a form of energy? (such a machine would be called '*perpetuum mobile* of the second kind')

- Why does a refrigerator need energy even though it is only required to *extract* energy from the inside?

The answer 'this follows from the second law' is not really satisfying. One has learned to accept the fact that energy generation from heat requires temperature *differences* and that, for the same reason, the generation of temperature differences requires energy (otherwise one could generate energy by generating temperature differences first).

As a physicist, one learns to accept that the impossibility of a perpetuum mobile of the second kind has to do with a 'mysterious quantity', called entropy which can only increase but never decrease. And, what makes all these statements even harder to believe, entropy quantifies information, or more precisely, the *missing* information about nature. The fact that the information about nature is decisive for the worth of the available energy seems a bit odd. However, it may be illustrated by the following:

A shop sells batteries and collects those which are used up for recycling purposes. After almost all the full batteries are sold, the shop contains almost only empty batteries. Unfortunately, the owner is sloppy and he forgets to keep the full and empty batteries separately. If only one full battery is among the empty, its energy is lost from the point of view of a lazy shop-owner who does not want to search all his shelves for the full one. However, it is not lost from the fundamental point of view. A thrifty shop-owner would measure the voltage of the batteries in order to identify the new one. But the measurement consumes a little bit of the energy. If we substitute the batteries with some energy storage on the atomic level the measurement may consume more energy than the atom contains and thus some energy is really lost by *forgetting* the location of the full battery.

In the following we will present models which provide a quantitative understanding that information is as valuable as energy. The standard model of quantum computing (see Section 1.1) also provides appealing models to explain this. One may be surprised why we have chosen to use quantum computers as models even though superposition states do not play any role in this section. The ability to process superpositions seems to be *the essential* feature of the quantum computer. However, the quantum computer has in fact *two* important features:

1. It preserves superpositions between logically different states
2. It does not copy any information from its register to the environment during the computation.

It is easy to see that (1) implies (2). This is essentially the no-cloning theorem [9]. As soon as a mechanism copies the classical states to the environment a superposition between different logical states could not survive. One tends to believe that (2) implies also (1). We will address this question in Subsection 2.2.8.

Now we will try to understand thermodynamics by scaling down thermo-dynamical machines to the quantum scale, since this makes the relation between information and energy more obvious. We recall that the main physical information about a system is represented by its Hamiltonian H. If a fixed reference temperature $\infty \neq T > 0$ is given, one can equivalently describe the energy level structure by the equilibrium state γ_T. Hence we will describe a system merely by this density operator γ_T. If a system is not in equilibrium it is described by some density operator $\rho \neq \gamma_T$. By specifying the pair (ρ, γ_T) (both density operators acting on a common Hilbert space $\mathcal{H}$) one has all thermodynamically relevant information. We define therefore:

Definition 8 (Thermodynamical Object)
A thermodynamical object is given by the pair $O := (\rho, \gamma_T)$ where γ_T is the equilibrium state and ρ the actual state.

An equilibrium object is a pair (γ_T, γ_T). Since the equilibrium state of non-interacting systems is simply the tensor product of equilibrium states one has a simple composition rule for describing two objects in one:

$$O \otimes \tilde{O} = (\rho, \gamma_T) \otimes (\tilde{\gamma}, \tilde{\gamma}_T) := (\rho \otimes \tilde{\rho}, \gamma_T \otimes \tilde{\gamma}_T)\,.$$

Now we consider the question which objects are energy resources, i.e., which objects can be subjected to a physical process in such a way that the process extracts energy in the statistical average. Here we allow only unitary transformations as processes and define:

Definition 9 (Energy Source)
An object (ρ, γ_T) is an energy source if there is a unitary U acting on the system Hilbert space $\mathcal{H}$ (where ρ and γ_T act) such that

$$E_{gain} := tr(\rho H) - tr(U\rho U^\dagger H) > 0\,,$$

where H is the Hamiltonian specified by γ_T.

The gained energy is transfered to the environment or the physical medium (e.g. control field) which drives the implementation of U. The reason why we allow only for unitary operations is that we want to include all possible entropy sinks in the description[3]. It is

[3]For a detailed justification see [56].

easy to characterize non-energy sources explicitly (see [57, 56]):

Lemma 1 (Condition for Energy Sources) *An object O is not an energy source if and only if ρ and γ_T commute*[4] *and the eigenvalues of ρ are non-increasing with decreasing eigenvalues of γ_T, i.e., with increasing eigenvalues of H.*

We conclude easily:

Theorem 1 (Impossibility of Perpetuum Mobile) *An equilibrium object is not an energy source.*

But of course there are also non-equilibrium objects which are not an energy source in the sense above. Nevertheless, there is a relation between the property of being an energy source and being the equilibrium state. This is shown by the following theorem[5]. We gave a slightly different proof in [56] which shows nicely that the Boltzmann distribution arises from geometric arguments.

Theorem 2 (Copies of Non-Equilibrium States are Energy Sources)
For an object O with different energy levels the following statements are equivalent:

1. *O is an equilibrium object for some $T \in (0, \infty]$ or ρ is such that the whole probability is concentrated in the lowest energy levels.*
2. *There is no $n \in \mathbb{N}$ such that the n-fold copy $O^{\otimes n}$ is an energy source.*

We conclude that every non-equilibrium object (ρ, γ_T) can be used as energy source if we have sufficiently many copies of it. This asymptotic statement coincides with conventional thermodynamics where every non-equilibrium system is an energy source. For large systems one can argue as follows. Given a state ρ which is not an equilibrium state for any temperature T. Then the average energy $tr(\rho H)$ of ρ coincides with the average energy $tr(\gamma_T H)$ of the equilibrium state for some temperature T. Since γ_T is the unique state with maximal entropy given its average energy the entropy of ρ is less than the entropy of γ_T. This means that one can lower the average energy of the system without extracting entropy and the system is thus an energy source. To discuss issues like that conventional thermodynamics introduces the concept 'free energy', formalizing the fact that the amount of usable energy depends on the entropy inherent in the system. We will come to that point later.

2.2.2 The Most Elementary Heat Engine

The essential difference between a heat engine like a steam engine in contrast to a perpetuum mobile of the second kind is that the former is driven by two or more equilibrium objects with different temperatures and not by one equilibrium object as is the latter. We would like to understand intuitively why two systems with different temperatures may be an energy source, whereas two systems with equal temperature are not. For doing so

[4] Note that this condition is equivalent to $[\rho, H] = 0$.

[5] see [57], for a generalization to infinite dimensional systems see also [58].

we take both systems to be the most elementary ones, namely two-level systems. Let system A have energy gap E_A and system B have gap $E_B < E_A$. Assume furthermore

$$\frac{T_A}{E_A} > \frac{T_B}{E_B}\,.$$

Label the basis states of the bipartite system (A, B) by $00, 01, 10, 11$. One observes easily that the state 10 has more energy than 01 even though the latter is more likely. This shows the following calculation:

$$\frac{p_{10}}{p_{01}} = e^{-\frac{E_A}{kT_A}} e^{\frac{E_B}{kT_B}} = e^{\frac{E_B}{T_B} - \frac{E_A}{T_A}}\,,$$

which is greater than 1 by assumption. Hence we gain energy by implementing the SWAP-gate, i.e., the permutation $10 \leftrightarrow 01$. The fact that the system is only an energy source if $E_A \neq E_B$ is a constraint which seems not directly to be related with the second law. It is something like an additional constraint for micro-physics. Those constraints become less relevant in larger systems since we can construct heat engines consisting of many two-level systems with equal energy gap.

2.2.3 The Most Elementary Refrigerator

The same system as above can be used as a refrigerator. We assume that both two-level systems have temperature T at the beginning. Due to $E_A > E_B$ the state 01 is more likely than 10 and has less energy than the latter. Therefore the permutation $10 \leftrightarrow 01$ reduces the (marginal) probability for the upper state in the right system by the amount $p_{01} - p_{10}$. Since every decreasing distribution on a two-level system is an equilibrium state for some temperature (see remarks in Subsection 2.2.1), the final marginal distribution of system B is an equilibrium state with lower T. The marginal distribution of A is a state with higher T. It is easy to see that the process $10 \leftrightarrow 01$ *requires* energy since now the higher-energy state is more likely. This nicely shows that this fridge needs energy in order to generate a temperature difference.

2.2.4 Computer Scientist's Fridge

Our toy models for heat engines and refrigerators with 2 two-level systems requires two systems with different energy gaps. Here we show that refrigerators are also possible for systems with equal energy gaps when at least 3 of them are available. Such a system has 8 basis states. Now we construct a cooling process for the third two-level system which maps the 4 states with least probabilities to those states $|b\rangle$ where b has suffix 1. This can be done by the map $|011\rangle \leftrightarrow |100\rangle$ which leaves all the other basis states invariant. A circuit[6] from C-NOT and TOFFOLI gates that implements this transformation can be seen in Fig. 2.1.

Let p, q the probabilities for the upper state of one qubit. They are given by (see Section 2.1)

$$\frac{p}{q} := e^{-E/(kT)}\,.$$

[6]Thanks to Joe Renes for this circuit.

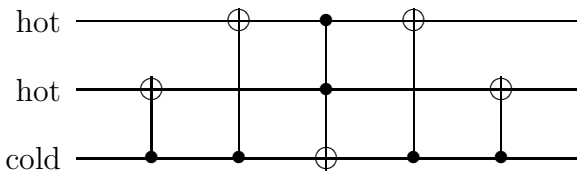


Figure 2.1: Quantum circuit with C-NOT and TOFFOLI gates which is a fridge and heat engine at the same time. If all three qubits start with identical temperatures the circuit lowers the temperature of the third qubit and the implementation *requires* energy. If the third qubit is initially already sufficiently colder than the others, the implementation *releases* energy.

After the cooling process the temperature T' of the third qubit is given by

$$e^{-E/(kT')} = \frac{p^3 + 3p^2}{q^3 + 3q^2},$$

where numerator and denominator are the total initial probabilities for the 4 initial states with Hamming weight at least 2 and less than 2, respectively.

In general one may wish to cool k systems by transferring the heat to the remaining $n-k$ systems. For combinatorial questions of this kind it is typically (as in coding theory) easier to make statements about the asymptotics than for any finite n. According to Shannon's theorem [29] it is possible that the probability for the state $0\ldots0$ in the k rightmost systems approaches 1 with arbitrary accuracy whenever k/n is at most $1 - S$ where the entropy S is defined by $S := -p\log_2 p - (1-p)\log_2(1-p)$. Note that there exists indeed a proposal for a cooling algorithm [35] where the entropy of one part of a quantum register is transfered to the other part in order to obtain almost pure initial states in NMR computing (see Subsection 4.2.1). The scheme shows that initialization and cooling are closely connected. This suggests an implication of the second law observed by Landauer [11]: because bit erasure requires the transfer of entropy to the environment, it requires some amount of energy in analogy to the energy consumption of a refrigerator. We will elaborate on this statement in Subsection 2.2.6.

2.2.5 Computer Scientist's Heat Engine

Not only a refrigerator but also the heat engine can be run with equal gap two-level systems. Consider, for instance $2n$ two-level systems, where n have the temperature T_A and n have temperature $T_B \neq T_A$. To see that this is an energy source for sufficiently large n we consider first the object O given by the composition of one system with temperature T_A with another one with temperature T_B. It can easily be checked that O is not an equilibrium object if $\Delta E \neq 0$. Hence there is an $n \in \mathbb{N}$ such that $O^{\otimes n}$ is an energy source due to Theorem 2. The smallest number n for which this is possible depends on T_A and T_B. Whenever we have $T_A > 2T_B$ one can also construct a heat engine with $2+1$ equal systems: Then the state 110 is more likely than 001 even though its energy is twice as much. Hence the process $110 \leftrightarrow 001$ which already appeared as a refrigerator extracts some energy (see Fig 2.1). This example suggests already that there is a strong analogy between computation and work extraction. In [56] we have tried to make this analogy more precise by showing that there are molecular systems where the

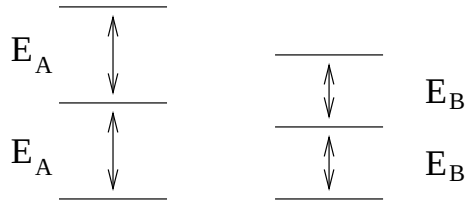


Figure 2.2: Two 3-level systems with energy gaps E_A and E_B satsifying $2E_B > E_A > E_B$. If the temperatures T_A and T_B satisfy $E_A T_B > 2E_B T_A$ the optimal heat engine implements a computationally universal transformation.

unitary U implementing an optimal heat engines is *necessarily* a transformation that can be used for computation.

One instance is a pair of 3-level systems where the unitary which implements the optimal heat engine is a transformation that is universal for classical boolean circuits. We rephrase our construction presented in [56]. We assume that system A and B have both equidistant levels $|0\rangle, |1\rangle, |2\rangle$ with energy gaps E_A and E_B, respectively. Up to an irrelevant factor the energy of a state $|n, m\rangle$ with $n, m = \{0, 1, 2\}$ is given by

$$E(n, m) = en + m$$

with $e := E_A/E_B$. The inverse logarithm of the probabilities is, up to irrelevant additive and multiplicative constants, given by

$$Q(n, m) = qn + m$$

with $q := E_A T_B/(E_B T_A)$. When e and q are not in $\{1/2, 1, 2\}$ the Hamiltonian as well as the density matrix of the bipartite system are non-degenerate and the optimal heat engine implements a unique reordering of basis states. The following choice of values e, q turns out to be useful: setting $1 < e < 2$ (as in Fig. 2.2) we induce an order on energy values of the pairs n, m which is a refinement of the degenerate order induced by $n + m$ such that for pairs with equal $n + m$ preference is given to the pair with smaller m.

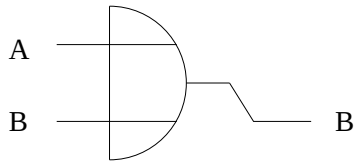
Explicitly, this is the order $00, 10, 01, 20, 11, 02, 21, 12, 22$. With $q > 2$ the probabilities are in the lexicographic order $00, 01, 02, 10, 11, 12, 20, 21, 22$. By comparing these orders one checks easily that the optimal heat engine implements the map

$$\begin{aligned} 00 &\mapsto 00 \\ 01 &\mapsto 10 \\ 02 &\mapsto 01 \\ 10 &\mapsto 20 \\ 11 &\mapsto 11 \\ 12 &\mapsto 02 \\ 20 &\mapsto 21 \\ 21 &\mapsto 12 \\ 22 &\mapsto 22\,. \end{aligned} \tag{2.3}$$

Assume we are given a collection of systems of type A and type B. If we are able to implement the heat engine above on every pair of 3-level systems consisting of one

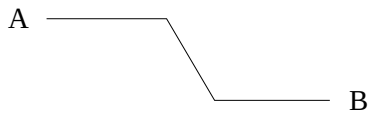
system of type A and one of type B we can also implement classical computation on the collection of these 3-level systems. In order to show this, we choose the encoding such that the logical states 0, 1 are the states $|1\rangle$ and $|2\rangle$, respectively and obtain a universal set of logical operations as follows:

1. **OR from A, B to B :**



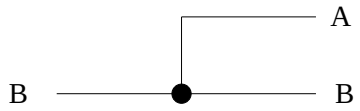
   Apply U once. One checks easily on tabular (2.3) that the second state is $|2\rangle$ if the input is one of the states $|12\rangle, |21\rangle, |22\rangle$ and $|1\rangle$ if the input is $|11\rangle$.

2. **WIRE from A to B :**



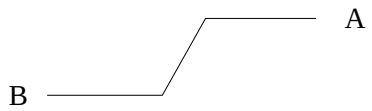
   Use our OR gate by initializing B to $|1\rangle$, i.e., the logical 0 state.

3. **FANOUT from B to A, B :**



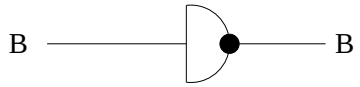
   Initialize system B to $|1\rangle$. Apply U four times. We get the mapping $12 \mapsto 20$ and $11 \mapsto 11$. The output on A coincides already with the input on B. The output on B is 1 or 0 according to whether the input on B was 1 or 2. Hence the information has already been copied to B but with the wrong encoding. For the decoding we initialize an additional system A' to the state $|1\rangle$ and apply U^4 to A', B. We get $10 \mapsto 02$ and $11 \mapsto 11$. Hence B agrees with the original input on B.

4. **WIRE from B to A :**



   Use the FANOUT.

5. **NOT from B to B :**



   Implement the first part of the FANOUT operation which changes the input state 2 to 0 and leaves 1 invariant. By initializing an additional system A' to $|2\rangle$ and

applying U once we decode and negate the information on B simultaneously : $20 \mapsto 21$ and $21 \mapsto 12$.

These operations allow obviously universal computation since every boolean function can be computed from circuits which consist only of NOR gates. We summarize :

Theorem 3 (Computing with Heat Engines on 3-level Systems)
Given two reservoirs of 3-level systems with temperature T_A and T_B and energy gap E_A and E_B, respectively, such that

$$2\frac{T_A}{E_A} < \frac{T_B}{E_B}$$

and

$$2 > \frac{E_A}{E_B} > 1\,,$$

then the ability to implement the optimal heat engine on any chosen pair consisting of one system of type A and one of type B implies the ability to implement universal classical computation on the 3-level systems.

The results above stimulate the following questions:

1. Is it possible that "algorithmic" heat engines of this kind can be implemented in future technology such that the energy *consumed* by the implementation is less than the *formal energy gain* according to Definition 9?

   The question whether this will ever be possible is strongly connected with to questions of the ultimate limits of computation: Is there a fundamental lower bound on the energy consumption for performing a logical operation? This question will be addressed in Subsection 2.2.9, even though it will not be answered there.

2. Consider $n + n$ two level systems as above with large n. How complex is the implementation of an optimal process π? In case the optima are complex, what is the trade-off between efficiency (i.e. the energy yield compared to the maximal yield) of the heat engine and the complexity of the quantum circuit?

We will discuss these questions partly in Chapter 4.

2.2.6 Classifying Thermodynamic Resources

The essential feature of the transformations considered above was that they are not energy conserving. The cooling mechanism requires energy for the implementation and the heat engine releases energy. In the first case the required energy is provided by the control fields which implement the transformations, in the second case the energy is transferred to the environment, e.g. some field mode. This can best be seen in the case of a simple two-level system: If the upper level is more likely occupied then the lower level the implementation of NOT is usually a *stimulated emission*, i.e., an emitted photon transfers the released energy to the environment. If the lower level is more likely occupied the energy required to implement the NOT gate is taken from the field which implements it. Now we present a thermodynamic setting where these energy sources

and sinks are explicitly included into the model. This could be interesting if in future technology control fields are scaled down in such a way that a microscopic description is appropriate. In particular, it is of interest whether the amount of *energy* for manipulating the microscopic world can be reduced arbitrarily. We have already mentioned one well-known bound given by the second law or Landauer's principle: Preparing certain pure quantum state from a mixed one requires energy since it is connected with the extraction of entropy. However, the amount S of entropy which has to be extracted for preparing single quantum systems is tiny (for a two-level system it is at most one bit!) hence the energy SkT is[7] negligible. For our purpose, we are more concerned with the fact that an energy source is required *at all* than about the tiny amount of energy which has to be supplied. The reason why this is important is that we show that it is not sufficient to have an energy source which can supply the required amount of energy, it is rather necessary that the source provides energy of *high quality* in a sense which is formalized by the quasi-order of resources explained below (for further details see [59]). This 'quality' depends in a sophisticated way on statistical properties of the energy source, in other words, it depends on *our knowledge* about the state of the resource system. It can be shown that *reliable* state preparation can only be driven by an energy source about which we have *reliable* knowledge in a certain sense. We consider this as a refinement of the second law of thermodynamics. To motivate this point of view we emphasize that the idea that the worth of an amount of energy depends on our knowledge about the state of the resources is one of the central statements of the Second Law: the knowledge about a system containing a certain amount of energy is minimal in its thermal equilibrium state. Accordingly, the internal energy of such a thermal heat bath is completely worthless without making use of other systems with deviations from equilibrium. This can be expressed by the quantity 'free energy': An energy source containing the energy E and the entropy S contains the free energy $E - SkT \ln 2$, i.e., the worthless part of the energy is subtracted. The amount $SkT \ln 2$ of energy would be necessary in order to bring the energy source into a perfectly known pure state. The quasi-order is a refinement of these statements in the sense that the usability of energy resources cannot be expressed by a single quantity like free energy: for the purpose of an exact state preparation it is possible that the free energy of a source may be sufficient, but other statistical properties of the source may make it worthless for achieving the desired accuracy. It is even possible that an energy source is able to drive *many unreliable* preparations but is not able to drive *a single reliable* one.

We can now formalize what it means that 'an energy source is able to drive' a preparation process and work with thermodynamical objects in the sense of Definition 8 where we drop the index of γ_T. Our model consists of three quantum systems:

- The *resource system* which is the energy source driving the process. It is initially in the mixed state ρ and therefore the object (ρ, γ). After the preparation procedure its state is closer to the thermal equilibrium state γ, because the free energy has partly been used up.

- The *environment* which can be any quantum system starting in its equilibrium state $\hat{\gamma}$. This is not a restriction of the theory, it expresses rather our point of view,

[7] where k denotes the Boltzmann constant and T is the temperature of the entropy sink.

since every system in a non-equilibrium state should be regarded as resources. The environment can be a large quantum system.[8] It is the object $(\hat{\gamma}, \hat{\gamma})$.

- The *target system* which is initially in its thermal equilibrium state $\tilde{\gamma}$ and should be driven into another state $\tilde{\rho}$. Hence it is initially the object $(\tilde{\gamma}, \tilde{\gamma})$ and should finally be $(\tilde{\rho}, \tilde{\gamma})$.

Now we define:

Definition 10 (Thermodynamic Quasi-Order)
The resource object (ρ, γ) is able to drive the preparation of $(\tilde{\rho}, \tilde{\gamma})$ in the target system if and only if there is an environment such that there is an energy conserving physical process involving the three systems described above such that the target system ends up in the state $\tilde{\rho}$. We write

$$(\rho, \gamma) \geq (\tilde{\rho}, \tilde{\gamma})$$

for short. Formally, this means the following. Let $\mathcal{H}$ and $\tilde{\mathcal{H}}$ be the Hilbert spaces of resource and target system, respectively. Then there exist Hilbert spaces $\hat{\mathcal{H}}_n$, a state $\hat{\gamma}$, and unitary transformations U_n on

$$\mathcal{H} \otimes \hat{\mathcal{H}}_n \otimes \tilde{\mathcal{H}}$$

commuting with $\gamma \otimes \hat{\gamma}_n \otimes \tilde{\gamma}$ such that

$$\lim_{n\to\infty} tr_{12}(U_n(\rho \otimes \hat{\gamma}_n \otimes \tilde{\gamma})U_n^\dagger) = \tilde{\rho}\,,$$

where tr_{12} denotes the partial trace over the first and the second system.

It is easy to check that the condition that U is energy conserving is equivalent to $[U, \gamma \otimes \hat{\gamma} \otimes \tilde{\gamma}] = 0$. One may wonder why γ appears in the notation although it appears neither in the initial nor in the final state of the composed system. Actually γ appears in a subtle way since the term 'energy conserving' refers to the Hamiltonian of the resource system which is (up to an irrelevant constant) determined by γ.

Due to the conservation of free energy, the final amount of free energy of the target system can never exceed the initial amount of the resource system. This gives a necessary condition for $(\rho, \gamma) \geq (\tilde{\rho}, \tilde{\gamma})$ which is well-known in standard thermodynamics. But this condition is by no means sufficient. Under the special assumption that $\tilde{\rho}$ or ρ are time-invariant states we can characterize the quasi-order in a quite explicit way [59]. When ρ and γ commute we characterize the object simply by the vectors p, g of eigenvalues of ρ and γ, respectively and denote the object by (p, g).

Theorem 4 (Explicit Description of Quasi-Order)
Let (p, g) and $(\tilde{p}, \tilde{g})$ be resource systems. Then

$$(p, g) \geq (\tilde{p}, \tilde{g})$$

if and only if there exists a stochastic matrix A such that

$$Ap = \tilde{p} \quad \textit{and} \quad Ag = \tilde{g}\,.$$

[8] It might possess an infinite dimensional Hilbert space. In our formalism, we work therefore with a sequence of finite dimensional Hilbert spaces.

In the limit of infinite temperature, g is the uniform distribution on the set of energy eigenstates and if g and $\tilde{g}$ have equal dimension the condition $Ag = \tilde{g}$ is simply the condition that A is a double stochastic matrix. Then $(p, g) \geq (\tilde{p}, \tilde{g})$ if and only if p *majorizes* $\tilde{p}$. We say that the vector p majorizes $\tilde{p}$ if

$$\sum_{j=1}^{l} p_j \geq \sum_{j=1}^{l} \tilde{p}_j \quad \forall j \leq n$$

whenever $p_1, \ldots, p_n$ and $\tilde{p}_1, \ldots, \tilde{p}_n$ are the entries of p and $\tilde{p}$, respectively, in non-decreasing order. Note that the density matrix γ corresponding to a vector g with uniform distribution is the maximally mixed state which commutes with every other matrix. Therefore we can diagonalize ρ and γ always simultaneously for this case. Hence we have an explicit description for the quasi-order:

Corollary 1 (Resource Order by Majorization)
Let resource and target system both have d dimensional Hilbert spaces. Then

$$(\rho, \mathbf{1}/d) \geq (\tilde{\rho}, \mathbf{1}/d)$$

if and only if the vector of eigenvalues of ρ majorizes the vector of eigenvalues of $\tilde{\rho}$.

Note that the spectral majorization criterion above is equivalent to stating that $\tilde{\rho}$ is in the convex span of unitary conjugates of ρ [60]. Then the only possible operations on resources are a deterministic or a random choice of unitary transformations, i.e., the environment can only serve as an entropy source but never as an entropy sink. The Hamiltonian of the system looses its relevance and the worth of resources are only determined by the eigenvalues of the density matrix. Then our theory of resources coincides with the manner in which M. Horodecki, P. Horodecki, and J. Oppenheim [61] classify thermodynamic resources: in their setting all ancillas in maximally mixed states are free, which is exactly the Gibbs state to infinite temperature or completely degenerate Hamiltonian. However, this limiting case cannot consistently describe infinite dimensional systems.

An important observation is that the quasi-order is not a linear order, i.e., we do not have necessarily $(\rho, \gamma) \geq (\tilde{\rho}, \tilde{\gamma})$ or $(\tilde{\rho}, \tilde{\gamma}) \geq (\rho, \gamma)$. Mathematically this can most easily be seen in the limit $\gamma = \mathbf{1}/d$. Chose two density matrices $\rho, \tilde{\rho}$. In 3 dimensions they can certainly be chosen such that they have the same von-Neumann entropy but different spectrum. If one would majorize the other its entropy would be strictly lower.

An interesting feature of the quasi-order is that one can compare also systems with different Hilbert spaces and different Hamiltonians. We want, for instance, to recover Landauer's principle [11] stating that the initialization of one bit requires the energy $\ln 2kT$, where k is Boltzmann's constant and T the reference temperature of the environment which absorbs the entropy. In order to simplify the discussion we represent the bit by a degenerated two-level system. The initialized bit is hence the object

$$\tilde{O} := (|0\rangle\langle 0|, \mathbf{1}/2)\,.$$

First we have to formally introduce free energy. We define it such that it is zero for equilibrium states.

Definition 11 (Free Energy)
The free energy of an object is given by

$$F(\rho) := tr(\rho H) - S(\rho)\ln 2\, kT - tr(\gamma H) + S(\gamma)\ln 2\, kT\,.$$

It can be written in terms of the Kullback-Leibler relative entropy between ρ and γ:

$$F(\rho) = kT\,\ln 2\, K(\rho||\gamma) := kT(tr(\gamma\log_2\gamma) - tr(\rho\log_2\gamma))\,.$$

The free energy of the initialized bit is $\ln 2\, kT$, i.e., its information in natural units multiplied by kT. Note furthermore that we have monotonicity of the quasi-order with respect to relative entropy distance:

Theorem 5 (Monotonicity of Relative Entropy)
Conversion of resources can never increase the relative entropy distance of the actual state from its equilibrium, i.e.,

$$(\rho,\gamma) \geq (\tilde{\rho},\tilde{\gamma})$$

implies

$$K(\rho||\gamma) \geq K(\tilde{\rho}||\tilde{\gamma})\,. \tag{2.4}$$

and

$$K(\tilde{\gamma}||\tilde{\rho}) \geq K(\gamma||\rho)\,. \tag{2.5}$$

The proof follows directly from the fact that the operation which converts ρ to $\tilde{\rho}$ would convert γ to $\tilde{\gamma}$, since equilibrium states can only create equilibrium states. Then the statement follows because such operations cannot increase the distance between ρ and γ [62]. We conclude:

Corollary 2 (Landauer's Principle)
Let

$$(|0\rangle\langle 0|,\, \mathbf{1}/2)$$

be the formal representation of an initialized bit. Every resource (ρ,γ) with

$$(\rho,\gamma) \geq (|0\rangle\langle 0|,\, \mathbf{1}/2)$$

satisfies

$$K(\rho||\gamma) \geq K(|0\rangle\langle 0|\,||\,\mathbf{1}/2)\,,$$

i.e. $F(\rho) \geq \ln 2\, kT$.

Recalling that we have here derived Landauer's principle from the monotonicity of $K(\rho||\gamma)$ according to ineq. (2.4) one may ask the implications of the monotonicity of $K(\gamma||\rho)$ according to ineq. (2.5). We derive a tightened version of Landauer's principle:

Corollary 3 (Tightening the Second Law) *There are objects $O, \tilde{O}$ where O has more free energy than $\tilde{O}$ but nevertheless*

$$O \not\geq \tilde{O}\,.$$

Another necessary condition is given by

$$K(\gamma||\rho) \geq K(\tilde{\gamma}||\rho)\,.$$

The latter condition is remarkable since the distance diverges for non-invertible ρ. This implies that one can never obtain resources with singular density matrices from resources with invertible density matrices. In [59] we argued that a natural energy source supplies always non-singular density matrices. This can, for instance, be a Gibbs state with higher temperature. With such an object one can never drive the initialization of a bit *reliably*. This is only possible in the limit of infinite copies:

Theorem 6 (Tightening Landauer's Principle)
Let O be an object with non-singular density matrix and

$$O_\epsilon := (\epsilon|1\rangle\langle 1| + (1-\epsilon)|0\rangle\langle 0|,\ \mathbf{1}/2)$$

be a bit which is initialized with probability $1-\epsilon$. Then one needs increasing number of copies of O to prepare $\tilde{O}_\epsilon$, i.e.,

$$O^{\otimes n} \geq \tilde{O}_\epsilon$$

holds only for $n \to \infty$ if ϵ tends to 0.

Actually there is an extremely simple idea behind the tightened version of Landauer's principle: For any two non-singular matrices ρ and γ there is no measurement that can reliably distinguish between them. Every rule specifying how we could estimate whether ρ or γ is the actual state could be formalized by a POVM with two operators P_ρ and P_γ. It is easy to see that no estimation which is neither always ρ or always γ can avoid errors of both kinds: estimating ρ when it was γ and vice versa. Assume (ρ, γ) could reliably drive the bit initialization. Then we could check the bit, if its value is 1 the resource system cannot have been in the state ρ. This would be a non-trivial decision rule with only one kind of error: one could take the state γ for the state ρ but not vice versa. We found [59]:

Theorem 7 (Resources for Reliable Bit Erasure)
The object $O = (\rho, \gamma)$ can drive the preparation of the state $(1-\epsilon)|0\rangle\langle 0| + \epsilon|1\rangle\langle 1|$ if and only if there is a measurement with outcomes ρ, γ described by a POVM (M_ρ, M_γ) such that $[\gamma, M_\rho] = 0$ with the error probabilities

$$F_1 := tr(\gamma M_\rho) = 1/2\,,$$

and

$$F_2 := tr(\rho M_\gamma) = \epsilon\,.$$

One direction of the proof is already given by the arguments above: The preparation procedure followed by a readout would provide a POVM with exactly these error probabilities. For the proof of the converse we refer to [59]. The same arguments as above hold for perfect cooling of non-degenerate two-level system:

Theorem 8 (Resource Requirements for Low Temperatures)
The object $O = (\rho, \gamma)$ can drive the cooling of a two-level system with energy gap $\tilde{E}$ down to the temperature $\tilde{T}$ if and only if there is a measurement with outcomes ρ, γ described by a POVM (M_ρ, M_γ) such that $[\gamma, M_\rho] = 0$ with the error probabilities

$$F_1 := tr(\gamma M_\rho) = \frac{e^{-\tilde{E}/(kT)}}{1 + e^{-\tilde{E}/(kT)}}$$

and

$$F_2 := tr(\rho M_\gamma) = \frac{1}{1 + e^{-\tilde{E}/(k\tilde{T})}} .$$

The theorem is a quantitative basis for the intuition that the generation of very low temperatures is extremely costly.

The hypothesis testing argument (see [31] for quantum estimation theory) explained here might be called the principle that 'distinguishability can never increase'. It shows that much more sophisticated information theoretic quantities than entropy play also a fundamental role in thermodynamics. But apart from this, one should expect that the worth of resources is given by quantities which do not have necessarily any obvious information theoretic meaning. Assume for instance that the resource state is time-invariant. Then it can never drive the preparation of a non-time-invariant state. This is formulated quantitatively in the quasi-order of clocks in Subsection 2.2.7.

The practical relevance of the quasi-order lies rather in stating which processes are not possible than to say which are possible. Even though the preparation of $\tilde{\rho}$ given the resource ρ would not require energy "in principle" in the sense of our setting whenever $(\rho, \gamma) \geq (\tilde{\rho}, \tilde{\gamma})$, it seems to be far from currently feasible to implement the resource conversion using only a negligible amount of energy. However, such an implementation would be given if the process was implemented by the autonomous time evolution of resource system + ancilla system + target system. As we have seen in Subsection 2.2.5 heat engines may require quite a few non-trivial logical operations. The problem of finding an interaction which implements transformations like those could be as hard as finding an interaction for autonomous universal quantum computing as needed for the ergodic quantum computer (see Subsection 2.2.9).

2.2.7 Timing Information as a Thermodynamic Quantity

Thermodynamics describes the interplay between *energy* and *information* and how this determines fundamental limits to physical processes. In our quasi-order of resources all systems in the Gibbs-state are free. One could be more generous and state that all systems with stationary states are free, i.e., all systems with $[\rho, H] = 0$ instead of $\rho = \exp(-H/(kT))$. Then the value of a resource lies alone in its dynamics, i.e., the off-diagonal elements with respect to the energy eigenbasis makes the resource valuable. The extent to which the states $\exp(-iHt)\rho\exp(iHt)$ are distinguishable for different t determines the worth of this Hamiltonian system when it is used as a *clock*. This is the main idea of the quasi-order of clocks introduced and investigated in [63]. In this setting, we characterize a quantum 'clock' abstractly by the pair (ρ, H), where ρ is the system's density matrix and H its Hamiltonian. We define formally:

Definition 12 (Quasi-Order of Clocks)
The clock (ρ, H) is able to prepare the clock $(\tilde{\rho}, \tilde{H})$), formally writing

$$(\rho, H) \geq (\tilde{\rho}, \tilde{H})$$

if there is a completely positive trace preserving map G such that $G(\rho) = \tilde{\rho}$ and G satisfies the covariance condition

$$G \circ \alpha_t = \tilde{\alpha}_t \circ G \quad \textit{for all } t \in \mathbb{R}$$

with the abbreviation

$$\alpha_t(\sigma) := \exp(-iHt)\,\sigma\,\exp(iHt)$$

for each density matrix σ (and $\tilde{\alpha}_t$ defined similarly).

We say that two clocks are equivalent if each of both majorizes the other in the sense of this quasi-ordering. The equivalence class of trivial (or 'worthless') clocks is given by all those clocks with $[\rho, H] = 0$. They are majorized by all the other clocks.

We illustrate the intuition of this quasi-order in the following protocol: assume Alice sends a clock (ρ, H) to Bob. Assume this system for the moment to be a spin 3/2 system rotating about its z-axis. Carol asks Bob to send her a clock showing as much about Alice's time as possible, but she complains: 'please do not send the spin 3/2 system, I am not able to read it out.' Since she can only deal with spin 1/2 particles, she asks Bob to put as much information as possible about Alice's time into a spin-1/2-particle. Bob's possible transformations are mathematically described by a completely positive trace preserving map G from the set of density matrices of spin 3/2-systems to the set of density matrices of spin-1/2-systems. To illustrate why G has to satisfy the covariance condition note that the state of the system received by Carol is given by

$$(\tilde{\alpha}_{t-\tau} \circ G \circ \alpha_\tau)(\rho)$$

if the time t has passed since the preparation of the resource state ρ and Bob ran the conversion process at time τ. Since Bob is not able to run the process at a well-defined time τ, he applies a mixture over all $0 \leq \tau \leq t$, whether he wants to or not. If we assume for simplicity that t is large compared to the recurrence time of the quantum dynamics, this mixture can equivalently be described by a time covariant map $\overline{G}$.

As already stated, the thermo-dynamical quasi-order is stronger than the quasi-order of clocks, i.e.,

$$(\rho, H) \geq_{thermodyn} (\tilde{\rho}, \tilde{H}) \quad \text{implies} \quad (\rho, H) \geq_{clock} (\tilde{\rho}, \tilde{H})\,.$$

The reason that the covariance condition is also a *necessary* condition (even though it is not sufficient) in the thermodynamic quasi-order is essentially the energy-time uncertainty principle: If a process is started by a negligible amount of switching energy it cannot be well-localized in time, i.e., the process cannot be started without using a clock.

One can easily extend the quasi-order in such a way that it also includes classical systems since the setting of operator-algebraic quantum theory [64] allows a unified description of quantum and classical aspects [63]. The observable algebras are abelian C^*-algebras for classical systems and non-abelian for quantum systems. The state of a classical clock is a moving probability distribution in its phase space[9]. The simplest example would be a probability distribution on the unit circle which rotates with constant angular velocity. The time evolution is in both cases a C^*-automorphism group. The completely positive operation which generates the state of the target clock from the state of the resource clock satisfies a covariance condition with respect to these automorphism groups.

The quasi-order of clocks (like the thermodynamic quasi-order) is not a linear order. This reflects the fact that timing information has many aspects, here we only mention

[9] which is the Gelfand space [65] of the abelian algebra

a few of them. First, the *periodicity* of two systems is essential. If the dynamics of one system has period 1 and the other has period 2 the first cannot majorize the other in the quasi-order since the latter distinguishes between the time instants 0 and 1. Furthermore, the *time resolution* matters. If ρ_t and $\rho_{t+\Delta t}$ act on mutually orthogonal subspaces whereas $\tilde{\rho}_t$ and $\tilde{\rho}_{t+\Delta t}$ have large overlap, the latter clock can certainly not majorize the former. Another interesting question about timing information is whether it is quantum or classical information. This is a deep question which will be addressed later. In the setting above, purely classical timing information exists since the setting includes propagating probability measures, i.e., the dynamics is a flow in a classical phase space. This is however, the limit of infinite energy since this dynamics a point measure moves to another point measure, perfectly distinguishable from the first one. Now we describe in which situations our quasi-order puts constraints on many physical processes. For the formal proofs in the C^*-algebraic setting we refer to [63].

1. **Time Transfer between Classical Clocks:** In this case the quasi-order is easy to understand. For simplicity we assume that both are probability measures $\mu, \tilde{\mu}$ on the unit circle Γ. The time evolution of these measures is generated by the rotation

   $$z \mapsto e^{-i\omega t}\, z\,.$$

   Then the first clock majorizes the second if and only if there is a measure ν on Γ such that

   $$\tilde{\mu} = \mu * \nu$$

   where $*$ denotes the convolution. The intuition is that the transfer of timing information from one clock to another may only add some independent error to the uncertainty of the first.

2. **Time Transfer from Classical to Quantum Clocks:** One can easily argue [63] that all such transfer processes are given by preparing a state

   $$\alpha_t(\rho)$$

   whenever the classical clock shows the time t. It is clear that this can be done without disturbing the classical clock and can therefore be used to generate arbitrarily many copies of quantum clocks.

3. **Time Transfer from Quantum to Classical Clocks:** Assume we have a quantum clock (ρ, α) rotating with the recurrence time 1, i.e., $\alpha_1(\rho) = \rho$ and we want to transfer its timing information on a classical clock with the same period. Consider for instance the classical clock mentioned above with $\omega = 2\pi$. Time transfer is done by *measuring* the quantum system and *initializing* the classical pointer to the position $\exp(-i2\pi t') \in \Gamma$ according to the estimated time t'. The set of possible values t' in this estimation might be discrete or continuous and a priori there is no restriction to the positive operator valued measure describing the possible measurements. Nevertheless one can show that each clock initialization can equivalently be performed by a time covariant measurement [63]. The classical clock generated by the covariant measurement $(M_t)_{t\in[0,1]}$ is given by the probability measure μ on Γ with density

   $$p(t) := tr(\rho M_t)\,.$$

It is clear that the transfer of timing information from quantum to classical clocks is irreversible. If one could obtain a quantum clock which is equivalent to the original one from the classical clock one could also get many copies of it, i.e., one could clone the quantum clock. The task of copying timing information is related to the problem of cloning quantum states with unknown phase [66, 67]. An example is the qubit state $|0\rangle + \exp(-i\omega t)|1\rangle$ with unknown t.

4. **Time Transfer from Quantum to Quantum Clocks:** In [63] we have shown, as an example, that the quasi-order helps to derive fidelity bounds when an optical mode in a coherent state drives the preparation of a superposition state $|+\rangle := (1/\sqrt{2})(|0\rangle + |1\rangle)$ in a two-level system. The coherent state does not, strictly speaking, majorize the pure superposition state in the quasi-order sense and one can only achieve increasing reliability with increasing field strength. We have argued in [63] that this is not just a consequence of phase uncertainty for finite field energy since superpositions of finite photon number states exist that allow the preparation of a pure $|+\rangle$ state.

5. **Copying Timing Information:** The following result gives some insight on the 'quantumness' of timing information. Given a quantum clock we would like to transfer as much timing information as possible to two other quantum clocks, whereas the timing information of the original clock may be destroyed. We quantify the timing information by the quantum Fisher information F. Using the quasi-order we proved the following bound [63]: Let the resource clock have initially the timing information F. Then for the timing information F_1, F_2 of the two target clocks we find:

$$\frac{1}{F_1} + \frac{1}{F_2} \geq \frac{2}{F} + \frac{1}{\langle E^2 \rangle}\,,$$

where $\langle E^2 \rangle$ denotes the expectation value of the square of the total energy of the two target clocks. This shows that the Fisher information F_1 and F_2 of the target clocks cannot both be as large as F. The bound allows this only in the limit of infinite $\langle E^2 \rangle$. This shows that perfect cloning of timing information is only possible in the limit of clocks with infinite energy. Of course such a quantum limitation is not specific to *timing* information, but time is one of the most natural applications where the estimation of a parameter by measurements on a continuous family of states is necessary.

A rather explicit characterization of the set of all covariant maps G mapping states on a Hilbert space $\mathcal{H}$ on states on $\mathcal{H}$ is given in [68]. This can help to better understand the quasi-order of clocks. The decomposition derived holds also for infinite dimensional Hamiltonian systems. Here we rephrase the finite dimensional case with non-degenerate Hamiltonian. A covariant map G can be written as a sum over CP maps G_σ where each G_σ is a concatenation of a *dephasing* map and an *energy shift*, i.e., CP map that increases or decreases the energy of the system by the amount $\sigma \in R$. The dephasing is described by the multiplication of the density operator (given with respect to the energy basis) with a positive operator. The Kraus operator of the energy shift is a partial isometry that generalizes the canonical shift on square-integrable functions on $\mathbb{R}$ to measure spaces with non-translational invariant measures. The fact that these operations can only destroy or

translate superpositions between different energy eigenstates shows that the classification of clocks with respect to their timing information considers coherent superpositions of energy states as the essential resource.

The most important drawback of the quasi-order above is probably that *absolute* timing information with respect to a unique reference clock is less relevant in applications than *relative* timing information. To put relative timing information between two parties, say Alice and Bob, into an analogous formal setting we define a *synchronism of clocks* A and B as a correlated joint state ρ which is invariant with respect to the joint evolution $\alpha_t \otimes \beta_t$ and not invariant with respect to $\alpha_t \otimes \beta_{-t}$, where α_t and β_t acts on Alice's and Bob's clock. The fact that the joint state is $\alpha_t \otimes \beta_t$-invariant formalizes that we consider only relative timing information, i.e., the state is seen from the point of view of an observer without any knowledge on the actual time. Consider the case that Alice and Bob share two-level systems with the same frequency ω. Both states are in a superposition state with the same phase:

$$\frac{1}{2}(|0\rangle + e^{-\omega t}|1\rangle) \otimes (|0\rangle + e^{-\omega t}|1\rangle)\,.$$

For the external observer they share a mixed state obtained by mixing this states over all t.

We denote a synchronism as defined above by the triple (ρ, α, β) and define:

Definition 13 (Quasi-Order of Synchronisms)
The bipartite system $A \times B$ is said to be better than or equally synchronized as the system $\tilde{A} \times \tilde{B}$, formally denoted by

$$(\rho, \alpha, \beta) \geq (\tilde{\rho}, \tilde{\alpha}, \tilde{\beta})$$

if there is a process G converting ρ to $\tilde{\rho}$ satisfying the covariance condition

$$G \circ (\alpha^{-1} \otimes \beta) = (\tilde{\alpha}^{-1} \otimes \tilde{\beta}) \circ G\,.$$

The covariance condition formalizes the requirement that Alice and Bob are not allowed to use additional synchronized clocks in order to run the process G. Of course *relative* timing information can be created in a synchronization scheme whereas absolute timing information is a fixed resource.

2.2.8 Coherence, Reversibility and Secrecy of Computation

As we have already mentioned in Subsection 2.2.1 there is roughly speaking a strong connection between the tasks of preserving coherence during a computation and of performing computation in a thermodynamically reversible way. In both cases one has to ensure that no information is copied to the environment. Consider for instance the following model where the logical state of a bit is represented by a qubit and the environment by a second qubit. Let the interaction between bit and environment implement a controlled not with the environment qubit as target (see Fig. 2.3).
Let the logical bit be encoded in the basis states of the qubit, with initial state γ. If the environment is in either of the states $\rho_0 := |0\rangle\langle 0|$ or $\rho_1 = |1\rangle\langle 1|$ the CNOT copies the logical state $\sigma_0 := |0\rangle\langle 0|$ or $\sigma_1 := |1\rangle\langle 1|$ to the environment since the logical state of the

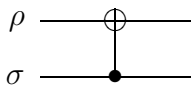


Figure 2.3: When is a bit a qubit? Toy model for the interaction of a qubit with its environment. The environment is also modeled by a single qubit. The interaction destroys alway coherence. Whether information is transferred to the environment depends on the initial state ρ of the latter.

environment is inverted whenever the qubit was in the state 1 and not inverted if it is in the state 0. It therefore destroys the superpositions in the qubit. This is also true if the environment starts in an equally weighted mixture of $|0\rangle\langle 0|$ and $|1\rangle\langle 1|$. The argument showing that superpositions are destroyed works as well since it is irrelevant whether ρ_1 or ρ_2 is present. On the other hand, the environment does not get any information about the classical state: it remains in its maximally mixed state, no matter whether the register was in state σ_0 or σ_1. This is like running a copy machine with black paper. The document which is to be copied is treated as if a real copy would be produced. But the copy is useless because the 'initial state of the paper' was bad. This toy model shows that there is no obvious reason for the statement that thermodynamically reversible computation should always keep quantum superpositions alive: The computation may leave the classical state unaffected but would nevertheless destroy superpositions if the input was in such a superposition state. Therefore the connection between physical reversibility and coherence of a computation is less tight than it may seem at first sight. However, the model above does not provide an example for a computation which runs in an energetically closed system. The operation is here triggered by an external signal which is not taken into account as a thermodynamical resource. The question to what extent thermodynamically reversible computation requires quantum coherence when *is driven by a Hamiltonian of a closed physical system* seems more difficult to answer. We will address it in Subsection 2.2.9.

We also want to mention a tight quantitative connection between thermodynamic reversibility and secrecy of a computation. Whereas cryptographic security refers usually more to coherence than to thermodynamic reversibility, there is an interesting sense in which secrecy is also connected with reversibility even though the remarks below may be of no practical interest for cryptography.

To explain the connection we consider an 1-bit computer[10] on which we implement a NOT gate. Let γ be the environment in its thermal equilibrium state for the temperature T. Let U be a unitary operation acting on $\gamma \otimes \sigma$ to implementing the NOT. The reduced state of the environment is

$$\rho_j := tr_2(U(\gamma \otimes \sigma_j)U^\dagger) \quad j = 0, 1\,.$$

The reduced state of the logical bit is

$$\sigma_{j\oplus 1}\,.$$

The Holevo information (see Section 1.3) of the environment about the logical state is

$$I := S(\frac{1}{2}(\rho_0 + \rho_1)) - \frac{1}{2}(S(\rho_0) + S(\rho_1))\,.$$

[10]We prefer to use 'bit' instead of qubit because it is a priori a bit. The fact that it is represented by a qubit is only due to its quantum mechanical description.

Since γ is the state with the least free energy, the average free energy of ρ_j is greater than the free energy of γ by at least the amount I due to

$$\frac{1}{2}(F(\rho_0) + F(\rho_1)) \geq I + F(\frac{1}{2}(\rho_0 + \rho_1)) \geq F(\gamma)\,.$$

Since the average free energy of the bit cannot have changed (because it is still either of the states σ_j with probability 1/2), the operation U has required at least the energy I. Of course the same argument works also for registers with several qubits and other reversible logical functions as NOT. We conclude:

Theorem 9 (Reversible Computation is Secret)
A logical operation which requires at most the average energy E has at most transferred the information $E/(\ln 2\, kT)$ to the environment.

2.2.9 Computing and Quantum Control in a Closed Physical System

As we have already mentioned, information processing necessarily consumes some energy whenever it involves logically irreversible operations. Furthermore it has been clarified [10, 69, 70] that computing is possible with logically reversible operations[11] and that the complexity overhead seems acceptable [71].

However, logically reversible operations are not necessarily implemented by thermodynamically reversible devices. Therefore it has been asked whether computers could exist which do not 'consume free energy' at all, i.e., do not generate entropy. It is clear that each bijective function f on $\{0,1\}^k$ defines in a canonical way a unitary operation on k qubits by the linear extension of $|b\rangle \mapsto |f(b)\rangle$. An important example is the TOFFOLI gate which allows universal classical computation. Unfortunately this does not show that the *whole computation process* can be thermodynamically reversible because the physical system which controls the implementation of the circuit has not yet been taken into account. We will therefore describe those problems arising in models of computers where the computation is the dynamics of a closed physical system and recall solutions in the literature as well as own contributions.

Clearly one can introduce a quantum system which controls the implementation of these reversible gates. A simple model of such a 'clock' is the wave packet of a particle moving in one spatial dimension. Its Hilbert space is $L^2(\mathbb{R})$, i.e., the square integrable functions. It would be convenient to have the time evolution

$$(S_t\psi)(x) = \psi(x+t)\,,$$

i.e., a simple shift. Then the wave packet moves without dispersion, i.e., given an initial state with compact support the support will remain compact. The interaction between this clock and the devices to be controlled could be such that it triggers the implementation of a certain operation j whenever the wave packet reaches a certain region R_j. This is one of the ideas behind Benioff's construction of a Hamiltonian computer [72].

[11] See also the simulation of NAND by TOFFOLI gates in Subsection 3.2.1

Unfortunately the spectrum of the Hamiltonian which generates the shift is unbounded from below and above since it is given by

$$H := \frac{1}{i}\frac{d}{dx}\,.$$

Since the spectrum of physical Hamiltonians are bounded from below the situation above can only be approached in the limit that the system energy is far above the lower bound such that it becomes irrelevant. The wave packet with compact support is therefore an infinite-energy limit, whereas the realistic physical Hamiltonian $-d^2/(dx)^2$ would lead to a dispersing wave packet. This would then lead to superpositions of computation states because different operations ('computation steps') j are performed simultaneously.

Feynman [73, 74] considered a discrete clock which is a linear chain of many qubits where the clock states are binary words with Hamming weight 1 and a symbol 1 at position j is the jth clock state. Before we explain Feynman's Hamiltonian we first discuss a naive construction of a clock in order to motivate his approach. To simplify mathematics we consider a cyclic chain with n qubits. One would like a Hamiltonian H which implements the cyclic shift

$$S|b_0,\ldots,b_{n-1}\rangle := |b_{n-1},b_0,\ldots,b_{n-2}\rangle$$

in the sense that $S = \exp(-iH)$. The simplest way to construct H is to choose a polynomial p which solves the interpolation problem

$$p(e^{-i2\pi j/n}) = 2\pi j/n \quad j = 0,\ldots,n-1\,.$$

Then $H := p(S)$ satisfies $\exp(-iH) = S$ by usual functional calculus of operators [65] because $\exp(-i2\pi j/n)$ are the eigenvalues of S. Unfortunately H contains all powers of S and involves rather unphysical interactions between long-distant qubits. Feynman preferred to choose a more physical Hamiltonian H involving only interactions between adjacent qubits:

$$H := \sum_{j=0}^{n-1} a_j a_{j+1}^\dagger + a_j^\dagger a_{j+1}\,,$$

where a_j is the annihilation operator $a := |0\rangle\langle 1|$ acting on qubit j. The restriction of H to the subspace spanned by basis states with Hamming weight 1 is the operator

$$H_r := S + S^\dagger\,,$$

where S is the cyclic shift on $\mathbb{C}^n$. The problem is that the dynamics generated by H_r leads to a spreading wave packet [75]. Hence the computer controlled by this clock is in a superposition between different computation steps. If one starts the clock in a basis state there would even part of the wave propagating backwards. Feynman argued that one can prepare superpositions of basis states such that the spreading is not too large and that the wave is mainly propagating in forward direction.

Margolus [76] constructed a cellular automaton which is implemented by a Hamiltonian that is strongly reminiscent of solid-state interactions due to its lattice-symmetric structure. Here we explain his two-dimensional version. The whole data register which

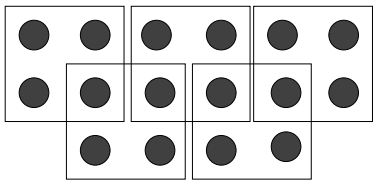


Figure 2.4: System of neighborhoods of the Margolus cellular automaton. The circles indicate the cells and the squares are the neighborhoods on which the updates are implemented.

carries the information to be processed is divided into a two-dimensional array of cells $(i, j) \in \mathbb{Z} \times \mathbb{Z}$, where each cell consists of several data qubits. First one thinks of a computation that consists of updating all cells in rows $j, j+1$ in time step j. The update rules consist of independent operations involving 4 cells. They are given by $(i, j), (i, j+1), (i+1, j), (i+1, j+1)$ for all i, j where $i + j$ is even. This system of overlapping neighborhoods is depicted in Fig. 2.4

The update rules are such that they propagate the information of the lower two cells in a neighborhood to the upper two cells after the two-qubit state of the lower ones has been subjected to a logical transformation which allows universal computation.[12]. Margolus observed that this update scheme does not really require global clocking; it only has to be ensured that each neighborhood is updated only if the two neighborhoods which intersect the lower cells have already been updated. It is like building a wall with bricks where each brick in row k can only be added if both bricks in row $k-1$ where it has to lie on are already there. Based on this observation, his local synchronization uses one additional clock qubit in each cell. The allowed states of the clock are given by those states where in each row exactly one clock qubit is in state $|1\rangle$. The local synchronization condition imposes (in analogy to the growth of a wall) that the positions of the symbol 1 in two adjacent columns can differ at most by one row. The Hamiltonian of the clock reads

$$H_C := \sum a_{i,j} a_{i+1,j} a^\dagger_{i+1,j} a^\dagger_{i+1,j+1} + h.c.\,, \tag{2.6}$$

where h.c. denotes the adjoint ('hermitian conjugate') of the first term. This term ensures that H is self-adjoint. The sum runs over all $i, j \in \mathbb{Z}$ with $i + j$ even and $a_{i,j}$ denotes the annihilation operator $|0\rangle\langle 1|$ on cell i, j. Whenever there are two symbols 1 in the lower two cells they are both propagated to the upper cells. Furthermore, the full Hamiltonian is defined such that this propagation of the 'clock wave front' is accompanied by an update of the neighborhood. In analogy to the Hamiltonian of the Feynman model one also has the problem that the dynamics leads to a clock wave which is a superposition of many wave front positions. Starting with a wave front which is localized in a well-defined row, part of the wave will also run backwards and trigger a computation in the wrong direction. However, Margolus argues that the state of the clock wave can (as in Feynman's proposal) be prepared such that it propagates mainly forward. These superpositions of different wave front positions are in general entangled states. Furthermore one has to read out the computation result during the time period where the clock wave is in the correct region of the clock register. If the clock register is finite the clock wave will be reflected and

[12]Note that already a lot of computationally universal reversible cellular automata are known [77, 78, 79, 80].

if the topology of the register is cyclic the computation will be repeated. A computer whose calculation stops would therefore need an infinite clock register. From the point of view of thermodynamics this is a problematic point: an initialized register with an infinity of mutually orthogonal states needs infinite resources since this corresponds to the erasure of infinitely many bits. This can be formulated as a general paradox in the thermodynamics of computation:

Lemma 2 (Computers with Finite Free Energy can Never Stop)
Let ρ be the initial state of a computer and the computation be driven by the Hamiltonian H on some Hilbert space $\mathcal{H}$. Let the computation be finished after the time t_f and all states $\alpha_t(\rho)$ for $t \geq t_f$ be perfectly distinguishable from ρ, i.e., the density operators act on mutually orthogonal subspaces of $\mathcal{H}$. Then the free energy of ρ is infinite.

Proof: It is clear that for all $n, m \in \mathbb{Z}$ with $n \neq m$

$$\alpha_{nt_f}(\rho) \perp \alpha_{mt_f}(\rho)\,, \tag{2.7}$$

where $\perp$ denotes orthogonality of density matrices, since

$$\rho \perp \alpha_{(n-m)t_f}(\rho)$$

by assumption. Let p be a probability measure on $\mathbb{Z}$. Define ρ_p by the mixture

$$\rho_p := \sum_j p_j \alpha_{tj}(\rho)\,.$$

Due to the orthogonality (2.7) the von-Neumann entropy of ρ_p satisfies

$$S(\rho_p) = S(\rho) + S(p)$$

where $S(p)$ is the Shannon entropy of p. For the free energy (see Definition 11) we find

$$F(\rho) = F(\rho_p) + S(p)$$

because the energy $tr(H\alpha_t(\rho))$ is the same for all t. The free energy of ρ_p is at least the free energy of the thermal equilibrium state (which was defined to be zero). Since the entropy of p can be chosen infinitely large the lemma follows. □

It is clear that the practical relevance of this result is limited. One will be satisfied if the initial state and the final state are only distinguishable with very high probability. Consider for instance a two-level system with energy gap $E \gg kT$ and assume that it is initialized to its excited state $|1\rangle$. Due to its coupling to the environment, usually to some field modes, one will observe relaxation, i.e., the system goes back to $|0\rangle$ with high probability. If t_f is chosen sufficiently large compared to the life time of the excited state the density matrix of the two-level system together after the time t_f is *almost* orthogonal to its initial density matrix. Nevertheless it should, for physical reasons, be possible to describe the decay by an appropriate Hamiltonian system including the field modes such that the free energy of the composed system is finite. It is important to note that the approximative statements like *almost orthogonal* may hold up to an error rate which

decreases exponentially in the amount of available energy. For instance, the probability to find a two-level system still in its initial state $|1\rangle$ after thermalization is exponentially small for increasing E. This toy model of a 'primitive computation' which switches both states $|1\rangle, |0\rangle$ to $|0\rangle$ shows that finite energy allows rather reliable computation despite Lemma 2.

Nevertheless we found it unsatisfactory that the Feynman and Margolus computer needed to be read out during a certain time interval. In order to determine the resource requirements of computation we wanted to avoid this since timing information is also a thermodynamic resource (see Subsection 2.2.7). Furthermore we did not want to start with *entangled* initial states of the clock since this already requires the solution of a non-trivial quantum control problem with unknown resource requirements. Therefore we modified Margolus' model of a Hamiltonian computer such that the computation result needs not to be read out during a certain time interval. It can rather be read out at a random time instant since the result is encoded in the time average of the state of the system. As an additional feature, we defined the update operations such that one obtains a 'programmable' *quantum* computer which can simulate every quantum circuit when it is appropriately initialized, whereas Feynman's and Margolus' models were designed for classical computation[13]. Therefore we add two program qubits to the center of each neighborhood. The 4 basis states of this system determine which one of the following 4 two-qubit gates is applied to the lower two data qubits before the information is propagated to the upper row: (0) identity (1) Hadamard on the first qubit (2) Hadamard on the second (3) controlled-$\sigma_z^{1/2}$ (see Fig. 2.5).

In order to have a finite system, we choose cylindric topology (see Fig. 2.6, left), i.e., a forward propagating clock wave circulates around the cylinder and triggers the implementation of the quantum gates according to the states of the program qubits. The quantum circuit can be thought of as wrapped around the cylinder (see Fig. 2.6, right). In general, the computation may require several rounds, i.e., the quantum circuit is applied several times. This feature ensures that the space requirement of this computation model does not necessarily linearly increase with the running time of the algorithm. In [50] we have shown that our computation model can solve all PSPACE problems on polynomial space, i.e., the space overhead compared to a Turing machine is only polynomial. After sufficiently many circulations of the clock wave the qubits in the output region (see Fig. 2.6, right) are changed from their initial state to the solution. Of course they cannot stay in this state forever since every finite dimensional dynamics is quasi-periodic. We have designed the circuit such that it is, after several more rounds, set back to their initial state. In other words, the output region oscillates between correct result and initial state as the wave front circulates and each is *half of the rounds* present. But both results are only visible when the wave is in the output region. Measuring the clock qubits in the output region at a random time one has good chance to find the wave there when the output area is large compared to the whole cylinder. Measuring the data qubits in the cells where the clock wave was found one would find the computation result or the initial state with probability 1/2 each. So far, we have described the computation process from the point of view of a forward propagating wave front. But the computer works also if we initialize the clock state such that the wave front starts at a well-

[13]The name 'quantum computer' for their models should only express that the computer is a closed quantum system and not that quantum gates are simulated.

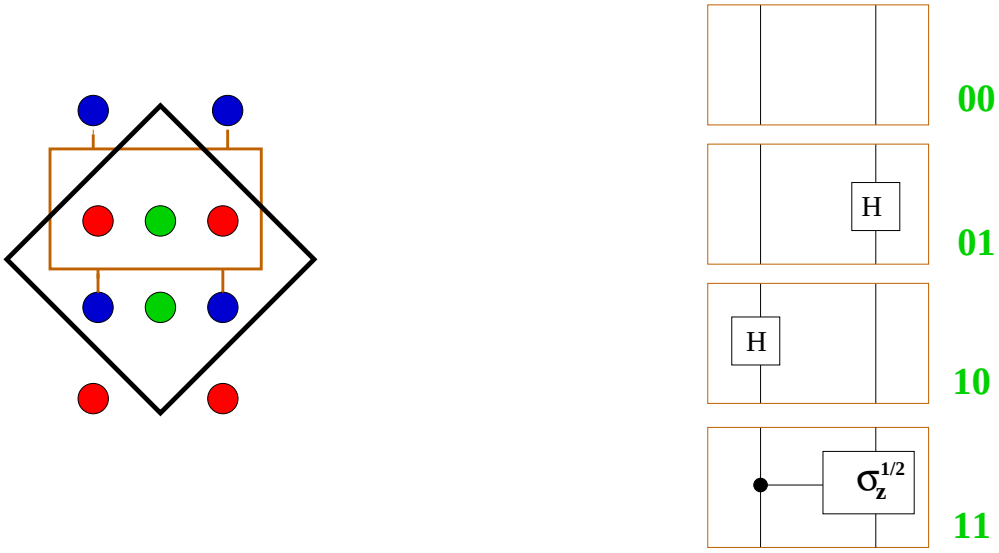


Figure 2.5: The spin wave which propagates over the clock register triggers the implementation of one of the four gates on the right. The program qubits specify which one of those is chosen. The gates are applied to the lower (blue) data qubits. Afterwards a SWAP between the upper and the lower pair of data qubits propagates the information one row upwards. The diagonal square indicate the Wigner-Seitz cell of the crystal which characterizes the symmetry of the interaction and does not coincide with the neighborhoods (see text).

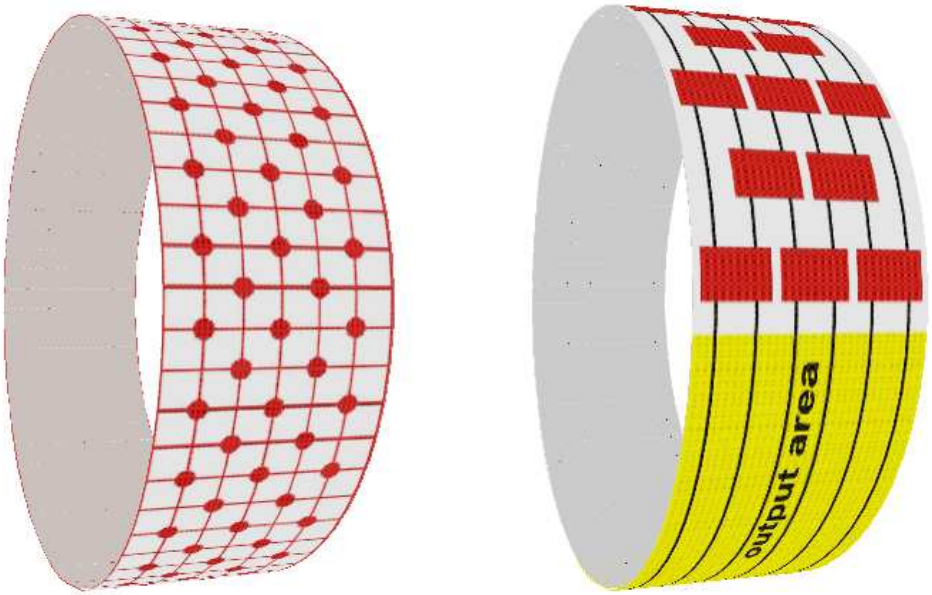


Figure 2.6: (Left) Cylindric crystal consisting of $c \times h$ cells. A pair of program qubits is located at the red points. The lines indicate the boundaries of a cell. (Right) The circuit wrapped around the cylinder. Every time when a two-qubit gate is applied the information of both qubits is propagated one row upwards. The output region consists only of trivial gates, i.e. the information is only propagated.

defined position. Then the wave propagates with rather different velocities and part of it travels backwards. The idea of ergodic quantum computing is that this 'probability-1/2-principle' holds even though the wave front does not propagate step by step but propagates forwards and backwards simultaneously. To show this formally and to analyze the required time interval such that the limiting time average is approximately attained is a technically complex task, worked out in [50]. The idea is that on the relevant subspace the Hamiltonian dynamics is isomorphic to several independent quantum random walks on one-dimensional circular chains.

To characterize the symmetry of our hypothetical crystal we have figured out how a so-called Wigner-Seitz (WS) cell [81] can be defined in analogy to crystallography. It is the diagonal square in Fig. 2.5 and contains 6 qubits. A translation from one WS-cell to the next is a symmetry operation of our interaction.

The Hamiltonian described above, also provides an interesting model in which to discuss the following questions. Looking at present day experiments with coherent quantum control (e.g. with lasers in quantum optics or high frequency fields in NMR) one could conclude that every coherent quantum control must necessarily use macroscopic fields. Otherwise the quantum uncertainty of the relevant field variables seems to create entanglement between control field and controlled system. Does the Hamiltonian model of a quantum computer above support this conjecture or does it disprove it? One could easily initialize the program qubits such that some desired unitary operations are applied to the data qubits. However, the data qubits are entangled with the clock qubits and the desired output state is only available given that the clock wave is in the desired region. Nevertheless we have found a spatially homogeneous Hamiltonian with finite range interactions which solves complex coherent quantum control tasks even though no macroscopic control fields are involved.

2.2.10 Saving Energy by Quantum Information Transfer?

The models for computing in a closed physical system are by definition *quantum coherent* models, i.e., they preserve superpositions – otherwise they would not be closed. Nevertheless it is a difficult question whether thermodynamical reversible computation requires the isolation of the computer such that it would even preserve superposition. Clearly, the destruction of superposition is an irreversible process. On the other hand the dephasing operation shown in Subsection 2.2.8 is thermodynamically reversible whenever it is applied to basis states and only generates entropy when it is applied to superpositions. Therefore one could think of a classical computer which *does not* generate entropy when it is run with classical states (i.e. basis states) but would generate entropy if one would put superposition into the register. In such a system one could avoid wasting resources by avoiding superpositions. However, it is remarkable that there are communication tasks where the ability to send quantum superpositions can save thermodynamical resources. We first want to rephrase some interesting observations by several authors about the thermodynamical relevance of quantum information in communication schemes and use them as basis for our own work.

In their paper "quantum non-locality without entanglement" [82] Bennett et al. consider the situation where two parties, Alice and Bob, want to prepare either of some

bipartite pure states

$$|\psi_j\rangle := |\alpha_j\rangle \otimes |\beta_j\rangle$$

with probabilities p_j. The states $|\psi_j\rangle$ are constructed such that they are mutually orthogonal but neither all $|\alpha_j\rangle$ nor all $|\beta_j\rangle$ are mutually orthogonal. If Alice and Bob want to prepare state $|\psi_j\rangle$ after they have agreed on the choice of j by classical communication, at least some of the information which was sent has to be erased since the classical messages must be distinguishable whereas the states $|\beta_j\rangle$ are not all perfectly distinguishable. A somewhat different situation has been considered by Zurek [83]. He considered bipartite mixed states of two parties, Alice and Bob, which have discord (see Section 1.3). Then we assumes that Alice and Bob both want to erase their systems, i.e., prepare an uncorrelated pure state. By Landauer's principle this initialization requires energy since the entropy has to be transfered to the environment. Consider first the case that they do not communicate at all. Then the entropy Alice has to transfer to the environment is $S(\sigma_A)$. The entropy which Bob has to transfer is $S(\sigma_B)$. Then consider the case that Alice and Bob communicate with a quantum channel. The entropy cost to initialize a pure state is only $S(\sigma)$, the von-Neumann entropy of the joint state. Hence they waste the information $I(A : B)$ when not using their channel. Zurek asked the question of how much information they waste if they use only a classical channel. Here we explain the argument according to our reformulation in [84]. Assume, for instance, that they use only classical one-way communication from Alice to Bob after a von-Neumann measurement (P_j) on Alice's system. There can be two reasons why they lose thermodynamic resources. First the measurement creates some entropy if the unselected post-measurement state is considered. This entropy production coincides with the Kullback-Leibler distance between the initial state and the post-measurement state [85]:

$$\Delta S = K(\sigma||\sum_j q_j\sigma_j) \tag{2.8}$$

where q_j is the probability for the measurement outcome j and σ_j is the post-measurement state $\sigma_j := (P_j \otimes \mathbf{1})\sigma(P_j \otimes \mathbf{1})/q_j$ given j. A second reason why Alice and Bob will in general waste some information by not using a quantum channel is that they do not make use of the fact that the selected post-measurement states σ_j may still be correlated. The average over the remaining mutual information given the message J with values j can be written as the average Kullback-Leibler distance

$$I(A : B|J) = \sum_j q_j K(\sigma_j||\sigma_j^A \otimes \sigma_j^B)\,, \tag{2.9}$$

where the states with superscripts denote the restrictions to A and B, respectively. We have shown [84] that the sum of (2.8) and (2.9) coincides with a modified definition of discord introduced by Zurek in [83]. Although it differs from the first discord (see Subsection 1.3.2) it vanishes if and only if σ can be written as a sum μ_j, ν_j with mutually orthogonal states μ_j such that

$$\sigma = \sum_j p_j\, \mu_j \otimes \nu_j\,.$$

Then Alice is able to distinguish between all possible product states $\mu_j \otimes \nu_j$ without any information loss and without any entropy generation by choosing a measurement (P_j) such that $P_j \mu_j = \mu_j$.

Horodecki, Horodecki, and Oppenheim [61] considered two-sided multi-step protocols and proposed to quantify quantumness of correlations by the so-called thermodynamic work deficit, i.e., the entropy cost to prepare pure product states when only a classical channel is available.

Even though these remarks may show the thermodynamical relevance of quantum communication one may ask why Alice and Bob should prepare those non-entangled quantum correlated bipartite states at all. In [84] we tried to answer this question by stating that the need to prepare quantum correlated separable states appears naturally in clock synchronization schemes (this is related with the remarks above saying that timing information is always to some extent quantum information except in the infinite energy limit). Our reasoning was that the joint state of two synchronized clocks, viewed by an external observer who does not have any knowledge about the actual time, always has some discord whenever both clocks are quantum. This statement refers to the following model: Given two independent quantum dynamical systems A and B, with Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ and time evolutions $U_t^A := \exp(-iH_A t)$ and $U_t^B := \exp(-iH_B t)$, respectively. Let α_t be the time evolution on the density operators given by

$$\alpha_t(\rho) = U_t^A \rho U_t^B$$

and let similarly β_t be the time evolution of Bob's clock. Then we proved that every joint state that is stationary with respect to the joint evolution $\alpha_t \otimes \beta_t$ and non-stationary with respect to the relative translation $\alpha_t \otimes \beta_{-t}$ has some discord (see Subsection 2.2.7). The state of synchronized clock is always non-invariant with respect to such a relative time translation. Our quantitative bound states, roughly speaking, that the joint state of two clocks which are synchronized up to an error of Δt necessarily have the discord

$$\delta(B|A) \geq \frac{1}{256(\Delta E \Delta t)^2},$$

where ΔE is the energy bandwidth of the bipartite system. The same bound applies certainly to the discord $\delta(A|B)$.

Whereas discord concerns here the cost for *resetting* synchronized clocks, we also considered the thermodynamic cost for *preparing* synchronized states, i.e., the entropy generated in a synchronization protocol. We showed for one-way protocols that synchronization can only be performed in a thermodynamical reversible way if a quantum channel is used. We briefly sketch the argument. The whole correlation between Alice's and Bob's clock is generated by sending a classical message from Alice to Bob. That is, given this message j, both clocks are in an uncorrelated state $\mu_j \otimes \nu_j$. If the states ν_j are not perfectly distinguishable Bob has erased some information since the different messages are represented by mutually distinguishable states. In our derivation of a lower bound on the thermodynamical information loss of a state $\sigma = \sum_j p_j \, \mu_j \otimes \nu_j$ we argued that the bipartite state of Alice on one side and Bob and his environment on the other side is a state without discord. This means that it has a decomposition into product states that are mutually orthogonal on Bob's side. Only the restriction to the two clock system

has no such orthogonal decomposition. We proved that for every two synchronized clocks the mutual information between Bob's clock and his environment is at least

$$\frac{\hbar^2}{(4\Delta E \Delta t)^2}$$

if the synchronization has accuracy Δt. Hence the entropy generation during the synchronization is at least this amount of entropy.

As we stated in [84] the result may have implications for low power computation. Since computation always involves some degree of synchronization among different components and devices, it always generates entropy when no quantum channels connect the devices. This raises the question to what extent purely classical reversible computation is possible at all if dynamical and aspects of clocking are taken into account.

2.3 Complexity Theory

Quantum Computing provides a computation model which is inequivalent to the classical Turing machine at least with respect to query complexity. It is known that Grover's algorithm searching for the pre-image of a function needs less oracle queries than the best possible classical would need. Meanwhile, quantum complexity theory has already attracted broad interest in the field [14]. Chapters 3 and 4 will address issues of complexity theory in a different context. Here we will only briefly mention why already the standard model of quantum computing made it necessary to rethink complexity theory.

2.3.1 Challenging the Strong Church-Turing Thesis

The so-called strong Church Turing thesis states that every process in nature where n components interact can be simulated on a Turing machine with some given precision such that the simulation time is polynomial in n. Implicitly, it has already been questioned by Feynman [86] whether this still holds true for the simulation of interacting quantum systems. The discovery of Shor's factoring algorithm strongly supported this doubt[14]. Hence one should work instead with the quantum version of the strong Church Turing thesis: every system in nature which consists of n components can be simulated on a quantum computer such that the number of required qubits k and required time steps is polynomial in n, in the simulated time T and in $1/\epsilon$ when ϵ is the demanded accuracy. In other words, the quantum version of the strong Church-Turing thesis states the following:

1. If there are fundamental laws which make the realization of the quantum computer impossible, the complexity measure given by the standard model of quantum computation is a lower bound for the 'complexity measure given by nature'.

2. If quantum computing is in principle possible the complexity measure given by nature is equivalent to the measure given by the standard model.

Later we will argue that the quantum version of the strong Church Turing thesis implies predictions on the efficiency of control mechanisms.

[14]Note, however, that in [87] the conjecture was stated that an efficient classical algorithm for factoring large numbers is possible.

2.3.2 Quantum Complexity Classes

Complexity classes are an essential part of complexity theory in classical computer science [88]. These classes are often associated with *languages*. These languages define computational problems with answer 'yes' or 'no' as determining whether an input string is in a language L or not. Probably the most important complexity class in classical computer science is $\mathcal{P}$, the set of problems which can be solved by algorithms where the running time increases only polynomially in the length of the input string. Formally one has:

Definition 14 (the Languages in the Class P)
A language L is in P if there is an algorithm on a classical computer which determines whether the input string x is in L such that the running time increases only polynomially in the length $|x|$ of x. For short: Its running time is in $O(poly(|x|))$.

There is also a probabilistic version of this [14]:

Definition 15 (the Languages in BPP)
Let $1/2 > \epsilon > 0$ be constant. L is in BPP if there exists an algorithm A with running time in $O(poly(|x|))$ such that

- *If $x \in L$ the output of A is 1 with probability at least ϵ*
- *If $x \notin L$ the output is 1 with probability at most ϵ.*

In quantum computing it is usual that the output is probabilistic. Hence it is more common to define the quantum analogue of BPP than that of P:

Definition 16 (the Languages in BQP)
A language L is in BQP if there is a quantum algorithm with running time in $O(poly(|x|)$ that computes probabilistically whether $x \in L$.

It is believed that $BPP \subset BQP$ (in the sense of proper inclusion), i.e., that there are problems which can be solved by a quantum computer in polynomial time but not by a classical one. For instance, no classical polynomial algorithm for factoring large numbers is known, whereas the Shor algorithm is polynomial (see Subsection 1.1.3). Feynman's conjecture that quantum dynamics cannot be simulated efficiently by classical computers suggests the definition of a class of problems which arise from the simulation of dynamics and which are in BQP but probably not in BPP: Given two k-local n-qubit interactions H_1 and H_2, decide whether the norm distance between

$$\exp(-iH_1t)|0\ldots0\rangle$$

and

$$\exp(-iH_2t)|0\ldots0\rangle$$

is either at least 2ϵ or at most ϵ with $t \in O(poly(|x|)$ and $1/\epsilon \in O(poly(|x|)$, where x is a string specifying H_1 and H_2. The following quantum algorithm decides this in polynomial time:

1. Prepare the state $|0\dots0\rangle$.
2. Simulate $\exp(-iH_1t)$ on a quantum computer. Due to [89] an accuracy of δ requires the time $O(t^2/\delta)$. Simulate $\exp(iH_2t)$.
3. Perform a measurement on the state $|\psi\rangle := \exp(iH_2t)\exp(-iH_1t)|0\dots0\rangle$ in the computational basis. Repeat the whole procedure such that $|\langle\psi|0\dots0\rangle|^2$ can be determined with the required accuracy.

Clearly the determination of the overlap up to an inverse polynomial error requires in this procedure only polynomial running time because of the promise that it is not between ϵ and 2ϵ. This example shows that BQP problems arise in a natural way from quantum control problems.

In Subsection 4.2.2 we will furthermore need the complexity class QMA which is the quantum analogue of NP. We shortly recall the definition of NP.

Definition 17 (the Languages in the class NP)
A language L is in NP if for every string $x \in L$ there is a witness, i.e., another string y_x, such that, given y_x, a classical algorithm can check in polynomial time that $x \in L$ holds. If $L \notin NP$ no such witness exists.

The probabilistic version for NP is MA:

Definition 18 (the Languages in MA)
A language L is in in MA if for every $x \in L$ there is a string y_x and a probabilistic program A with input x, y_x such that the output is 1 with probability at least $1-\epsilon$ and for every $x \notin L$ the output of A is 0 with probability at least $1-\epsilon$ for all y_x.

The string y_x which leads to the positive result with probability $1-\epsilon$ is called a witness. There are at least two possible quantum analogues of MA:

Definition 19 (the Languages in QMA/QCMA)
A language L is in QMA (QCMA, respectively) if for every $x \in L$ there is a quantum state (respectively a basis state) $|\psi_x\rangle$ and a probabilistic quantum algorithm A with input $|x\rangle \otimes |\psi_x\rangle$ such that the output is 1 with probability at least $1-\epsilon$ and for every $x \notin L$ the output of A is 0 with probability at least $1-\epsilon$ for all y_x, respectively $|\psi_x\rangle$.

In other words, the witness for QMA is a quantum state, for QCMA it is a classical string. In Subsection 4.2.2 we will rephrase an interesting example of a QMA-complete problem which appears in the theory of low temperature physics. We have given [90] an example of a QMA-complete problem which appears in the design of quantum circuits, namely to decide whether two circuits which are specified by a sequence of gates implement unitaries that are not almost equivalent. We rephrase the completeness statement for a special instance, namely to decide whether a circuit is almost the identity:

Theorem 10 (Non-Identity-Check is QMA-Complete)
Let x be a classical description of a quantum circuit U_x of complexity polynomial in $|x|$. Then the following problem is QMA-complete.

Decide whether U_x is up to a global phase close to the identity in the following sense: Decide which of the two following cases is true given the promise that either of the conditions 1. or 2. is satisfied:

1. *for all* $\phi \in [0, 2\pi)$ *one has* $\|U_x - e^{i\phi}\mathbf{1}\| \geq \delta$ *or*

2. *there exists an angle* $\phi \in [0, 2\pi)$ *such that* $\|U_x - e^{i\phi}\mathbf{1}\| \leq \mu$.

Assume furthermore that $1/(\delta - \mu) \in O(poly(|x|))$.

This shows that problems of optimal circuit design for implementing a desired quantum process can lead easily to problems which are not only NP-hard but even QMA-hard.

The complexity classes NP, MA, QMA, and QCMA are believed to include problems which require exponential running time. Nevertheless they can all be solved with polynomial space resources. In other words, they belong to the complexity class PSPACE, which is usually defined with respect to the Turing machine model [91, 92]:

Definition 20 (the Languages in PSPACE) *PSPACE is the class of all languages recognizable by polynomial space bounded deterministic Turing machines that halt on all inputs.*

It is known [93] that the class of problems which a quantum computer could solve in polynomial space is not greater than PSPACE. Therefore there is no need for a specific quantum version of PSPACE. The class PSPACE will occur in Subsection 4.1.1 where we will show that accurate von-Neumann measurements of some natural types of observables could solve problems in this class in polynomial time. The fact that PSPACE is an extremely large class will be interpreted as an indicator for limits of complexity theoretic limits of measurement technology. This shows that limits of computation and limits of other technologies controlling the nanoscopic world may have a common origin in the law of physics which all processes must respect. In Section 4.2.2 we will consider the complexity of a state preparation problem and define a complexity class which will be related to complexity classes of computation.

2.3.3 Complexity Measures from Nature?

The statement that the laws of physics determine complexity of computation [86, 94] seems meanwhile to be widely accepted. In a vague sense, the conventional computational models like Turing machines or cellular automata already reflect some features of physics. The fact that the head of the Turing machine only reads and writes on the cells at its current position can be seen as an idealization of the statement that its interaction with other cells decreases strongly with distance. A similar kind of locality condition appears in the update rules of cellular automata which refer only to a neighborhood of the updated cell. The spatial homogeneity of a cellular automaton can furthermore be seen as an analogue of the crystal structure of solid states. To make this analogy even closer, one should modify the cellular automaton to continuous-time dynamics. Since classical dynamics on cells with finitely many states would not allow deterministic dynamics, it is natural to replace the classical automaton by a quantum system. In this sense, a finite range Hamiltonian on a periodic array of finite dimensional quantum systems (like those studied in [76, 50]) is the most natural continuous-time analogue of a cellular automaton. In one dimension, for instance, they have the form

$$H = \sum_{R \subset \mathbb{Z}} H_R$$

where H_R is an operator acting on a region R of k adjacent qudits $j, j+1, \ldots, j+k-1$ and the sum runs over all intervals of length k. Each operator H_R can be considered as the continuous-time analogue of a cellular automaton update rule acting on region R. Due to the continuity of the dynamics, the update in each region R is only an 'infinitesimal' change of the state $|\psi\rangle$:

$$|\psi\rangle \mapsto |\psi\rangle - idt\, H_R|\psi\rangle \,,$$

after the infinitesimal time dt. Note that this type of continuous time cellular automaton is indeed a type of Hamiltonian which is used to describe the theory of crystals [95, 96]. However, to study complexity issues using such physical Hamiltonians could be a hard task. Thus, even in quantum computing one prefers therefore to consider *discrete* cellular automata [97, 98]. Note that the results in [89] imply that the standard quantum computer can efficiently simulate the time evolution of the 'crystal Hamiltonian' above. Therefore it can also be simulated by every discrete quantum cellular automaton which is able to simulate the quantum computer.

It seems reasonable to assume that the interaction Hamiltonians available in nature determine quantum complexity. In a world where every interaction among the qubits could be achieved one could, for any given unitary U on n qubits, choose a Hamiltonian H such that

$$\exp(-iH) = U \,,$$

i.e., the system needs only a time period of length 1 to perform U and H can be chosen such that its operator norm is only π. Hence one needs not even a strong interaction for the computation. Physics suggests to restricting the Hamiltonians as follows.

Definition 21 (k-local Hamiltonian)
A k-local Hamiltonian H is a sum of operators which act on only k qudits, i.e.,

$$H := \sum_{j_1,\ldots,j_k} H_{j_1,\ldots,j_k}$$

where each summand acts on the qudits $j_1, \ldots, j_k$.

If the qudits coincide with fermionic particles, the fundamental forces are 2-local. Sometimes it is convenient to represent the state of a particle by several qubits. Then one can also have $k > 2$, meaning effective interactions which appear in a phenomenological treatment of real physical situations. However, large k are certainly unphysical. Even though physical reality would also suggest excluding interactions between distant particles it follows from [99, 100] that this constraint would not change the computational power of the Hamiltonian up to a polynomial time overhead as long as additional single-qudit operations are available. For details we refer to the broad literature on simulation of Hamiltonians ([101] and references therein) and to Section 3.1.

2.4 Quantum Communication and Causal Reasoning

Quantum theory has challenged human intuition about causality since its early days. In their famous objection against quantum mechanics, Einstein, Podolsky, and Rosen [102]

argued that it cannot be true that measurements performed on one part of an entangled state of a bipartite system affect the quantum state of the other subsystem. Bell realized [103] that this paradox should have testable implications: entangled systems show correlations between potential measurements on each of the subsystems which could only explained by a classical theory if there was a signal running from one system to the other. In other words, from the classical point of view, the correlations would require causal connections between the subsystems. This lead to experiments where causal connections could excluded with high probability and strong correlations could nevertheless be observed. More interesting for current quantum information is the situation in which there is some signal between two parties and the entanglement enables some communication tasks to be better performed than would be possible without entanglement. It allows, for instance, the creation of a secret key or the computation of a certain function with less communication complexity. If one considers the amount of information exchange between two parties as a measure for the strength of a causal interaction between their systems, it follows that entanglement allows some phenomena on interacting systems which required 'more causal influence' among the systems if entanglement would not be there. Here we will explain a situation where one have to rethink rules of causal reasoning in classical statistics. It should be emphasized that the paradox below does not require entanglement between distant systems. We only use a system with bipartite entanglement because it is easier to explain within the chosen quantum communication setting.

Causal reasoning in every-day life is mainly based on humans observations of correlations between events, respectively variables. A rather explicit formulation of this was given by Reichenbach [44] when he stated the "Common Cause Principle": Given the observation that two events A and B are correlated, i.e., the probability $P(A \wedge B)$ differs from the product $P(A)P(B)$, there is always a causal influence from A to B, from B to A or there is a common cause C influencing both. In order to consider C as a complete explanation for the correlations between A and B, Reichenbach formulated the condition that A and B must be independent, given C. This approach turned out to be a clear axiomatic basis to derive causal statements from statistical data. In [104, 105, 106] one can find rules to infer cause-effect relations between n variables from the structure of the correlations among them. Here we want to explain why these rules for causal reasoning do not remain the same when the indeterminism of the observations are partly caused by *quantum* randomness.

2.4.1 Dense Coding and Quantification of Causal Effects

The fact that quantum correlated systems behave different from classically correlated systems is closely related to the fact that shared prior entanglement helps to reduce the amount of information exchange which is necessary for certain communication tasks. An example is dense coding where the transfer of one qubit allows to transfer two classical bits [107].

Before we explain the idea, we first mention briefly several methods to quantify the strength of a causal effect in the theory of causal reasoning. Assume that a variable X (cause) affects the variable Y (effect). Assume that the whole correlation between X and Y stems from the cause-effect relation, i.e., there is neither a backwards-directed causal

effect from Y to X nor an effect from a confounding variable Z effecting both X and Y. First of all, one often considers linear models [106], in which the cause-effect relation is given by a so-called structural equation [105] of the form

$$Y = \beta X + S\,,$$

where $\beta \in \mathbb{R}$ is a structural parameter and S is an independent variable which introduces some indeterminism. It is straightforward to regard $|\beta|$ as the strength of the causal effect. If X and Y are binary variables with values x_0, x_1 and y_0, y_1 one can also define the *average causal effect* [105] by

$$ACE(X \to Y) := P(y_1|x_1) - P(y_1|x_0)\,. \tag{2.10}$$

For a general non-linear model, concepts of information theory help: In [108] Granger uses the mutual information

$$I(X:Y) = -\sum_y P(y)\log_2 P(y) - \sum_{xy} P(y|x)\log_2 P(y|x)P(x)\,,$$

to quantify the causal strength. In [109] we proposed to consider the capacity of the channel $X \to Y$ defined by the stochastic matrix $P(y|x)$ as measure for the causal strength. It is defined by the maximum over the mutual information over all distributions of X. The method is a little more complicated if there are confounding variables. Assume, for instance, that the correlations between X and Y are partly due to a variable Z effecting both X and Y. Then the strength of the stochastic dependencies between X and Y is not relevant, but only that part which is due to the cause-effect relation. To formalize this difference, Pearl [105] introduced the so-called do-calculus. We will briefly sketch this concept. Given that the causal relations between a set of n random variables $X_1, \ldots, X_n$ is known and formalized as a directed acyclic graph with the random variables as vertices. Pearl then gives rules for calculating the probability of $X_j = x_j$ given that variable X_k *was set* to x_k. These rules are important whenever it is not possible to do the experiment setting X_k to x_k. He calls this $P(x_j|\text{ do } X_k = x_k)$ and explains in detail how it has to be distinguished from the usual conditional probability $P(x_j|x_k)$. Here we prefer to demonstrate this difference using the example with Z as common effect. If X is set to x, the probability distribution of y is not $P(y|x)$ but

$$P(y|\text{ do } X = x) = \sum_z P(y|x,z)P(z)\,.$$

Formally $P(y|\text{ do } X = x)$ defines also a classical-to-classical channel and its capacity C would again be a good measure to quantify the effect of X on Y.

It is an interesting question how the concepts in [104, 105, 106] could be generalized to the quantum setting. This is non-trivial even for Reichenbach's principle of the common cause [110]. Here we do not claim to present an answer to these deep questions; we only want to show that the possibility of dense coding can be a difficulty for the quantification of causal effects. Dense coding works as follows. Alice and Bob share a maximally entangled state

$$|\phi^+\rangle := \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)\,,$$

which is one of the four Bell states $|\phi^{\pm}\rangle, |\psi^{\pm}\rangle$ [3]. It is easy to see that they can be converted into each other by local transformations $\mathit{1}, \sigma_x, \sigma_y, \sigma_z$ on Alice's or Bob's subsystem alone. Let the classical variable X determine which of the 4 operations Alice applies. Then she sends her qubit to Bob. He can distinguish between all 4 Bell states with certainty. By such a measurement Bob could define 4 values of a random variable Y. They are completely determined by the values of X and therefore without loss of generality the same. Even though we have no formal do-calculus in the quantum setting, we can set X to specific values x and verify that $P(y|\text{ do } X = x) = \delta_{x,y}$. Hence we have a two-bit classical-to-classical channel. This means that the causal effect of Alice's transformations on Bob's measurement results is 2 classical bit. On the other hand, one may like to define the strength of a causal arrow with the amount of *information flow* (through space-time) in a reasonable sense. And there is no doubt that the information which propagated through the space from Alice to Bob was only *one* qubit. One may resolve this conflict between these two reasonable ways to quantify the causal strength by saying that one qubit is more than one bit. However, here it should only be stressed that problems of this kind occur when causal strength is defined in information-like terms.

2.4.2 Hidden Variable Models in Clinical Drug Testing

Now we explain another interesting phenomenon about causal reasoning in every-day life in the presence of quantum correlations. The paradox shows that a classical model may fail in describing correlations between classical observations even when the causal effect from one system to the other is a purely classical signal. The potential relevance of this phenomenon for every-day life statistical reasoning can be explained in a clinical drug test.

To test whether or not a drug helps to recover from a certain disease one may direct some patients to take the drug and some not to take it. It is clear that allowing the patients to freely decide whether or not to take it, the patients who take it may no longer be representative since for instance their age, their personality, or some other confounding variables may be different in the group of volunteers compared to the control group. Therefore one has to decide randomly who should take the drug and who shouldn't. However, it is a well-known problem in clinical tests that not all patients comply. Some may take the drug even though they were instructed not to take it and some could not take it even though they were directed to take it. Due to the arguments above it is clear that the error caused by the non-compliers cannot be avoided if one could find out the non-compliers (for instance by a blood-test): the compliers may no longer be a representative group and correlations between taking or non-taking the drug and recovery are not necessarily caused by the drug; they could also be caused by confounding variables. On the other hand it is clear that an extremely small compliance rate could not fake arbitrarily high correlations between taking the drug and recovery. This is the idea of the analysis below where we follow mainly Pearl [105].

First, let us rephrase his most intuitive conclusion. Imagine that an inexperienced researcher, unaware of the non-compliance issue, observes the two binary variables Z and Y. The values $Z = z_1$ or $Z = z_0$ mean that the patient had or had not been selected to take the drug and $Y = y_1$ or $Y = y_0$ means that the patient recovered or not, respectively. The inexperienced researcher would then take the difference between

the probability of recovering if the patient was instructed to take the drug minus the probability of recovering if the patient was not instructed to take the drug, i.e., the value

$$P(y_1|z_1) - P(y_1|z_0)\,,$$

as the (positive) causal effect of the drug.[15] Now imagine that a more experienced statistician asks for the maximal error that this naive conclusion can cause. He observes the variable X where $X = x_1$ or $X = x_0$ means that the patient has or has not taken the drug, respectively. He concludes that in the worst case of overestimating the effect, all those who had been instructed to take the drug but did not comply and have recovered, would have stayed ill had they complied. On the other hand, all those patients who had been instructed not to take it and took it nevertheless and did not recover may have recovered if they would have complied. By this intuition he concludes that the causal effect of the drug is to increase the probability of recovering at least by

$$P(y_1|z_1) - P(y_1|z_0) \;\; - P(y_1, x_0|z_1) - P(y_0, x_1|z_0)\,, \tag{2.11}$$

where $P(y_1, x_0|z_1)$ denotes the conditional probability of the event 'no drug taken and recovered', given that the instruction was to take it. The other definitions are similarly constructed.

By the same kind of reasoning, one can find bounds on the underestimation of the causal effect. The result is that the recovery rate is increased by at most

$$P(y_1|z_1) - P(y_1|z_0) \;\; - P(y_0, x_0|z_1) - P(y_1, x_1|z_0)\,. \tag{2.12}$$

Note that the increase in the recovery rate we are describing is the increase that would happen if all patients took the drug (including those who have decided not to take it). Therefore the definition of the causal effect relies on the hypothetical result of an experiment where all patients are forced to comply.

For a formal proof of statements of this kind we need a precise model in which terms such as "the recovery rate if all patients take the drug" make sense. Remarkably, the proofs which can be found in the literature [111, 105] refer to a hidden variable model of mental and physical behavior of the patients. Fig. 2.7 shows the graphical model of the causal structure.

It is assumed that the mental and physical state which determines whether he decides to take the drug or not and whether he recovers or not can be described by a hidden variable U. As we have briefly sketched above, the causal effect of Z on Y could be identified if the confounding variable U could be observed. We have then

$$P(y|\text{ do } x) = \sum_u P(y|x,u)P(u)\,.$$

Due to the existence of a confounder, the formula (2.10) for the average causal effect has to be redefined in terms of *do*-probabilities instead of usual conditional probabilities:

$$\begin{aligned} ACE(X \to Y) &:= P(y_1|\text{ do } x_1) - P(y_1|\text{ do } x_0) \\ &= \sum_u (P(y_1|x_1,u)P(u) - P(y_1|x_0,u)P(u)\,. \end{aligned}$$

[15] Here we use the large sampling assumption, i.e., the sample size is large enough to estimate the joint distribution P on all observed random variables. Issues of significance of correlations can therefore be neglected here.

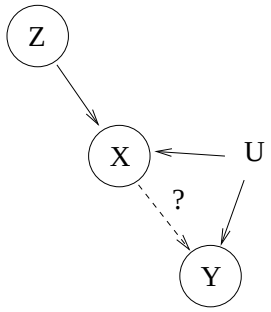


Figure 2.7: Graphical model of causal structure of the compliance problem: Z is the instruction to take the drug, X is patient's decision to take it, Y is his physical response (recovered or not), and U represents all relevant latent factors influencing decision and response.

The interesting insight of the analysis in [105, 111] is that one can derive bounds on the causal effect even without observing U at all. They argue that U, no matter how complex its influence on behaviors and health is, can be represented by a variable with 16 values without loss of generality. Based on this hidden-variable model they derive 8 lower and 8 upper bounds on ACE in terms of conditional probabilities with the observable quantities X, Y, Z. One of them reads, for instance:

$$\begin{aligned} ACE \geq P(y_1, x_1|z_0) - P(y_1, x_1|z_1) - P(y_1, x_0|z_1) \\ -P(y_0, x_1|z_0) - P(y_1, x_0|z_0)\,. \end{aligned} \tag{2.13}$$

These 16 bounds are tighter than the more intuitive bounds (2.11), (2.12). But what happens if the confounder U is not a classical variable but a quantum state? Even though we have chosen not to digress into the controversial subject of to what extent quantum mechanical effects play a crucial role for mental and physiological processes here, we cannot assume, as yet, that the quantum effects are irrelevant. The toy model below shows that neglecting these quantum effects can lead to causal misconclusions. Assume that mental and physical conditions of the patient could be described by a state ρ of two qubits, where the left is thought to belong to the mental part of the patient and the right to his physical constitution. Assume furthermore that the instruction to take the drug influences his mind in such a way that it implements $V \otimes \mathbf{1}$. The patient's decision is determined by a measurement of the left qubit in the computational basis (with projectors P_0 and P_1). The effect of the drug is that it implements $\mathbf{1} \otimes W$. Then the outcome of a measurement on the right qubit determines whether the patient recovers or not.

Obviously the causal effect of the drug is given by

$$P(y_1|\text{ do } x_1) - P(y_1|\text{ do } x_0) = tr(\rho(\mathbf{1} \otimes U^\dagger P_1 U)) - tr(\rho(\mathbf{1} \otimes P_1))\,,$$

since the instruction is irrelevant for the right qubit. The strongest violation of the ACE bounds that we found is given as follows. Define

$$\rho := (H^{1/2} \otimes H^{-1})|\Phi^-\rangle\langle\Phi^-|(H^{1/2} \otimes H^{-1}) = (H^{3/2} \otimes 1)|\Phi^-\rangle\langle\Phi^-|(H^{3/2} \otimes 1),$$

with the Bell state

$$|\Psi^-\rangle := \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)\,,$$

and the Hadamard gate H. The second identity is due to the invariance of $|\Psi^-\rangle$ with respect to common rotations on both qubits. Let V and W both be the Hadamard gate H. The specific feature of $|\Phi^-\rangle$ is that measurements in the same basis performed on both qubits lead always to different outcomes. Let P_0^α, P_1^α be measurement operators with respect to the basis

$$|\psi_0\rangle := \cos\alpha|0\rangle + \sin\alpha|1\rangle$$

and

$$|\psi_1\rangle := \sin\alpha|0\rangle - \cos\alpha|1\rangle\,.$$

If the measurements P_0^α, P_1^α and P_0^β, P_1^β are performed on the left and the right side, respectively, then the probability for obtaining the same outcome is given by $\sin^2(\alpha-\beta)$. Reinterpreting the transformations H and $H^{3/2}$ as a change of the measurement basis the instruction to take or not to take the drug leads to measurement angles $\alpha_1 = 22,5^o, \alpha_0 = 67,5^o$, respectively. The taking or not-taking of the drug leads to $\beta_1 = 0^o, \beta_1 = -45^o$, respectively.

The average causal effect of the drug is zero. To see this recall that ACE measures the increase of the recovery rate in a hypothetical experiment where all patients take the drug compared to the case that nobody takes it. In this hypothetical experiment, the taking of the drug is decoupled from the patient's decision to take it. This means that the angle of the right hand filter is not determined by the outcome of the left hand measurement. In this experiment there is certainly no correlation between the taking of the drug and the recovery since the change of the measurement angle by the drug is irrelevant for the fact that half of the patient's recover. However, the ACE bounds claim a causal effect with our specific choice of the angles. The decisive conditional probabilities are given by

$$P(y_j, x_k|z_l) = \frac{1}{4}\Big(1 - (-1)^{j+k}\cos(2\alpha_l - 2\beta_k)\Big)\,.$$

This can be seen as follows: The indices l (instruction to take/ not to take) and k (taken or not) determine the angles α and β of the filters. The probability that the patient's decision (0 or 1) coincides with his response to the drug (0 or 1) is given by $(1-\cos(2\alpha_l - 2\beta_k))/2$. The probability that the results disagree is $(1+\cos(2\alpha_l - 2\beta_k))/2$. The probabilities for the results 1 and 0 in the first measurement are $1/2$ each, regardless of the measurement angle. This gives an additional factor $1/2$. In contrast to the ACE bound (2.13) we find

$$P(y_1, x_1|z_0) = 1/4(1 + 1/\sqrt{2}) =: a^+$$

and all the other terms on the right hand side are

$$1/4(1 - 1/\sqrt{2}) =: a^-\,.$$

The ACE-bound would therefore imply

$$ACE \geq 1/4(-3 + 5/\sqrt{2}) \approx 0.134\,.$$

Hence the third ACE-bound claims the average causal effect to be at least about 13% although it is zero. Due to the symmetry of the problem we can violate three of the 8 inequalities similarly (for details see [112]).

Note that in this setting it was essential that there indeed exists a causal effect of the drug on the recovery - sometimes negative and sometimes positive depending on the patient's decision. Hence the conclusion that the drug influences patient's health is true nevertheless. In [112] we have furthermore shown that the other bounds as well as the intuitive bounds (2.11),(2.12) are even true in the quantum setting. The formal setting for investigating the violation of the classical ACE-bounds by latent quantum factors is based on the following assumptions.

1. The instruction to take or not to take the drug is perfectly randomized and independent of all other factors.

2. All relevant latent factors influencing the patient's decision and his response behavior to the drug are described by the state of a physical system in the sense of algebraic quantum theory. This state includes the patient's mental and physical state as well as noise that influences the decision or response or both. The state is the state ρ of a physical system described by an observable algebra, i.e., a C^*-algebra $\mathcal{A}$ (see [64]). This generalizes the concept of states used so far, where it is a density operator. Explicitly, it is a positive linear functional of norm 1 on $\mathcal{A}$. The system is either purely quantum, purely classical, or a mixture of both. Although this may be too a materialistic view on mind and consciousness, this approach is more general than any hidden variable model in the literature.

3. To take or not to take the drug is a classical event that either occurs or does not occur but there is no quantum superposition between both. The process of human decision is therefore like a measurement process in its broadest sense, as explained in Subsection 4.1.2. This instrument is described by CP-maps D_1, D_0 acting on $\mathcal{A}$. Therefore, the state ρ is transformed to $\rho \circ D_1/\|\rho \circ D_1\|$ if the patient has decided to take the drug and $\rho \circ D_0/\|\rho \circ D_0\|$ otherwise. If the decision itself is ignored, the process of decision making is described by the process $\rho \mapsto \rho \circ D$ with $D := D_0 + D_1$.

4. The instruction to take or not to take the drug is a classical signal that influences the patient's internal state. The instruction to take or not to take transforms the state to $\rho \circ G_1$ or $\rho \circ G_0$, respectively. Here G_j are CP-maps on $\mathcal{A}$.

5. The effect of the drug is to transform the internal state ρ to $\rho \circ E_1$, whereas the natural evolution without the drug in the considered time interval changes the state according to the operation E_0. The operations E_j are CP-maps on $\mathcal{A}$.

6. Whether the patient recovers or not is a classical event and is therefore equivalent to a measurement process in the sense above. It corresponds to a yes-no-experiment described by a positive operator $m \in \mathcal{A}$.

7. The instruction to take or not to take the drug has no direct causal influence on the health. It influences the probability of recovery only by indirectly influencing the decision. This corresponds to the fact that the graphical model Fig. 1 for the classical setting has no arrow from Z to Y.

One may think that it would be more appropriate to assume that the operations G_j and E_j act on different systems: G_j acts on the mind and E_j on the body. But we do not

want to restrict our proofs to this assumption. In particular, we emphasize that there may be a part of the body with the property that its quantum state influences the decision and the recovery. This may, for instance, be a cell that influences the production of some hormone that has a causal effect on both mood and health. It would be rather speculative to discuss whether such a phenomenon is likely to happen; nevertheless cautious scientific reasoning should not rely to hidden-variable models of phenomena like mental processes which are not yet well-understood.

Chapter 3

Extending the Definition of Algorithms

Here we will argue that the usual understanding of algorithms is too strict when all existing proposals of quantum algorithms are taken into account. First, the finiteness of the number of steps which is essential in classical computer science does not give consideration to the continuity of quantum logical operations. Continuous algorithms require their own language based on Lie-algebraic terms. Second, quantum algorithms for non-computational problems deserve other specifications than algorithms for computation since input and output are not necessarily classical strings.

3.1 Continuous Algorithms

In the preface we have quoted a definition of algorithms which states that an algorithm consists of finitely many steps. The idea is that this is necessary in order to have finite running time since every machine has a minimal time to perform a logical operation. For classical computation, there are clear physical reasons supporting this point of view. As shown in [113] this is due to the finiteness of the available energy since the change of the logical state of a bit within a time interval Δt requires at least an amount of $E \geq h/(4\Delta t)$ of available energy. For a qubit there is, however, no such a bound. The state space of a qubit or a quantum register is a continuum. Changing the state only a little bit may require an arbitrarily short time period.

Of course, all processes in a classical computer are also continuous processes when the computation is considered on the *physical level*. But the computer science description of a computation is restricted to the logical level and does not consider the period of the switching process itself where some bits may be in a logically undefined state.

In quantum computing, the quantum logical state *can* be defined even during the implementation of a unitary operation. However, it is important to note that this is not necessarily the case. Consider a qubit being in the state $|0\rangle$. In order to switch to $|1\rangle$ we implement the σ_x gate. Consider two quantum computers with different hardware. The first implements σ_x by the time evolution $\exp(i\sigma_y t)$. We could call this evolution a *continuous algorithm* for implementing σ_x. In contrast, the second computer creates entanglement with the environment during the implementation. Then the time evolution of the qubit is described by a family of completely positive maps instead of a family of

unitaries. But this is not the essential point. One could easily generalize the term 'continuous algorithm' to CP-maps. But the family of CP-maps which describe the evolution of the qubit density matrix is not a semi-group and is therefore not a concatenation of infinitesimal CP-maps. Therefore the latter hardware would *not* use a continuous algorithm for the σ_x-implementation.

The implementation of the continuous algorithm for implementing the NOT operation is usually controlled by a classical field. Applying a magnetic field with field vector $B := (B_x, B_y, B_z)$ yields, for instance, a Hamiltonian dynamics of a spin-1/2 particle given by

$$H = \sum_{\alpha=x,y,z} B_\alpha \sigma_\alpha \,. \tag{3.1}$$

An algorithm could be a description of a function $t \mapsto B(t)$, i.e., the time-dependence of the field. It does not make sense to talk about *time steps* here. Therefore, the basic resource for quantum computation is in many cases given instead by the available physical interactions than by a discrete set of gates. We briefly want to mention that there is also another type of algorithm which does not fit in the discrete standard model of quantum computing at all. In *adiabatic quantum computing* a method which will be addressed in Subsection 3.1.4, a computational problem is encoded into an interaction in such a way that the ground state of the system (i.e. the zero temperature state) gives the solution.

3.1.1 Languages for Continuous Algorithms

Motivated by the remarks above we may define an algorithm with running time T as a mapping

$$t \mapsto H(t)$$

with $t \in [0, T]$ and $H(t)$ the Hamiltonian at time t. We will assume that the map is integrable in an appropriate sense such that for every $t \in [0, T]$ there is a unique unitary transformation U_t as solution of the time-dependent differential equation

$$\frac{d}{dt} U_t = -iH(t)U_t \,.$$

Furthermore, there must be an efficient algorithm that computes the function $t \mapsto H(t)$ in analogy to the requirement that the circuits used for a discrete quantum algorithm can be computed efficiently.

Above we have argued that the time-dependence of H may stem from a time-dependent field which controls the system. Another instance of such a continuous algorithm is given by models with a fixed Hamiltonian H and the additional ability to implement so-called bang-bang control. These control operations are usually unitaries in some set $\mathcal{U}$ and it is assumed that the implementation time for each $U \in \mathcal{U}$ is extremely short compared to the time scale of the natural evolution $\exp(-iHt)$. This is also called the 'fast-control limit'. One obtains algorithms which consist of unitaries U_j implemented at time t_j and the natural evolution during the time interval between. This implements the transformation

$$\exp(-iHt_n)U_n \exp(-iHt_{n-1})U_{n-1} \ldots \exp(-iHt_1)U_1 \,.$$

By defining $V_j := U_j \cdots U_1$ this is equal to

$$V_j^\dagger \exp(-iV_j H V_j^\dagger t_j)\,,$$

i.e. up to the transformation $V_j^\dagger$ the dynamics is as if it was generated by a time-dependent Hamiltonian with

$$H(t) := V_j H V_j^\dagger \quad \text{for } t \in [t_j, t_{j+1}]\,.$$

If $\mathcal{U}$ is a group, the set of available Hamiltonians is the set of conjugates of H under all available control operations $\mathcal{U}$. Since these different Hamiltonians can (according to the fast-control limit) be applied on short time intervals one can easily generate further Hamiltonians: Using concatenations of evolutions

$$\exp(-iH_1\epsilon)\exp(-iH_2\epsilon) = \exp(i(H_1+H_2)\epsilon) + O(\epsilon^2) \tag{3.2}$$

simulates the sum of two available Hamiltonians H_1, H_2. But one can also simulate $i[H_1, H_2]$ using the rule

$$e^{iH_1\epsilon}e^{iH_2\epsilon}e^{-iH_1\epsilon}e^{-iH_2\epsilon} = e^{-[H_1,H_2]\epsilon^2} + O(\epsilon^3)\,. \tag{3.3}$$

This is also possible if H_1, H_2 are already given by simulations. In other words, procedures to generate a Hamiltonian $\tilde{H}$ using a set of natural Hamiltonians $H_1, H_2, \ldots$ could be formulated in terms of subroutines simulating "intermediate" Hamiltonians in order to obtain the final one. This Lie-algebraic methods offer, so to speak, a 'higher programming language' for simulating $\tilde{H}$ by a given set of Hamiltonians. Keeping in mind the subroutines above one may use terms like 'simulate $[H_1, H_2] + H_3$ using simulations of H_1, H_2, H_3'. This language was widely used in the theory of simulating Hamiltonian and for proving that different many-particle interactions have the same computational power [99]. We have formulated a measurement algorithm for the energy of n qubits which are subjected to an unknown interaction in this language [114] and an algorithm to distinguish between a set of n Hamiltonians on $\mathbb{C}^n$ [115]. Note that every explicit description of the steps, i.e., the time intervals at which the Hamiltonians are switched on, refers to a specific approximation accuracy. Hence the number of necessary steps for these simulations cannot be defined. The running time, however, is in the limit of arbitrary accuracy infinite, only becoming finite if only first order simulation according to rule (3.2) is used.

3.1.2 Comparing Discrete to Continuous Complexity

In the complexity theory of classical algorithms the time complexity plays a crucial role and an algorithm is called efficient if its running time increases only polynomially in the size of the input string. It seems therefore straightforward to consider the running time T of the continuous algorithm as the analogue of the classical time complexity. However, this does not make sense if the set of available Hamiltonians is a vector space. This is shown by the example in eq. (3.1) where the magnetic field can be made arbitrarily strong. Then one could obtain arbitrarily small running time by rescaling the Hamiltonian. Sometimes it will make sense to bound the norm of all available Hamiltonians. However, one can also obtain interesting computation models if the norms of some available Hamiltonians

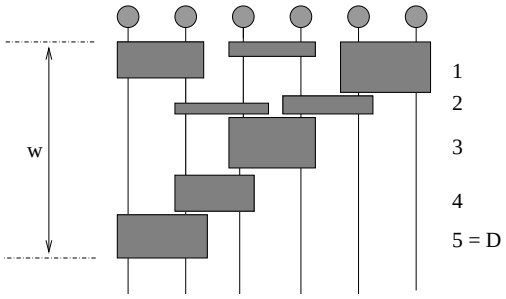


Figure 3.1: Depth (D) and weighted depth (w) of a circuit: D counts only the number of layers (time steps) while w takes also into account their weights.

are unbounded. Actually, the bang-bang control (fast-control limit) in Subsection 3.1.1 can be interpreted as a time-dependent Hamiltonian model where on some intervals $H(t)$ is only the natural Hamiltonian H active and on other small intervals the superposition of H with a strong control Hamiltonian[1]. If the control Hamiltonian is much larger than the natural one, the bang-bang model is obtained.

In [117] we have tried to find a complexity measure for continuous algorithms which is as close to gate complexity as possible. For doing so, we modify the complexity measure *depth* (see Definition 4) such that it better meets the continuity of physical processes: in order to take into account that gates which are close to the identity can possibly be implemented more quickly, we introduced the concept *weighted depth*. First we defined the *angle* of a unitary U as the minimal norm over all self-adjoint A with $\exp(iA) = U$. Then we define (see Fig. 3.1 for an illustration):

Definition 22 (Weighted Depth of a Circuit)
The weighted depth w of a quantum circuit of depth d is the sum

$$w := \sum_{j=1}^{d} w_j$$

where the weight w_j is the maximal angle of all gates implemented in step j.

It order to define complexity of continuous algorithms we define the complexity of an interaction. We will restrict to pair-interactions partly for physical reasons and partly because this fits well with the discrete model with two-qubit gates. Let H on n qubits be given by

$$H := \sum_{jk \leq n} H_{jk}\,,$$

where each H_{jk} acts on the qubit pair jk. Our following definition of complexity is lead by the intuition that H is more complex if many H_{jk} are non-zero for *overlapping* qubit pairs and is not complex if it acts only on mutually disjoint pairs. Furthermore, due to the remarks above, we call an interaction more complex if the norms of H_{jk} are large. A complexity measure which takes both aspects into account is given as follows. First

[1]Simulation of Hamiltonians using bounded control fields are considered in [116].

define a weighted graph with the qubit numbers $1, \ldots, n$ as vertices and $c_{jk} := \|H_{jk}\|$ as weights. Two vertices j, k are connected if and only if $H_{jk} \neq 0$. As generalization of the chromatic index of a graph (i.e. the least number of colors needed to color the edges) we define [117]:

Definition 23 (Weighted Chromatic Index)
The weighted chromatic index W of a weighted graph with weights $c_{jk} \geq 0$ is given as follows. Let n_r be the chromatic index of the graph which consists of all edges (j, k) for which $c_{jk} > r$. Then

$$W := \int_0^\infty n_r dr \,.$$

Now we can straightforwardly define the complexity of a continuous algorithm as the time integral of the complexity of the interaction over the whole running time:

Definition 24 (Continuous Complexity)
Let $t \mapsto H(t)$ with $t \in [0, T]$ a continuous algorithm and $W(t)$ its weighted chromatic index at time t. Then its complexity is

$$C := \int_0^T W(t)\, dt \,.$$

In [117] we have shown that this continuous complexity measure coincides with the weighted depth, i.e., we have:

Theorem 11 (Continuous Complexity Equals Weighted Depth)
There is a continuous algorithm with complexity C which implements U if and only if there is a sequence of discrete algorithms implementing unitaries U_n converging to U such that their weighted depth converges to C.

Remarkably our theory of simulation of Hamiltonians justifies the point of view that interactions with low weighted chromatic index are easy to obtain. This is explained in the next subsection.

3.1.3 Mutual Simulation of Hamiltonians

We have already argued that qubits, in contrast to classical bits, allow 'logical' operations which change the state arbitrarily little. Hence a quantum computer may not necessarily have a 'smallest', most elementary logical operation. Therefore a set of available Hamiltonians may replace the elementary steps. According to common language of physicists, they represent 'infinitesimal' operations. In classical computer science one compares different computing devices by establishing the overhead when one device simulates the other. Therefore it is natural to compare quantum computers by asking for the complexity overhead when one Hamiltonian simulates another. But, apart from this straightforward analogy, one may use mutual simulation of Hamiltonians in order to use a quantum computer for simulating an arbitrary physical system which is not a computer.

To explain our results on these issues we follow the setting for mutual simulation of Hamiltonians in [118] which refers to the first order approximation of eq. (3.2). Let H

and $\tilde{H}$ be the Hamiltonian of the system which simulates and is simulated, respectively. Furthermore the first system is endowed with a set of some additional group $\mathcal{U}$ of unitary control operations. As already explained in Subsection 3.1.1 one can simulate every Hamiltonian in the convex span of all $UHU^\dagger$ for $U \in \mathcal{U}$ if the implementation time for each U is neglected. If $\tilde{H}$ is not in the *convex* span but in the *positive* span, i.e.,

$$\tilde{H} = \sum_j c_j U_j H U_j^\dagger \,,$$

with $\sum_j c_j = \tau > 1$, the simulation has time overhead τ since this is exactly the slow-down factor of the simulation compared to the dynamics to be simulated. In [119] one can find the theory of mutual simulation of two-qubit interactions where H and $\tilde{H}$ are Hamiltonians of the form

$$H = \sum_{\alpha,\beta} c_{\alpha,\beta} \sigma_\alpha \otimes \sigma_\beta \,,$$

where $\sigma_x, \sigma_y, \sigma_z$ denote the Pauli matrices and the set of allowed unitary operations is given by products of the form

$$U = u_1 \otimes u_2$$

acting on $\mathbb{C}^2 \otimes \mathbb{C}^2$. Then the convex optimization problem of simulating $\tilde{H}$ is worked out in detail in [119]. We considered the case where H and $\tilde{H}$ are n-qubit interactions of the form

$$H := \sum_{k<l} \sum_{\alpha\beta} J_{kl;\alpha\beta} \sigma_\alpha^{(k)} \sigma_\beta^{(l)} \,,$$

with an appropriate symmetric $3n \times 3n$ matrix J. Note that the symmetry of the coupling matrix J does not imply any physical symmetry of the interaction. It is a consequence of our redundant representation which considers ordered qubit pairs. This turns out to be very useful below. The coupling matrix J consists of 3×3-blocks, where the 3×3-matrix J_{kl} given by the block at position (k, l) describes the coupling between the qudits k and l, and the blocks $k = l$ are zero. We will henceforth characterize H by J and $\tilde{H}$ by $\tilde{J}$. The available unitaries are also products

$$U := u_1 \otimes u_2 \otimes \cdots \otimes u_n \,. \tag{3.4}$$

This model is a strongly idealized version of real NMR-physics (e.g. [120, 47]). To express the effect of the control operations on the coupling matrix J one should note that any unitary operation $u \in SU(2)$ corresponds to a rotation on the Bloch sphere via the relation

$$u^\dagger \big(\sum_\alpha c_\alpha \sigma_\alpha \big) u = \sum_\alpha \tilde{c}_\alpha \sigma_\alpha \,,$$

where the vector $\tilde{c} = (\tilde{c}_1, \tilde{c}_2, \tilde{c}_3)$ is obtained by applying a rotation $U \in SO(3)$ on the vector $c = (c_1, c_2, c_3)$. It is straightforward to verify that conjugation of H_J by $U := u_1 \otimes u_2 \otimes \cdots \otimes u_n$ corresponds to conjugation of J by a block diagonal matrix of the form

$$V := U^{(1)} \oplus U^{(2)} \oplus \cdots \oplus U^{(n)} \in \bigoplus_{k=1}^{n} SO(3) \,.$$

Hence the condition for correct simulation is given by

$$\tilde{J} = \sum_j t_j V_j J V_j^T \,, \tag{3.5}$$

where the orthogonal matrix V_j corresponds to the unitary U_j which conjugates the time evolution in the time interval j. Obviously H can only simulate $\tilde{H}$ if its interaction graph is a subgraph of the interaction graph of H; the latter connects all qudit or qubit pairs j, k for which H contains an interaction term $H_{j,k}$. This requirement is stronger than in simulations which do not use average-Hamiltonian theory but refer to higher-order terms as eq. (3.3). Then it is sufficient that H has a connected interaction graph [99]. We showed that H can simulate all $\tilde{H}$ even if not all product unitaries are allowed. Moreover, one can also restrict the set $\mathcal{U}$ to all product unitaries where each tensor component is in a sufficiently large group [121] which we called transformer groups. They are subgroups of $SU(2)$ (or $SU(d)$ for qudits) which act irreducibly in their adjoint representation $U \mapsto UAU^\dagger$ on the self-adjoint traceless operators.

We first focus on our time overhead bounds for simulating the coupling $\tilde{J}$ by J. With $\tau := \sum_j t_j$ and $p_j := t_j \tau$ one can easily check that $\tilde{J}$ is a convex sum of orthogonal conjugates of the matrix τJ. Using elementary results of linear algebra we concluded in [118]:

Theorem 12 (Lower Bound on Time Overhead)
Let J and $\tilde{J}$ be the coupling matrices of two pair-interactions H and $\tilde{H}$ among n qubits. If H simulates $\tilde{H}$ with time overhead τ then τJ majorizes $\tilde{J}$.

Obviously we know also that H must majorize $\tilde{H}$ (a condition which was used in [119] for the 2-qubit case) but this condition seems less helpful for the n-qubit case since the Hamiltonians themselves can have exponential size whereas the coupling matrices J and $\tilde{J}$ are $3n \times 3n$-matrices.

The time overhead is not the only reasonable complexity measure. The number of time steps is also important. Both measures may differ significantly [46] as will briefly be reported below. To derive bounds on the number of time steps it will also turn out to be useful to compare spectra of J and $\tilde{J}$. We restrict our attention to interactions with an additional symmetry, namely Hamiltonians of the following form

$$H := \sum_{k<l} w_{kl} \sum_{\alpha\beta} c_{\alpha\beta} \sigma_\alpha^{(k)} \sigma_\beta^{(l)} \,. \tag{3.6}$$

The matrix $W := (w_{kl})$ is a real symmetric $n \times n$-matrix with zeros on the diagonal. It describes the coupling strengths and the signs of the interactions between all qudits. The matrix $C = (c_{\alpha\beta})$ is a real symmetric 3×3-matrix characterizing the type of the coupling. This means that all qudits interact with each other via the same interaction and that only the coupling strengths and the signs vary. It is important that in this special case the coupling matrix J can be expressed as a tensor product of W and C, i. e., $J = W \otimes C$.

To derive a general lower bound for simulating arbitrary interactions $\tilde{J}$ by a tensor product interaction $J = W \otimes C$ it is useful to observe that the decisive condition in eq. (3.5) is invariant with respect to the following *rescaling* of interactions: Multiply each 3×3-block k, l of J and $\tilde{J}$ with the same factor r_{kl}. In the case that J is a tensor product

$W \otimes C$ where W has no zero entries except from the diagonal, we can therefore replace W with a matrix K which has only 1 as non-diagonal entries after we have rescaled $\tilde{J}$ in the same way.

If the coupling $\tilde{J}$ should be simulated, we obtain

$$\sum_{j=1}^{N} t_j V_j (K \otimes C) V_j^\dagger = \tilde{J}\,.$$

Set $R := \sum_{j=1}^{N} t_j V_j (\mathbf{1} \otimes C) V_j^T$. By adding R to both sides we obtain

$$\sum_{j=1}^{N} t_j V_j \big((K + \mathbf{1}) \otimes C\big) V_j^T = J' + R\,. \tag{3.7}$$

The rank of the matrix $(K + \mathbf{1})$ is 1 since all its entries are 1. The rest uses some linear algebra using arguments on the rank of both side of (3.7). After relating this back to the initial scaling we obtain (see [122] for a detailed proof):

Theorem 13 (Lower Bound on Time Steps)
Let $J := W \otimes C$ be the coupling matrix of the system Hamiltonian, $\tilde{J}$ an arbitrary coupling matrix of the interaction that is simulated, and μ the time overhead. Denote the minimal and maximal eigenvalues of C by $\lambda_{\min}$ and $\lambda_{\max}$, respectively and its rank by $r(C)$. Let I be the $m \times m$-matrix whose all entries are 1. Let s be the number of eigenvalues of $\tilde{J}/(W \otimes I)$ that are not contained in the interval

$$\mathcal{I} := [\,\mu\lambda_{\min}, \mu\lambda_{\max}\,]\,.$$

Then the number of time steps required to simulate $H_{\tilde{J}}$ by H_J is at least $s/r(C)$.

More concrete statements can be given when the $\tilde{J}$ is homogeneous, i.e., also a tensor product:

Theorem 14 (Lower Bound on Time Steps; Homogeneous Case) *Let $W \otimes C$ be the coupling matrix of the natural Hamiltonian and $\tilde{W} \otimes C$ the coupling that we want to simulate. Assume all non-diagonal entries of W are non-zero.*

1. *Let C be a positive semidefinite matrix. Then the number of time steps is at least the number of positive eigenvalues of $\tilde{W}/W$.*

2. *Let $C = diag(1, 1, \ldots, 1)$ be the $m \times m$-identity matrix. Then the number of time steps is at least $n - k$, where k is the multiplicity of the smallest eigenvalue $\mu_{\min}$ of $\tilde{W}/W$.*

3. *Let the natural coupling be $C := diag(0, 0, 1)$, i.e., we have $\sigma_z \otimes \sigma_z$ interactions between all spin-1/2-particles. Let the set of local control operations be restricted to $i\sigma_x$–transformations. Then one requires at least $n-k$ time steps with k as in Case 2. If $\mu_{\min}$ is irrational then at least n steps are necessary. In any case, $n(n-1)/2+1$ time steps are always sufficient.*

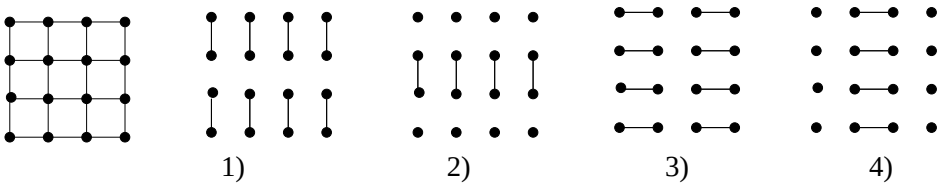


Figure 3.2: Simulation of the square lattice interaction with 4 subroutines

Interesting examples for the complexity bounds (time steps and time overhead) for the case $J = W \otimes C$ are the following:

1. **Time Inversion:** Simulate $-H$ by H. A detailed presentation is given in [46]. A relevant application is the refocussing problem in NMR. The following cases can occur:

   (a) C is traceless. For instance, the dipole-dipole coupling in NMR is described by $C = diag(1, 1, -2)$.
   All spins can be subjected to the same transformations in each time step, the number of time steps and the time overhead are at most 2.

   (b) C has negative and positive eigenvalues but $tr(C) \neq 0$. In NMR one obtains this case by a combination of dipole-dipole coupling with scalar coupling (see example 3) [123].
   The spins have to be addressed separately, the number of time steps necessarily grows for increasing n. But the time overhead does not depend on n. It depends only on the eigenvalues of C.

   (c) C is either positive or negative semidefinite, i. e., the non-zero eigenvalues have the same sign. An example in NMR is the strong scalar coupling where C is the identity.
   Then the spins have to be addressed separately, the number of time steps is at least $n - 1$, and the time overhead is also at least $n - 1$.

2. **Simulating a Square Lattice:** We consider a quantum system of $n = l^2$ spins located on a two-dimensional square lattice. For simplicity assume that l is even. We want to simulate a lattice with only nearest neighbor interactions.

   The desired interaction graph is shown on the left of Fig. 3.2. This kind of interaction can for instance be used for preparing the initial state in the 'one-way quantum computer' proposed in [23]. The eigenvalues of the corresponding adjacency matrix A are known in graph theory [124]:

   $$2\cos\left(\frac{\pi}{l+1}\,i\right) + 2\cos\left(\frac{\pi}{l+1}\,j\right), \quad i, j = 1, \ldots, l\,. \tag{3.8}$$

   We first consider the time overhead. An upper bound is given by 4 since this is the chromatic index of the graph (see, for instance, [117]). It is easy to see that the minimal eigenvalue of A is given by

   $$\lambda_{\min} = 2\cos\left(\frac{\pi}{l+1}\,l\right) + 2\cos\left(\frac{\pi}{l+1}\,l\right). \tag{3.9}$$

By Theorem 14 (Case 3) the lower bound on the number of time steps is n since the smallest eigenvalue is irrational. Note that this example shows that the complexity measures *time overhead* and *number of time steps* may differ significantly.

An upper bound on the number of time steps can be obtained as follows. The graph has $2(l-1)l$ edges. We can partition the edges into 4 sets of edges such that each set contains only disjoint interacting pairs. These 4 partitions are shown in Fig. 3.2. The simulation consists of 4 subroutines simulating one of the interactions in one of the 4 classes. For each subroutine we choose selective decoupling schemes (which will be described in more detail in Subsection 4.3.1) using Hadamard matrices. Since the numbers of cliques are $l^2/2$ or $l^2/2 + l$ in each subroutine, the square lattice graph can always be simulated in $O(l^2) = O(n)$ time steps.

Generalizations of the theory for qudits, i.e., higher dimensional units of quantum information, can be found in [101]. The common idea of all the bounds above is that the spectrum of the adjacency matrix tells us to what extent the simulation schemes can be parallelized. Remarkably, the continuous model of quantum computing lead to discrete combinatorial parallelization problems which are strongly analogous to parallelization problems in usual computer science.

The theory above shows that there are rather similar-looking interactions where one is more powerful. Consider for instance the two n-qubit Hamiltonians

$$H_{\pm} := \sum_{j,k} \sigma_x^{(j)} \sigma_x^{(k)} \pm \sigma_z^{(j)} \sigma_z^{(k)} .$$

The time overhead to simulate H_+ by H_- is at most 2 since one could cancel all z-interactions without time overhead. For half of the time period we convert this xx-interaction into a zz-term. But simulating H_- by H_+ requires at least time overhead $(n-1)/2$. The argument is that H_- can simulate $-H_-$ without any time overhead. Given a simulation of H_- by H_+ with time overhead τ, we can construct a time inversion scheme for H_+ with time overhead 2τ. On the other hand, H_+ needs time overhead $n-1$ to simulate its own inverse because it belongs to case (c).

Another message of the complexity theory of mutual simulation of Hamiltonians is the following. Given a fixed interaction, one may define a complexity of the topology of the interaction graph of another interaction. This can most nicely be demonstrated if H is a zz-interaction between all qubits and we only want to cancel the interaction between some qubit pairs. Let A be the adjacency matrix given by the remaining interaction graph. Then the theory above implies (for details see [118]) that the time overhead is the least number $\tau > 0$ such that the vector $(n-1, -1, -1 \ldots, -1)$ majorizes the spectrum of A. In particular, τ is at least the modulus of the smallest eigenvalue of A. On the other hand, the selective decoupling techniques explained in Subsection 4.3.1 allow one to cancel (without time overhead) all interactions between cliques of qubits after one has chosen an arbitrary partition of the qubits into arbitrarily many cliques. By choosing cliques which consist of at most 2 qubits it follows that an upper bound on the time overhead is given by the chromatic index of the simulated interaction graph, i.e., the number of colors required to color its edges. In Fig. 3.3 we show simple and complex coupling topology according to the time overhead.

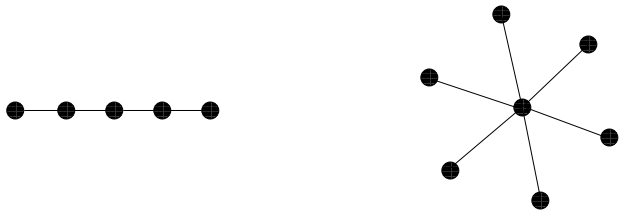


Figure 3.3: Interactions with 'simple' and 'complex' topology: The linear chain has chromatic index 2 and the star n for n exterior nodes

This lower bound on the simulation time overhead generalizes even to the *weighted chromatic index* (see Definition 23) when the given interaction is a zz-interaction between all pairs and with equal strength ('complete zz-Hamiltonian') and an arbitrary pair-interaction is simulated.

Theorem 15 (Relevance of Weighted Chromatic Index)
The time overhead to simulate an arbitrary pair-interaction on n qubits by the complete zz-Hamiltonian is at most its weighted chromatic index.

This result (shown in [118]) gives not only a further justification to our definition of weighted chromatic index, it supports also the *integrated* chromatic index (Definition 24) and the weighted depth (see Theorem 11) as reasonable complexity measures:

Theorem 16 (Relevance of Weighted Depth)
The time to implement a unitary U by the complete zz-Hamiltonian is at most the weighted depth of a discrete quantum circuit implementing U.

It is clear that the complete zz-Hamiltonian is unphysical since interactions between distant qubits will generally be smaller. However, qualitative ideas of the theory above apply also for interactions with deceasing strength between distant qubits. Partly we have stated those generalizations already above.

The average Hamiltonian model is of course only an approximation of the time evolution when the natural dynamics is interspersed by external unitary control. To generate an arbitrary n-qubit or n-qudit Hamiltonian in an optimal way from local unitaries and a given interaction is a hard problem. Even the derivation of bounds is difficult when the first order approximation is dropped. For the special case of $n = 2$ qubits the optimal implementation can be found in [125]. In [126] it was shown that the transport of quantum information from the qubit at one end of a 2-qubit chain to the qubit at the other end is faster by directly using the interaction compared to the usual approach where the interaction is selectively decoupled in order to implement two-qubit gates. For n qubits one can at least give some lower bounds [127, 128] on the implementation time for general unitaries.

3.1.4 Adiabatic Quantum Computing

Here we sketch a proposal for a type of continuous quantum algorithms which had attracted broad interest (see e.g. [129, 130, 131]). It shows that continuous algorithms do

not necessarily just simulate discrete circuits to solve a computation problem. Therefore we mention it in order to show how general any reasonable definition of 'algorithm' has to be in order to include only every present day proposal for quantum computing. Furthermore it is a nice potential application for mutual simulation of Hamiltonians. As in subsection 3.1.1 an adiabatic algorithm is also described by a time dependent Hamiltonian

$$H(t) := (1 - \frac{t}{T})H_B + \frac{t}{T}H_P \,.$$

The essential point is that H_B is here some Hamiltonian with knwon ground states which are easy to prepare; whereas the ground state of H_P is not known and encodes the solution of a problem. The idea is that the problem can even be NP-complete. Provided that $H(t)$ is changed sufficiently slowly compared to the inverse of the energy gaps between ground state and first excited states for all $H(t)$, the quantum adiabatic theorem says that the system will at any time be with good reliability in its ground state. Whether this 'adiabatic change' condition could imply a running time which is not better than the running time of 'conventional' algorithms has been an issue of controversial discussions [131]. This is not our subject.

It is known [132] that the ground state of the following H_P encodes the solution of the NP-complete problem max independent set. We recall [92]:

Definition 25 (Max Independent Set)
The NP-complete problem 'Independent Set' reads: Given a graph $G := (V, E)$ and a positive integer $K \leq |E|$. Does G contain an independent set of size K or more, i.e. a subset $V' \subset V$ such that $|V'| \geq K$ and such that no two vertices in V' are joined by an edge in E?

We define

$$H_P := \sum_{(k,l)\in E} \sigma_z^{(k)}\sigma_z^{(l)} + \sum_{k\in V} \sigma_z^{(k)} \,.$$

The problems are (1) that H_P may involve interactions between arbitrarily distant qubits with equal strength and (2) that the solution of every problem instance requires a different interaction. It seems hence unsatisfying to realize such Hamiltonians as *hardware*. Therefore we considered in [133] the problem of simulating Hamiltonians which could solve NP-complete problems. We modified H_P to $\hat{H}_P$ such that it (1) involves only nearest neighbor interactions on a 2-dimensional grid (a so-called orthogonal planar embedding) and (2) solves still MAX INDEPENDENT SET. Then $\hat{H}_P$ is simulated by a nearest neighbor interaction which is either already given by nature or by cancelling unwanted interactions between more distant qubits using the scheme in the second example of Subsection 3.1.3.

3.2 Non-Computational Problems

The specification of algorithms for non-computational problems contains statements which do not only refer to its 'input' and its 'output'. Optimal thermodynamic machines should, for instance, process the information in such a way that no information is transfered to the environment since this would imply energy loss. This requirement is only included in

the specification of the demanded output if the whole 'state of the universe' is considered as output. One should rather consider it as an additional requirement which refers to neither the input nor to the output. Note that the 'no-information to the environment'-condition is not necessarily a restriction to the set of available operations but rather refers to the whole algorithm since information could be stored on an additional memory and erased afterwards. Requirements like this can pose a considerable difficulty in finding algorithms for non-computational problems as the next subsection shows.

3.2.1 Difference between Implementing and Computing a Boolean Function

First we define formally what it means to calculate a boolean function

$$f : \{0,1\}^n \to \{0,1\}^k$$

on a quantum computer. We define such a calculation as a unitary transformation U on a register of $n+m$ qubits such that the partial trace of

$$tr_{n+m-k}(U|a\,0\ldots0\rangle\langle a\,0\ldots0|U^\dagger) \tag{3.10}$$

is the state

$$|f(a)\rangle\langle f(a)|\,.$$

If $n=k$ and f is known to be bijective one could also demand that U satisfy

$$U|a\rangle\langle a|U^\dagger = |f(a)\rangle\langle f(a)|\,, \tag{3.11}$$

or the weaker condition

$$U|a\,0\ldots0\rangle\langle a\,0\ldots0|U^\dagger = |f(a)\,0\ldots0\rangle\langle f(a)\,0\ldots0|\,. \tag{3.12}$$

From the computational point of view, it does not make sense to demand (3.11) or (3.12). However, from the thermodynamical point of view the difference to (3.10) is important. Restoring the ancilla qubits in (3.10) after the implementation of U requires energy resources according to Landauer's principle (see Section 2.2). There could, in principle, also be cryptographic reasons to prefer (3.11) or (3.12) to (3.10) because the erasure operation transfers information to the environment which could be useful for a potential eavesdropper.

In order to explain why there is a difference from the point of view of complexity theory whether one demands (3.11) or (3.12) in contrast to demanding only (3.10), we rephrase some ideas from the theory of reversible computation [70, 71]. Assume one wants to compute $f(x)$ for an arbitrary boolean function from n to k bits. Then one can first decompose f into elementary boolean gates like AND, NAND, NOR, OR, NOT. Except from the NOT gate they are not bijective and therefore not allowed for reversible computation. However, they can be embedded into reversible functions [10]. One example of such a reversible extension of a NAND gate by a Toffoli gate is given in Fig. 3.4. Using such a reversible embedding one obtains a bijective function

$$F : \{0,1\}^m \to \{0,1\}^m$$

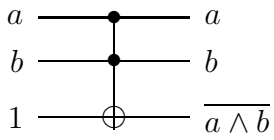


Figure 3.4: Simulation of a NAND gate by a TOFFOLI gate.

such that

$$F(0, \ldots, 0, x) = (f(x), y)\,,$$

where y is some data garbage on the remaining $m-k$ qubits. To remove this garbage, one can copy the result $f(x)$ to k additional ancilla qubits and implement F^{-1} afterwards. Unfortunately this trick does not help if we want to implement $|a\rangle \mapsto |f(a)\rangle$ since this requires not only the removal of the garbage but also the initial state. The method above implements

$$|0 \ldots 0\rangle \otimes |a\rangle \otimes |b\rangle \mapsto |0 \ldots 0\rangle \otimes |a\rangle \otimes |b \oplus f(a)\rangle\,.$$

where $|0 \ldots 0\rangle$ denotes additional ancilla qubits. They are used during the computation in order to simulate boolean functions like NAND using TOFFOLI-gates and restored afterwards. The symbol $\oplus$ denotes bitwise XOR.

Now assume that we have also a classical boolean circuit consisting of NAND gates that computes f^{-1}. We could use the same scheme as above to implement

$$|0 \ldots 0\rangle \otimes |a \oplus f^{-1}(c)\rangle \otimes |c\rangle$$

with a possibly different number of ancillas. We can use this algorithm to erase $|a\rangle$ in the middle register since we obtain $|a \oplus f^{-1}(f(a))\rangle = |0\rangle$. Using a SWAP-operation between the middle and the right register we obtain

$$|0 \ldots 0\rangle \otimes |a\rangle \otimes |0 \ldots 0\rangle \mapsto |0 \ldots 0\rangle \otimes |f(a)\rangle \otimes |0 \ldots 0\rangle\,.$$

This shows that an implementation of f on n qubits is possible using representations of f and f^{-1} as boolean circuits (in [70] one can find similiar arguments in the context of reversible Turing machines). It is easy to see that the reverse is also true: Given a sequence of unitary transformations which implement transformation

$$|a\rangle \mapsto |f(a)\rangle$$

we can certainly reverse the order of the gates and replace each gate by its inverse in order to have an implementation of

$$|a\rangle \mapsto |f^{-1}(a)\rangle\,.$$

Roughly speaking we conclude:

Theorem 17 (Implementing f is as Hard as Computing f and f^{-1})
Given a bijective functions on n bits. To find a gate sequence which 'implements f' in the sense that it performs a unitary operation U on $n+k$ qubits for some k such that

$$U(|a\rangle \otimes |0\rangle^{\otimes k}) = |f(a)\rangle \otimes |0\rangle^{\otimes k}$$

is as hard as finding classical boolean circuits which compute f and others which compute f^{-1}.

Implementations of bijective boolean functions will play a crucial role in the theory of algorithmic thermodynamic machines like refrigerators and heat engines in Section 4.2. For these purposes we have used computer algebra systems to decompose a given permutation on $\{0,1\}^n$ into elementary permutations induced by TOFFOLI and C-NOT and NOT gates.

3.2.2 Algorithms with Quantum Input or Output

Quantum algorithms like Shor's and Grover's receive classical strings as input and generate classical strings as solutions. Subroutines of such algorithms do not necessarily work with classical in- and output. Consider, for example the phase estimation procedure. Given an eigenstate of a unitary U it computes the corresponding eigenvalue. Apart from using it as a building block in algorithm, phase estimation is therefore also a measurement procedure for the projector-valued measurement (P_j) defined by the spectral projections of U. In other words, measurement algorithms are an example for algorithms with *quantum* input. This input cannot necessarily be described by classical variables with reasonable amount of resources since the description of an n-qubit quantum state up to a certain accuracy (with respect to the Hilbert space norm) requires an exponential number (in n) of classical bits. But sometimes a measurement is also an algorithm with relevant quantum *output*. This is, for instance, the case if measurements should be implemented in the sense of the Lüders postulate (see Subsection 2.1.1). In [40] we have shown that the additional demand that a measurement should *project* into the eigenspaces of the measured observable can increase the complexity of a von-Neumann measurement. Consider for instance the n-qubit observable

$$Z := \sigma_z^{\otimes n} .$$

If only the measurement result matters, one could simply measure Z by measuring every qubit and computing the parity of the obtained string. To reduce a von-Neumann measurement of Z to a single-qubit measurement in order to obtain 2^{n-1}-fold degenerated eigenspaces one could implement a unitary U such that

$$UZU^\dagger = \sigma_z \otimes \mathbf{1}_{n-1} .$$

It follows from our results in [40] that circuits which implement Z in the sense of a projection onto its 2 eigenspaces have at least the depth

$$k \geq \frac{1}{2}(1 + \log_2 n) .$$

To show this we define the macroscopic observable (see Subsection 2.1.2):

$$\overline{\sigma_x} := \sum_j \sigma_x^{(j)} ,$$

where $\sigma_z^{(j)}$ is the Pauli matrix σ_α acting on qubit j. We observe that its commutator with Z satisfies

$$i[\overline{\sigma_x}, Z] = 2\overline{\sigma_y} .$$

We can apply inequality (2.2) to the projection $P := (Z + \mathbf{1})/2$ and find

$$\| [\overline{\sigma_x}, P] \| = \frac{1}{2} \| [\overline{\sigma_x}, Z] \| = 1 \,,$$

which implies

$$k \geq \frac{1}{2}(1 + \log_2 n) \,.$$

Note that this gives a logarithmic lower bound for the complexity of the von-Neumann measurement, whereas the arguments show that a measurement which does not care about the output state is given by independent single-qubit measurements. This demonstrates that the specification of the post-measurement state does indeed change the complexity of the measurement algorithm.

An algorithm with classical input and quantum output is, for instance, a state preparation procedure. The cooling problem, i.e., to prepare the ground state of a physical system with good reliability is an example for a specification of such an algorithm.

Note that the goal of thermodynamic reversible computation defines a problem which is somewhat at the border between a computation problem and a non-computational control problem. The latter consists of implementing the computation steps in such a way that no information is copied to the environment. The specification of this type of algorithm therefore does not only contain statements on its logical in- and output but also on the in- and output of physical resources.

Chapter 4

Algorithmic Approach to Natural Non-Computational Problems

In the context of quantum computing many non-computational problems like state preparation and measurements are treated in the literature. However, the usual intention is to use these procedures as building blocks of a quantum computation or for quantum cryptography. Here we want to describe problems which arise from other applications. Nevertheless the discussions will remain on an abstract level; for presentations of the problems which describe experimental facts we refer to the literature. We only want to describe briefly why many algorithmic control problems stem from interesting (potential or real) applications.

4.1 Measurements

The term 'measurement' is usually used for a process which has numerical values as outcomes describing the state of a system. There is no doubt that implementing this type of measurements in molecular systems is an important task: measuring the state of an atomic clock in order to get information about the actual time [134] is only one example. Another example would be measurements of the energy of a molecule which give insight in its level structure. However, the abstract concept of quantum measurements introduced already in the 70s [28] is general enough to include non-numerical outcomes. It describes every process where the interaction of a quantum object with some apparatus generates a visible effect on the macroscopic scale. The measurement outcome could therefore also be an *image* that is obtained when particles interact with a screen. Algorithmic measurement theory tries to use a set of available operations to ensure that the interaction between system and apparatus extracts the relevant information about the quantum system in an optimal way.

To develop a complexity theory of measurements is an ambitious goal. First of all one has to ask what the *elementary* measurements are. But it is by no means clear which observables can be 'directly' measured. Furthermore, one could object that a complexity theory of observables would necessarily strongly depend on the considered physical system since there is no canonical correspondence between observables in different physical systems. However, if one would therefore claim that a complexity theory of observables would not make sense one could not accept a complexity theory for computational prob-

lems as being device-independent either: the latter relies on assumptions about which observables are easy to measure. Given an appropriate measurement, one could even save the whole computation. In other words, lower bounds on the computational complexity require assumptions on the complexity of measurements or restrictions on the set of available measurements. We will be content to (1) give algorithms for certain types of interesting measurements and (2) connect complexity of observables with complexity of computational problems.

4.1.1 Von-Neumann Measurements and their Complexity

In the standard model of the quantum computer the only possible measurements are given by reading out single qubits. There is no clear physical justification for this restriction. On the other hand, it would trivialize all algorithms if one could implement any arbitrary measurement, since the implementation of a unitary circuit U followed by a diagonal observable D (with respect to the computational basis) could be replaced with a measurement of $U^\dagger DU$. This shows that lower complexity bounds for solving a certain computational problem can only exist where the set of observables is either restricted or there is a notion of "complexity of observables". It is clear that every measurement of A can in principle be implemented by diagonalizing A (see Subsection 1.2.1) whenever the post-measurement state does not matter. Simpler algorithms can be expected if additional ancillas are allowed.

These remarks relate the complexity for measuring a self-adjoint observable A with the complexity of the *diagonalizing* transformation U. But there is also another simple idea that relates complexity of measurements for self-adjoint operators A to the complexity of the unitaries $\exp(-iAt)$ generated by A. To explain this well-known idea [135] we start by reducing measurements to unitaries in analogy to the observations in using the phase-estimation procedure [136]:

Algorithm 1 (Measuring A using exp(iAt))
Given descriptions of quantum circuits $U_j := \exp(i\epsilon T A 2^j)$ for $j = 1, \ldots, k$ where $T \in [-\pi/\|A\|, \pi/\|A\|)$.

1. *Initialize an ancilla register with k qubits in the equal superposition of all binary words.*
$$\frac{1}{\sqrt{2^k}} \sum_b |b\rangle = \frac{1}{\sqrt{2^k}} (|0\rangle + |1\rangle)^{\otimes k} .$$

2. *Implement a controlled $-U_j$ conditioned at the state of the j-th ancilla.*

3. *Apply the Fourier transform to the ancilla register.*

4. *Measure the logical state of the ancilla register. If the first register is in an eigenstate of A with eigenvalue λ the result l satisfies*
$$|\lambda T - l| \leq \epsilon \tag{4.1}$$

with at least probability

$$1 - \frac{1}{2(\epsilon - 1)} .$$

Hence we get a binary digit representation of the corresponding eigenvalue as measurement result. Due to inequality (4.1) the accuracy increases exponentially in the number of ancillas. Note that the algorithm does not work with black-box implementations of U_j. It needs the circuit descriptions in order to build the controlled-U_j gates by replacing each single gate V with a controlled-V gate. For the case that A is an unknown pair-interaction Hamiltonian I have in [114] given a 'continuous time algorithm' (in the sense of Subsection 3.1.1) to convert $\exp(-iAt)$ into a controlled-$\exp(-iAt)$ evolution.

An interesting question arises as to how the required resources increase with the demanded accuracy. The error decreases exponentially in the number of ancilla qubits, and the running time of the Fourier transformation increases only as $O(k^2)$. The bottleneck comes from the implementation of $\exp(-iAT2^j)$ for $j = 1, \ldots, k$ typically requiring a running time which is exponentially increasing in k. One may interpret the result above by stating that measurements of A up to an accuracy ϵ have the time complexity at most

$$O(1/\epsilon^2) + \sum_{j \leq -\log_2 \epsilon} C_j$$

where C_j is the time complexity for implementing $\exp(iA2^j)$.

To explain the converse statement which relates the complexity of observables with the complexity of unitaries[1], we assume first that we have given a so-called pre-measurement for an n-qubit observable A with some natural numbers of eigenvalues. We assume further that it is given by a unitary operation U on $n + k$ qubits such that

$$U(|\psi\rangle \otimes |0 \ldots 0\rangle) = \sum_b P_b |\psi\rangle \otimes |b\rangle \quad \forall |\psi\rangle \in (\mathbb{C}^2)^{\otimes n} ,$$

where P_b are the spectral projections of A corresponding to the eigenvalue b. Given this subroutine the following algorithm implements $\exp(iAt)$ for any desired t:

Algorithm 2 (Implementing exp(iAt) using A-pre-measurements)

1. *Initialize the k ancilla qubits to $|0 \ldots 0\rangle$.*

2. *Implement U on the $n + k$ qubit register.*

3. *Implement*

$$\exp(it2^j)|1\rangle\langle 1| + |0\rangle\langle 0|$$

on the jth ancilla qubit of the ancilla register. On the whole ancilla register this is the phase shift $|b\rangle \mapsto \exp(itb)|b\rangle$ for $0 \leq b < 2^k$.

4. *Implement $U^\dagger$.*

[1]This direction should also be well-known even though I did not find an explicit remark in the literature.

Of course, a measurement apparatus is not necessarily a quantum computer. We now assume that it is some physical mechanism which is efficient in any reasonable sense. Then the reduction above is not directly possible since we need a coherent implementation of a pre-measurement instead of a real measurement which contains an irreversible destruction of superpositions. Furthermore we also need an implementation of the reverse process $U^\dagger$.

However, if the strong quantum Church-Turing Thesis holds in a sufficiently general sense, these requirements should be possible to fulfill: without loss of generality we assume that the results of our measurement are displayed in the form of a binary number since it should always be possible to convert any other display of a measurement instrument into such a form. Then there should exist an efficient simulation of the whole process on a quantum computer such that the digits of the display are represented by some qubits. But this simulation would then already provide the pre-measurement above! On the quantum computer we can also efficiently invert the simulation after we have given an appropriate phase to the display qubits. Note that the complexity of implementing $\exp(-itA)$ does not depend on t in the example above. This is not surprising, since the dynamics $\exp(-itA)$ is periodic as it has integer spectrum. One could object that this example is also specific in another respect. It can neither be expected that a realistic measurement projects perfectly onto the eigenspaces nor that its simulation on a quantum computer does. This inaccuracy will be increasingly relevant for increasing t. If the spectral values of A consist of a finite set of rational numbers we could rescale and shift it to obtain natural numbers as eigenvalues. The number of required ancillas would then depend on the least common divisor of the original spectrum. For an observable A with the spectrum $\lambda_1, \ldots, \lambda_{2^n}$ if any efficient implementation of

$$|l\rangle \mapsto \exp(i\lambda_l t)|l\rangle$$

is available one could also simulate the measurement in such way that the lth eigenvalue corresponds to the lth state of the k ancilla qubits. This shows that exact pre-measurements of A together with appropriate diagonal unitaries can simulate long-time behavior of a system with Hamiltonian A in such a way that the running time of the simulation would not increase with the running time of the dynamics. This surprising feature already suggests that *precise* von-Neumann measurements can be hard. We will return to this issue later in this subsection.

The relation between the complexity of implementing the unitaries

$$\exp(iAt)$$

and a von-Neumann measurement of A appears also in a control-theoretic setting. In [48] we have considered a bipartite quantum system, the *controller* and the the *system*. Their Hilbert spaces are denoted by $\mathcal{H}_C$ and $\mathcal{H}_S$, respectively. Then we assume that a fixed interaction Hamiltonian H on $\mathcal{H}_C \otimes \mathcal{H}_S$ is given and that this interaction is the only possible medium to access the system S. The only way to implement transformations and measurements in this model is to operate on the controller. On the other hand, we assume that we can implement arbitrary unitary transformations and von-Neumann measurements on the controller and that these operations take an negligible amount of time. One of the main intention of this model was to show that H determines which measurements on S are simple and which ones require complex control sequences. If one

allows some unitary transformations or measurements on S, the complexity of observables depends strongly on this set of control operations. Here we want to show that in this model with optimal access to the controller the interaction between system and controller 'supports' some unitaries and some observables on $\mathcal{H}_S$ and that there is a tight relation between both. We can always write H as

$$H := \sum_j A_j \otimes B_j \,,$$

with self-adjoint operators A_j, B_j such that the set (A_j) is linearly independent and the set (B_j) is, too. For simplicity we have assumed that all A_j, B_j can be chosen to be traceless. The idea to use this interaction for measurements of an observable D is to simulate an interaction of the form

$$C \otimes D \tag{4.2}$$

with some operator C. This interaction allows (1) measurements of D without disturbing the eigenstates of D during the measurement and (2) implementation of the transformations $\exp(iDt)$. The latter is done by initializing the controller to some eigenstate of C. To use $C \otimes D$ for a D-measurement we choose eigenvectors $|\psi_0\rangle$ and $|\psi_1\rangle$ of C with different eigenvalues λ_0, λ_1 (this is always possible because C is traceless) and consider their span as the state space of a qubit. The effect of the interaction (4.2) on this qubit is equivalent to the effect of a controlled-$\exp(-iDt(\lambda_1 - \lambda_0))$ gate after the time t. Now we assume that the remaining space of $\mathcal{H}_C$ is sufficiently large to perform a measurement of D by Algorithm 1. Note that this measurement does not change the state of the system when it is in an D-eigenstate even during the procedure is running. This is a specific feature of this measurement scheme. If one allows state changes on the system during the procedure, the complexity of any arbitrary observable is bounded from above by two times the complexity to implement a SWAP between $\mathcal{H}_S$ and an isomorphic subspace of $\mathcal{H}_C$. Similarly, the time of this SWAP operation is an upper bound for the implementation time of any $\exp(-iDt)$. Only if we restrict the attention to algorithms which do not change the eigenstates of D at all, we need to simulate an interaction $\tilde{H}$ which allows a continuous evolution $\exp(-iDt)$ on the system. The question is for which D we can simulate the interaction (4.2) easily and for which it requires complex control sequences.

1. **First Order Simulation:** Let D be in the span of all B_j. For $D = \sum_j c_j B_j$ we have to choose a linear map L on the real vector space of traceless self-adjoint operators such that $L(A_j) = c_j D$. We can always find [115] a set of unitaries $U_1, \ldots, U_k$ and positive numbers $t_1, \ldots, t_k$ such that

$$\sum_l t_l U_l A_j U_l^\dagger = c_j C \;\; \forall j \,.$$

This clearly defines a simulation scheme that transforms H into

$$\sum_j \sum_l t_l U_l A_j U_l^\dagger \otimes B_j = \sum_j c_j C \otimes B_j = C \otimes \sum_j c_j B_j \,.$$

The space of observables spanned by all B_j plays therefore a distinguished role in our setting.

2. **Higher Order Simulation:** By the remarks above it is clear that one can simulate $A_j \otimes B_j$ for each j in first order approximation schemes. Using the anti-commutator $\{F, G\} := FG + GF$ we have

$$[A_j \otimes B_j, A_k \otimes B_k] = \frac{1}{2}([A_j, A_k] \otimes \{B_j, B_k\} + \{A_j, A_k\} \otimes [B_j, B_k]) .$$

The remarks above suggest that not only commutators of B_j but also anti-commutators can be simulated (for details see [48]). By concatenating Lie-brackets in this way we find simulation schemes for $C \otimes D$ for every D in the *matrix algebra* (and not only in the *Lie-algebra*) generated by all B_j.

This gives an interesting structure to the set of observables as we will briefly sketch. Let the system S consist of n qubits and the interaction between system and controller be

$$H := \sum_{j,\alpha} A_{j,\alpha} \otimes \sigma_\alpha^{(j)} ,$$

where $\alpha_\alpha^{(j)}$ denotes the Pauli matrix σ_α acting on qubit j. We said that all observables D in the span of the operators $\sigma_\alpha^{(j)}$ allow a first order simulation. This includes, for instance all single qubit observables as well as the 'macroscopic' observables (see Subsection 2.1.2)

$$\frac{1}{n} \sum_l \sigma_\alpha^{(l)} .$$

This explains why both types of observables are very natural ones. To implement a measurement for products of single qubit observables like $\sigma_\alpha^{(j)} \sigma_\beta^{(k)}$ by simulating an interaction of the form $C \otimes \sigma_\alpha^{(j)} \sigma_\beta^{(k)}$ requires a second order simulation scheme using commutators. To implement measurements for higher order spin correlations uses even higher order approximations in this setting. This gives an intuition about how the interaction supports measurements of some observables more than measurements of others.

Now we explain how the complexity of observables could be compared to complexity classes of computational problems. A natural non-trivial observable is the energy of an interacting many-particle system. The fact that the long-term dynamics of such a system is hard to predict, already suggests that precise energy measurements may be hard since the remarks above show that both complexities are related. Let us consider a k-local n-qubit Hamiltonian H as a model for a natural interaction. Recall that k-locality means that H consists of operators acting on k qubits. Efficient *approximative* measurements for k-local observables can be obtained by Algorithm 1 since k-local dynamics have efficient simulations due to [89]. We have mentioned that the complexity increases with the desired accuracy but it was not clear whether this shortcoming is specific to this particular algorithm or whether accurate measurements require in general long running times. For an unknown Hamiltonian H it is certainly true that accurate measurements of H require a long measurement time since two Hamiltonians with small operator-norm distance can only drive a state into mutually orthogonal states after evolving the state for a sufficiently long time. Clearly this argument does not hold for measurements which are specified by the classical description of an observable. Otherwise one could not measure any observable in a degenerate two-level system. But the following arguments suggest

that accurate measurements of k-local observables are already hard for $k = 4$: In [137] we have shown that precise measurements for $k = 4$ correspond already to the complexity class PSPACE. Before we rephrase our main result we must specify more clearly what we mean by 'accuracy of a measurement'. We do not need a precise definition of accuracy, but only demand that the following condition is satisfied:

Postulate 1 (Measurement Accuracy)
Let A be a self-adjoint operator on a finite dimensional Hilbert space and $\sum_j \lambda_j P_j$ its spectral resolution where λ_j are the eigenvalues and P_j are the corresponding spectral projections.

A measurement of A with accuracy $\Delta\lambda$ has the following property: For all density matrices ρ the probability to obtain an outcome in the interval $I := [\lambda_j - \Delta\lambda, \lambda_j + \Delta\lambda]$ is at least $(3/4)\,\mathrm{tr}(\rho P_j)$.

The results below are not sensitive to the particular definition of accuracy. However, it is convenient to work with the formulation above. We found [137]:

Theorem 18 (Measuring 4-Local Observables is PSPACE-Hard)
Let M be a hypothetical machine with the following properties:

1. *M receives a classical description of a 4-local n-qubit observable A.*
2. *M receives a state $|\psi\rangle$ on a quantum register with state space $(\mathbb{C}^2)^{\otimes n}$.*
3. *M implements an A-measurement on $|\psi\rangle$ and gives the result as output*
4. *The accuracy of the output is sufficient to distinguish all different eigenvalues of A.*

Then M can be used to solve any decision problem in the complexity class PSPACE in polynomial time.

To prove this we need a characterization of PSPACE with respect to quantum circuits. In particular, we need the result that every PSPACE language can be recognized by applying an appropriate circuit many times. This is stated by the following lemma.

Lemma 3 (Solving PSPACE with one Circuit)
For every language L in PSPACE there is a polynomial-time uniformly generated family of quantum circuits $(V_l)_{l\in\mathbb{N}}$, l denoting the length of the input x such that the following conditions hold:

Each V_l consists of $s_l = poly(l)$ elementary quantum gates and acts on $m_l = poly(l)$ many qubits. The circuit V_l decides whether an input string x is an element of L in the following sense.

There is a polynomial-time computable natural number r_l such that the r_l-fold concatenation of V_l solves the corresponding PSPACE problem, i.e.

$$V_l^{r_l}(|x\rangle \otimes |y\rangle \otimes |00\ldots0\rangle) = |x\rangle \otimes |y \oplus f(x)\rangle \otimes |00\ldots0\rangle\,,$$

where f is the characteristic function of L, that is $f(x) = 1$ if $x \in L$ and $f(x) = 0$ otherwise. The vector $|x\rangle$ is the basis state given by the binary word $x \in \{0,1\}^l$, the

vector $|y\rangle$ is the state of the output qubit and $|00\ldots0\rangle$ is the initial state of $m_l - l - 1$ ancilla qubits.

Moreover, the circuits V_l can be chosen such that they only permute computational basis states.

The details of the proof can be found in [137]. Let us only sketch the idea. By definition, every problem in the complexity class PSPACE can be solved on a classical Turing machine using a polynomial number of bits. It is furthermore possible to replace the initial Turing machine by a reversible one at the cost of some polynomial space overhead [138]. The operations of the latter Turing machine can be simulated by a quantum circuit acting on a polynomial number of qubits when the position of the simulated head is represented by the state of an addition register *tape index*. But then the problem arises that the concatenated application of the same circuit can never simulate a *terminating* computation[2]. To overcome this difficulty we observe that there is an upper bound on the running time from the number of states due to the bounded tape. Instead of simulating the termination, the circuit U performs idle steps which consist in incrementing a counter until the upper bound on the running time is reached. Hence the formal end of the computation is after a fixed number of steps. Unfortunately we demand that U has to restore the initial state (except for the output qubit). Therefore it has to count backwards on the idle step counter and reverse the whole computation. The switching between these 4 operation modes (computation, idle counts, reverse idle counts, reverse computation) has to be done by the circuit itself. The whole circuit V can be seen in Fig. 4.1. The circuit U simulates the behavior of the Turing machine. It acts on the cells (each consisting of a few qubits) of the Turing machine, the head and an accumulator, and is controlled by the operation mode. The circuits INC and DEC increase or decrease the counters, the lower counter counts always forward or backwards. The idle counter is activated as soon as the circuit U is finished. It is decreased afterwards. Now we come to the main idea of the proof of Theorem 18:

1. Given a binary function f in the complexity class PSPACE. Construct a quantum network V which computes f after it is applied r times to the input string x:
$$V^r(|x\rangle \otimes |y\rangle \otimes |0\cdots0\rangle) = |x\rangle \otimes |y \oplus f(x)\rangle \otimes |0\cdots0\rangle$$
2. If $f(x) = 1$ the initial state is restored after $2r$ applications, if $f(x) = 0$ it is restored after r applications.
3. The decomposition of $|x\rangle \otimes |y\rangle \otimes |0\cdots0\rangle$ with respect to the eigenspaces of V depends on $f(x)$ due to the different period length r or $2r$.
4. Construct an observable A such that the spectral decomposition of $|x\rangle\otimes|y\rangle\otimes|0\cdots0\rangle$ with respect to A coincides (almost) with the decomposition with respect to V.

Now we shall sketch the idea how A can be constructed based on the circuit V above. Let $U_1, \ldots, U_l$ be a sequence of two-qubit gates implementing V. Set
$$A := \sum_j V_j \otimes U_j + V_j^\dagger \otimes U_j^\dagger$$

[2]analogue to 'Halting-problems' in the context of other chapters

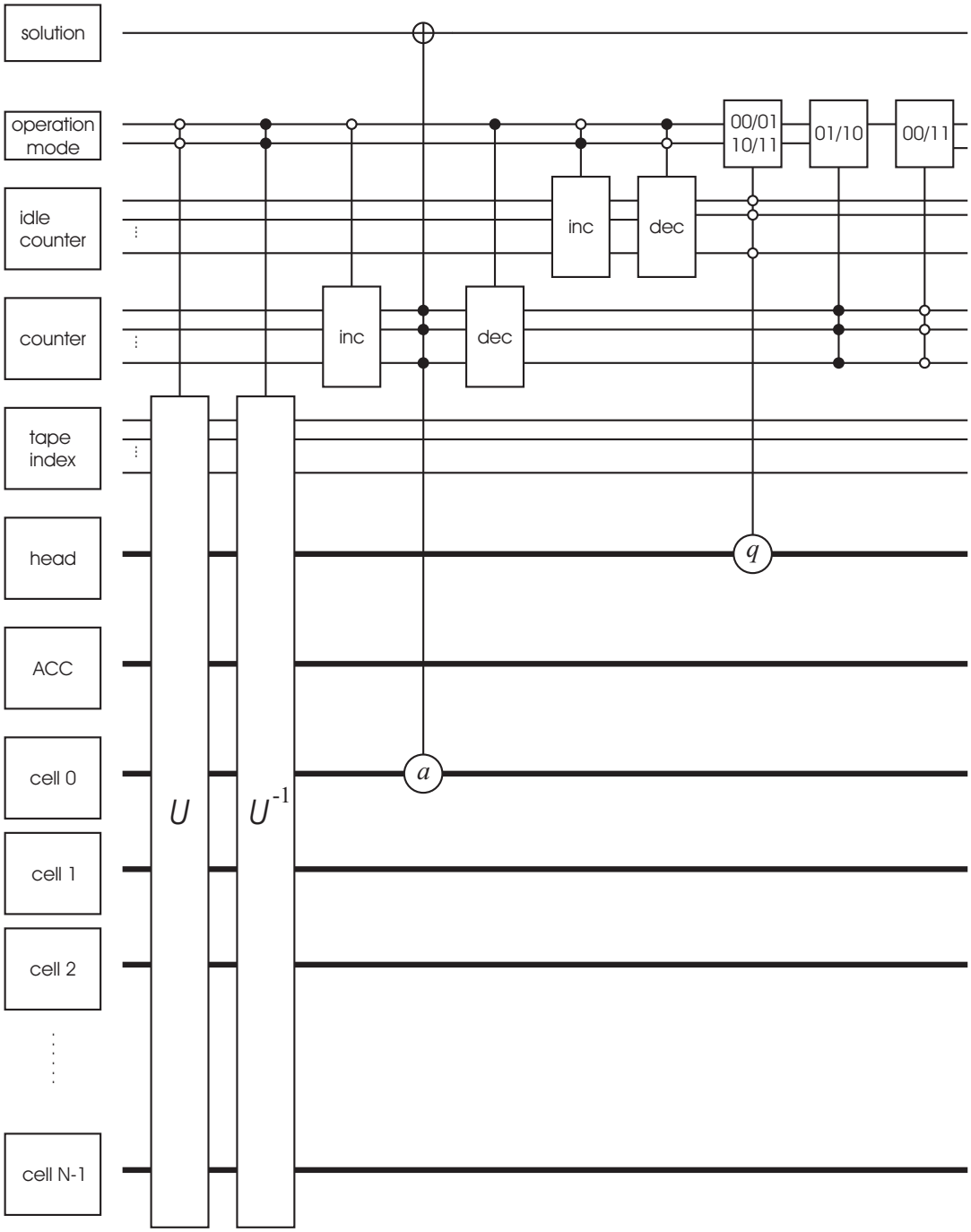


Figure 4.1: Quantum circuit computing $f(x)$ using a known number of steps

where each V_j implements the propagation of the state of an additional clock register. The interaction with the clock register encodes the order of the unitaries U_j. Since V_j and U_j are chosen as two-qubit operators, A is by construction 4-local.

4.1.2 Generalized Measurements (POVMs)

Even though textbook quantum mechanics refers to observables usually as self-adjoint operators, the most natural measurements in nature are rather proper POVM measurements since the description of measurement inaccuracies already requires POVMs. But it has also been noted that in some situations POVM measurements extract more information about an unknown state from a quantum system than any von-Neumann measurement could do [139]. Apart from this, POVMs allow approximate simultaneous measurements of physical quantities which are actually incompatible: For two non-commuting self-adjoint operators A and B there is no common spectral resolution. Therefore there is no von-Neumann measurement that measures A and B simultaneously. Nevertheless there can be a POVM which can be interpreted as an approximative simultaneous measurement for A and B. An example is position and velocity of a particle where appropriate POVM measurements allow pretty good simultaneous estimations of both quantities [28] as long as the relevant scale is not too close to the Heisenberg limit. Another problem which requires POVMs is the estimation of the direction of a nuclear spin. Since the operators for the angular momenta in x, y, and z-direction do not mutually commute there is no measurement which determines the spin direction. On the other hand, it is clear that the angular momentum of a sufficiently large physical object is an almost well-defined vector in $\mathbb{R}^3$. However, even though it is rarely explicitly stated, this macroscopic limit requires POVMs [53] which allow appropriate estimations of the spin direction.[3] To respect the symmetry of this estimation problem it is natural to use a $SO(3)$-covariant POVMs. D'Ariano [140] has described a method to implement the $SO(3)$-symmetric POVM using an interaction between the magnetic moment of the spin with three field modes.

Here we will describe general design principles for measurement algorithms for group symmetric POVMs. First we consider POVMs on a single qubit. We assume that this qubit is part of a quantum register. Since we can apply every unitary to the register we can, by Naimark's theorem [42], reduce the POVM measurement to a read out of the computational basis. Clearly an n-qubit register only allows the implemention of measurements with 2^n outcomes. We must therefore restrict our attention to POVMs with finitely many outcomes in contrast to [140]. Furthermore, we only consider POVMs (M_j) where all M_j are rank one operators. For the further discussion one should recall that such a POVM can be characterized by a set of vectors in $\mathbb{R}^3$ by

$$v_j := (tr(M_j\sigma_x), tr(M_j\sigma_y), tr(M_j\sigma_z))\,.$$

The condition $\sum_j v_j = 0$ corresponds to $\sum_j M_j = \mathbf{1}$.

Formally, an implementation of $(M_j)_{j=1,\ldots,N}$ means the following: To measure the qubit it is embedded in an n qubit quantum register such that $N \leq 2^n$. Let the other $n-1$ qubits be initialized in the state $|0\ldots0\rangle$. Then implement a unitary transformation

[3] POVMs therefore define the natural measurements which intermediate between micro- and macrophysics on a mesoscopic scale between the quantum and the classical regime.

U on the register such that a readout of the logical state has N outcomes which occur with probabilities $tr(M_j\rho)$ when the qubit was in the state ρ.

However, the problem to implement a POVM on one qubit consisting of only rank-one operators is still too general since it can be rewritten as follows. Let $|\psi_1\rangle, \ldots, |\psi_N\rangle$ be vectors with $|\psi_j\rangle\langle\psi_j| = M_j$. Define the $2 \times N$ matrix

$$\mathcal{M} := \left(\begin{array}{cccc} |\psi_1\rangle & |\psi_2\rangle & \cdots & |\psi_N\rangle \end{array} \right) .$$

Then the implementation of the POVM requires to implement some $N \times N$ unitary $U^\dagger$ such that U extends $\mathcal{M}$ up to some global phases for each column. As a short remark, we note the formal analogy to the state preparation problem to find a unitary U such that $U|0\ldots0\rangle$ is a desired state $|\phi\rangle$. The latter problem requires the construction of a unitary extension of the $2^n \times 1$ matrix $|\psi\rangle$ and finding an implemention for U by circuits. Since the problem to extend rows or columns seems still too general to establish general rules we restricted our attention to *group covariant* POVMs [141, 142]:

Definition 26 (Group-Covariant POVMs)
A POVM (M_j) with finitely many operators is called covariant with respect to a unitary representation $g \mapsto U_g$ of a group G if the set (M_j) is invariant with respect to the group action

$$M_j \mapsto U_g M_j U_g^\dagger .$$

Still a bit stronger is the requirement that the group acts transitively, i.e., one can obtain all M_j by operating on a single operator M_1. If one has furthermore

$$U_g M_1 U_g^\dagger \neq U_{\tilde{g}} M_1 U_{\tilde{g}}^\dagger$$

for $g \neq \tilde{g}$ the elements g can be used as index characterizing the element in (M_j) and the possible measurement outcomes.

Remarkably, the quantum Fourier transform turns out to be a useful tool for implementing group covariant POVMs. This supports the general statement of this thesis that algorithms for non-computational quantum control problems use often the same tools as those solving computational problems. To give a rough idea about how the Fourier transform comes in we consider a POVM with vectors $v_0, \ldots, v_N$ in the equatorial plane with equal length and angle $2\pi/N$ between neighbors. One can choose the vectors $|\psi_j\rangle$ such that $\mathcal{M}$ consists of the first two rows of the quantum Fourier matrix DFT_N [141]. The Fourier transform is also an essential building block in POVMs corresponding to the symmetry groups of platonic solids. We described [141], for instance, a quantum circuit to measure a POVM where the vectors $v_1, \ldots, v_N$ are the edges of a dodecahedron as seen in Fig. 4.2. The circuit for the implementation is seen in Fig. 4.3. The matrices B and C in Fig 4.3 are defined by

$$B := \left(\begin{array}{cc} u_- & -u_+ \\ u_+ & u_- \end{array} \right), \quad C := \left(\begin{array}{cc} v_- & v_+ \\ v_+ & -v_- \end{array} \right)$$

with

$$u_\pm := \sqrt{\frac{1}{2} \pm \sqrt{\frac{3+\sqrt{5}}{24}}}, \quad v_\pm := \mp\sqrt{\frac{1}{2} \pm \sqrt{\frac{\sqrt{5}-1}{8\sqrt{5}}}} .$$

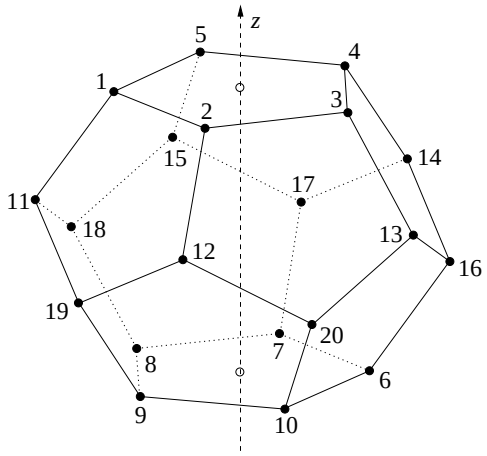


Figure 4.2: The dodecahedron with two faces perpendicular to the z-axis.

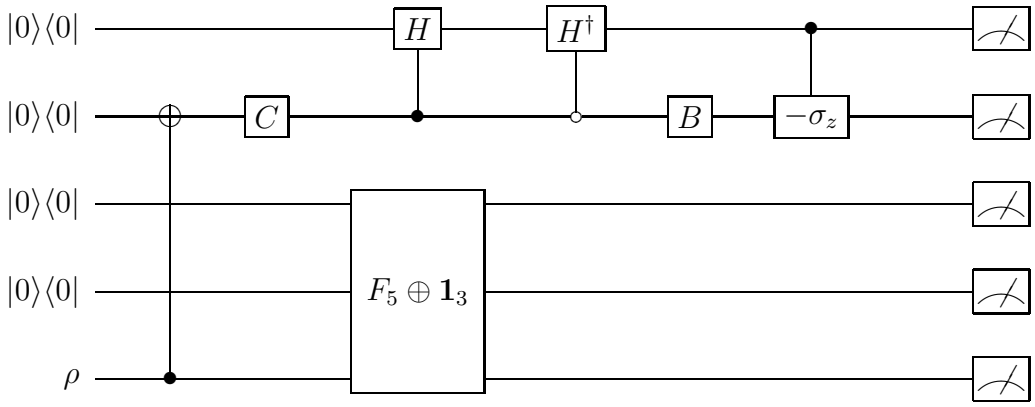


Figure 4.3: Circuit to implement the dodecahedral POVM. The Fourier transformation in dimension 5 is an important constituent. The unitaries B and C are defined in the text.

Note that it is completely new compared to classical computer science that the *readout of a single quantum bit* is a non-trivial task.

Using the equatorial POVM above as an example, we sketch now the general principles of [142] to implement group covariant POVMs. The POVM above is group-covariant with respect to a representation of the abelian group $\mathbb{Z}_N$, the integer addition modulo N. Therefore the unitary W which implements a $2\pi/N$ rotation in the equatorial plane shifts the POVM operators $M_0, \ldots, M_{N-1}$ cyclically. If the phase factors of each $|\psi_j\rangle$ are chosen appropriately, W will shift the vectors $|\psi_j\rangle$ in a cyclic way:

$$W|\psi_j\rangle = |\psi_{j\oplus 1}\rangle\,.$$

Therefore $\mathcal{M}$ satisfies

$$W^j\mathcal{M} = \mathcal{M}S^j \quad \forall j \in \mathbb{Z}$$

where S is the cyclic shift on $\mathbb{C}^N$. The maps $j \mapsto W^j$ and $j \mapsto S^j$ are both representations of $\mathbb{Z}_N$ and $\mathcal{M}$ is an *intertwiner* [142] between these two representations. The first is the representation which defines the POVM and the second is the *regular* representation of the group $\mathbb{Z}_N$.

In general $\mathcal{M}$ intertwines between the representation defining the group covariance and a *monomial* representation, i.e., a representation consisting only of unitary matrices which have only one non-zero entry in each column. In [142] we decided to choose the unitary extension $\tilde{\mathcal{M}}$ of $\mathcal{M}$ such that it intertwines $j \mapsto S^j$ with an appropriate $N \times N$ extension of $j \mapsto W^j$. To implement the obtained $\tilde{\mathcal{M}}$ we can make use of a well-developed theory on decomposing unitaries which intertwine between group representations [143]. Implementations for unitary intertwiners with quantum circuits can be found in [144]. An essential tool are Fourier transforms and their generalization to non-abelian groups [145]. Note that the feature of generalized Fourier transforms is that they block-diagonalize the regular representation which makes it an important tool in our work [142].

A natural example where the implementation of a group covariant POVM could be relevant to future explorations of microscopic and mesoscopic physics has already been mentioned above: The observables position and momentum[4] of a quantum particle are incompatible, i.e., the self-adjoint operators X and P on $L^2(\mathbb{R})$ defined by

$$(X\psi)(x) := x\psi(x) \qquad \text{and} \qquad (P\psi)(x) := \frac{-i\,d}{dx}\psi(x)$$

have no common spectral resolution. Simultaneous measurements of momentum and position exist in an approximative sense if one allows POVM-measurements [28]. In infinite dimensions such a POVM is e.g. given by the operator-valued density

$$M_{s,t} = \frac{1}{2\pi}e^{i(Ps+Xt)}|\psi\rangle\langle\psi|e^{-i(Ps+Xt)}\,.$$

Here we interpret

$$p(s,t) := tr(\rho M_{s,t})$$

as the probability density that the particle has position s and momentum t, or, a bit more correctly, that we *estimate* 'position s and momentum t'.

[4] which is the velocity up to the factor 'mass'

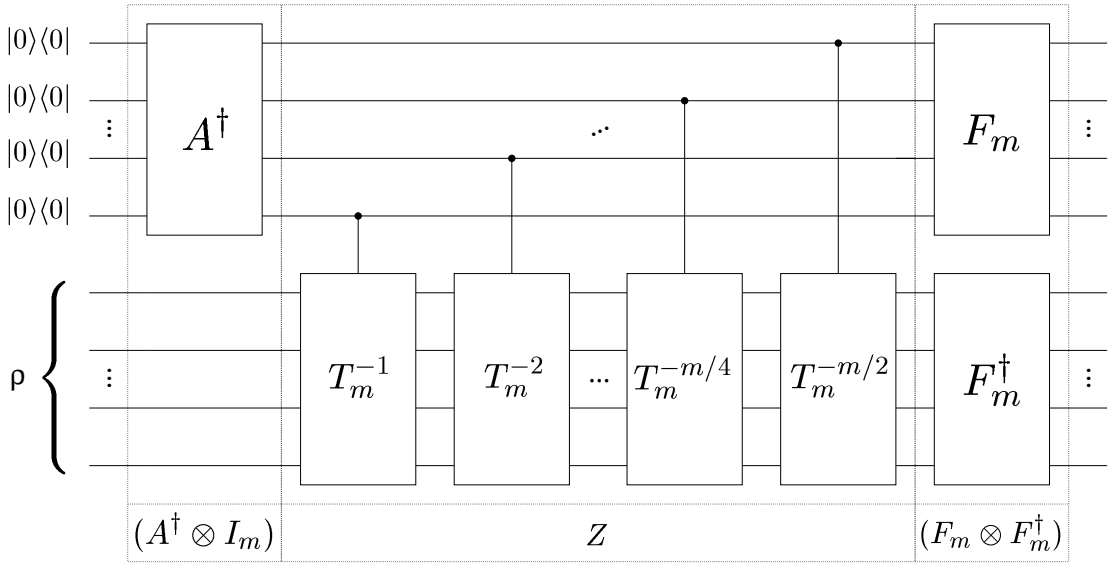


Figure 4.4: Circuit for the implementation of the POVM with respect to the Weyl-Heisenberg group and the vector $|\Psi\rangle = (v_1, \ldots, v_m)^T$. The vector $|\Psi\rangle$ determines the matrix $A^\dagger$.

Now we will give an implementation of a finite dimensional analogue of this POVM which has furthermore finitely many outcomes. They can be seen as a simultaneous measurement of discretized position and momentum. The physical situation would be a particle on a cyclic chain which can be at the discrete positions $j = 0, 1, \ldots, m-1$. Its crystal momentum [81] can take the values in the 1-dimensional Brillouin zone $[-\pi, \pi]$. Due to the cyclic structure of the chain there are only m possible values $-\pi + 2\pi l/m$ with $l = 0, \ldots, m$. A POVM that allows reasonable simultaneous estimations of j and l is given by the orbit of an appropriate rank-one operator $|\psi\rangle\langle\psi|$ under the Weyl-Heisenberg group:

$$M_{j,l} := \frac{1}{m} S_m^j T_m^l \, |\psi\rangle \, \langle\psi| \, T_m^{-l} S_m^{-j} \,,$$

where S is the cyclic shift on $\mathbb{C}^m$ on $T := diag(1, \omega, \omega^2, \ldots, \omega^{m-1})$ with the root of unity $\omega := \exp(-i2\pi/m)$. To implement it on a quantum computer we assume that $m = 2^k$ and that there is an interface between the cyclic chain and the $2k$-qubit quantum register such that the position states $|0\rangle, |1\rangle, \ldots, |m-1\rangle$ of the chain can be mapped onto the binary words $|j\rangle$ of an k-qubit sub-register. We will also investigate the efficiency of the implementation. In analogy to computational problems we shall call an algorithm 'efficient' if its complexity increases only polynomially with the length of the input string. Here we do not have an input *string* but rather a k-qubit quantum state as input, and therefore consider k as the input size. The circuit in Fig. 4.4 implements the POVM on $2k$ qubits for a k-qubit input. The gate T_m is the phase shift

$$T_m := \begin{pmatrix} 1 & 0 \\ 0 & \omega^{m/2} \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & \omega^{m/4} \end{pmatrix} \otimes \ldots \otimes \begin{pmatrix} 1 & 0 \\ 0 & \omega \end{pmatrix} .$$

One can easily show that for all l we can implement a controlled-T_m^l gate efficiently. The gate F_m is the DFT_m, which is efficient if m is a power of two [3]. The gate $A^\dagger$ is

any unitary that has $F_m|\overline{\psi}\rangle$ as first coefficients where $\overline{\psi}$ is obtained by conjugating the coefficients of ψ. Note that we have to choose $|\psi\rangle\langle\psi|$ such that uncertainty of position and momentum are both not too large. Otherwise the POVM would be inappropriate to estimate these quantities. We found a vector $|\psi\rangle$ satisfying this for which we could find an efficient circuit $A^\dagger$. The above arguments show that the whole scheme is efficient.

At the moment it is unclear which other useful applications POVM measurements could allow. Meanwhile Renes [146] has, for instance, proposed a protocol for cryptographic key distribution which uses the *tetrahedral* POVM (with Bloch vectors pointing on the vertices of a tetrahedron) for which an implementation was given in [141]. We discuss further potential applications of POVMs in Subsection 4.3.2.

4.2 Thermodynamic Machines

Now we will address other non-computational control problems and develop a complexity theory in analogy to complexity theory of computation problems.

We have already used toy models to illustrate the thermodynamics of refrigerators and heat engines. However, one would not expect that future cooling technologies and molecular heat engines would really be based on quantum gate operations on qubits. This might be true for some exceptions like the algorithmic cooling proposed in the context of NMR quantum computing, but it deserves some more justification why we consider the complexity of algorithmic heat engines and refrigerators here. The main motivation is the hope that the models presented here could nevertheless give an intuition about the complexity of molecular thermodynamical machines if they should be, for instance, a *maximally efficient* heat engines or a cooling apparatus leading to the *minimal achievable* temperature. Whether one shares this hope or not, one may in any case agree that it is only a generalized version of the strong Church Turing principle to believe that complexity of non-computational tasks performed on a non-computing device could also be studied in quantum computing models.

4.2.1 Cooling Weakly Interacting Systems

Cooling microsystems can be considered as key technology in quantum control since many interesting quantum effects can only be observed in sufficiently cold systems. For the Cirac-Zoller proposal [20] (which is an ion trap quantum computer), the thermal motion of the ions is a major obstacle and recent developments of cooling these degrees of freedom are considered a decisive step towards the realization. Despite of the practical relevance of these laser cooling technologies, we prefer to sketch a cooling algorithm for molecules which can be described as an algorithm in the standard model quantum computer. Whereas the laser cooling schemes mentioned above transfer the entropy to the environment, the algorithm below transports the entropy only within the quantum computer[5]. This makes it a nice setting in which to study it from an information-theoretic point of view.

[5]Note, however, that a mechanism which is cooling one part of a register for the cost of heating the other part can also lower the total entropy of the register because the hot part will transfer some heat to the colder environment [147, 148]. This method has even be demonstrated experimentally.

In Subsection 2.2.4 we have described the idea of cooling with logical gates on 3 qubits. Here we want to discuss algorithms for arbitrarily large qubit numbers and the corresponding complexity issues. Consider n qubits in a thermal state γ_T for temperature $\infty > T > 0$. For *optimal* cooling we would look for a unitary transformation U that maximizes the probability for the ground state $|0\dots0\rangle$ for $k < n$ qubits. This means that

$$tr(U\gamma_T^{\otimes n}U^\dagger\ (\mathbf{1_{n-k}}\otimes|0\dots0\rangle\langle0\dots0|))$$

is minimal. One would minimize the entropy of the k-qubit system by transferring as much entropy into the remaining $n-k$ qubits. In order to get efficient algorithms we probably have to relax this demand.

The cooling algorithm in [35] has been proposed in the context of NMR quantum computing. In this context one would rather call it 'initialization' than 'cooling'. However, it is the equivalence of both terms which makes the example so nice. The major problem in all realizations of NMR quantum computing so far is the initialization of the register. The system starts in its equilibrium state and each atom can be considered as a two-level system. The energy of the lower and upper level $|0\rangle$ and $|1\rangle$ differ only slightly compared to kT which implies (see subsection 1.2.1) that both are almost equally likely. There is only a small statistical bias ϵ, i.e., the upper state has probability $(1-\epsilon)/2$ and the lower state $(1+\epsilon)/2$. Hence all computation works with a state which is almost the maximally mixed state. Only the extremely large size of the ensemble makes it possible at all to readout the computational results despite of the noise. As we have already argued, thermodynamics allows the transfer of entropy from k qubits into the remaining $n-k$. However, every algorithm doing this has to be implemented without using initialized ancillas, making the problem that much harder. Note that even a new complexity class has been introduced consisting of those problems which can be solved using only one initialized qubit when all the rest of the register is in its maximally mixed state [149]. In order to give a rough idea how many clean bits we can expect we recall that the binary entropy function

$$p\mapsto -p\log_2 p-(1-p)\log_2(1-p)$$

has a maximum at $p=1/2$. Therefore the entropy decreases in the order of ϵ^2 with the bias ϵ, so we can expect that asymptotically the fraction of clean qubits is of order ϵ^2.

The first phase in the algorithm [35] is a simple bias amplification. We simplify the setting to the case that all qubits are uncorrelated. Consider two qubits with bias ϵ. Assume we could measure whether both are in the same state or not and discard both whenever the states are different and discard one when the states are equal. Whenever they are in the same state, they are with probability

$$p_{00}=\frac{1+\frac{2\epsilon}{1+\epsilon^2}}{2}\quad\text{and}\quad p_{11}=\frac{1-\frac{2\epsilon}{1+\epsilon^2}}{2}$$

in the state 00 and 11, respectively. Given that one qubit survived the procedure, its state has therefore the new bias $2\epsilon/(1+\epsilon^2)$. The expected number of bits that survive is $n(1+\epsilon^2)/4$. In order to implement the discarding unitarily, a C-NOT is applied to the pair such that the target qubit is 1 if and only if the logical values were different. Then the target qubit can be used to control SWAP-operations which transport the logical state of the qubit to be 'discarded' into an irrelevant part of the register. This procedure can

then be concatenated. However, as argued in [35] an arbitrary concatenation of this bias amplification could not provide a fraction of clean qubits that is constant in n. Therefore this was only phase 1 of the 3 phases of the algorithm. In phase 2 they discard a collection of l bits $b_1, \ldots, b_l$ whenever its parity is odd. This is done by adding the parity of the bit string $b_2, \ldots, b_l$ on b_1. The authors argue that this allows an exponential decrease of the probability for the upper state. After this probability has been sufficiently decreased they perform phase 3, in which a collection of bits is discarded if the Hamming weight modulo 4 of the string is not zero. The authors analyze the performance of this scheme and come to the conclusion that $\Omega(\epsilon)n$ bits can be initialized with arbitrary reliability.

The proposal above shows that cooling *can* be done by logical operations. Now we will describe some circumstances under which cooling *necessarily* implements logical gates:

Theorem 19 (Optimal Refrigerator Computes MAJORITY)
Given $2n+1$ two-level systems with finite temperature $T \neq 0$. Let U be an optimal unitary cooling process on the corresponding $2n+1$-qubit Hilbert space in the sense that it lowers the temperature of the first two-level system as much as possible. Then U computes the logical function MAJORITY, i.e.,

$$U|b\rangle = |f(b)\rangle \otimes |\psi_b\rangle \quad b \in \{0,1\}^{2n+1},$$

where $f(b) = 1$ if and only if the Hamming weight of b is larger than n. $|\psi_b\rangle$ is an arbitrary state (not necessarily a basis state) in the remaining $2n$-qubit Hilbert space.

Proof: The optimal cooling process has to reduce the probability for the $|1\rangle$ state of the first qubit as much as possible. Let $\rho := \gamma_T^{\otimes 2n+1}$ be the density matrix of the initial state. The eigenspace of ρ corresponding to its 2^{2n} smallest eigenvalues have to be mapped to $|1\rangle$. This eigenspace is clearly spanned by the words with Hamming weight greater than n since the probability for the upper state is smaller than for the lower state in each two-level system. □.

Even though MAJORITY is not a particularly complex logical function, it is treated as an interesting example in the theory of small-depth classical circuits [19], because it cannot be computed by circuits with bounded depth [19, 88]. Note that the construction of U from elementary quantum gates is probably more difficult than the computation of MAJORITY because by assumption the fridge consists of unitary gates. Initialized ancilla qubits (which would be required to simulate classical logical operations like AND and OR with TOFFOLI-gates) are not allowed in the setting above (compare Subsection 3.2.1). Note that this restriction is, for instance in NMR, given by physical reality since initialized qubits would be cold two-level systems. Hence the results in [19] can only provide *lower* bounds on the complexity of the optimal refrigerator.

4.2.2 Complexity of Cooling Strongly Interacting Systems

In the last subsection we have described cooling procedures for weakly interacting systems. In the NMR cooling proposal the interaction between the qubits must, of course, be used for the implementation of the algorithm. However, the interaction is so weak that the

level structure of the system is dominated by the energy gap of each two-level system. In other words, the energy eigenstates are the logical states of the quantum register and cooling means initializing the system to the ground state $|0\ldots0\rangle$ with high fidelity. This is completely different when the interactions change the level structure. Then the states with least energy may be any other states, not necessarily basis states. It is not even clear whether they have an explicit short description. In statistical physics it was known for many years that one can construct classical spin-spin interactions such that the computation of the ground state energy is an NP-hard problem [150]. Kitaev, Shen, and Vyalyi [14] found a quantum generalization of this statement which states, roughly speaking, that (even an approximative) determination of the ground state energy of interacting quantum systems is quantum-NP-complete. Here we rephrase the improved version of Kempe, Kitaev, and Regev [151]:

Theorem 20 (2-local Hamiltonian is QMA-Complete)
Let H be an n-Qubit Hamiltonian which consists of pair-interactions only, i.e.,

$$H = \sum_{j,k} H_{jk}\,,$$

where each $H_{j,k}$ acts only on the qubit pair (j,k). Then the decision problem 'determine whether the smallest eigenvalue of H is smaller than b or greater than a with $1/(a-b) \in O(poly(n))$' is QMA-complete.

In a previous article [152] this was only shown for 3-local Hamiltonians. We have modified [153] this problem in such a way that it leads to a QCMA-complete problem: Decide whether all states which can be prepared with at most k gates have energy at least a or whether there exists at least one state which can be prepared using at most k gates that has energy at most b with $1/(a-b) \in O(poly(n))$. The following idea should show the potential relevance of this result. One could be tempted to assume that a physical system needs a lot of time 'to find' ground states which have a large preparation complexity in terms of elementary quantum gates. It seems reasonable to consider this as an implication of a generalized strong Quantum Church-Turing thesis. Then one may only be interested in low energy states with low complexity since the system 'would not find' the others on the relevant time scale, it would rather remain in less complex 'meta-stable' states.

In order to show that complexity classes also for non-computational problems as state preparation makes sense, we first recall that the following classical computational problems are closely related:

Definition 27 (SAT)
Given a family of functions $f_n : \{0,1\}^n \to \{0,1\}$ defined by a Boolean expression of polynomial size. Decide whether there exists an $x \in \{0,1\}^n$ with $f(x) = 1$.

This problem is NP-complete [154]. Strictly speaking, this problem has to be distinguished from the problem to *find an* x with $f(x) = 1$, the problem FSAT. However, given an oracle which tells us whether a Boolean expression has a solution, one can easily find a solution: Define functions $f^0, f^1 : \{0,1\}^{n-1} \to \{0,1\}$ that are obtained from f if the first input bit is set to 0 or 1, respectively. Asking the oracle whether f^0 and/or f^0 is

satisfiable, one already has the first bit of a solution x. Proceeding in the same way, one obtains an x with less than $2n$ queries. In general one can define FNP as the class that consists of the problems to find a proof of an NP decision problem. Such a reduction of FNP to NP cannot be transferred to the "Quantum-NP" class QMA. The problem of finding a state $|\psi\rangle$ which works as a proof for a QMA problem cannot be reduced to the problem of finding the restriction of $|\psi\rangle$ to all n qubits. Therefore the problem to construct the proof has to be distinguished from the problem to find the yes/no-answer. The problem to provide the proof $|\psi\rangle$ is a *quantum state preparation problem* and not a computational problem. Based on this observation we defined FQMA as a complexity class for channels [155]:

Definition 28 (FQMA)
Fix $\delta = 1/r(|x|)$ for an arbitrary polynomial r. A sequence of channels with classical input x and quantum output ρ_x is in FQMA if there is a language L in QMA with a verifier U_x and ϵ as in Definition 19 such that

$$tr(U_x\,(\rho_x \otimes |0\ldots0\rangle\,\langle 0\ldots0|)\,U_x^\dagger\,P_1) \geq 1-\epsilon-\delta\,,$$

whenever $x \in L$. For $x \notin L$ the output is allowed to be arbitrary.

Kempe and Regev have shown that for every circuit U with output yes/no, one can construct a 2-local Hamiltonian H such that H eigenvalues smaller than or equal to b if there is a state which is accepted by U and all eigenvalues of H are at least $a > b$ if there is no state accepted by U with high probability. Consider a machine M that prepares states with energy about a. Certainly, M generates proofs for the 2-local Hamiltonian problem. However, it is not clear that the generated states are accepted by U with high probability. If two different circuits U and V define the same language $L \in QMA$ it may make an important difference to prepare states that are accepted by U or by V. This demonstrates the following quantum control problem.

Given n qubits with k-local interaction Hamiltonian H, assume they are initially in the state $|0\ldots0\rangle$ and that we are only able to access $m < n$ of them. Therefore we have to use these m qubits to control the other $n-m$. For instance, we would like to prepare the m "controller"-qubits in such way that the rightmost qubit is with high probability in its 1-state after the natural time evolution $\exp(-iHt)$ was active for the time T. This leads to the following problem:

- Is there a state $|\psi\rangle \in (C^2)^{\otimes m}$ such that the rightmost qubit is with high probability 1 in the state
$$\exp(-iHT)(|\psi\rangle \otimes |0\ldots0\rangle)\,.$$
This problem is in QMA for constant T, since the time evolution can efficiently be simulated by a quantum circuit U such that the running time increases only with $O(T^2/\epsilon)$ for the error ϵ [89].

- Prepare the state $|\psi\rangle$. This problem is in FQMA.

Clearly, one could prove that such a $|\psi\rangle$ exists by constructing another circuit U' which has an accepted state if and only if U has. However, we would not be satisfied to have

a channel which prepares proofs for the circuit U' since this would not solve our original control problem. Therefore it does not directly follow from [151] that a channel preparing low-energy states could always prepare proofs for each circuit U defining a QMA-problem. However, in [155] we have shown that this is indeed the case. This means that every machine which is able to prepare low energy states and low temperature states could be used to prepare proofs for *all* QMA-problems. Furthermore we have estimated the temperature which would be sufficient to do this, finding:

Theorem 21 (Temperature which is Sufficient for QMA)
To prepare proofs for a QMA-problem with r qubit input defined by a circuit U acting on $r+m$ qubits which is composed of L gates, a cooling procedure satisfying the following requirement is sufficient:

Cool a system with $n := r+m+L$ qubits with an appropriate 3-local Hamiltonian to the temperature T_n such that

$$T_n < \frac{1-2\epsilon}{4\ln 2k(L+1)n}\,, \tag{4.3}$$

where k is Boltzmann's constant.

For the classical complexity class NP we can prove a higher lower bound on the necessary temperature [155]:

Theorem 22 (Temperature sufficient for NP)
For NP-problems it is sufficient to have

$$T_n < \frac{1}{kn\ 2\ln 2\ q(n)} =: \frac{1}{k\,\tilde{q}(n)}\,, \tag{4.4}$$

with polynomials q and $\tilde{q}$. In other words, the temperature must only decrease as the reciprocal of a polynomial in n.

The fact that our bound is lower for QMA than for classical NP is illustrated in Fig. 4.5. It is not clear whether lower temperatures could even solve PSPACE-hard problems. An indicator supporting this conjecture is maybe the fact that our result mentioned in Subsection 4.1.1 showed that the fine-structure of the spectrum of a 4-local Hamiltonian H can encode the solution of PSPACE-problems. However, the problem is that this statement refers to the spectrum of H in a specific H-invariant subspace and not to the whole spectrum of H.

4.2.3 Complexity and Efficiency of Molecular Heat Engines

In Subsection 2.2.5 we have explained the idea that molecular heat engines may be implemented by logical transformations. We have considered quite simple molecular systems as hold and cold reservoirs. It is natural to ask how the complexity of logical transformation required for work extraction increases for larger reservoirs. We will not expect that heat engines using the heat of many particles are necessarily complex in the sense that they

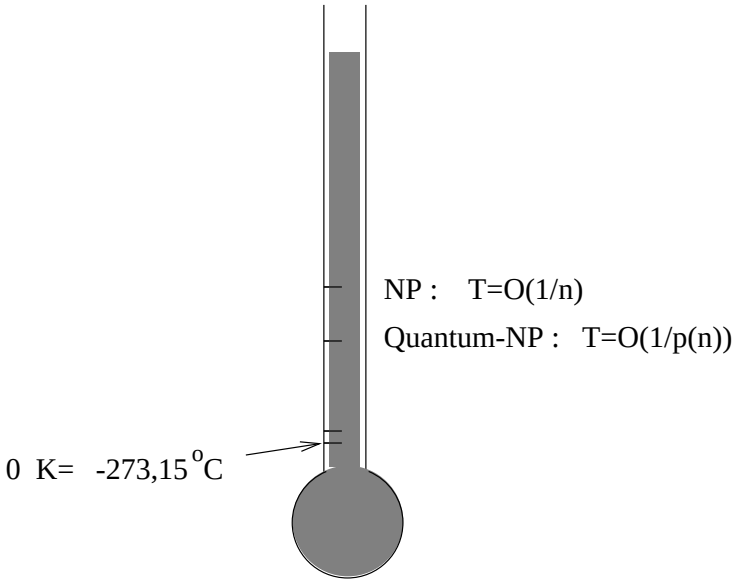


Figure 4.5: The thermometer indicates the different upper temperature bounds for solving NP and quantum-NP: our bound for the latter is lower.

would require circuits of large logical depth. This is for the following two reasons: first one can obviously group the particles and act on only a few hot and a few cold particles at once. Second we know that heat engines acting on reservoirs with more than 10^{23} particles exist since the beginning of the industrial revolution. Due to the strong Church-Turing thesis, we should expect that there is an efficient simulation of these macroscopic heat engines on a quantum computer.

However, we will argue that there are two main conditions which require *complex* heat engines. First, if we demand *optimal* heat engines in the sense that they extract maximal amount of work. By relaxing the demand on optimal efficiency, the complexity decreases substantially and for very large numbers of particles the loss of extracted work is negligible. Second, it is complex to extract work from reservoirs having *almost* the same temperature.

Before we will provide theorems on the complexity of heat engines we will first consider a system that helps to understand intuitively the complexity of *optimal* heat engines[6].

A very natural system in physics is a quantum harmonic oscillator. Its Hilbert space $l^2(\mathbb{N}_0)$ is spanned by the number states $|0\rangle, |1\rangle, |2\rangle, \ldots$ with $0, 1, 2, \ldots$ quanta. Such a system can be a quantum optical mode or a mechanical oscillator. A state with j quanta of frequency ω has the energy $E(j) = j\hbar\omega$ and the system Hamiltonian is therefore

$$H := \hbar\omega \sum_{j=0}^{\infty} j|j\rangle\langle j| \,.$$

The bipartite system on which our heat engine will be defined consists of two modes with different frequencies ω_A and ω_B.

Now we assume that the ratio $e := \omega_A/\omega_B$ is irrational. This ensures that the Hamiltonian of the composite system is non-degenerate. Up to irrelevant constants, the energy

[6]The following ideas are taken from my article [56].

of a state with n_A quanta in mode A and n_B quanta in mode B is

$$E(n_A, n_B) = en_A + n_B$$

with $e \in \mathbb{R} \setminus \mathbb{Q}$. We define a bijective function $k : \mathbb{N}_0^2 \to \mathbb{N}_0$ such that $k(n_A, n_B)$ indicates the number of the pair (n_A, n_B) when all pairs are put into an increasing order with respect to $E(n_A, n_B)$. Now we choose the temperatures $0 \neq T_A \neq T_B \neq 0$ such that

$$q := \frac{E_A/T_A}{E_B/T_B}$$

is also irrational which holds for instance when T_A/T_B is rational. It follows that the density operator $\rho_A \otimes \rho_B$ is also non-degenerate. Up to an additive constant and a negative factor, the logarithm of the probability for a state $|n_A\rangle \otimes |n_B\rangle$ is given by

$$Q(n_A, n_B) := qn_A + n_B \,.$$

A larger value $Q(n_A, n_B)$ indicates that the state is less likely. In analogy to the map k we define a bijective function $l : \mathbb{N}_0^2 \to \mathbb{N}_0$ indicating the order of the pairs (n_A, n_B) with respect to their values $Q(n_A, n_B)$. Define a permutation π on $\mathbb{N}_0^2$ by

$$\pi := k \circ l^{-1} \,.$$

This permutation of basis states $|n_A, n_B\rangle$ defines a unitary U_π by linear extension[7]. The density operator of the whole system after having implemented the heat engine U_π is

$$U_\pi(\rho_A \otimes \rho_B)U_\pi^\dagger \,.$$

The heat engine permutes the eigenvalues such that they are reordered according to the corresponding energy values. We have computed the corresponding reordering of states for the values $e = \sqrt{2}$ and $q = 1/\sqrt{3}$. The mapping is depicted in Fig. 4.6, showing that the heat engine defines a quite *complex* flow in the discrete two-dimensional plane. Here complexity is understood in a rather intuitive sense.

In Subsection 2.2.5 we have introduced heat engines which act on two-level systems. This setting makes a computer science approach to complexity possible since the analogy to logical operations is more obvious.

First we discuss the complexity of heat engines which are able to extract work from two systems with *almost* the same temperature. The required number of two-level systems which are necessary in order to make a heat engine possible at all increases whenever the temperature quotient gets closer to 1 :

Theorem 23 (Complexity of Using Small Temperature Gaps)
A heat engine on n_A hot and n_B cold qubits with temperatures T_A and T_B, respectively, and equal energy gaps, is possible if and only if

1. *(for $n_A \leq n_B$)*

$$\frac{T_A}{T_B} \geq \frac{n_A}{n_A - 1}$$

[7]Note that the ordering of pairs given by E or Q is a term order in the sense of [156].

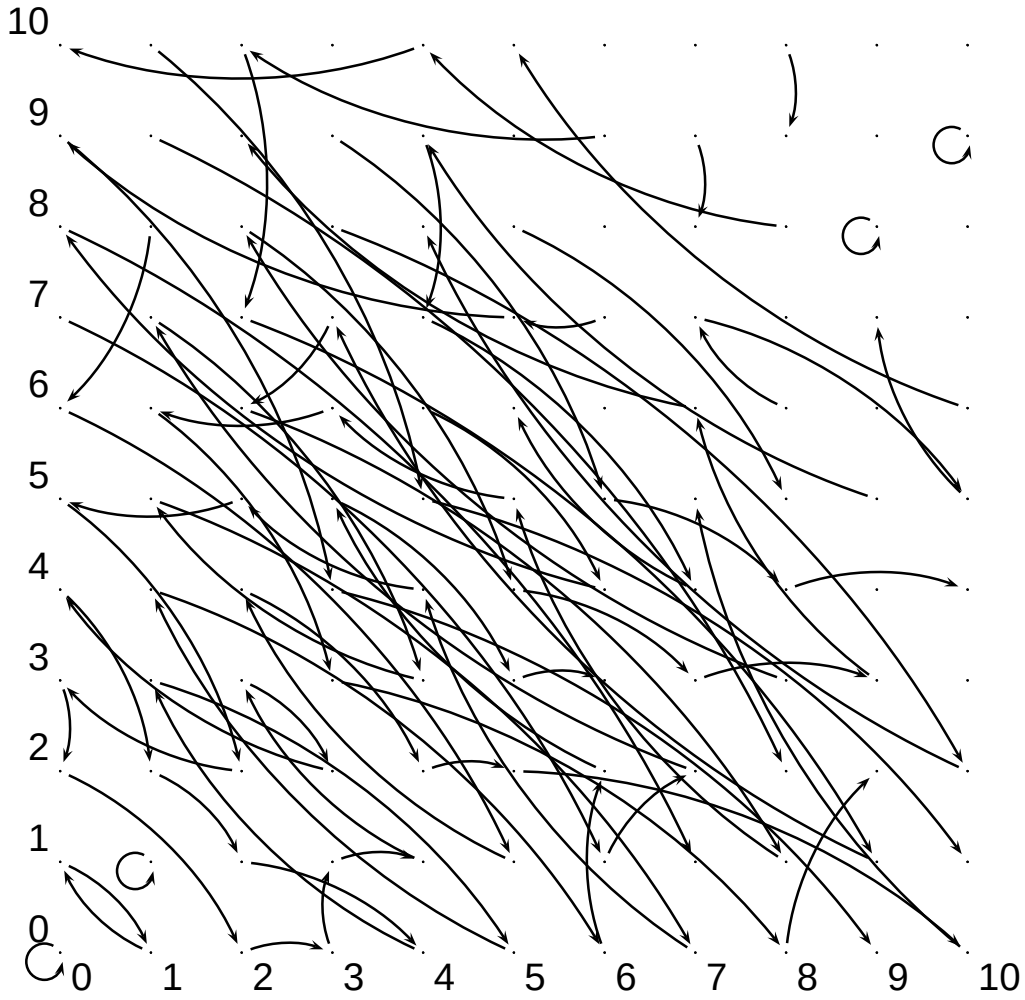


Figure 4.6: Optimal heat engine acting on two harmonic oscillators with frequency ratio $\omega_A/\omega_B = \sqrt{2}$ and temperature ratio $T_B/T_A = \omega_B/(\sqrt{3}\omega_A)$. A point in row n and column m is a basis state with n quanta in mode A and m in mode B. An arrow $(n, m) \to (\tilde{n}, \tilde{m})$ indicates that a state with n quanta in mode A and m in mode B has to be converted into a state with $\tilde{n}, \tilde{m}$ quanta, respectively. Points which have their image or pre-image outside the depicted area obtain no arrow.

2. *(for $n_A \geq n_B$)*

$$\frac{T_A}{T_B} \geq \frac{n_B + 1}{n_B}$$

Furthermore, every heat engine acting on an infinite reservoir of hot and cold qubit level systems must use operations which connect at least n_A hot and n_B cold systems such that the above conditions hold.

Proof: We note that a heat engine can work if and only if a pair of states exist such that the first has more energy even though it is more likely. Let (l_A, l_B) denote the Hamming weights (i.e. the number of symbols 1) of a basis state in the $n_A + n_B$ qubit system. The pair (l_A, l_B) and (k_A, k_B) satisfies this condition if

$$(l_A - k_A) - (l_B - k_B) > 0$$

and

$$(l_A - k_A)T_A - (l_B - k_B)T_B < 0$$

Elementary computation shows that this implies

$$\frac{T_A}{T_B} > \frac{l_A - k_A}{l_B - k_B} > 1\,.$$

Clearly the modulus of the numerator and the denominator are at most n_A and n_B, respectively. The smallest possible quotient which is still greater than 1 is therefore $n_A/(n_A - 1)$ or $(n_B + 1)/n_B$, respectively. This shows that the conditions (1), respectively (2) are necessary in order to make a heat engine possible.

For the converse we observe that in case (1) a permutation of the states $(n_A, 0)$ and $(0, n_A - 1)$ extracts some amount of energy. In case (2) one extracts energy by permuting $(n_B + 1, 0)$ and $(0, n_B)$. □.

Fig. 4.7 illustrates how the complexity of heat engines on two-level systems with equal energy gap increases when the temperature gaps decrease in the sense that more qubits have to be involved. Note that Fig. 4.7 furthermore suggests a simple method to obtain *suboptimal* heat engines on many particles by independently applying few-qubit heat engines.

In order to present now a computer science approach to the complexity of heat engines we first mention an instance where the optimal heat engine requires at least as many gates as computing the boolean function MAJORITY:

Consider $2n$ two-level systems with temperature $T_A = \infty$ and 1 system with $T_B = 0$. The basis states of the system are binary words of length $2n + 1$. The joint Hamiltonian of the system is given by

$$H := E \sum_b wgt(b) |b\rangle\langle b|\,,$$

where E is the energy gap of each two-level system and $wgt(b)$ denotes the Hamming weight of the binary word b. Let the suffix of each of this binary words indicate the state of system B. Then all binary words with suffix 0 have probability $1/2^n$ and words with suffix 1 never occur. Every optimal heat engine U has to map the subspace spanned by

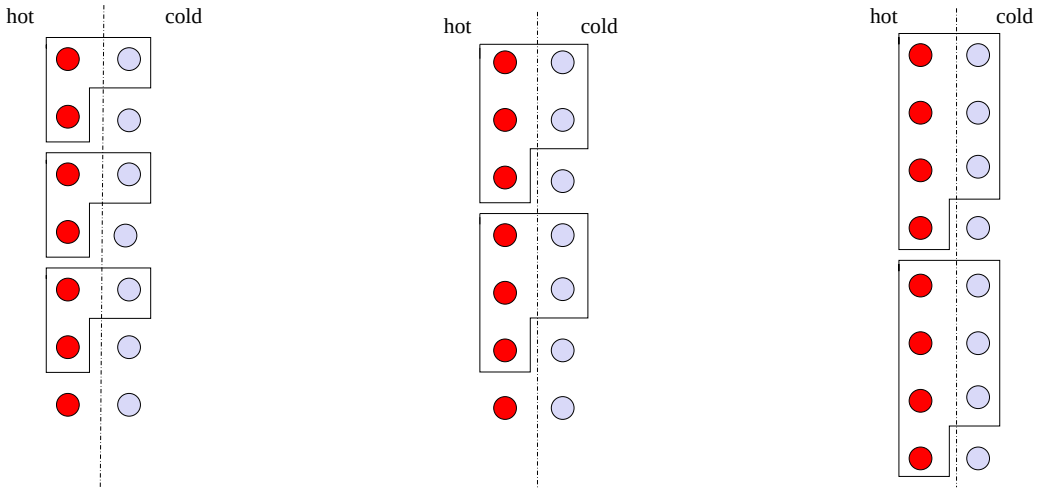


Figure 4.7: Heat engines with $T_A/T_B > 2$ can be implemented with joint operations on 2 hot and 1 cold qubit (left). For $2 \geq T_A/T_B > 3/2$ operations on 3 hot and 2 cold qubits are needed (middle), and heat engines for $3/2 \geq T_A/T_B > 4/3$ must involve 4 hot and 3 cold qubits (right)

the former 2^n words onto the subspace corresponding to the 2^n smallest eigenvalues of H. It is the space spanned by all words with Hamming weight at most n. Therefore the inverse of the heat engine, i.e., U^{-1} computes the boolean function MAJORITY in the sense that the rightmost qubit in the state

$$U^{-1}|b\rangle$$

is 1 if and only if $wgt(b) > n$, i.e., the majority of the qubits are in the 1 state. We would like to estimate the gate complexity of U when it is implemented by elementary gates. If the set of elementary gates contains with every gate also its inverse the complexity of U and U^{-1} coincide. To obtain a lower bound on the circuit complexity we could therefore use bounds on the circuit complexity of MAJORITY. In [19] one can find bounds for classical circuits with bounded depth consisting of AND and OR with arbitrary fan-in. We can give a lower bound on the circuit depth which holds for arbitrary k-qubit gates. Our reasoning is as follows. The observable which measures whether the suffix of a binary word is 1 or 0 is $A := \mathbf{1}_{2n} \otimes \sigma_z$. This is obviously a 1-qubit observable since A acts only on the rightmost qubit non-trivially. The observable $UAU^\dagger$ which measures whether the majority of qubits are 1 is a proper $2n+1$-qubit observable because the logical states of all qubits are relevant. In [40] we have argued that a circuit of depth l can convert a 1-qubit observable at most into a k^l-qubit observable. Therefore we obtain

$$l \geq log_k(2n+1)$$

as lower bound on the depth. This shows after all that the depth must necessarily increase with n even though logarithmic growth would be quite slow. We summarize :

Theorem 24 (Lower Bound on the Depth)
Let U be an optimal heat engine on $2n$ two-level systems with temperature $T_A \neq 0$ and one two-level system with $T_B = 0$ where all $2n+1$ systems have the same energy gap. Then the implementation of U with k-qubit gates requires at least a circuit of depth $\log_k(2n+1)$.

Now we will describe an instance of heat engines where the required complexity becomes serious: Given n two-level systems with different energy gaps, the optimal heat engine could solve the NP-complete problem KNAPSACK. First we recall an instance which is already NP-complete [92]:

Definition 29 (KNAPSACK)
Given a sequence of some positive integers $v_1, \ldots, v_k$ and two natural numbers K, B, is there a subset $S \subset \{1, \ldots, k\}$ such that

$$K \leq \sum_{j \in S} v_j \leq B \,.$$

If there is such a subset the optimal heat engine will always find it:

Theorem 25 (Optimal Heat Engine solves KNAPSACK)
Let

$$E_0, \ldots, E_n$$

be the energy gaps of $n+1$ two-level systems. Let T_0 be the temperature of the 0th system and T of the remaining n. Let the values be such that there is no $b \in \{0,1\}^n$ such that

$$\langle b|E\rangle = E_0 \frac{T}{T_0} \,.$$

Let U acting on $\mathbb{C}^2 \otimes (\mathbb{C}^2)^{\otimes n}$ be an optimal unitary heat engine for this system. Then U solves a KNAPSACK problem in the following sense. Perform a measurement in the computational basis on the rightmost n two-level systems in the state

$$U(|1\rangle \otimes |0 \ldots 0\rangle) \,.$$

Let $b \in \{0,1\}^n$ be the obtained result. Then b satisfies

$$E_0 > \langle b|E\rangle > E_0 \frac{T}{T_0} \tag{4.5}$$

if and only if such a binary word b exists.

For the detailed reduction of KNAPSACK to this heat engine see [56]. It is seen that the hard instances of KNAPSACK (with $B - K$ small) correspond to small temperature differences. One may ask whether this theorem indicates complexity theoretic bounds on the efficiency of nanoscopic heat engines. Admittedly, the setting is very special since it is more likely to have a collection of two-level systems with *equal* gap (when the temperature difference is not chosen as in the example above). However, we leave the complexity of such a heat engine as an open question.

One could object that it is not natural to implement heat engines by quantum gates and real interactions could potentially solve the heat engine problem in a much more natural way than our artificial toy models. But this raises, again, the question whether one should extend the strong Church-Turing principle to non-computational problems. It could read roughly as follows: Given a physical system with Hilbert space $\mathcal{H}$ which has an interface to a quantum computer such that the quantum state on $\mathcal{H}$ can be swapped

with the register state. Assume there is an efficient implementation for a transformation U on $\mathcal{H}$ which achieves some non-computational task. Then there is always an efficient quantum circuit V such that

$$U = W^\dagger \, V \, W \, ,$$

where W transfers the state of $\mathcal{H}$ to some subspace of the register. Applied to heat engines, this would mean: If there is any efficient way to extract energy from a family of increasing systems and there is an interaction with a quantum computer which is powerful enough to implement the SWAP, then one can always realize the heat engine by a quantum computer.

4.3 Imaging and Material Analysis

We have already mentioned that imaging can also be considered as a generalized measurement. This will only play a role in a part of this section. The aim of this section is rather to sketch some examples where imaging may lead to algorithmic control problems. The type of algorithm discussed in the next subsection is from the mathematical point of view simply a special case of the simulation of Hamiltonians (Section 3.1). However, according to the general message of this thesis, I preferred to mention it here in order to emphasize that simple forms of these algorithmic control techniques have already been applied in NMR spectroscopy for decades without any computational application in mind. The remaining subsections sketch only vague ideas about perspectives for future imaging technologies. It is certainly not possible to discuss experimental feasibility at this stage.

4.3.1 Decoupling Strategies and their Complexity

It is maybe a matter of opinion to judge when the era of quantum information processing began. However, there are decades-old standard techniques in NMR imaging involving algorithmic control of many-body quantum systems. We rephrase these decoupling strategies and sketch the complexity theory which we have developed [46].

The principle of NMR imaging is the following. The nuclear spins are subjected to a static magnetic field, whose direction is referred to as the z-axis. Then the Hamiltonian of the system is a scalar multiple of the Pauli matrix σ_z, i.e.,

$$H = Bc\sigma_z \tag{4.6}$$

where B is the strength of the field and c an appropriate constant. The initial state is a thermal equilibrium state where the probability for the upper state $|1\rangle$ is slightly lower than for the lower state $|0\rangle$. Now we neglect that this bias is only small (and take this only into account by the remark that it causes a bad signal to noise ratio) and assume that the initial state is $|0\rangle$. By applying an oscillating magnetic field the qubit is brought into the superposition state $(1/\sqrt{2})(|0\rangle + |1\rangle)$. The natural time evolution according to H is the phase rotation $(1/\sqrt{2})(|0\rangle + \exp(-iBct))|1\rangle)$ which is, classically speaking, a precession. This motion causes the system to emit radiation with the frequency $\omega = Bc$. For different molecules these frequencies are different, giving some information about the chemical structure. Using magnetic fields with different strength at different positions

one could also get information about the distribution of molecules at different positions. With classical post-processing one obtains NMR images.

The Hamiltonian (4.6) was oversimplified; in reality, the Hamiltonian of a molecule with several spins involves interactions between the spins. Assuming that we have n spin-1/2 particles (which allows the interpretation of the molecule as an n-qubit quantum register) the Hamiltonian of the n-spin molecule could have the form

$$H := \sum_j \sigma_z^{(j)} + \sum_{jk} J_{jk,\alpha,\beta} \sigma_\alpha^{(j)} \sigma_\beta^k$$

with an $3n \times 3n$ coupling matrix J. For many tasks, the interactions are unwanted, therefore one would like to cancel them by applying appropriate unitary transformations. Using the language of mutual simulation of Hamiltonians, the task of decoupling is to simulate the Hamiltonian $\tilde{H} = 0$ using H. It has been observed [157] that the combinatorial concept of orthogonal arrays [158, 159] offers a systematic way to construct decoupling schemes. We rephrase this concept for the case that the coupling involves only $\sigma_z \otimes \sigma_z$ terms between the qubits.

Assume we apply the transformation σ_x to qubit j and let the system evolve due to its natural evolution and apply σ_x again. Then the whole dynamics is as if the system was subjected to a modified Hamiltonian, namely the operator obtained by adding a minus sign to all interactions that involve qubit j due to $\sigma_x \sigma_z \sigma_x = -\sigma_z$. This is a very simple way to remove all interactions with qubit j. Certainly we cannot use it to remove the interactions between *all* pairs in a single step because the interaction $\sigma_z \otimes \sigma_z$ is unchanged if the σ_x-transformation is applied to *both* qubits. To cancel all interactions one has to use more than one time step. The whole problem can be formulated with graphs: Let each qubit be a vertex. Each time-step is represented by a labeling of the vertices with labels ± 1. To conjugate the time evolution by a σ_x transformation on qubit j corresponds to labeling node j with -1, not to apply it corresponds to $+1$. The product of the labels of nodes j, k determines the value that the edge (j, k) obtains, showing whether or not the interaction between qubit j and k is inverted. For each edge the interaction of the whole procedure is the sum over all time steps when they are chosen to last equally long. To cancel the interaction the values $+1$ and -1 must occur equally often for each vertex. The whole scheme is represented by a matrix where each row shows the sequence of signs over the time steps. The inner product between row j and k is the total interaction strength between j and k over all time steps. Hence the matrix is a decoupling scheme if and only if all rows are mutually orthogonal. This is for instance the case when for each two rows each of the 4 pairs $(\pm 1, \pm 1)$ occur equally often over the columns. In the language of combinatorics the matrix is an orthogonal array "of strength 2". For an extensive study and overview of decoupling schemes see [160] and references therein. In [161] it is shown that not only does every orthogonal array define a decoupling scheme, but also the converse.

The fact that orthogonal arrays exist for every power of 2 implies that the interaction above can be cancelled within 2^k time steps whenever $2^k \geq n$. Hence the decoupling schemes have time step complexity $O(n)$. In [46] we show that this bound is optimal for many interactions. If one would restrict the attention to the example above where one has to find n orthogonal vectors with ± 1 as entries, one had obviously the lower bound n, since the number of time steps is the dimension of the vectors. General decoupling

schemes could apply arbitrary unitary transformations to each qubit. We briefly explain how to derive a complexity bound in this case. First we could derive it from complexity bounds for the time inversion problem to simulate $-H$ by H since this problem is closely related [46]. Here we will describe a direct derivation using the spectral methods explained in Subsection 3.1.3. We assume that the coupling matrix J is

$$J = K \otimes C$$

where

$$C := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

and K is the adjacency matrix of the complete graph. A decoupling scheme with N time steps of duration $t_1, \ldots, t_N$ using the orthogonal transformation $O_j = O_j^1 \oplus O_j^2 \oplus \cdots \oplus O_j^n$ in time step j satisfies

$$\sum_{j=1}^{N} t_j O_j J O_j^T = 0\,.$$

We add $R := \sum_j t_j O_j (1 \otimes C) O_j^T$ on both sides and obtain

$$\sum_j t_j O_j (I \otimes C) O_j^T = R\,.$$

The left hand side has at most rank N. To see that the right hand side has at least rank n we note that R is block diagonal and its restriction to block l is $\sum_j t_j O_j^{(l)} C O_j^{(l)T}$ which has at least rank 1 due to $0 \neq C \geq 0$. We conclude that decoupling schemes for the zz-interaction require at least n time steps. If n is a power of two this can exactly be achieved using an orthogonal array of length n. It is easy to see that the decoupling schemes sketched above work also for *selective* decoupling, where the spins are arbitrarily partitioned into cliques and only connections between spins in different cliques are removed.

4.3.2 Is an Image a Covariant POVM-Measurement?

The remarks below should only give a vague idea about how general the tool 'POVM-implementation' is as a theoretical concept. Traditionally, the term 'measurement' suggests numbers as outcomes, but the Platonic solid POVMs [141] are a nice example where it is more natural to use non-numerical labeling (for instance the vertices of the Platonic solids) for the measurement outcomes. For a POVM which is used to estimate which state in a number of potential states is present one may label the outcomes simply by the estimated state. Here we argue that it also could make sense to describe an image as the result of a POVM measurement whenever the imaged object is quantum.

An article with headline "electrons seen in its orbit" [162] showed pictures of the electron density in the electron orbitals of a crystal. The images were obtained using the diffraction of an electron beam. This raises the question of the ultimate quantum limits for obtaining pictures of wave functions of Schrödinger particles. First one has to clarify what a picture of an orbital *is*. Clearly, the unknown wave function of a single

electron cannot be pictured in the sense that we have depict real and imaginary part of the Schrödinger wave. Otherwise one could prepare the same wave function again and would have violated the no-cloning principle. Of course we could *estimate* the wave function after an appropriate measurement. Given that the electron is, for instance, in an energy eigenstate of an Hydrogen atom one could clearly measure the energy and depict the corresponding orbital. But this is obviously not what one means by an *image* because one has used too much prior information on the possible orbital shapes for the illustration. The imaging process is only useful if it can be applied to *unknown* orbitals. A reasonable method could be to use only the prior information that the electron is confined to some area by a potential with unknown shape and that its energy is at most some given value E. Then the remaining state space is of finite dimension d since the momentum satisfies $p^2 \leq 2Em$ and also the position is in a certain circle. On this subspace one could perform an estimation using a $SU(d)$ covariant POVM. This formalizes the demand that no prior information on the state is used apart from the confinement to $\mathbb{C}^d$. Given several copies of the same orbital $SU(d)$, covariant POVMs on the joint space allow better estimation of the wave function. If one prefers to measure a finite POVM which is a pretty good approximation for the uniform distribution of operators of the d-sphere one has to solve problems which are related to distributing points on the surface of a unit sphere such that they are pretty much uniformly distributed (problems which appear also in coding theory). We do not mathematically analyze this situation here, we only emphasize that covariance conditions may formalize the idea that an imaging process should by definition not require too much prior information on the depicted object. This can already be seen in usual photography: The photographer does not have to tell his camera where the object is which should be shot. If the object is translated, it appears simply at another position on the picture. As an idealization, one may assume that $\mathbb{R}^3$-translations of the object lead to $\mathbb{R}^2$-translations on the picture. This is certainly a $\mathbb{R}^3$-covariant measurement: if one neglects the boundaries of the image and the fact that it consists of pixels the measurement outcomes are functions on $\mathbb{R}^2$ indicating the color distribution. If the depicted object is quantum, such an analysis could indeed make sense in order to investigate ultimate quantum limits for the resolution. Whether covariant POVMs will really play a crucial role in the construction of optimal imaging techniques has to be left to the future.

4.3.3 Microscopy with Pre-Processing the Input Beam

Interesting proposals to improve microscopy and lithography by using entangled photon states can be found in the literature [163, 164]. The idea is that light which consists of entangled photons behaves in some respect like photons with smaller wave length. Theoretical analysis has shown that one therefore obtains images with sub-wavelength resolution. We will not explain the physics of this phenomenon here, but only want to mention that the proposal shows a natural algorithmic state preparation problem.

Microscopy and also lithography with entangled photons can use the so-called NOON state. It is a state of two light modes which should interfere and is defined as follows. The Hilbert space is $l^2(\mathbb{N}_0) \otimes l^2(\mathbb{N}_0)$ where $|n, m\rangle$ is a state with n photons in the first

mode and m in the second. Then one defines

$$NOON := \frac{1}{\sqrt{2}}(|N,0\rangle + |0,N\rangle)\,.$$

Then the proposals explain [164] how the NOON state can be prepared from the Fock state $|N,N\rangle$ by $N/2$ operations. These operations $\Phi_1,\ldots,\Phi_{N/2}$ are, however, not unitary transformations since they are obtained by an optical interferometric device with two beam splitters, two detectors, and one phase plate where the output state is only used whenever there are coincident clicks in the detectors. Given such a coincidence, the device Φ_k has transformed the state $|n,m\rangle$ to

$$(a^2 + e^{i\phi_k}b^2)|n,m\rangle\,,$$

where a and b are annihilation operators on the first and second mode, respectively, i.e., they decrease the photon numbers by 1. Using a and b, the NOON state can be written as

$$NOON = (a^N + b^N)|N,N\rangle\,,$$

which can be decomposed into

$$(a^2 + e^{i\phi_1}b^2)(a^2 + e^{i\phi_2}b^2)\cdots(a^2 + e^{i\phi_{N/2}}b^2)|N,N\rangle$$

when the phase factors ϕ_k are chosen as the $N/2$-th roots of unity, i.e.,

$$\phi_k = \frac{4\pi k}{N}\,.$$

The authors of [164] admit that the performance of the scheme scales badly with increasing N since its success probability decreases exponentially. They conjecture that this problem could possibly be solved by producing NOON states off-line and storing them in memories.

4.3.4 Microscopy with Quantum Post-Processing

The intention of the proposal above was to improve microscopy by generating quantum states of light which would allow higher resolution. One could say that the input beam is therefore subjected to a 'pre-computation' which prepares the desired state. It is also interesting to think about schemes which apply quantum transformations to the post-object beam in order to convert the information which is inherent in the quantum state of the output beam to classical information in an optimal way. Here we describe a model which is formally *quantum process tomography*. By this one means algorithms to gain information about the unknown dynamics of a quantum system. We first explain the usual setting for process tomography. Given a Hilbert space $\mathcal{H}$ and an unknown completely positive map $G \in \mathcal{G}$ where $\mathcal{G}$ is some set of possible maps, identify G by subjecting several initial states ρ to G and measuring the final states. This setting has been a widely studied subject (see e.g. [165, 166, 167]). It has been emphasized in [87] that interferometry experiments, for instance in the context of microscopy, are formally equivalent to interference in quantum computing. Here we argue that there are standard

methods in electron holography which have direct interpretation as quantum process tomography. Then we want to use the formal setting of quantum information processing in order to show that it offers interesting perspectives. Actually, the main part of this section deals with ideas which are standard in quantum interferometry. The intention is only to show that quantum information language could help to further develop these ideas.

We first describe so-called off-axis electron holography (see e.g. [168, 169]) in a simplified way. An electron source with a lens generates plane electron waves. The object is located in the electron beam such that only part of each wave front passes the object, the other part is later used as reference beam. The object changes the electron wave with respect to its phase and its amplitude. For simplicity we assume that the amplitude is not affected, i.e., all electrons pass the object. Such objects are called *phase objects*. The object beam and reference beam are diffracted by a biprism in such a way that they interfere. According to the phase change caused by the object, one observes constructive or destructive interference. Consider the following simplified mathematical model. The Schrödinger wave of a single electron is described by a vector in $L^2(\mathbb{R}^2)$ when the degree of freedom in the direction of propagation is not explicitly taken into account. We assume that the object beam is confined to some region A and the reference beam to some region A'. Therefore the wave function is actually confined to the subspace $L^2(A) \oplus L^2(A')$ where A' is obtained by translating A by the vector l, i.e., $A' = A + l$.

The phase object can be described by a diagonal unitary which maps the wave function $|\psi\rangle$ onto $U|\psi\rangle$ with $U\psi(x) = u(x)\psi(x)$, where $u(x) = \exp(i\phi(x))$ and $u(x) = 1$ for all $x \notin A$. When both beams interfere we obtain a probability distribution to detect the electron on a screen which is described by

$$|\psi(x) + u(x)\psi(x)|^2$$

which gives certainly some insight on the phase function $u(x)$ after sampling with sufficiently many electrons. The task to obtain as much information about U as possible is therefore an issue of quantum process tomography.

Now we describe a situation where terminology and tools of quantum information processing may help to create new methods in microscopy. Assume one has two objects and would like to know whether they differ considerably or not. We could split a plane Schrödinger wave such that one part passes object O and the other passes object O'. Assuming that the input waves, before they pass the objects, have constant probability amplitude and phase, the joint wave behind the objects is given by

$$\psi := c(u \oplus u')\,,$$

where c is some positive normalization factor. In order to check whether the wave functions u and u' are substantially different one could use a usual interference device: A beam splitter could transfer the function to

$$\widetilde{\psi} := \frac{c}{\sqrt{2}}((u + u') \oplus (u - u'))\,.$$

The probability for finding the particle in the left beam is given by

$$\frac{c^2}{2}(1 + 2Re(\int u(x)\overline{u'(x)}\,dx))\,,$$

where Re denotes the real part of a complex number. For $u = u'$ we obtain 1. After a small number of runs of the experiment one can estimate whether the real part of the integral is considerably smaller than 1 which would indicate a difference of the phase functions u and u'.

Another possibility to compare the objects is to choose *two* electrons, one which passes object O and one for object O'. The output state of the bipartite system is therefore

$$\frac{1}{c^2}(u \otimes u') \,.$$

The controlled SWAP permutes

$$u \otimes u' \mapsto u' \otimes u$$

conditioned on the ancilla state. The ancilla is prepared and measured as above. The probability for a positive measurement result is now given by

$$\langle u \otimes u' | P^+ | u \otimes u' \rangle = \frac{c^2}{2}(1 + | \int u(x)\overline{u'(x)}\, dx|^2) \,.$$

This measurement is in some respects more sensitive to small differences between the two wave functions: If the overlap between u and u' is only slightly below 1 the probability for the P^+ measurement is more reduced than above since the term appears quadratically. On the other hand the two-electron experiment is insensitive to global phase differences between u and u'. If $u' = \exp(i\lambda)u$ with $\lambda \in \mathbb{R}$, only the one-electron experiments detects that the functions are different. Note that the terms 'symmetric' and 'antisymmetric' should not be confused with symmetry or antisymmetry of the whole wave function of two particles. Since electrons are fermions the total wave function of two electrons is always antisymmetric. However, here we can treat the particles as if they were distinguishable since we identify them by being in different beams.

Perhaps one could think of applications where it is important to use as few electrons as possible in order to disturb or damage the object as little as possible. It should be noted that this goal is formally equivalent to minimizing the number of queries of an oracle for answering a certain question. The oracle is here given by the unitary $U \otimes U'$. Each passing electron is one query. We can formally consider the two object as one 'bipartite object' where the decision problem consists of 'similarity'. For one object there could be other features of interest like: 'Is there a region with area at least a where the phase $\phi(x)$ satisfies $|\phi(x)| > d$ for some $0 < d < \pi$?' This could, for instance indicate that the object is thicker in that region. We summarize:

Observation 1 (Radiation Exposure is Query Complexity)
The problem to answer a question about a phase object with as few electrons as possible is formally equivalent to the construction of an algorithm for determining some analogous property of a black-box diagonal unitary such that the query complexity is minimal.

This relation should show that a quantum algorithmic approach to microscopy could in principle help to reduce the radiation of objects when only a certain aspect of the object is relevant. This connection between physical resources and computer scientific resources shows that Landauer's statement 'information is physical' [170] has many interesting aspects.

Chapter 5

Conclusions

David Deutsch emphasized [94] that the laws of physics determine which computational problems can be solved efficiently. The approach here seems to reverse his statement by claiming that the laws of computation determine which physical processes can be performed efficiently. However, this point of view ignores the fact that the laws of physics determine which operations are reasonable to consider as the elementary operations of quantum computing. But given these operations, the question how to concatenate them in order to generate a complex process is certainly an issue of a generalized computer science.

A complexity theory for physical processes will always have the problem that it is not clear which parts of the whole experimental setup has to be taken into account. If the controlled-not in an ion trap is considered as an *elementary* operation, one neglects all the operations which were necessary to cool the vibrational modes. Furthermore the laser has already solved the non-trivial quantum control problem of generating a many-photon coherent light beam. Hence elementary operations on a quantum register may require a huge experimental setup which solves already complex non-computational control problems to enable the 'elementary' operation. But this does not show that complexity theory does not make sense here; it rather leads to a complexity theoretical aspect of reliability. In other words, it suggests that we need of a complexity theory that takes into account the resource requirements specific to *precise* control. One would like to know how the complexity of the controlling devices increases if a 'basic' operation should be performed more and more reliably. This connection meets very well the spirit of this thesis: We have discussed, for instance, how the complexity of a measurement can increase with the demanded accuracy and have illustrated in toy models that the complexity of a heat engine grows with its efficiency. Furthermore, our toy models suggest that the resource requirements of cooling processes diverge for information theoretic reasons when approaching the zero temperature limit. However, it is not clear whether the asymptotic point of view which is usual in complexity and information theory will ever be as useful in physics as in computer science. Perhaps the finiteness of all resources is much more relevant for difficult control tasks. Nevertheless I hope that the considered toy models can give some intuition and hints on the real physical limitations.

Bibliography

[1] Wikipedia. Free encyclopedia: Algorithm. *http://en.wikipedia.org/wiki/Algorithm*, version found at some day in 2004.

[2] E. Horowitz and S. Sahni. *Fundamentals of Data Structures*. Computer Science Press, 1976.

[3] M. Nielsen and I. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[4] P. Dirac. *The Principles of Quantum Mechanics*. Clarendon Press, Oxford, 1949.

[5] D. Janzing. *Quantum Computing Models as a Tool Box for Controlling and Understanding the Nanoscopic World*. Informatik in Forschung und Entwicklung, special issue Quantum Information Technology. Springer Verlag, Berlin, 2006.

[6] L. Hardy. Quantum theory from five reasonable axioms. *http://xxx.lanl.gov/abs/quant-ph/0101012*.

[7] L. Hardy. Why quantum theory? *http://xxx.lanl.gov/abs/quant-ph/0111068*.

[8] D. DiVincenzo. Two-qubit gates are universal for quantum computation. *Phys. Rev A*, 51:1015–1022, 1995.

[9] W. Wootters and W. Zurek. A single quantum bit can not be cloned. *Nature*, page 802, 1982.

[10] T. Toffoli. Reversible computing. *MIT Report MIT/LCS/TM-151*, 1980.

[11] R. Landauer. Irreversibility and heat generation in the computing process. *IBM J. Res. Develop.*, 5:183–191, 1961.

[12] A. Barenco, A. Derthiaume, R. Cleve, D. DiVincenzo, N. Margolus, P. Shor, T. Leatir, J. Smolin, and H. Weinfurter. Elementary gates for quantum computation. *Phys. Rev. A*, 52(5):3457–3467, 1995.

[13] A. Kitaev. Quantum measurements and the abelian stabilizer problem. *Electronic Colloquium on Computational Complexity*, (TR96-003), 1996. see also *http://xxx.lanl.gov/abs/quant-ph/9511026*.

[14] A. Kitaev, A. Shen, and M. Vyalyi. *Classical and Quantum Computation*, volume 47. Am. Math. Soc., Providence, Rhode Island, 2002.

[15] J. Myers. Can a universal quantum computer be fully quantum? *Phys. Rev. Lett.*, 78:1823, 1997.

[16] N. Linden and S. Popescu. The halting problem for quantum computers. *http://xxx.lanl.gov/abs/quant-ph/9806054*.

[17] Omnès, R. *The interpretation of quantum mechanics*. Princeton Series in Physics. Princeton University Press, 1994.

[18] G. Miller. Riemann's hypothesis and test for primality. *Journ. Comp. System Sci.*, 13:300–317, 1976.

[19] J. Håstad. *Computational Limitations for Small-Depth Circuits*. MIT Press, Cambridge, Massachusetts, 1987.

[20] I. Cirac and P. Zoller. Quantum computation with cold trapped ions. *Phys. Rev. Lett.*, 74:4091–4094, 1995.

[21] N. Gershenfeld and I. Chuang. Bulk spin-resonance quantum computation. *Science*, 275:350–356, 1997.

[22] A. Steane. Error correcting codes in quantum theory. *Phys. Rev. Letters*, 77:793–797, 1996.

[23] R. Raussendorf and H. Briegel. Quantum computing via measurements only. *Phys. Rev. Lett.*, page 5188, 2000.

[24] S. Lloyd and S. Braunstein. Quantum computation over continuous variables. *Phys. Rev. Lett.*, 82(8):0031–9007, 1999.

[25] B. Sanders, S. Bartlett, and H. de Guise. From qubits to continuous-variable quantum computation. *CLEO/Europe Focus 2000*, Munich 2001. See also *http://xxx.lanl.gov/abs/quant-ph/0208008*.

[26] F. Schmüser and D. Janzing. On quantum analogue-to-digital and digital-to-analogue conversion. *Phys. Rev. A*, 72:042324, 2005.

[27] Josef M. Jauch. *Foundations of quantum mechanics*. Addison-Wesley, Reading, Mass., 1968.

[28] E. Davies. *Quantum theory of open systems*. Academic Press, London, 1976.

[29] T. Cover and J. Thomas. *Elements of Information Theory*. Wileys Series in Telecommunications, New York, 1991.

[30] H. Ollivier and W. Zurek. Quantum discord: A measure for the quantumness of correlations. *Phys. Rev. Lett.*, 88:017901, 2002.

[31] C. Helstrom. *Quantum Detection and Estimation Theory*. Academic Press, New York, 1976.

[32] R. Schumann. Developping discord. Unpublished manuscript, 2001.

[33] W. Zurek. Private communication, 2004.

[34] D. DiVincenzo. Quantum computation. *Science*, 270:255, 1995.

[35] L. Schulman and U. Vazirani. Scalable NMR Quantum Computers. *http://xxx.lanl.gov/abs/quant-ph/9804060*, 1998.

[36] V. Shende, S. Bullock, and I. Markov. A practical top-down approach to quantum circuit synthesis. *to appear in Proc. ACM/IEEE Asia and Pacific Design Automation Conf., Shanghai, January 2005*, *http://xxx.lanl.gov/abs/quant-ph/0406176*.

[37] Mottonen et. al. Transformation of quantum states using uniformly controlled rotations. *http://xxx.lanl.gov/abs/quant-ph/0407010*.

[38] C. Mora and H. Briegel. Algorithmic complexity of quantum states. *http://xxx.lanl.gov/abs/quant-ph/0412172*.

[39] A. Soklakov and R. Schack. Efficient state preparation for a register of quantum bits. *http://xxx.lanl.gov/abs/quant-ph/0408045*.

[40] D. Janzing and Th. Beth. Remark on multi-particle states and observables with constant complexity. *http://xxx.lanl.gov/ abs/quant-ph/0003117*.

[41] G. Lüders. Über die Zustandsänderung durch den Meßprozeß. *Ann. Phys.*, 8:322, 1951.

[42] A. Peres. Neumark's theorem and quantum inseparability. *Found. of Physics*, 12:1141–1453, 1990.

[43] A. Peres. *Quantum Theory: Concepts and Methods*. Kluwer Academic, 1993.

[44] H. Reichenbach. *The direction of time*. Dover, 1999.

[45] S. Hawking. *A brief history of time*. Bantam, 1995.

[46] D. Janzing, P. Wocjan, and Th. Beth. Complexity of decoupling and time-reversal for n spins with pair-interactions: Arrow of time in quantum control. *Physical Review A*, 66:042311, 2002.

[47] W. K. Rhim, A. Pines, and J. S. Waugh. Violation of the spin-temperature hypothesis. *Phys. Rev. Lett.*, 25:218–220, 1970.

[48] D. Janzing, F. Armknecht, R. Zeier, and Th. Beth. Quantum control without access to the controlling interaction. *Phys. Rev. A*, 65:022104, 2002.

[49] H.D. Zeh. There are no quantum jumps, nor there are particles! *Phys. Lett. A*, 172:189–192, 1993.

[50] D. Janzing and P. Wocjan. Ergodic quantum computing. *Quant. Inf. Process.*, 4(2):129–158, 2005.

[51] D. Janzing. Spin-1/2 particles moving on a 2D lattice with nearest-neighbor interactions can realize an autonomous quantum computer. *http://xxx.lanl.gov/abs/quant-ph/0506270*.

[52] D. Janzing and T. Beth. Fragility of a class of highly entangled states with n qubits. *Phys. Rev. A*, 61:052308, 2000.

[53] D. Poulin. Macroscopic observables. *Phys. Rev. A*, 71:022102, 2005.

[54] D. Guilini, E. Joos, C. Kiefer, J. Kupsch, I.-O. Stamatescu, and H.D. Zeh. *Decoherence and the Appearance of a Classical World in Quantum Theory*. Springer, Berlin, 1996.

[55] F. Reza, editor. *An Introduction to Information Theory*. McGraw Hill, New York, 1961.

[56] D. Janzing. On the computational power of molecular heat engines. *J. Stat. Phys.*, 122(3):531–556, 2006.

[57] A. Lenard. Thermodynamical proof of the Gibbs formula for elementary quantum systems. *Journ. Stat. Phys.*, 19:575, 1978.

[58] W. Pusz and L. Woronowicz. Passive states and KMS states for general quantum systems. *Comm. Math. Phys.*, 58:273–290, 1978.

[59] D. Janzing, P. Wocjan, R. Zeier, R. Geiss, and Th. Beth. Thermodynamic cost of reliability and low temperatures : Tightening Landauer's principle and the Second Law. *Int. Jour. Theor. Phys.*, 39(12):2217–2753, 2000.

[60] M. Nielsen. Chracterizing mixing and measurements in quantum mechanics. *http://xxx.lanl.gov/abs/quant-ph0008073*.

[61] M. Horodecki, P. Horodecki, and J. Oppenheim. Reversible transformations from pure to mixed states, and the unique measure of information. *Phys. Rev. A*, 67:062104, 2003.

[62] M. Ohya and D. Petz. *Quantum entropy and its use*. Springer Verlag, 1993.

[63] D. Janzing and T. Beth. Quasi-order of clocks and their synchronism and quantum bounds for copying timing information. *IEEE Trans. Inform. Theor.*, 49(1):230–240, 2003.

[64] O. Bratteli and D. Robinson. *Operator algebras and quantum statistical mechanics*, volume 1. Springer, New York, 1987.

[65] G. Murphy. *C^*-algebras and operator theory*. Academic Press, Boston, 1990.

[66] H. Fang, K. Matsumoto, X. Wang, and M. Wadati. Quantum cloning machnines for equatorial qubits. *http://xxx.lanl.gov/abs/quant-ph/0101101*.

[67] D. Bruss, M. Cinchetti, G. D'Ariano, and C. Machiavello. Phase covariant quantum cloning. *Phys. Rev. A*, 62:12302, 2000.

[68] D. Janzing. Decomposition of time-covariant operations on quantum systems with continuous and/or discrete energy spectrum. *Journ. Math. Phys.*, page 122107, 2005.

[69] E. Fredkin and T. Toffoli. Conservative logic. *Int. Journ. Theor. Phys.*, 21(219), 1982.

[70] C. H. Bennett. Logical reversibility of computation. *IBM J. Res. Develop.*, 17:525–532, 1973.

[71] C. H. Bennett. Time/space trade-offs for reversible computation. *SIAM J. Computing*, 18(4):766–776, 1989.

[72] P. Benioff. The computer as a physical system: A microscopic quantum mechanical model of computers as represented by Turing machines. *J. Stat. Phys.*, 22(5):562–591, 1980.

[73] R. Feynman. Quantum mechanical computers. *Opt. News*, 11:11–46, 1985.

[74] R. Feynman. Quantum mechanical computers. *Found. Phys.*, 16(6):503–531, 1986.

[75] T. Gramss. On the speed of quantum computers with finite size clocks. *Santa Fe Institute Working Papers*, 1995.
http://www.santafe.edu/sfi/publications/wpabstract/199510086.

[76] N. Margolus. Parallel quantum computation. In W. Zurek, editor, *Complexity, Entropy, and the Physics of Information*. Addison Wesley Longman, 1990.

[77] K. Imai and Morita K. A computation-universal two-dimensional 8-state triangular reversible cellular automaton. In *Proc. Second Colloquium on Universal Machines and Computations*, volume II, pages 90–99, Metz, 1998.

[78] K. Morita and M. Harao. Computation universality of one-dimensional reversible (injective) cellular automata. *Trans. EICE Japan*, 1989.

[79] T. Toffoli. Computation and construction universality of reversible cellular automata. *J. Comp. Syst. Sci.*, 15:213–231, 1977.

[80] K. Morita and K. Imai. Number-conserving reversible cellular automata and their computation-universality. In *Proceedings of MFCS'98 Workshop on Cellular Automata*, Brno, 1998.

[81] J. Ziman. *Principles of the Theory of Solids*. Cambridge University Press, 1972.

[82] C. Bennett, D. DiVincenzo, Fuchs C., T. Mor, E. Rains, P. Shor, J. Smolin, and W. Wootters. Quantum nonlocality without entanglement. *Phys. Rev. A*, 59:1070–1091, 1999.

[83] W. Zurek. Quantum Discord and Maxwell's Demons. *Phys. Rev. A*, 67:012320, 2003.

[84] D. Janzing and T. Beth. Synchronizing quantum clocks with classical one-way communication: Bounds on the generated entropy. *http://xxx.lanl.gov/abs/quant-ph/0306023v1*.

[85] D. Janzing and T. Beth. Bounds on the entropy generated when timing information is extracted from microscopic systems. *http://xxx.lanl.gov/abs/quant-ph/0301125*.

[86] R. Feynman. *Feynman Lectures on Computation*. Perseus Pr., 1996.

[87] H. Lee, P. Kok, and J. Dowling. A quantum Rosetta stone for interferometry. *http://xxx.lanl.gov/abs/quant-ph/0202133*.

[88] J. van Leeuwen, editor. *Algorithms and Complexity*. Elsevier, Cambridge, MA, 1990.

[89] S. Lloyd. Universal quantum simulators. *Science*, 273:1073, 1996.

[90] D. Janzing, P. Wocjan, and T. Beth. "Non-Identity check" is QMA-complete. *Int. Journ. Quant. Inf.*, 3(3):463–473, 2005.

[91] J. E. Hopcroft and J. D. Ullman. *Introduction to Automata Theory, Languages, and Computation*. Addison-Wesley series in computer science. Addison-Wesley, 1979.

[92] M. Garey and D. Johnson. *Computers and Intractibility*. Freeman and Company, New York.

[93] J. Watrous. PSPACE has 2-round quantum interactive proof systems. *Theoretical Computer Science, Algorithms, automata, complexity and games*, 292:575–588, 2003.

[94] D. Deutsch. *The Fabric of Reality*. The Penguin Press, 1997.

[95] E. Lieb. The Hubbard model: Some rigorous results and open problems. *http://xxx.lanl.gov/abs/cond-mat/9311033*.

[96] M. Biafore. Can quantum computers have simple Hamiltonians? In *Proc. Workshop on Physics of Comp.*, pages 63–86, Los Alamitos, CA, 1994. IEEE Computer Soc. Press.

[97] A. Ambainis and J. Watrous. Two-way finite automata with quantum and classical states. *Theoretical Computer Science*, 287:299–311, 2002.

[98] B. Schumacher and R. Werner. Reversible quantum cellular automata. *http://xxx.lanl.gov/abs/quant-ph/0405174*.

[99] J. Dodd, M. Nielsen, M. Bremner, and R. Thew. Universal quantum computation and simulation using any entangling hamiltonian and local unitaries. *Phys. Rev. A*, 65:040301, 2002.

[100] M. Bremner, D. Bacon, and M. Nielsen. Simulating Hamiltonian dynamics using many-qudit Hamiltonians and local unitary control. *Phys. Rev. A*, 71:052312, 2005.

[101] P. Wocjan, M. Rötteler, D. Janzing, and Th. Beth. Simulating Hamiltonians in quantum Networks: Efficient schemes and complexity bounds. *Phys. Rev. A*, 65:042309, 2002.

[102] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 47:777–780, 1935.

[103] J. Bell. On the Einstein-Podolsky-Rosen paradox. *Physics*, 1:195–200, 1964. *reprinted in* J. Bell: Speakable and unspeakable in quantum mechanics, Cambridge University Press, Cambridge, 1987.

[104] J. Pearl. Bayesian networks: A model of self-activated memory for evidential reasoning. In *Proceedings, Cognitive Science Society*, pages 329–334, Greenwich, Albex, 1985.

[105] J. Pearl. *Causality.* Cambridge University Press, 2000.

[106] P. Spirtes, C. Glymour, and R. Scheines. *Causation, Prediction, and Search.* Lecture Notes in Statistics. Springer, New York, 1993.

[107] C. Bennett and S. Wiesner. Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states. *Phys. Rev. Lett.*, 69(20):2881, 1992.

[108] C. Granger and J. Lin. Using the mutual information coefficient to identify lags in nonlinear models. *Journal of Time Series Analysis*, 15(4):371–384, 1994.

[109] D. Janzing. Kann Statistik Ursachen beweisen?, Lecture notes Sommersemester 03, Universität Karlsruhe, online available (November 2006) *http://iaks-www.ira.uka.de/home/janzing/kausalskriptum03.ps*

[110] J. Henson. Comparing causality principles. *http://xxx.lanl.gov/ abs/quant-ph/041005.*

[111] D. Heckermann and R. Shachter. Decision-theoretic foundations for causal reasoning. *Journ. Art. Intelligence Res.*, (3):405–30, 1995.

[112] D. Janzing and T. Beth. on the potential influence of quantum noise on measuring effectiveness in clinical trial. *Quant. Inf., & Comp.*, vol. 4, (2):347, 2006.

[113] N. Margolus and L. Levitin. The maximum speed of dynamical evolution. *Physica D*, 120:188–195, 1998.

[114] D. Janzing. Quantum algorithm for measuring the energy of n qubits with unknown pair-interactions. *Quant. Inform. & Comp.*, 2(3):198–207, 2002.

[115] D. Janzing and Th. Beth. Distinguishing n Hamiltonians on C^n by a single measurement. *Phys. Rev. A*, 65:022303, 2002.

[116] L. Viola and E. Knill. Robust dynamical decoupling with bounded controls. *Phys. Rev. Lett.*, 90:037901, 2003.

[117] D. Janzing and Th. Beth. Complexity measure for continuous time quantum algorithms. *Phys. Rev. A*, 64(2):022301, 2001.

[118] P. Wocjan, D. Janzing, and Th. Beth. Simulating arbitrary pair-interactions by a given Hamiltonian: Graph-theoretical bounds on the time complexity. *Quant. Inform. & Comp.*, 2(2):117–132, 2002.

[119] C. Bennett, J. Cirac, M. Leifer, D. Leung, N. Linden, S. Popescu, and G. Vidal. Optimal simulation of two-qubit Hamiltonians using general local operations. *Phys. Rev. A*, 66:012305, 2002.

[120] C. P. Slichter. *Principles of Magnetic Resonance*. Springer, 3rd edition, 1990.

[121] P. Wocjan, M. Rötteler, D. Janzing, and Th. Beth. Universal simulation of hamiltonians using a finite set of control operations. *Quant. Inform. & Comp.*, 2(2):133–150, 2002.

[122] D. Janzing, P. Wocjan, and T. Beth. On the computational power of physical interactions: Bounds on the number of time steps for simulating arbitrary interaction graphs. *Int. Jour. Found. Comp. Science, special issue for Quantum Computation*, 14(5):889–903, 2002.

[123] B. Luy and S. Glaser. Superpositions of scalar and residual dipolar couplings: Analytical transfer of the spin 1/2 under cylindrical mixing conditions. *J. Magn. Res.*, 148:169–181, 2001.

[124] Cvetkovic D., M. Doob, and H. Sachs. *Spectra of Graphs*. Johann Ambrosius Barth Verlag, 3rd edition, 1995.

[125] Khaneja, N. and Brockett, R. and Glaser, S. Time optimal control in spin systems. *Phys. Rev. A*, 63(3):032308, 2001.

[126] N. Khaneja, S. Glaser, and R. Brockett. Sub-Riemannian geometry and time optimal control of three spin systems: Quantum gates and coherence transfer. *Phys. Rev. A*, 71:039906, 2005.

[127] G. Vidal, K. Hammerer, and J. Cirac. Interaction cost of nonlocal gates. *Phys. Rev. Lett.*, 63:237902, 2002.

[128] R. Zeier, M. Grassl, and T. Beth. Gate simulation and lower bound on the simulation time. *Phys. Rev. A*, 70:032319, 2004.

[129] E. Farhi et al. A quantum adiabatic evolution algorithm applied to random instances of an NP-complete problem. *Science*, 292(5516):472–475, 2001.

[130] W. Kaminsky and S. Lloyd. Scalable architecture for adiabatic quantum computing of NP-hard problems. In *Quantum computing & Quantum Bits in Mesoscopic System*. Kluwer, Dordrecht, 2003.

[131] D. Mitchell, C. Adami, W. Lue, and C. Williams. A random matrix model of adiabatic quantum computing. *http://xxx.lanl.gov/abs/quant-ph/0409088*.

[132] P. Wocjan and Th. Beth. The 2-local Hamiltonian problem encompasses NP. *Int. Journ. Quant. Inf.*, 1(3):349–357, 2003.

[133] P. Wocjan, D. Janzing, and Th. Beth. Treating the Independent Set Problem by 2D Ising Interactions with Adiabatic Quantum Computing. *Quant. Inf. Proc.*, 2(4):259–270, 2003.

[134] Oblak et. al. Quantum noise limited interferometric measurement of atomic noise: towards spin squeezing on the Cs clock transition. *Phys. Rev. A*, 71:033803, 2005.

[135] B. Travaglione and G. Milburn. Generation of eigenstates using the phase estimation algorithm. *http://eprint.uq.edu.au/archive /00000251/, see also http://xxx.lanl.gov /abs/quant-ph/0008053.*

[136] R. Cleve, A. Ekert, C. Macchiavello, and M. Mosca. Quantum algorithms revisited. *Proc. Roy. Soc. London A*, 454:339–354, 1998. See also *http://xxx.lanl.gov/abs/quant-ph/9708016.*

[137] P. Wocjan, D. Janzing, Th. Decker, and Th. Beth. Measuring 4-local n-qubit observables could probabilistically solve PSPACE. *Proceedings of the WISICT conference, Cancun 2004.* See also *http:// xxx.lanl.gov/abs/quant-ph/0308011.*

[138] K.-J. Lange, P. McKenzie, and A. Tapp. Reversible space equals deterministic space. *Journal of Computer and System Sciences*, 60:354–367, 2000.

[139] P. Shor. The Number of POVM Elements Needed for Accessible Information. In *Proc. of the Fifth Int. Conf. on Q. Comm., Meas. and Comp.*, pages 107–114, New York 2001. Kluwer.

[140] G. D'Ariano, P. Lo Presti, and M. Sacchi. A quantum measurement of the spin direction. *Phys. Lett. A.*, 292(233), 2002.

[141] T. Decker, D. Janzing, and T. Beth. Quantum circuits for single-qubit measurements corresponding to platonic solids. *Int. Journ. Quant. Inform.*, 2(3):353–377, 2004.

[142] T. Decker, D. Janzing, and M. Rötteler. Implementation of group-covariant positive operator valued measures by orthogonal measurements. *Journ. Math. Phys.*, 46:012104, 2005.

[143] S. Egner and M. Püschel. Symmetry-based matrix factorization. *Journ. Symb. Comp.*, 37(2):157–186, 2004.

[144] Rötteler, M. *Schnelle Signaltransformationen für Quantenrechner.* PhD thesis, Universität Karlsruhe, 2001. (In German).

[145] M. Püschel, M. Rötteler, and T. Beth. Fast fourier transforms for a class of non-abelian groups. *Proc. Applied Algebra, Algebraic Algorithms and error Correcting Codes (AAECC-13), Lecture Notes in Comp. Sc.*, 1719, 1999. Spinger.

[146] J. Renes. Spherical code key distribution protocols for qubits. *http://xxx.lanl.gov/abs/quant-ph/0402135*.

[147] P. Boykin, T. Mor, V. Roychowdhury, F. Vatan, and R. Vrijen. Algorithmic cooling of spins and scalable NMR quantum computing. In *Proc. of the National Academy of Sciences*, volume 99, pages 3388–3393, 2002.

[148] J. Fernandez, S. Lloyd, T. Mor, and V. Roychowdhury. Algorithmic cooling of spins: A practicable method for increasing polarization. *Int. Journ. Quant. Inf.*, 2(4):461–467, 2004.

[149] R. Schack. Simulation on a quantum computer. Informatik in Forschung und Entwicklung, special issue Quantum Information Technology. Springer Verlag, Berlin, 2006.

[150] F. Baharona. On the computational complexity of Ising spin models. *J. Phys. A: Math. Gen.*, 15:3241–3253, 1982.

[151] J. Kempe, A. Kitaev, and O. Regev. The complexity of the local Hamiltonian problem. *Proc. 24th FSTTCS, accepted to SICOMP*, 2004.

[152] J. Kempe and O. Regev. 3-local Hamiltonian is QMA-complete. *Quantum Computation and Information*, 3(3):258–264, 2003.

[153] P. Wocjan, D. Janzing, and T. Beth. Two QCMA-complete problems. *Quant. Inf. & Comp.*, 3(6):635–643, 2003.

[154] Ch. Papadimitriou. *Computational Complexity*. Addison Wesley, Reading, Massachusetts, 1994.

[155] D. Janzing, P. Wocjan, and Th. Beth. Cooling and low energy state preparation for 3-local Hamiltonians are FQMA-complete. *http://xxx.lanl.gov/abs/quant-ph/0305050*.

[156] Becker. T. and V. Weispfenning. *Gröbner Bases*. Springer, New York, 1993.

[157] M. Stollsteimer and G. Mahler. Suppression of arbitrary internal coupling in a quantum register. *Phys. Rev. A*, 64:052301, 2001.

[158] T. Beth, D. Jungnickel, and H. Lenz. *Design Theory*. Bibliographisches Institute/Cambridge University Press, 1993.

[159] A. Hedayat, N. Sloane, and J. Stufken. *Orthogonal Arrays*. Series in Statistics. Springer, 1999.

[160] P. Wocjan. *Computational Power of Hamiltonians in Quantum Computing*. PhD thesis, Universität Karlsruhe, 2003.

[161] M. Rötteler and P. Wocjan. Equivalence of decoupling schemes and orthogonal arrays. *http://xxx.lanl.gov/abs/quant-ph/0409135*.

[162] C. Humphrey. Electrons seen in orbit. *Nature*, 401:21–22, 1999.

[163] Kok P. Lee, H. and J. Dowling. Quantum imaging and metrology. *http://xxx.lanl.gov/abs/quant-ph/0306113*.

[164] C. Williams, P. Kok, H. Lee, and J. Dowling. Quantum lithography: a non-computing application of quantum information. 2006.

[165] A. Childs, I. Chuang, and D. Leung. Realization of quantum process tomography in NMR. *Phys. Rev. A*, 64:012314, 2001.

[166] D. Leung. Choi's proof and quantum process tomography. *J. Math. Phys.*, 44(2):528–533.

[167] C. Fuchs, R. Schack, and P. Scudo. A de Finetti Representation Theorem for Quantum Process Tomography. *Phys. Rev. A*, 69:062305, 2004.

[168] A. Tonomura. Applications of electron holography. *Rev. Mod. Phys.*, 59(3):673–669, 1987.

[169] H. Lichte. Electron holograpgy approaching atomic resolution. *Ultramicroscopy*, pages 293–304, 1986.

[170] R. Landauer. Information is physical. *Physics Today*, pages 23–29, May 1991.

Danksagungen

Zunächst einmal möchte ich mich bei Herrn Prof. Thomas Beth bedanken, dass er meine Habilitation von Anfang an unterstützt hat und mit seinen Ideen und seiner Begeisterung für die Ideen anderer ein so "querverbindungsfreudiges" Klima schuf. Vielen Dank auch an die Professoren Roland Vollmar und Hans Briegel für die Begutachtung dieser Arbeit.

Meinem Freund und langjährigen Bürogenossen Pawel Wocjan verdanke ich fachlich und privat einfach eine "bärenmäßig" gute Zeit am IAKS. Danke auch für's Korrekturlesen, womit mir auch Joe Renes und Frank Schmüser sehr geholfen haben.

Auch die gute Zusammenarbeit mit Thomas Decker hat zu dieser Arbeit schöne Beiträge geliefert; meinen anderen (aktuellen und ehemaligen) Kollegen ein herzliches Dankeschön für ihre Hilfsbereitschaft, ausführliche und interessante Diskussionen über viele Aspekte der Informatik, sowie über Algebra und Physik bis hin zur handfesten Computerei – insbesondere Markus Grassl, Martin Rötteler, Jörn Müller-Quade, Rainer Steinwandt, Jörg Moldenhauer und Ingo Boesnach.

Unsere wissenschaftlichen Hilfskräfte Anja Groch und Khoder El-Zein gaben mir mit ihren Computeralgebra-Experimenten anschauliche Beispiele für "algorithmische Wärmekraftmaschinen".

ISBN-13: 978-3-86644-083-8
ISBN-10: 3-86644-083-9

www.uvka.de