Quantum Algorithms
and
Quantum Entanglement

Quantum Algorithms and Quantum Entanglement

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam,
op gezag van de Rector Magnificus
prof. dr J.J.M. Franse
ten overstaan van een door het college voor promoties ingestelde
commissie in het openbaar te verdedigen in de Aula der Universiteit
op woensdag 3 november 1999 te 11.00 uur

door

Barbara Maria Terhal

geboren te Leiden

Promotor: prof. dr ir P.M.B. Vitányi
Co-Promotor: prof. dr B. Nienhuis

Promotiecommissie: dr C.H. Bennett
prof. dr A.E. Brouwer
dr H. Buhrman
prof. dr R.D. Gill
dr N.P. Landsman
prof. dr ir J.E. Mooij
prof. dr C.J.M. Schoutens
dr L.G. Suttorp

Paranimfen: Paulien Terhal
Monica Terhal

Faculteit der Wiskunde, Informatica, Natuur- en Sterrenkunde (WINS) en
Centrum voor Wiskunde en Informatica (CWI).

ISBN 90-9013009-8
Frontcover: Moiré pattern, illustrating the phenomenon of interference.
Backcover: ‘digitized’ Moiré pattern.

Parts of this thesis are based on material contained in the following papers:

B.M. Terhal and J.A. Smolin,
Single Quantum Querying of a Database, *Phys. Rev A* **58**, 1822 (1998), `quant-ph/9705041` → Chap. 2.

B.M. Terhal, I.L. Chuang, D.P. DiVincenzo, M. Grassl and J.A. Smolin,
Simulating Quantum Operations with Mixed Environments, *Phys. Rev. A* **60**, 881 (1999), `quant-ph/9806095` → Chap. 3.

B.M. Terhal and D.P. DiVincenzo,
The Problem of Equilibration and the Computation of Correlation Functions on a Quantum Computer, to appear in *Phys. Rev. A*, `quant-ph/9810063` → Chap. 4.

C.H. Bennett, D.P. DiVincenzo, T. Mor, P.W. Shor, J.A. Smolin and B.M. Terhal,
Unextendible Product Bases and Bound Entanglement, *Phys. Rev. Lett.* **82**, 5385 (1999), `quant-ph/9808030` → Chap. 5.

D.P. DiVincenzo, B.M. Terhal and A.V. Thapliyal,
Optimal Decompositions of Barely Separable States, to appear in *Journal of Modern Optics*, `quant-ph/9904005` → Chap. 5.

D.P. DiVincenzo, T. Mor, P.W. Shor, J.A. Smolin and B.M. Terhal,
Unextendible Product Bases, Uncompletable Product Bases, and Bound Entanglement, `quant-ph/9908070` → Chap. 5.

B.M. Terhal,
A Family of Indecomposable Positive Linear Maps based on Entangled Quantum States, `quant-ph/9810091`, submitted to Lin. Alg. and Its Appl. → Chap. 5.

B.M. Terhal,
Bell Inequalities and the Separability Criterion, in preparation → Chap. 5.

Other papers to which the author contributed:

D.P. DiVincenzo and B.M. Terhal,
Decoherence: The Obstacle to Quantum Computation, *Physics World* Vol.11, No. 3 (1998).

H. Barnum, J.A. Smolin and B.M. Terhal,
The Quantum Capacity is Properly Defined without Encodings, *Phys. Rev. A* **58**, 3496 (1998), `quant-ph/9711032`.

Voor mijn ouders Piet en Leonie

Le savant n'étudie pas la nature parce que cela est utile;
il l'étudie parce qu'il y prend plaisir et il y prend plaisir
parce qu'elle est belle. Si la nature n'était pas belle,
elle ne vaudrait pas la peine d'être connue, la vie
ne vaudrait pas la peine d'être vécue.

—Henri Poincaré (1854-1912), *Science et Méthode*.

The second benefit to be expected from giving to women
the free use of their faculties, by leaving them the free
choice of their employments, and opening to them the same
field of occupation and the same prizes and encouragements
as to other human beings, would be that of doubling the mass
of mental faculties available for the higher service of humanity.

—John Stuart Mill (1806-1873), *On the Subjection of Women*.

Contents

Chapter 1

Quantum Information and Computation

1.1 The Emergence of a New Field

The central issue in the emerging field of quantum computation and quantum information theory is the application of quantum mechanics in the domain of computation and information processing. Interest in the field has partially been created by the idea that at the present rate of miniaturization, magnetic storage and silicon technology will reach their limits around the year 2020. Any alternative technology for which the units of computation would be on an atomic (10^{-10} m) scale would face the fact that nature is ultimately only correctly described by quantum mechanics.

Richard Feynman [1] observed that there exists no fundamental limit imposed by quantum mechanics on the scale at which computation can take place. His view on a quantum computer was positive; a quantum mechanical computer could have a strong advantage over a classical device in simulating the dynamics of quantum mechanical systems.

While thinking about the physical basis of computation, Rolf Landauer [2] realized that an intrinsic lower bound on the amount of heat that is generated in a computer, is given by the heat generation due to the erasure of bit registers. An amount of $kT\ln(2)$ calories of heat is generated by the erasure of one bit of information. Although this contribution to heat generation is negligible for present-day computers, it could result in unacceptable levels of heat generation as the miniaturization of computers continues.

Notational conventions:

- $g(x) = O(f(x))$ if there are positive constants c and x_0 such that for all $x \geq x_0$ $|g(x)| \leq c|f(x)|$.
- $g(x) = \Omega(f(x))$ if there is a positive constant c such that $|g(x)| \geq c|f(x)|$ for infinitely many x.
- $f(x) = \mathrm{Poly}(x)$ if there are positive constants c and x_0 such that for all $x \geq x_0$ $|f(x)| \leq c|\mathrm{Poly}(x)|$ where $\mathrm{Poly}(x)$ is some polynomial in x.
- $\log$ denotes the binary logarithm unless stated otherwise.

In 1973 Charles Bennett [3] found that computation, unlike physical work, does not need to be an irreversible heat-generating process. He showed that every computation can be done in a logically reversible manner. In 1989 he showed that there is a universal way of reversibly simulating every irreversible computation that uses space S and time T, in space $c_\epsilon S \log T$ and time $T^{1+\epsilon}$ for all $\epsilon > 0$, where c_ϵ is a constant depending on ϵ. These ideas opened the way to the exploration of new ways of intrinsically reversible computation such a quantum computation.

In 1982 David Deutsch [4] raised the question of digital quantum mechanical computers and formulated a quantum mechanical version of the Turing machine—a logical model for classical computation. He realized that this computer might have computational advantages over a classical device. Together with Richard Jozsa [5] he formulated the first quantum algorithm that presented a major gain over any classical algorithm. The problem was to determine whether a function $f\colon \{0,1\}^n \to \{0,1\}$ with even n was *constant*, that is, $f(x) = 1$ for all x or $f(x) = 0$ for all x, or *balanced* meaning that $f(x) = 1$ for half of the inputs x. Classically in the worst case, the function has to be evaluated for $n/2+1$ inputs to decide which property the function has. With the quantum algorithm of Deutsch and Jozsa it is possible to decide between the two options with certainty with only *one* evaluation of the function and $O(\log(n))$ extra steps in handling inputs and outputs.

Computation however was not the only task for which researchers had been considering the help of quantum mechanics. As early as 1970 Wiesner [6] formulated the idea of 'quantum money'. The idea is to produce banknotes whose physical make-up contains, say, a photon that is either circularly or linearly polarized. Anyone who tries to counterfeit such notes, would run into the problem of having to determine the polarization of the photon. This cannot be done without disturbance of the quantum state, as the counterfeiter does not know the basis in which the photon was polarized. This idea became the basis of quantum cryptography, of which the first protocol (BB84) was formulated in 1982 by Bennett and Brassard [7].

An important breakthrough in the field of computation came in 1994 with Peter Shor's polynomial time factoring algorithm [8]. We will discuss this algorithm briefly in the next chapter, section 2.1.1.

In Figure 1.1 we present an overview of the most striking efforts and achievements in quantum computation and quantum information theory as they stand right now. The overview in the figure is by no means comprehensive (and is quickly outdated). In the following sections we will discuss some of these results.

1.2 Applications of Quantum Tools

1.2.1 Quantum Key Distribution

The goal of the BB84 quantum key distribution protocol is for two people, usually called Alice and Bob, to enlarge a shared random secret bit string. This bit string can serve at a later time

Topic	Theoretical Achievements	Experimental Realization
1. Quantum cryptography	BB84 [7] protocol (and others) proven secure by Mayers [10] (and others)	Polarized photons over 23 km fiber [12]; 1 km open night air [13]
2. Quantum algorithms	Shor's [8] factoring in polynomial time; Grover's $O(\sqrt{n})$ search [14] algorithm etc.	Various 2 qubit algorithms implemented on NMR quantum computer [18]; Deterministic GHZ and
3. Physical implementations of a quantum computer	Internal levels of trapped ions [15]; Nuclear spins in NMR [16]; Electron spins on quantum dots [17] etc.	EPR preparation [19]; Error correction schemes in NMR [20]
4. Quantum communication over noiseless and noisy channels	Noiseless quantum coding theorem [21]; Quantum teleportation [22]; Quantum error correcting code theory and fault tolerant computation [23]	Teleportation of a polarized photon [24, 25] and nuclear spin (in NMR) [26]; see also 2-3.
5. Classical comm. complexity, using quantum communication	$O(\sqrt{n}\log n)$ **set intersection** and other speed-ups [28]; $O(\log(n))$ sampling complexity [29] etc.	QC
6. Classical capacity of quantum channels	Capacity of some channels is achieved by encoding with nonorthogonal states [30]; Capacity enhanced by sharing of entanglement [31]	QC
7. Simulation of physical systems	Efficient implementation of unitary evolution [32]; Finite temperature simulations (see Chap. 4)	2-qubit NMR unitary evolution [33]

Figure 1.1: An overview of the various applications and achievements of the use of quantum mechanics in computation and information processing tasks. QC denotes that no experiments have been performed for the implementation of the protocol/task, but the implementation is of similar hardness as the building of a quantum computer for which the efforts are summarized at items 2. and 3.

as an encryption key for messages between Alice and Bob. Their task is to enlarge their key by sending each other messages over a classical or quantum channel, on which an eavesdropper (Eve) might be listening. The BB84 protocol provides a way to perform this task by sending quantum data.

Alice sends a random sequence of qubits in two different bases, the $\{|0\rangle, |1\rangle\}$ basis and the $\{\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\}$ basis over a public quantum channel to Bob. She privately writes down the choices for bases and states that she made. Bob measures these qubits, in one of the two different bases at random. He announces publicly (over a classical channel) which bases he chose, but not which measurement outcomes he found. Alice responds publicly by telling him which bases were correct, that is, which measurements were done in the same basis in which Alice prepared the qubits. They then discard the qubits on which Bob did the wrong measurement. The outcomes of these measurements for which Bob's basis differs from Alice's, are uncorrelated with the states that Alice has sent; a $|0\rangle$ has a 50% of being measured as $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and a 50% chance of being measured as $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle))$.

If there is no eavesdropper Eve, Alice and Bob now would share a random string of bits. These are the outcomes of Bob's measurements in the correct basis. An eavesdropper Eve could try to copy the quantum states that Alice sends. The no-cloning theorem [9] makes it impossible for Eve to perfectly copy quantum states in the basis $\frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$ *and* $|0\rangle, |1\rangle$. More generally, the quantum key distribution protocol relies on the fact that the information that Eve gets about the quantum states is visible to Alice and Bob as a disturbance in the correlation between the sets of outcomes that they obtain when Bob's measurement is aligned with Alice's preparation. To determine their degree of privacy in the presence of an eavesdropper, Bob and Alice go through a protocol of selecting a random subset of the remaining bits and see how well they correlate publicly. After these bits are sacrificed Alice and Bob go through a procedure of error correction on the remaining qubits and finally privacy amplification to distill a set of shared bits about which they are highly confident that Eve has little information. Mayers [10] has proven the security of quantum key distribution (BB84) under all possible attacks of Eve, that is, he has shown [11] that the mutual information of Eve with a final key which has $\frac{n}{4}(1 - H_2(2\delta) - H_2(\delta))$ bits where n is the initial set of qubits that Alice sends, can be made arbitrarily small for large enough n. Here δ is the one bit error probability for bits of the key; it is the probability that a randomly selected bit on which Alice's and Bob's bases coincide is different. $H_2(p)$ is the binary entropy function $H_2(p) = -p \log p - (1-p)\log(1-p)$.

The first quantum cryptography apparatus was built by Charles Bennett and John Smolin at IBM in 1989. The qubits were polarized photons and the channel between Alice and Bob was 30 cm of air. At present quantum cryptography takes place over 23 km in optical (telecom) fiber [12] and 1 km in open night (and also day) air [13].

1.2.2 Classical Communication over Quantum Channels

Channels with a capacity to transmit classical information as described in classical information theory, are called *classical* channels in the broader context of quantum information theory.

These channels are described by a set of errors, –bitflips–, and their probabilities on a set of orthogonal bit strings. A complete description of any physical channel requires a description of the action of the channel on any quantum mechanical state. We then speak of quantum channels. The study of the capacity of quantum channels for transmitting classical information has resulted in an explicit expression for the classical capacity of a quantum channel [34]. In this more general setting it has been shown [30] that there are quantum channels for which optimal classical information transmission is achieved by encoding the data in non-orthogonal quantum states. Such an encoding thus surpasses any 'classical' encoding scheme in which only orthogonal sets of states are being used. In Ref. [31] a scenario was considered in which a sender and receiver share an unlimited amount of entanglement. Sender and receiver are also connected by a noisy quantum channel. It was shown that the classical capacity of the quantum channel was enhanced by the use of the shared entanglement.

Another topic in which interesting advances have been made is the problem of classical communication complexity. Two parties wish to compute the value of a function $f(x,y)$ on some inputs x and y. One of the parties, Alice, holds bit string x and the other party, Bob, holds bit string y. The goal is for the two parties to determine the value $f(x,y)$ by using the *minimal* amount of classical communication. If we replace the classical communication with the communication of quantum bits, it has been found that less communication is needed for certain functions. Buhrman, Cleve and Wigderson [28] have shown such a reduction in communication costs for the computation of the **set intersection** function on n-bit strings, $f(x,y) = \vee_{i=1}^{n}(x_i \wedge y_i) \in \{0,1\}$. With their protocol Alice and Bob can determine the value of the function with bounded probability of error using an amount of quantum communication of $O(\sqrt{n}\log n)$ qubits. Classically, in the worst case, $\Omega(n)$ bits are needed for Alice and Bob to find the correct answer.

A different approach was pursued by Ambainis *et al.* [29]. The goal now is for Alice and Bob to set up a joint probability distribution $\pi(x,y)$ over bit strings x and y, where Alice can draw bit string x and Bob can draw bit string y. Setting up this distribution can form the first part of a protocol in which they sample a function $f(x,y)$ over this joint distribution $\pi(x,y)$. The authors considered for example how much quantum communication it would cost to set up the uniform distribution over all n-bit *disjoint* strings x and y with Hamming weight $O(\sqrt{n})$. They found a quantum protocol that enabled Alice and Bob to sample from such a distribution using only $O(\log n)$ quantum communication. This presents an exponential reduction from the classical costs where the number of bits that have to be communicated is $\Omega(\sqrt{n})$.

A consequence of these gains with quantum protocols is that as the amount of communication is reduced, the evaluation of a joint function $f(x,y)$ or the sampling from a joint distribution is carried out in a more discreet way; Alice and Bob obtain less information about their mutual bit strings x and y. For example, in the **set intersection** problem, in the worst case, $\Omega(n)$ bits need to be communicated in order to solve the problem with classical communication. In the quantum protocol $O(\sqrt{n}\log n)$ quantum bits are sufficient. Holevo's theorem

[35] states that k qubits can carry no more than k classical bits of information. Therefore Alice will not learn more than $O(\sqrt{n}\log n)$ bits of information about Bob's bit string y.

1.2.3 Quantum Communication and Teleportation

In the previous paragraph we briefly discussed the use of quantum data in problems of classical data transmission. Central to the field of quantum information theory is the problem of transmission of *quantum* data over quantum channels.

For a *noiseless* quantum channel Schumacher [21] has established the quantum equivalent of Shannon's classical noiseless coding theorem. Given is a quantum source characterized by a set of states and their probabilities: $\{p_i, |\psi_i\rangle\}$. Let

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|, \tag{1.2.1}$$

and $S(\rho) = -\mathrm{Tr}\,\rho\log\rho$, the von Neumann entropy of the source. Schumacher showed that the source messages can be transmitted with arbitrary high fidelity (arbitrary low probability of error) in on average $nS(\rho)$ quantum bits, where n is the number of qubits emitted in the source message, for sufficiently large n. In contrast, Shannon's expression for the capacity of a *noisy* classical channel reads

$$C = \max_X H(X;Y), \tag{1.2.2}$$

where X is the input source and Y represents the output signal. The input source is given by a set of a messages and their probabilities $\{x_i, p(x_i)\}$ with $\sum_i p(x_i) = 1$ and similarly for the output signal, $\{y_j, p(y_j)\}$. $H(X;Y)$ is the mutual information between X and Y:

$$H(X;Y) = H(X) - H(X|Y), \tag{1.2.3}$$

where $H(X)$ is the Shannon entropy of the input source,

$$H(X) = -\sum_i p(x_i)\log p(x_i), \tag{1.2.4}$$

and $H(X|Y)$ is the conditional entropy

$$H(X|Y) = -\sum_{i,j} p(x_i, y_j)\log p(x_i|y_j). \tag{1.2.5}$$

The $\max$ in Eq. (1.2.2) denotes the maximum over all possible input sources $\{x_i, p(x_i)\}$. Noisy quantum channels have turned out to be much richer in structure and much harder to characterize. We have not yet obtained an expression for the quantum channel capacity that allows us to calculate the capacity of a quantum channel, such as Eq. (1.2.2).

Although determining the capacity of a quantum channel remains an open problem for almost all quantum channels, much progress has been made in developing a theory of quantum error correction. The goal of quantum error correction is to encode a quantum state that is to be protected from noise on a quantum channel or decoherence in a quantum memory, in a quantum state in a larger Hilbert space such that an error process that corrupts this encoded quantum state can be undone by performing an error correcting procedure.

Many workers were initially doubtful of the very existence of error correction for quantum states. Most of the objections which were raised were centered on two points: (1) the process of decoherence, seen as a measurement process, will irreversibly destroy the information that is contained in the quantum state, and (2) the quantum state is analog (it is, after all, specified by a set of complex numbers); thus, errors caused by decoherence come in an almost infinite variety (perhaps limited by the precision with which states are specified), and at least some of the errors simply rotate the system into a different legal quantum state and thus would not be detectable as errors.

The quantum error correcting codes, first discovered by Peter Shor [23] and subsequently elucidated by many others, overcome both of these objections. Nonetheless the first objection *is* correct when the decoherence rate is high, i.e., when the error rate is large ; it is typical of any error correction scheme, quantum or classical, that it can be overwhelmed if errors occur faster than they can be corrected. When we consider such a quantum channel, we are led to say that the channel has zero quantum capacity at these high error rates.

Perhaps the greatest surprise which emerges from the error correcting procedures is that the way quantum error correction works is basically digital and not analog. The essential digitizing steps are quantum measurements, which force the whole continuum of possible errors to effect the quantum state, for example a qubit, in just one of three possible canonical ways; once one of these "stereotyped" errors are detected, it can be undone by one of a discrete set of unitary transformations.

Teleportation

When a sender Alice wishes to send a unknown quantum state over a quantum channel to a receiver Bob, she cannot keep a copy of this state to herself. This is a consequence of the "no-cloning theorem". There exists an alternative to a straightforward transmission of the quantum state that can be used when sender and receiver share entangled states. This protocol, called teleportation [22], exhibits some of the essential features of quantum data.

Assume that Alice has a single unknown qubit that she wants to send to Bob:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \tag{1.2.6}$$

with $|\alpha|^2 + |\beta|^2 = 1$. Alice starts by making an entangled state

$$|\Psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle). \tag{1.2.7}$$

She sends half of the state, one qubit, to Bob, over the quantum channel, which we assume to be noiseless. On the unknown state, Eq. (1.2.6) and her half of the entangled state, she performs a measurement in the "Bell basis" given by the following four orthogonal states:

$$\left\{|\Phi^{\pm}\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle), |\Psi^{\pm}\rangle = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle)\right\}. \tag{1.2.8}$$

The measurement outcomes represent two bits of classical data which she sends to Bob over a classical channel. Bob rotates his half of the original entangled state in the following manner

$$\begin{array}{ll} \text{Alice's outcome} & \text{Bob applies} \\ |\Phi^-\rangle & \sigma_x, \\ |\Phi^+\rangle & \sigma_y, \\ |\Psi^+\rangle & \sigma_z, \\ |\Psi^-\rangle & \mathbf{1}, \end{array} \tag{1.2.9}$$

where σ_x, σ_y and σ_z are the three Pauli matrices:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.2.10}$$

The protocol has the effect that the state of Bob's half of the entangled pair is now identical to the state that Alice wanted to send him, Eq. (1.2.6).

There is an advantage in using the teleportation protocol over a straightforward encoding of the qubit for certain noisy channels. When the channel is noisy, Alice and Bob will not share a pure entangled state such as Eq. (1.2.7), as half of the entangled state that Alice prepared went through the noisy channel. They will start out with a mixed entangled state. In Refs. [36, 37] it was shown how Alice and Bob can distill pure (maximally) entangled states out of a special family of such mixed states by using classical communication and local quantum operations only. It was shown that when they perform such a "purification" protocol with the help of 1-way classical communication only, from Alice to Bob, that the purification protocol and the subsequent teleportation is equivalent to the sending of the quantum data with the use of a quantum error correction code. However when one allows communication between Alice and Bob in both directions, then the amount of quantum information that can be sent can be larger than in a 1-way protocol. This implies that there exist quantum channels for which sending quantum data via teleportation is more efficient than via error correcting codes.

The teleportation of a qubit has been accomplished in several experiments in 1998, see Refs. [24, 25, 26, 27].

1.3 Decoherence and Physical Implementations

Between Alan Turing's paper of 1936 that laid the foundation of the modern computer and the building of the first modern electronic computer in 1946 (the ENIAC, capable of arithmetic manipulations of 10-digit numbers) stood a period of 10 years of research. It has now been 14 years since David Deutsch in 1985 accomplished his feat of imagination by formulating a model of a quantum mechanical Turing machine. While one can expect that building a prototype quantum computer with the computational capabilities of the ENIAC is still some time in the future, important progress has been made in the last two years in surmounting what had been considered a major, perhaps impassable hurdle on the way to building a quantum computer: the phenomenon of decoherence. A method for limiting decoherence in a quantum computer has emerged in the form of quantum error correcting codes and fault-tolerant schemes for computation. These will permit a quantum computer to operate in the presence of decoherence, given that the noise level is sufficiently low. The present estimate of the rate below which quantum computation can take place is in the neighborhood of an error probability of 10^{-4} per qubit per clock cycle. The assumption on which such an estimate is based is that these errors occur independently on individual qubits (or other small subsystems). If a clock cycle, the time it takes to implement elementary gates, is of the order of 1 μs, then a decoherence time of the order of 0.01 sec. is required for the successful implementation of fault-tolerant computation.

DiVincenzo [38] has identified five basic requirements that must be met in order for a physical system to serve as a candidate quantum computer. These are:

1. Control of the Hilbert space: (1) the computationally available states in the Hilbert space can be enumerated and (2) it is possible to extend the available Hilbert space with a polynomial amount of qubits without an exponential increase in physical resources. This last requirement ensures that the computation is scalable.

2. State preparation: one must be able to prepare the quantum computer in a known initial state with low constant probability of error.

3. Limited decoherence: the amount of decoherence that can be tolerated, will often depend on the specifics of the physical setup. There will usually be a non-uniformity in errors, a mutual dependence of errors and the error rates can critically depend on the size of the quantum computer.

4. Controlled unitary transformations: one has to be able to implement a *universal* set of elementary quantum gates that operate on few qubits. By using elements of a universal set of quantum gates it is possible to build any unitary transformation on n qubits. An example of a universal set of a quantum gates is the following set. We have a CNOT gate:

$$\mathrm{CNOT}\colon |a\rangle \otimes |b\rangle \to |a\rangle \otimes |a \oplus b\rangle, \tag{1.3.1}$$

where $\oplus$ denotes the XOR operation, and any one qubit phase-shift P:

$$P\colon |b\rangle \to \begin{cases} |b\rangle & \text{if} \quad b = 0, \\ e^{i\phi}|b\rangle & \text{if} \quad b = 1, \end{cases} \tag{1.3.2}$$

where $\frac{\phi}{\pi}$ is irrational. In addition we have a one qubit Hadamard transform:

$$H\colon |b\rangle \rightarrow \begin{cases} \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) & \text{if} \quad b=0, \\ \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) & \text{if} \quad b=1. \end{cases} \tag{1.3.3}$$

5. Readout: one should be able to efficiently perform measurements on local subsystems (for example qubits) of the computer in the computational basis.

Currently existing proposals and physical implementations each have their way of addressing these requirements. We will discuss one of them here. Cirac and Zoller [15] have proposed a model quantum computer in which two internal levels of an ion (for example a beryllium ion Be^+) in a trap serve as a single qubit. By putting several ions in a linear trap a small quantum processor could be made (requirement 1). The one-dimensional vibrational mode of the string of ions in the trap forms a last 'helper' qubit; it is a qubit that mediates the interaction between the ion qubits. The trap is cooled down to the regime where kT is lower than the energy ν_x of the lowest vibrational mode (which is the center-of-mass mode) of the string of qubits. This ensures that the helper qubit state is prepared in its ground state (requirement 2).

By applying laser pulses on each ion individually at resonance with ω, the energy difference between the $|0\rangle$ and $|1\rangle$ state, for an appropriate amount of time, every single qubit rotation can be performed. A CNOT operation between two qubits involves the helper qubit. The idea is to first perform a SWAP between the helper (or 'bus') qubit and the state of the control ion:

$$\mathrm{SWAP}\colon |a\rangle \otimes |b\rangle \rightarrow |b\rangle \otimes |a\rangle, \tag{1.3.4}$$

on basis states $|a\rangle$ and $|b\rangle$. Then one performs a CNOT, Eq. (1.3.1), between the helper qubit and the 'target' ion, where the helper qubit now functions as a control. Finally, one swaps the helper qubit state back with the original control qubit. The SWAP between the helper phonon and the ion-qubit is enacted by tuning the laser to be resonant with $\omega - \nu_x$. This has the effect that an excited ion state $|1\rangle$ de-excites to $|0\rangle$ and emits a phonon, that is, the center of mass mode is excited. In this way a set of universal quantum gates is constructed (requirement 4).

Finally, for the read-out Cirac and Zoller propose to use the technique of quantum jumps. The idea is to tune the laser to a sharp transition between say the $|0\rangle$ state of the ion and a high-lying unstable level. If the state of the qubit is $|0\rangle$ the transition to the unstable level will be made and when the level decays an emitted photon can be observed. If the state of the qubit is $|1\rangle$ no photon will be observed (requirement 5).

Heating of the vibrational mode has been estimated to be the limiting factor in the performance of the trap, as the timescales of other sources of decoherence, such as transitions in the internal state of the 'qubit' ions and inaccuracies in the performance of the laser, will typically be longer. A decoherence time of $1ms$ has been reported. In a review paper about the ion trap model [39] Andrew Steane estimated that without the use of error correction a quantum information processor with 10 qubits on which about 200 elementary gate operations can be performed coherently, lies within the reach of experiment.

The cooling to the quantum regime and the operation of CNOT gates have been demonstrated experimentally by the group of Monroe and Wineland at the National Institute of Standards and Technology. In Ref. [19] this group presented an experiment in which an entangled state between two trapped ions was prepared deterministically. This is the first experiment in which an entangled state was created 'on demand' and not by means of post selection on a probabilistic process. It is clear that the ability to create entanglement and implement gates deterministically is one of the achievements that may help us eventually to build a quantum computer.

Although the overview of Table 1.1 shows that the theoretical efforts are still far ahead of what has been achieved in the laboratory, continuing progress is being made on both fronts. We hope that the results presented in the following chapters will contribute to this progress.

Chapter 2

Quantum Algorithms

2.1 Introduction

In this chapter we discuss various quantum algorithms. In the first two sections 2.1.1 and 2.2 we review the two classics: Shor's factoring and Grover's search algorithm. In section 2.2.1 we will present an application of a generalization of the quantum counting algorithm. In section 2.3 we exhibit a problem for which a quantum computer has a considerable advantage over a classical device. Finally, in section 2.4 we present a problem for which a quantum computer has almost no advantages over a classical device.

2.1.1 The Factoring Algorithm

The best classical algorithm that factors a composite number N in its prime factors takes time $O(e^{c(\log N)^{1/3}(\log\log N)^{2/3}})$ with a fixed constant c. The algorithm is thus exponential in the size of the problem $\log N$. The decision variant of factoring, –are there integers $m_1, m_2 > 1$ such that $N = m_1 m_2$–, lies inside the complexity class $NP \cap co{-}NP$, and hence is very likely not NP-complete.

In 1994 Peter Shor [8] exhibited an algorithm that factors a number N in polynomial time on a quantum computer. His algorithm takes time $O((\log N)^2(\log\log N)(\log\log\log N))$ plus a polynomial amount of postprocessing on a classical computer and $O(\log N)$ space. The importance of factoring stems from the fact that the popular RSA public-key crypto-system is based on the hardness of factoring. As a benchmark to assess the state of the art, the inventors have published a challenge list of RSA numbers of increasing length. The current record holder is the CWI. At the CWI currently 512 bit RSA numbers are being factored using 300 computers with an average clockspeed of 300 MHz for two months [40]. This amounts to a total of approximately 10^{17} elementary operations. If a perfectly operating (hence no error correction) quantum computer would be available such a 512 bit number can be factored with around 10^{11} elementary operations, which would take (with a 300 MHz quantum processor) 5 minutes. This quantum computer would use approximately 2560 qubits.

The problem that Shor's quantum algorithm solves is finding the order of an integer x

$\bmod N$ which is coprime to N, i.e. $\gcd(x, N) = 1$. The order of x is the smallest number r such that $x^r = 1 \mod N$. It can be shown that when the order of a random $x \mod N$ (coprime to N) is determined one can find with polynomial extra effort a non-trivial factor of N with probability larger than $1/2$.

Let us review the steps of the quantum algorithm. We choose an x randomly (coprime to N). We choose an integer q as a power of 2 such that $N^2 < q < 2N^2$. The first step is to prepare an equal superposition indexed with $a = 0, \dots, q-1$:

$$\frac{1}{\sqrt{q}} \sum_{a=0}^{q-1} |a\rangle \otimes |x\rangle \otimes |0\rangle. \tag{2.1.1}$$

Then in the last register, that initially is prepared in the state $|0\rangle$, we compute the function $f(a) = x^a \mod N$:

$$\frac{1}{\sqrt{q}} \sum_{a=0}^{q-1} |a\rangle \otimes |x\rangle \otimes |x^a \mod N\rangle. \tag{2.1.2}$$

The next step is to measure the last register. We can distinguish two cases: (1) the order r of x divides q and (2) r does not divide q. Let us see here what happens in the simplest case when r divides q. When the measurement on the last register results in the value $x^k \mod N$, this amounts to selecting in the superposition in the first register all the values $a = k + jr$ for $j = 0, \dots, \frac{q}{r} - 1$. The remaining (leaving out the last two registers) normalized state then reads:

$$\frac{1}{\sqrt{q/r}} \sum_{j=0}^{q/r-1} |k + jr\rangle. \tag{2.1.3}$$

Now a quantum discrete Fourier transform DFT_q is performed on the first register

$$DFT_q \colon |a\rangle \to \frac{1}{\sqrt{q}} \sum_{c=0}^{q-1} e^{i\,2\pi ac/q} |c\rangle, \tag{2.1.4}$$

leading to the state

$$\frac{1}{\sqrt{q}} \frac{1}{\sqrt{q/r}} \sum_{c=0}^{q-1} \left[\sum_{j=0}^{q/r-1} e^{i\,2\pi(k+jr)c/q} \right] |c\rangle. \tag{2.1.5}$$

The amplitude in each state $|c\rangle \otimes |x\rangle$ is a Fourier sum which we can rewrite using the identity

$$\frac{1}{q/r} \sum_{j=0}^{q/r-1} e^{i\,2\pi(k+jr)c/q} = e^{i\,2\pi kc/q} \delta_{c,\frac{mq}{r}} \quad m = 0, \dots, r-1. \tag{2.1.6}$$

Thus we may rewrite the state of Eq. (2.1.5) as

$$\frac{1}{\sqrt{r}} \sum_{m=0}^{r-1} |c = mq/r\rangle. \tag{2.1.7}$$

When finally the first register $|c\rangle$ is observed, the values of c that occur with non-zero probability are multiples of $\frac{q}{r}$. We have performed a Fourier analysis on the periodic amplitude of the state in Eq. (2.1.3). With a sufficient number, $O(\log\log N)$, of repetitions of this preparation and measurement routine, we can approximate the period q/r between these peaks of the Fourier spectrum and thereby determine r. When r does not divide q the state in Eq. (2.1.3) will not exactly correspond to a periodic function. This will introduce some smearing of the peaks in the Fourier spectrum. With some additional analysis it is possible to show that also in this case a reliable estimate of the period r can be found efficiently.

The effectiveness of Shor's algorithm relies on the efficient implementation of two crucial subroutines:
1. the computation of the function $f(a) = x^a \mod N$ for l bit numbers a and N can be implemented in $O(l^3)$ time and,
2. the quantum Fourier transform DFT_q over $\mathbf{Z}_q$ for $q = 2^l$ can be performed on a quantum computer in $O(l^2)$ time.

It would be of great interest to find other functions $f(a)$ for which period-finding is hard on a classical computer, but the implementation of $f(a)$ for exponentially large a in the size of the problem is efficient.

2.2 Generalized Quantum Searching and Counting

We review Grover's quantum search algorithm [14] and its generalizations. In Grover's original setting a function $f\colon \{0,1\}^n \to \{0,1\}$ is given as an oracle. Upon an input $|x\rangle \otimes |b\rangle$, the oracle returns

$$|x\rangle \otimes |b\rangle \overset{\text{Query}}{\to} |x\rangle \otimes |b \oplus f(x)\rangle. \tag{2.2.1}$$

It is promised that there is exactly one input x for which $f(x) = 1$, for all other inputs $f(x) = 0$. The problem is to determine x such that $f(x) = 1$ with the least number of oracle calls. To find the input for which $f(x) = 1$ with bounded probability of error on a classical device takes $\Omega(N)$ $(N = 2^n)$ evaluations of the function f. Grover's quantum algorithm [14] uses $O(\sqrt{N})$ oracle calls to find the input x for which $f(x) = 1$ with bounded probability of error. In Ref. [41] (see also Ref. [42]) the quantum algorithm was generalized to search problems in which the function f takes the value 1 for a possibly unknown number of inputs.

The search algorithm was subsequently generalized in Refs. [43] and [44] to quantum counting and amplitude amplification. We will discuss the amplitude amplification algorithm of which the search algorithm is a special case. We show how the amplitude amplification algorithm can be used to estimate the size of a matrix element of a unitary matrix.

We define a unitary transformation $G\colon \mathcal{H}_N \to \mathcal{H}_N$ as the Grover transform:

$$G = -UI_jU^\dagger I_i, \tag{2.2.2}$$

where $I_i = \mathbf{1} - 2|i\rangle\langle i|$ with $|i\rangle \in \mathcal{H}_N$ and similarly I_j. In Grover's original search algorithm $U = U^\dagger$ and

$$U|0\rangle = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} |k\rangle. \tag{2.2.3}$$

This is the first step of Grover's algorithm. Furthermore, in that algorithm we choose $|j = 0\rangle$ and $|i\rangle$ of I_i in Eq. (2.2.2) is the state for which $f(i) = 1$. The unitary transformation I_i is implemented with a single function call: We use an additional register prepared in the state $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ and query the oracle:

$$|j\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \overset{\text{Query}}{\to} (-1)^{f(j)}|j\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle). \tag{2.2.4}$$

In the generalized algorithm we start in the state

$$U|j\rangle = U_{ij}|i\rangle + \sum_{k\neq i} U_{kj}|k\rangle. \tag{2.2.5}$$

Here the states $|i\rangle$ and $|j\rangle$ are orthogonal. Now the Grover transform is applied repeatedly. This leads to a quantum state of the form

$$G^k U|j\rangle = \alpha_k U_{ij}|i\rangle + \beta_k \sum_{k\neq i} U_{kj}|k\rangle, \tag{2.2.6}$$

where α_k and β_k are the solutions of a recurrence relation:

$$\begin{pmatrix} \alpha_{k+1} \\ \beta_{k+1} \end{pmatrix} = \begin{pmatrix} 1-2|U_{ij}|^2 & 2(1-|U_{ij}|^2) \\ -2|U_{ij}|^2 & 1-2|U_{ij}|^2 \end{pmatrix} \begin{pmatrix} \alpha_k \\ \beta_k \end{pmatrix}. \tag{2.2.7}$$

The solution for these recurrence relations is

$$\begin{aligned} \alpha_k &= \tfrac{1}{|U_{ij}|} \sin((2k+1)\theta), \\ \beta_k &= \tfrac{1}{\sqrt{1-|U_{ij}|^2}} \cos((2k+1)\theta), \end{aligned} \tag{2.2.8}$$

where

$$\sin^2\theta = |U_{ij}|^2. \tag{2.2.9}$$

In the original Grover search $|U_{ij}|^2 = \frac{1}{N}$. The amplitude of the state $|i\rangle$ which is the input for which $f(i) = 1$, is an oscillating function of the number of Grover iterations which gets close to 1 for the first time when $k = O(\sqrt{N})$. Thus when a measurement is performed after $O(\sqrt{N})$ Grover transforms G, the probability of finding state $|i\rangle$ is high.

In the general case the probability $|U_{ij}|^2|\alpha_k|^2$ is a periodic function in k and its period is determined by the size of $|U_{ij}|^2$ via the relation of Eq. (2.2.9). In Ref. [41] it was observed that by performing a Fourier transform as in Shor's factoring algorithm, Eq. (2.1.4), it is possible to estimate the period of this function and thereby indirectly estimate $|U_{ij}|^2$. In the case of Grover searching with an unknown number of inputs for which the function value is 1, this algorithm results in an approximate counting of the number of such inputs. In that case $|U_{ij}|^2 = \frac{t}{N}$ where t is the number of inputs for which f takes the value 1. For the general case, the following lemma has been proved:

Lemma 1 *[44] Let $|i\rangle \in \mathcal{H}_N$ and $|j\rangle \in \mathcal{H}_N$ be two known orthogonal quantum states. Let U be a unitary transformation and let $|U_{ij}|^2$ be given by*

$$U|j\rangle = U_{ij}|i\rangle + \dots . \tag{2.2.10}$$

Let the Grover transform be given by

$$G = -UI_jU^\dagger I_i, \tag{2.2.11}$$

with $I_i = \mathbf{1} - 2|i\rangle\langle i|$. There exists a quantum algorithm that estimates the value of $p = |U_{ij}|^2$ as $\tilde{p}$ such that

$$|p - \tilde{p}| \leq \frac{2\pi}{L}\sqrt{p} + \frac{\pi^2}{L^2}, \tag{2.2.12}$$

with probability at least $\frac{8}{\pi^2}$ using $L \geq 4$ Grover transforms.

We compare this result with estimating $|U_{ij}|^2$ in a 'classical' way; we prepare the state $|j\rangle$, then we apply U and measure in the basis which contains the vector $|i\rangle$. We repeat this process in order to estimate the probability $p = |U_{ij}|^2$ with which we obtain outcome $|i\rangle$. The number of times one has to repeat the process in order to get an estimate for $|U_{ij}|^2$ with precision δ and probability ϵ is $O(\ln(\epsilon^{-1})\delta^{-2})$ [52]. In the quantum algorithm a precision δ and a probability of at least $\frac{8}{\pi^2}$ is obtained by using only $L = O(\delta^{-1})$ Grover transforms. This illustrates the efficiency of the quantum algorithm.

2.2.1 An Application: Mean Estimation

It is possible to generalize the Grover transform of Eq. (2.2.2) and the quantum algorithm for matrix element estimation one step further. Consider the Grover transform in Eq. (2.2.2). Let U be a unitary transformation on $\mathcal{H}_1 \otimes \mathcal{H}_2$, a quantum system that is composed of two subsystems. Let $|j\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$. Let $|i\rangle \in \mathcal{H}_2$ and let I_i be of the form $\mathbf{1}_{\mathcal{H}_1} \otimes (\mathbf{1}_{\mathcal{H}_2} - 2|i\rangle\langle i|)$. We write

$$U|j\rangle = |\psi\rangle \otimes |i\rangle + |(\psi \otimes i)^\perp\rangle, \tag{2.2.13}$$

where $|\psi\rangle$ is some (unnormalized) state and $|(\psi \otimes i)^\perp\rangle$ is a state orthogonal to $|\psi \otimes i\rangle$. We also have

$$\langle j|\psi \otimes i\rangle = 0. \tag{2.2.14}$$

We can estimate the matrix element

$$|U_{j,\psi\otimes i}|^2 = \langle\psi|\psi\rangle, \tag{2.2.15}$$

with the quantum algorithm that was outlined in the last section (Lemma 1).

This kind of algorithm can be used in the following problem. A function $f\colon \{0,1\}^n \to [0,1]$ is given as an oracle. We would like to estimate the mean

$$\langle f \rangle \equiv \frac{1}{N} \sum_{x=0}^{N-1} f(x), \tag{2.2.16}$$

where $N = 2^n$, with a certain accuracy with the minimum number of function evaluations. The quantum algorithm that we propose is an application of the generalized quantum counting algorithm and provides a square-root speed up in the number of function evaluations compared to a classical algorithm. In Ref. [45] Grover presented a quantum algorithm for mean estimation which achieves a square-root speed up. The quantum algorithm that we present here has the benefit that it is a clear application of the generalized counting algorithm.

Let k be the number of bits of precision of f. Let $|j\rangle = |00\dots0\rangle$, a register with $n+1$ qubits. Let $|i\rangle = |1\rangle$ for the last qubit.

The following sequence of transformations is the unitary operation U on the state $|j\rangle \otimes \underbrace{|00\dots0\rangle}_{k} = |00\dots0\rangle$:

- **Prepare** a uniform superposition of inputs in the first n $(N = 2^n)$ qubits:

$$|00\dots0\rangle \to \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} |x\rangle \otimes |0\rangle \otimes \underbrace{|0,\dots,0\rangle}_{k}. \tag{2.2.17}$$

- **Query** the oracle and compute $f(x)$:

$$\frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} |x\rangle \otimes |0\rangle \otimes \underbrace{|0,\dots,0\rangle}_{k} \overset{\text{Query}}{\to} \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} |x\rangle \otimes |0\rangle \otimes |f(x)\rangle. \tag{2.2.18}$$

- **Rotate** the single (middle) qubit such that

$$|x\rangle \otimes |0\rangle \otimes |f(x)\rangle \to |x\rangle \otimes \left(\sqrt{f(x)}|1\rangle + \sqrt{1-f(x)}|0\rangle\right) \otimes |f(x)\rangle, \tag{2.2.19}$$

  and **uncompute** f to obtain

$$\frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} |x\rangle \otimes \left(\sqrt{f(x)}|1\rangle + \sqrt{1-f(x)}|0\rangle\right) \otimes \underbrace{|0,\dots,0\rangle}_{k}. \tag{2.2.20}$$

Let us consider the computational costs of the implementation of the Grover transform based on the transformations U, I_i and I_j that we have built. The implementation of the transformation U uses the following steps:

1. n 1-qubit Hadamard transforms of the form

$$H_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{2.2.21}$$

to create a uniform superposition of all inputs (step 1).
2. Two function calls of f, one to obtain f and one to uncompute f.
3. The rotation of Eq. (2.2.19) which can be implemented with $O(k)$ elementary operations. The k-bit representation of $f(x)$ is first translated in a k-bit representation of an angle $\phi(x)$ such that $\cos\phi(x) = \sqrt{f(x)}$. Then a sequence of k controlled rotations is executed on the middle qubit depending on the k bits of the angle $\phi(x)$.

The transformation $U^\dagger$ is implemented by reversing the operations of U and thus takes an equal number of elementary steps. The implementation of $I_{i=1} = \mathbf{1}_{2^n} \otimes (\mathbf{1}_2 - 2|1\rangle\langle 1|)$ takes a single qubit phaseshift. The implementation of $I_{j=00\ldots 0} = \mathbf{1}_{2^{n+1}} - 2|00\ldots 0\rangle\langle 00\ldots 0|$ takes $n+1$ single qubit phaseshifts.

We have constructed the unitary transformation U such that a matrix element estimation routine, in this case the estimation of the norm of

$$|\psi\rangle = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} \sqrt{f(x)}|x\rangle, \tag{2.2.22}$$

as in Eq. (2.2.15) is equal to the estimation of $\langle f\rangle$ since

$$\langle\psi|\psi\rangle = \frac{1}{N} \sum_{x=0}^{N-1} f(x) = \langle f\rangle. \tag{2.2.23}$$

Therefore by Lemma 1 we find that with probability at least $\frac{8}{\pi^2}$ we estimate $\langle f\rangle$ as $\langle f\rangle'$ such that

$$|\langle f\rangle - \langle f\rangle'| \le \frac{2\pi}{L}\sqrt{\langle f\rangle} + \frac{\pi^2}{L^2}, \tag{2.2.24}$$

where L is the number of Grover transforms. Thus an estimate of $\langle f\rangle$ can be obtained with precision δ in a total number of function calls $O(\delta^{-1})$ and an additional processing overhead of $O(n, k, \delta^{-1})$.

This result could be compared with a straightforward use of the counting algorithm in estimating $\langle f\rangle$. One can estimate $\langle f\rangle$ by running an approximate counting algorithm on each bit of $\langle f\rangle$. However, in this bit-by-bit procedure, one must run the approximate counting routine k times, one time for each bit. Although, one does not need to estimate the lesser significant bits with the same precision as the most significant bits, an analysis shows that when all bits are relevant, the total number of function calls will be $O(k, \delta^{-1})$. Thus, when function calls are 'expensive' and high precision is required, the quantum algorithm that is presented here outperforms a straightforward use of the counting algorithm.

2.3 Single Query Information Retrieval

In this section we present a quantum algorithm for a coin-weighing problem, a problem of information retrieval. This is an application of Bernstein and Vazirani's parity problem [46, 47] (to which we will refer as the **BV parity** problem) which provides a strong illustration of the power of quantum computation. We will point out the limitations of classical information-theoretic bounds applied to quantum computation.

Information theory is a useful tool for analyzing the efficiency of classical algorithms. Problems involving information retrieval are particularly amenable to such analysis. Examples of such problems can be parlor games such as Mastermind, or "The Twenty Question Game". We have a 'hider' who has hidden a piece of information from a questioner. In the case of Mastermind this could be an array of colored pins, in the case of the Twenty Question Game it could be some word that the hider has in mind. The questioner is allowed to ask certain questions in order to determine what information the hider is hiding. The goal is to ask those questions that minimize the total number of questions that are needed to determine the piece of information.

Here we will consider the situation in which the hider has an n bit string which we will call y. The interaction with this bit string, the kind of question that one may ask, is given by the following. Upon presenting a query x, which is a bit string of length n, the answer

$$a(x,y) = x \cdot y \equiv \left(\sum_{i=1}^{n} x_i y_i \right) \mod 2, \tag{2.3.1}$$

is returned. Here x_i and y_i are the i^{th} bits of x and y. An example of such a problem is the database search problem (see section 2.2). In this case the unknown bit string y of length n has Hamming weight one (y has exactly one "1"). When x is restricted to bit strings with Hamming weight 1, this problem becomes the problem of searching a marked item, the "1", in a database. Grover [14] has shown (see section 2.2) that there is a quantum algorithm which is faster than any classical search and finds the marked item with high probability in $O(\sqrt{n})$ *quantum* queries. Grover's algorithm does not, however, violate the information theoretic lower bound on the minimal number of queries M.

The information-theoretic lower bound [48] on M is given by the amount of information to be retrieved divided by the maximal amount of information retrieved by a query which has A possible answers, i.e.

$$M \geq \frac{\log \#y}{\log A}, \tag{2.3.2}$$

where $\#y$ is the total number of different bit strings y. For example, when y is promised to have Hamming weight 1, then $\log \#y = \log n$ and the lower bound for the database search problem is $\log n$.

How does the quantum querying take place? The 'quantum hider' acts on two input registers: register X containing the query state $|x\rangle$ and register B, an output register of dimension 2 initially containing state $|b\rangle$. We define the operation of querying as

$$|x, b\rangle \stackrel{Query}{\rightarrow} |x, [b + a(x,y)] \mod 2\rangle, \tag{2.3.3}$$

where $a(x, y)$ is the answer to query x. In a protocol of information retrieval by quantum queries, the queries can be presented to the 'hider' in arbitrary superpositions. Because of this the information that is retrieved by a single quantum query is not bounded by $\log_2 A$, a feature that is exploited in the Bernstein Vazirani algorithm. The relevant quantity in the quantum setting is the accessible information in the registers X and B (together called XB) and the rest of the quantum computer Φ (with which the queries could be entangled) about the bit string y. The accessible information about y is bounded by the Holevo bound [49]

$$I_{\mathrm{acc}}(\Phi XB) \leq S(\Phi XB) - \sum_y p_y S(|\psi_y\rangle\langle\psi_y|), \tag{2.3.4}$$

where $S(\Phi XB) = -\mathrm{Tr}\,\rho_{\Phi XB} \log_2 \rho_{\Phi XB}$ is the von Neumann entropy of ΦXB and $\rho_{\Phi XB} = \sum_y p_y |\psi_y\rangle\langle\psi_y|$. Here p_y is the probability for bit string y and $\sum_y p_y = 1$. The state $|\psi_y\rangle$ is the state of the register X, B and the rest of the computer Φ when the hider has bit string y. Since $|\psi_y\rangle$ is a pure state, $S(|\psi_y\rangle\langle\psi_y|) = 0$. In the case of a classical query, the von Neumann entropy $S(\Phi XB)$ is less than or equal to $\log_2 \dim(B) = \log_2 A$ where A is the possible number of answers, which gives rise to the classical bound (2.3.2). For a quantum algorithm this bound can formally be replaced by $I_{\mathrm{acc}}(\Phi XB) \leq S(\Phi XB) \leq \log \dim \Phi XB$. This bound is too weak though to be of any use. The quantum algorithm of Bernstein and Vazirani violates the classical information-theoretic bound by extracting extra information in the phases of the query register X, thus using the larger available Hilbert space.

2.3.1 Coin Weighing

In the **BV parity** problem we consider a hider which has hidden an arbitrary n-bit string y. The answer to queries represented by n-bit strings x to the hider is the parity of the bits common to x and y given by $a(x, y) = x \cdot y$, as in Eq. (2.3.1). The problem is to determine y in its entirety. Bernstein and Vazirani have shown that y can be determined in only two queries to the database. But by preparing the outputregister B in an initial superposition $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ [50] the algorithm can be simplified to comprise of a single query. We prepare the following superposition:

$$\frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle). \tag{2.3.5}$$

Then we make the query which results in

$$\frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} (-1)^{x \cdot y} |x\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) = |\psi_y\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle). \tag{2.3.6}$$

We have $\langle\psi_y|\psi_{y'}\rangle = 0$ since

$$\frac{1}{2^n} \sum_{x=0}^{2^n-1} (-1)^{x \cdot y} (-1)^{x \cdot y'} = \delta_{y,y'}. \tag{2.3.7}$$

Therefore we can uniquely determine y.

This algorithm can be directly applied to the coin weighing problem. Coin weighing problems are a group of problems in which a set of coins of less weight is to be identified in a total set of coins of otherwise identical weight [48]. The good coins have weight a and the defective coins have weight b. We can weigh arbitrary sets of coins with a spring-scale (which gives the weight of the set of coins directly, as opposed to a balance which compares two sets of coins). All sets of coins are equiprobable. A set of n coins is represented as a bit string y of length n where $y_i = 1$ indicates that coin i is defective. A weighing can be represented by a query string x of length n, where x_i specifies whether coin i is included in the set to be weighed. The result of a classical weighing is the Hamming weight of the bitwise product of x and y, $\sum_{i=1}^{n} x_i y_i$. For this problem the information theoretic bound (2.3.2) gives

$$M \geq \frac{n}{\log_2(n+1)}. \tag{2.3.8}$$

This is $1/2$ of what the optimal predetermined classical algorithm which perfectly identifies the set of coins achieves [48]

$$\lim_{n\to\infty} M_{\mathrm{pre}}(n) = \frac{2n}{\log_2(n)}. \tag{2.3.9}$$

If one has a spring scale capable of performing weighings in superposition, then, one can use the Bernstein Vazirani algorithm to identify the defective coins perfectly with a single weighing by using only the parity of the Hamming weight answer.

2.4 Limits to Quantum Computation

In order to understand the power of quantum computation, it is important to find problems for which the use of a quantum computer is not advantageous. In the work of Beals *et al.* [53] important progress has been made in this direction. In this work a tight lower bound was derived for the **parity** problem [1]. The problem is the following. Given is a function $g\colon \{0,1\}^n \to \{0,1\}$ as an oracle. Upon an input $|x\rangle \otimes |y\rangle$ where $x \in \{0,1\}^n$ and $y \in \{0,1\}$, the oracle returns

$$|x\rangle \otimes |y\rangle \overset{\mathrm{Query}_p}{\to} |x\rangle \otimes |y \oplus g(x)\rangle. \tag{2.4.1}$$

The goal is to compute the function parity $p = \oplus_i\, g(x_i)$ with the minimum number of queries. If we require a (2-sided) bounded error, that is, one always gets the correct answer with probability larger than 2/3, the lower bound on the number of oracle calls on a quantum computer is $\Omega(2^{n-1})$. The bound is tight as there is a quantum algorithm that computes the parity in 2^{n-1} steps by repeating the computation of the parity of 2 bits with a single query 2^{n-1} times.

[1] This is *not* the **BV parity** problem.

One can use this result to derive a lower bound for the following problem [2]. Consider a function $f\colon \{0,1\}^n \to \{0,1\}^n$. The function is presented to us as a black box:

$$|x\rangle \otimes |b\rangle \overset{\text{Query}}{\to} |x\rangle \otimes |b \oplus f(x)\rangle. \tag{2.4.2}$$

The problem is to determine $f^k(x)$ given k and x with the least number of queries. Call this problem the **chain** problem [3]. In the classical setting it is not possible to determine $f^k(x)$ without evaluating the function f at least k times. It would be quite extraordinary if a quantum computer could perform this task with less than k queries for arbitrary functions f. We will show here that there are functions for which at least $\frac{k}{2}$ calls are required to determine $f^k(x)$ for some inputs x. The result is derived by a reduction to **parity**; we take a function f such that when $f^k(x)$ is computed, the parity of some k-bit string is determined. For convenience we write $f(x,y)$ where y denotes the last bit of the input (x,y) and x the first $n-1$ input bits. Let g be a function $g\colon \{0,1\}^n \to \{0,1\}$. We choose f such that

$$f(x,y) = (x+1 \mod 2^{n-1}, g(x) \oplus y). \tag{2.4.3}$$

Hereby we ensure that

$$f^k(x,y) = (x+k \mod 2^{n-1}, y \oplus_{l=0}^{k-1} g(x+l \mod 2^{n-1})). \tag{2.4.4}$$

For the input $(x,y) = (0,0)$ we have $f^k(0,0) = (k \mod 2^{n-1}, \oplus_{l=0}^{k-1} g(l \mod 2^{n-1}))$. Let Q be the number of function calls to the **chain** oracle of a quantum algorithm that computes $f^k(x,y)$ given (x,y) with bounded probability of error. This quantum algorithm can be used to solve the **parity** problem. The answer of each query to the **parity** black box

$$\underbrace{|0,\ldots,0\rangle}_{n-k} \otimes \underbrace{|x\rangle}_{k} \otimes |y\rangle \overset{\text{Query}_\text{P}}{\to} \underbrace{|0,\ldots,0\rangle}_{n-k} \otimes \underbrace{|x\rangle}_{k} \otimes |g(x) \oplus y\rangle \tag{2.4.5}$$

can be interpreted as the answer to a query to the **chain** oracle, Eq. (2.4.3) by adding a bit $(\mathrm{mod}\ 2^{n-1})$ after the query to x:

$$|x\rangle \otimes |g(x) \oplus y\rangle \to |x+1\rangle \otimes |g(x) \oplus y\rangle. \tag{2.4.6}$$

Running the **chain** algorithm for an input $(x,y) = (0,0)$ will solve the **parity** problem in Q queries to the parity black box. Thus we find that Q, the number of oracle calls that are required to determine $f^k(x,y)$ given (x,y) for every k with bounded probability of error, is $\Omega(k/2)$.

[2] This result has been found independently by Farhi *et al.* [54].

[3] A different technique for proving a lower bound on the number of queries for the **chain** problem has been used by Ozhigov [55].

Chapter 3

Simulating Quantum Operations with Mixed Environments

3.1 Introduction

Future quantum computers may be useful in studying the behavior of open quantum systems and the nature of decoherence [32, 56, 57, 58]. Instead of performing real experiments on quantum systems, a single quantum computer can be used as an efficient, multiple-purpose simulator for a wide variety of physical systems. In general, an important goal of such experiments will be understanding the effects arising from interactions between the system of interest S and another quantum system E. For example, S could represent the states of a molecule, which couple to other molecules E through long-range electronic dipolar interactions.

In this chapter, we study the amount of physical resources that will generally be required to perform simulations with quantum computers. Suppose S exists in a Hilbert space $\mathcal{H}_n$ of dimension n, and E is in $\mathcal{H}_r$ of dimension r. It is well known that any *quantum operation* [59] on $\mathcal{H}_n$, resulting from some interaction with E in $\mathcal{H}_r$ with arbitrary r, can be performed by appending a state in $\mathcal{H}_{n^2}$, evolving unitarily, and then tracing over $\mathcal{H}_{n^2}$. The difference between r and n^2 can represent a significant reduction, since E is often the "environment", or a large bath (for example, of harmonic oscillators), and r can be arbitrary large.

Can a general quantum operation be implemented with an environment even smaller than n^2 dimensions? Lloyd suggested [32] that it may be possible to implement a general quantum operation on k quantum bits (qubits) with a k-qubit environment – if one prepares the environment not in a pure state, but rather in an arbitrary *mixed* state.

Here we provide a specific counterexample to this conjecture for $k = 1$, although we find that at least for some operations, fewer resources are required than was previously known. Our counterexample is part of a class known as the generalized depolarizing channels, for which we show that a three-dimensional environment is sufficient for simulation. The proof of the counterexample is established by the technique of computing Gröbner bases.

Our results also address the following question: suppose an environment E and the form of its interaction with another system S, is given as a black box. We prepare a system S in a known initial state, we let it evolve for a fixed period of time through the unknown interaction with E and finally measure the state of S. During this procedure the environment E is not accessible to us. A method to completely determine the quantum operation χ performed by this black box on S is known [60]. But we may also ask: what is the smallest environment E with which S could have interacted? This work shows that knowledge about χ can be used to find bounds on the nature of E.

3.2 Quantum Operations and Measurements

We begin by summarizing the mathematical formalism of quantum operations and measurements. The most general transformation on a quantum system is a linear, trace-preserving, completely positive map. These maps give a mathematical description of a general process that can take place on a quantum system S. Let ρ_S be the density matrix of an initially isolated n-dimensional quantum system S. This quantum system interacts unitarily with some other quantum system E of dimension k that starts in state ρ_E. Assume that after the interaction we can no longer access a part of this system, say, a subsystem D of dimension d. Then we will represent the state of the remaining system as

$$\rho_{final} = \mathrm{Tr}_D\, U\, \rho_S \otimes \rho_E\, U^\dagger = \chi(\rho_S). \tag{3.2.1}$$

The map χ is an example of a linear, trace-preserving completely positive map that maps a density matrix in a n-dimensional Hilbert space onto a density matrix in a kn/d dimensional Hilbert space. The mathematical description of these maps is the following. A positive linear map $\chi\colon B(\mathcal{H}_n) \to B(\mathcal{H}_m)$, where $B(\mathcal{H}_n)$ is the algebra of linear operators on a Hilbert space $\mathcal{H}_n$, maps positive semi-definite operators in $B(\mathcal{H}_n)$ onto positive semi-definite operators in $B(\mathcal{H}_m)$. A positive linear map χ is called completely positive iff for all $k = 1, 2, \ldots,$ the map

$$\chi \otimes \mathrm{id}_k\colon B(\mathcal{H}_n \otimes \mathcal{H}_k) \to B(\mathcal{H}_m \otimes \mathcal{H}_k), \tag{3.2.2}$$

is positive, where id_k is the identity map on $B(\mathcal{H}_k)$. This ensures, if the map is also trace-preserving, that a density matrix of which only a subsystem undergoes the action of the map remains a density matrix. Therefore the set of trace-preserving completely positive linear maps corresponds to the set of maps that can be physically implemented. Every completely positive linear map $\chi : B(\mathcal{H}_n) \to B(\mathcal{H}_m)$ can be decomposed into a set of at most nm many $m \times n$ matrices A_i [61] (which we shall refer to as "operation elements") as follows

$$\chi(\rho) = \sum_{i=1}^{nm} A_i \rho A_i^\dagger. \tag{3.2.3}$$

If the map is trace-preserving then the operation elements A_i obey the additional constraint

$$\sum_{i=1}^{nm} A_i^\dagger A_i = \mathbf{1}_n. \tag{3.2.4}$$

with $\mathbf{1}_n$ the identity matrix on $\mathcal{H}_n$. Following Choi [61], we call the set of all trace-preserving completely positive linear maps $\chi : B(\mathcal{H}_n) \to B(\mathcal{H}_m)$ $\mathbf{TCP}[n,m]$. A possible physical implementation of these maps is represented in Fig. 3.1: A unitary operation on the state $\rho \otimes |0\rangle\langle 0|$ (where $|0\rangle$ represents some pure state in an m^2-dimensional environment) is performed and then nm "degrees of freedom" are traced out:

$$\chi(\rho) = \sum_{k=1}^{nm} \langle e_k|\, U \left[\rho \otimes |0\rangle\langle 0|\right] U^\dagger\, |e_k\rangle. \tag{3.2.5}$$

Here $\{|e_k\rangle\}_{k=1}^{nm}$ is a set of basis vectors for $\mathcal{H}_{nm}$. As there are at most nm operation elements, it follows that one can implement any map in $\mathbf{TCP}[n,m]$ with an environment of dimension m^2.

General quantum measurements can be described within the same formalism. A general quantum measurement, a Positive Operator Valued Measurement (POVM) [62], on a density matrix $\rho \in B(\mathcal{H}_n)$ can be described by a collection of completely positive linear maps $\mathcal{S}_i\colon B(\mathcal{H}_n) \to B(\mathcal{H}_{m_i})$, $i = 1, \ldots, k$. Note that the output dimension can depend on i. Each measurement outcome i corresponds to the application of one of these maps $\mathcal{S}_i$ on the state ρ. As these maps are completely positive the action of $\mathcal{S}_i$ can be given by its operation elements A^i_j, or

$$\mathcal{S}_i(\rho) = \sum_{j=1}^{nm_i} A^i_j\, \rho {A^i_j}^\dagger. \tag{3.2.6}$$

The probability for measurement outcome i is then given by

$$\mathrm{Prob}(i) = \mathrm{Tr}\, \mathcal{S}_i(\rho) = \sum_{j=1}^{nm_i} \mathrm{Tr}\, {A^i_j}^\dagger A^i_j\, \rho. \tag{3.2.7}$$

The sum of the probabilities of the different measurement outcomes must be equal to one. This implies that

$$\sum_{i,j=1}^{k,nm_i} {A^i_j}^\dagger A^i_j = \mathbf{1}_n. \tag{3.2.8}$$

When we obtain outcome i the state itself is mapped onto

$$\frac{\sum_{j=1}^{nm_i} A^i_j\, \rho\, {A^i_j}^\dagger}{\sum_{j=1}^{nm_i} \mathrm{Tr}\, {A^i_j}^\dagger A^i_j\, \rho}. \tag{3.2.9}$$

The von Neumann measurements form a special subclass of POVM measurements for which there is a single operation element $A^i_j = \pi_i \delta_{ij}$ for each outcome i. The operator π_i is a projector. These projectors π_i are mutually orthogonal,

$$\forall\ i \neq j,\ \ \pi_i \pi_j = 0, \tag{3.2.10}$$

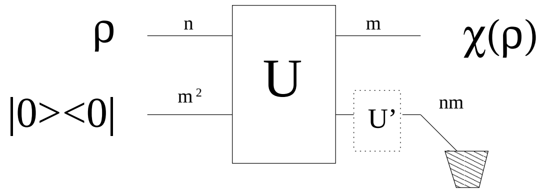

Figure 3.1: Implementation of the map χ using a pure state environment.

and they span the full Hilbert space:

$$\sum_i \pi_i = \mathbf{1}_n. \tag{3.2.11}$$

In a *complete* von Neumann measurement the projectors π_i have rank 1, i.e. they project onto pure states. For an *incomplete* von Neumann measurement the projectors π_i can have rank more than 1. Often, the POVMs that one considers are such that a measurement outcome i corresponds to a single operation element A_i. From such measurement, one can always construct the more general type by grouping measurement outcomes together. The essential difference between a quantum operation χ as in Eq. (3.2.5) and a quantum measurement is that in a quantum measurement we realize one of the outcomes corresponding to an operation element and acquire information about the state of the system. In a quantum operation as in Eq. (3.2.5) we do not acquire information about the system; we update our representation of the system corresponding to the possible interactions with the environment and the state of the environment.

To determine the dimension of the parameter space of $\mathbf{TCP}[n,m]$ we note that the map χ does not uniquely determine the set $\{A_i\}_{i=1}^{nm}$. Any set of $m \times n$ matrices $\{B_i\}_{i=1}^{nm}$ and $\{A_j\}_{j=1}^{nm}$ that are related by a unitary transformation

$$B_i = \sum_{j=1}^{nm} U'_{ij} A_j, \tag{3.2.12}$$

implement the same map χ. This freedom corresponds to a unitary rotation U' (see Fig. 3.1) of the environment qubits after the completion of the interaction U. It was shown in Ref. [59] that this unitary equivalence is the *only* freedom in the choice for the set of operators $\{A_i\}_{i=1}^{nm}$.

The dimension of the parameter space of all maps in $\mathbf{TCP}[n,m]$ that can be implemented with a d-dimensional pure environment will therefore be

$$D^{n\to m}_{\mathsf{pure},d} = \overbrace{2n^2 d}^{\text{parameters in } \{A_i\}} - \overbrace{(nd/m)^2}^{\text{unitary freedom}} - \overbrace{n^2}^{\text{constraint (3.2.4)}}, \tag{3.2.13}$$

since there are nd/m operation elements A_i (d is such that m divides nd). Thus we have $D_{\mathbf{TCP}[n,m]} = D^{n\to m}_{\mathsf{pure},m^2} = n^2(m^2-1)$.

In a more general physical implementation, however, the initial state of the environment can be an arbitrary density matrix. Consider the set of completely positive trace-preserving linear maps $\chi : B(\mathcal{H}_n) \to B(\mathcal{H}_m)$ that are implemented by an environment that is initially in some d-dimensional density matrix. We call this set $S_{\text{mix}}[d, n, m]$. The action on the input state ρ is

$$\chi(\rho) = \sum_{j=1}^{d} \lambda_j \sum_{k=1}^{dn/m} \langle e_k| \, U \left[\rho \otimes |j\rangle\langle j|\right] U^\dagger \, |e_k\rangle \,, \tag{3.2.14}$$

where $\{\lambda_j, |j\rangle\}_{j=1}^d$ are now the eigenvalues and eigenvectors of the mixed environment state. We identify a set of $m \times n$ matrices $\{A_{jk}\}_{j=1,k=1}^{d,dn/m}$ in the representation of Eq. (3.2.3):

$$A_{jk} = \sqrt{\lambda_j} \langle e_k|U|j\rangle. \tag{3.2.15}$$

Unitarity implies that these matrices are constrained,

$$\sum_{k=1}^{dn/m} A_{ik}^\dagger A_{jk} = \delta_{ij} \lambda_i \mathbf{1}_n. \tag{3.2.16}$$

There is a residual unitary freedom in choosing the set of matrices $\{A_{jk}\}_{j=1,k=1}^{d,nd/m}$. The set $\{B_{jm}\}_{j=1,m=1}^{d,nd/m}$ with $B_{jm} = \sum_k U'_{mk} A_{jk}$, where the $d\,n/m$-dimensional unitary matrix U' does not depend on the label j, implements the same quantum operation and also obeys constraint (3.2.16). As before, this freedom corresponds to a unitary transformation on the environment after the completion of the operation. The dimension of the parameter space of $S_{\text{mix}}[d, n, m]$ can be bounded as

$$D_{\text{pure},d}^{n\to m} \leq D_{\text{mix},d}^{n\to m} \leq D_{\text{pure},d^2}^{n\to m}. \tag{3.2.17}$$

The upper bound is given by the fact that one can always simulate a d-dimensional mixed environment with a d^2-dimensional pure environment: Let ρ be the density matrix of the d-dimensional environment and let $\{\lambda_i, |\psi_i\rangle_1\}_{i=1}^d$ be its eigenvalues and eigenvectors. The partial trace of the d^2-dimensional pure state $|\Phi\rangle = \sum_{i=1}^{d^2} \sqrt{\lambda_i} |\psi_i\rangle_1 \otimes |\psi_i\rangle_2$ is equal to ρ, i.e. $\rho = \mathrm{Tr}_2 \, |\Phi\rangle\langle\Phi|$.

From Eq. (3.2.13) and Eq. (3.2.17) it follows that an environment of dimension $d < m$ cannot be used to implement *all* maps in $\mathbf{TCP}[n, m]$. A map χ that is decomposable in m or fewer linearly-independent operation elements is extremal [61] in $\mathbf{TCP}[n, m]$. The extremal maps in $\mathbf{TCP}[n, m]$, cannot be simulated with $d < m$. These maps can be implemented with a pure-state environment of dimension m; moreover, we prove that there does not exist a more efficient implementation of these maps using a mixed-state environment:

Extremality implies that the map χ cannot be written as a convex combination of linearly independent maps χ^i that each have operation elements A_j^i for which $\sum_j {A_j^i}^\dagger A_j^i = \mathbf{1}_n$ for each i. This ensures that only one of the eigenvalues in constraint (3.2.16) is non-zero, but this in fact corresponds to a pure-state environment of dimension m. An example of such an extremal map is a complete von Neumann measurement on a n-dimensional system. The set of projection operators $\{\pi_i\}_{i=1}^n$ can be implemented minimally by using an n-dimensional pure state.

3.3 Generalized Depolarizing Channels

We now turn to the question of whether all maps in $\mathbf{TCP}[n,m]$ can be implemented with $d=m$. Note that our parameter count does not exclude this. In the following, we restrict ourselves to the case $n=m=2$. We study which maps can be implemented using a single-qubit environment and provide a proof that a particular qubit channel, the two-Pauli channel, cannot be implemented in this way.

We consider a special set of maps, the generalized depolarizing channels [37], which are described by the set $\{(\epsilon_i, A_i)\}_{i=1}^{4}$ where

$$\chi(\rho) = \sum_i \epsilon_i A_i \rho A_i^\dagger, \tag{3.3.1}$$

such that $\epsilon_1+\epsilon_2+\epsilon_3+\epsilon_4=1$ and the operators A_i are given by $A_1=\mathbf{1}_2$, $A_2=\sigma_x$, $A_3=\sigma_y$, $A_4=\sigma_z$. One can represent this family of maps geometrically as a tetrahedron, which is embedded in a cube with vertices at $(1,-1,-1)$, $(-1,1,-1)$, $(1,1,1)$ and $(-1,-1,1)$. The transformation that relates the parameters $\epsilon_1,\epsilon_2,\epsilon_3,\epsilon_4$ to the (x,y,z) coordinates is given by $x=\epsilon_1+\epsilon_2-\epsilon_3-\epsilon_4$, $y=\epsilon_1-\epsilon_2+\epsilon_3-\epsilon_4$, and $z=\epsilon_1-\epsilon_2-\epsilon_3+\epsilon_4$. The vertices of the tetrahedron correspond to a single-operator map. Its edges are two-operator maps, the four faces represent all three-operator maps, and the points in the interior of the tetrahedron are all the four-operator maps of Eq. (3.3.1).

A (nonexhaustive) computer search shows that only a subset of these maps can be simulated by using a qubit environment. For this subset we are able to construct an explicit qubit solution. At web address [63] one can find pictures of the three-dimensional volume that is described by the solution set and a picture of the solution set as generated by the computer search. The computer work also revealed that the dimension of $S_{\mathsf{mix}}[2,2,2]$ is equal to the upper bound of Eq. (3.2.17), namely $\mathbf{TCP}[2,2]=12$. Thus there is enough "room" for a solution, but it is not in the right place, as we will see.

This solution is constructed in the following way. We start with the center of mass of the tetrahedron, the point $(\epsilon_1,\epsilon_2,\epsilon_3,\epsilon_4)=(1/4,1/4,1/4,1/4)$. This channel has the property that it maps every input state ρ onto $\frac{1}{2}\mathbf{1}_2$. It can thus be easily implemented by performing a SWAP gate on a environment qubit that is initially in the $\frac{1}{2}\mathbf{1}_2$ state and the input qubit. The SWAP gate on two registers $|a\rangle\otimes|b\rangle$ gives $|b\rangle\otimes|a\rangle$. Then one considers the line that departs from a vertex, say the point $(\epsilon_1,\epsilon_2,\epsilon_3,\epsilon_4)=(1,0,0,0)$, and goes through the center of mass. This one-dimensional set of channels is characterized by $\epsilon_2=\epsilon_3=\epsilon_4$ and represents the regular depolarizing channel [37]. Performing a $\sqrt[m]{\mathrm{SWAP}}$ on a $\frac{1}{2}\mathbf{1}_2$ environment and the input qubit implements these channels, up to $\epsilon_1=1/4$. The integer m is related to the ϵ parameters by $\epsilon_2=\epsilon_3=\epsilon_4=\sin^2(\frac{\pi}{2m})/4$. One extra step of generalization gives us an even larger set of channels. The unitary matrix is a somewhat generalized form of $\sqrt[m]{\mathrm{SWAP}}$,

$$U = \begin{pmatrix} e^{i\theta}\cos\phi_1 & 0 & 0 & ie^{i\theta}\sin\phi_1 \\ 0 & \cos\phi_2 & i\sin\phi_2 & 0 \\ 0 & i\sin\phi_2 & \cos\phi_2 & 0 \\ ie^{i\theta}\sin\phi_1 & 0 & 0 & e^{i\theta}\cos\phi_1 \end{pmatrix}, \tag{3.3.2}$$

and the environment is again prepared in state $\frac{1}{2}\mathbf{1}_2$. We can determine the operation elements and express these as linear (unitary) combinations of the Pauli matrices. This leads to an expression of the parameters ϵ_i in terms of $(\theta, \phi_1, \phi_2) \in [0, 2\pi] \times [0, 2\pi] \times [0, 2\pi]$:

$$\begin{aligned} \epsilon_1 &= \tfrac{1}{4}(\cos^2\phi_1 + \cos^2\phi_2 + 2\cos\phi_1\cos\phi_2\cos\theta), \\ \epsilon_2 &= \tfrac{1}{4}(\sin^2\phi_1 + \sin^2\phi_2 + 2\sin\phi_1\sin\phi_2\cos\theta), \\ \epsilon_3 &= \tfrac{1}{4}(\sin^2\phi_1 + \sin^2\phi_2 - 2\sin\phi_1\sin\phi_2\cos\theta), \\ \epsilon_4 &= \tfrac{1}{4}(\cos^2\phi_1 + \cos^2\phi_2 - 2\cos\phi_1\cos\phi_2\cos\theta). \end{aligned} \tag{3.3.3}$$

Alternatively the solution set can be expressed as a set of inequalities on the parameters ϵ_i: $S = S_1 \cup S_2 \cup S_3 \cup S_4$ where

$$S_1 = \{(\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4) | (\epsilon_1\epsilon_2 \geq \epsilon_3\epsilon_4) \wedge (\epsilon_1\epsilon_3 \geq \epsilon_2\epsilon_4) \wedge (\epsilon_1\epsilon_4 \geq \epsilon_2\epsilon_3)\}, \tag{3.3.4}$$

and $S_2 = S_1(\epsilon_1 \leftrightarrow \epsilon_2)$, $S_3 = S_1(\epsilon_1 \leftrightarrow \epsilon_3)$ and $S_4 = S_1(\epsilon_1 \leftrightarrow \epsilon_4)$. In Appendix 3.A we provide a proof that S and the volume parameterized by Eq. (3.3.3) coincide.

3.3.1 Two-Pauli Channel

We now turn to a subset of these maps, the two-Pauli channels, which are given by the three operation elements

$$A_1 = \mathbf{1}_2\sqrt{x},\ A_2 = \sigma_x\sqrt{(1-x)/2},\ A_3 = i\sigma_y\sqrt{(1-x)/2}, \tag{3.3.5}$$

and A_4 is the null matrix. We will prove that for $0 < x < 1$, there is no qubit environment which simulates this channel. For $x = 0$ or $x = 1$ there is a two-dimensional environment that can simulate the channel as the channel has two (nonzero) operation elements when $x = 0$ and only one operator when $x = 1$.

Any unitary linear combination of A_1, A_2 and A_3 may be written as

$$B_k = \begin{pmatrix} b_k\sqrt{x} & (c_k - a_k)\sqrt{\frac{1}{2}(1-x)} \\ (c_k + a_k)\sqrt{\frac{1}{2}(1-x)} & b_k\sqrt{x} \end{pmatrix}, \tag{3.3.6}$$

with appropriate constraints resulting from unitarity on the coefficients a_k, b_k, c_k. This new set of operators $\{B_k\}_{k=0}^{3}$ will implement the same channel due to Eq. (3.2.12). Furthermore, these operators B_k are constrained through Eq. (3.2.16); we can relabel them as follows $B_0 = B_{11}$, $B_1 = B_{12}$, $B_2 = B_{21}$ and $B_3 = B_{22}$. For notational convenience, we define

$$\begin{aligned} |u_0\rangle &= \tfrac{1}{\sqrt{2}}\begin{pmatrix} a_0 + c_0 \\ a_1 + c_1 \end{pmatrix}, & |u_1\rangle &= \tfrac{1}{\sqrt{2}}\begin{pmatrix} a_2 + c_2 \\ a_3 + c_3 \end{pmatrix}, \\ |w_0\rangle &= \tfrac{1}{\sqrt{2}}\begin{pmatrix} c_0 - a_0 \\ c_1 - a_1 \end{pmatrix}, & |w_1\rangle &= \tfrac{1}{\sqrt{2}}\begin{pmatrix} c_2 - a_2 \\ c_3 - a_3 \end{pmatrix}, \\ |v_0\rangle &= \begin{pmatrix} b_0 \\ b_1 \end{pmatrix}, & |v_1\rangle &= \begin{pmatrix} b_2 \\ b_3 \end{pmatrix}. \end{aligned} \tag{3.3.7}$$

Using the assumption $0 \neq x \neq 1$ and by linearly combining all the equations we obtain:

$$\langle v_0|w_0\rangle + \langle u_0|v_0\rangle \;=: g_1 = \; 0 \tag{3.3.8}$$
$$\langle v_1|w_1\rangle + \langle u_1|v_1\rangle \;=: g_2 = \; 0 \tag{3.3.9}$$
$$\langle v_0|w_1\rangle + \langle u_0|v_1\rangle \;=: g_3 = \; 0 \tag{3.3.10}$$
$$\langle w_0|v_1\rangle + \langle v_0|u_1\rangle \;=: g_4 = \; 0 \tag{3.3.11}$$
$$\langle u_0|u_0\rangle - \langle w_0|w_0\rangle \;=: g_5 = \; 0 \tag{3.3.12}$$
$$\langle u_1|u_1\rangle - \langle w_1|w_1\rangle \;=: g_6 = \; 0 \tag{3.3.13}$$
$$\langle u_0|u_0\rangle + \langle u_1|u_1\rangle - 1 \;=: g_7 = \; 0 \tag{3.3.14}$$
$$\langle v_0|v_0\rangle + \langle v_1|v_1\rangle - 1 \;=: g_8 = \; 0 \tag{3.3.15}$$
$$\langle u_0|v_0\rangle + \langle u_1|v_1\rangle \;=: g_9 = \; 0 \tag{3.3.16}$$
$$\langle u_0|w_0\rangle + \langle u_1|w_1\rangle \;=: g_{10} = \; 0 \tag{3.3.17}$$
$$\langle u_0|u_1\rangle - \langle w_0|w_1\rangle \;=: g_{11} = \; 0 \tag{3.3.18}$$

Writing each of the coefficients a_k, b_k, and c_k in the form $x_j + ix_{j+1}$ (where $i^2 = -1$), we get a system of polynomial equations $\mathrm{Re}(g_1) = \mathrm{Im}(g_1) = \ldots = \mathrm{Im}(g_{11}) = 0$, where $\mathrm{Re}(g_k)$ and $\mathrm{Im}(g_k)$ are polynomials in the variables $x_1, \ldots, x_{24}$ with real coefficients. To show that this system of equations has no solution we make use of Gröbner bases (see e. g. [64]). The computation of a Gröbner basis with Buchberger's algorithm generalizes the Euclidean algorithm to compute the greatest common divisor (GCD) of univariate polynomials $p_1(x)$ and $p_2(x)$. In that case, the GCD $g(x)$ can be written as a "linear" combination

$$g(x) = f_1(x)p_1(x) + f_2(x)p_2(x). \tag{3.3.19}$$

The two univariate polynomials p_1 and p_2 have a common root if and only if their GCD is non-trivial, i. e., $g(x) \neq 1$.

For multivariate polynomials, a common solution exists iff the Gröbner basis of the ideal generated by them is non-trivial, i. e., does not contain a constant. In our case, using the computer algebra system Magma [65] we have shown that there exist polynomials $f_1, \ldots, f_{11}$ such that

$$\sum_{j=1}^{11} f_j(x_1, \ldots, x_{24}) g_j(x_1, \ldots, x_{24}) = 1, \tag{3.3.20}$$

i. e., the Gröbner basis contains 1 and there is no solution of the equations (3.3.8)–(3.3.18). □

3.3.2 Qutrit Solution

Despite the above proof, it turns out that the class of channels we have been studying does not require a two qubit environment ($d = 4$) for their simulation; a mixed *qutrit* ($d = 3$) suffices. For generalized depolarizing channels, there will be nine operators, $\{A_{ij}\}_{i,j=1}^{3,3}$. We set one eigenvalue $\lambda_3 = 0$ and thus

$$A_{31} = A_{32} = A_{33} = 0. \tag{3.3.21}$$

If $\epsilon_1\epsilon_2 \geq \epsilon_3\epsilon_4$ the solution is

$$\begin{array}{ll} A_{11} = 0, & A_{21} = \sqrt{\epsilon_2 - \epsilon_3\epsilon_4/\epsilon_1}\sigma_x, \\ A_{12} = \sqrt{\epsilon_3}\sigma_z, & A_{22} = \sqrt{\epsilon_4}\sigma_y, \\ A_{13} = \sqrt{\epsilon_1}\mathbf{1}_2, & A_{23} = -i\sqrt{\epsilon_3\epsilon_4/\epsilon_1}\sigma_x. \end{array} \tag{3.3.22}$$

Otherwise, we take

$$\begin{array}{ll} A_{11} = 0, & A_{21} = \sqrt{\epsilon_4 - \epsilon_1\epsilon_2/\epsilon_3}\sigma_y, \\ A_{12} = \sqrt{\epsilon_1}\mathbf{1}_2, & A_{22} = \sqrt{\epsilon_2}\sigma_x, \\ A_{13} = \sqrt{\epsilon_3}\sigma_z, & A_{23} = i\sqrt{\epsilon_1\epsilon_2/\epsilon_3}\sigma_y. \end{array} \tag{3.3.23}$$

One can check that this set implements any generalized depolarizing channel and satisfies Eq. (3.2.16).

On the basis of the computer work we conjecture that *any* map in $\mathbf{TCP}[2,2]$ can be simulated with a qutrit environment. Also, the numerics indicate that one can always set one eigenvalue to zero. Furthermore, we have numerical evidence that channels that have three linearly independent operation elements can never be simulated with a qubit environment.

3.4 Discussion

Our results provide new bounds on the size of an environment needed to simulate certain quantum operations on single qubits. However, we have only addressed simple mappings on the smallest input space. Many questions now arise: how do these results generalize to mappings on n-dimensional systems? A relevant scenario might be n uses of the generalized depolarizing channel, where the environment can be shared between the channels. In such a case, might a qubit environment per channel suffice for large n? A nice extension of the generalized depolarizing channels are the channels that are defined with the Heisenberg group elements [66]. These channels on n-dimensional inputs are mixtures of a set of n^2 unitary matrices $U(i,j)$. However, it is not straightforward to construct solutions, as in the qutrit case, for a general "Heisenberg channel", and we have no insight at the moment of what gain one can get by using mixed states here. We do expect that there will be channels on an n-dimensional input for which the dimension of the environment is $\Omega(n^2)$. The questions we have formulated also apply to the construction of generalized measurements: how large an environment is needed for the minimal-size construction of arbitrary generalized measurements on an n-dimensional system? We hope our results and the questions they motivate will be useful in future quantum computing applications, and provide fundamental insights into the properties of quantum systems.

3.A Proof of Coincidence of Volumes

The solution set that is characterized by the set of inequalities on the parameters ϵ_i reads

$$S = S_1 \cup S_2 \cup S_3 \cup S_4, \tag{3.A.1}$$

where

$$S_1 = \{(\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4) | (\epsilon_1\epsilon_2 \geq \epsilon_3\epsilon_4) \wedge (\epsilon_1\epsilon_3 \geq \epsilon_2\epsilon_4) \wedge (\epsilon_1\epsilon_4 \geq \epsilon_2\epsilon_3)\}, \tag{3.A.2}$$

and $S_2 = S_1(\epsilon_1 \leftrightarrow \epsilon_2)$, $S_3 = S_1(\epsilon_1 \leftrightarrow \epsilon_3)$ and $S_4 = S_1(\epsilon_1 \leftrightarrow \epsilon_4)$.

We prove that the volume given by the parameterization of (x, y, z) in terms of (ϕ_1, ϕ_2, θ), Eq. (3.3.3), which we call P_1, is equal to the volume described by the inequalities of Eq. (3.A.1) and (3.A.2), which we call P_2.

The volume P_1 is generated by a mapping of a three-dimensional torus specified by coordinates (ϕ_1, ϕ_2, θ) to (x, y, z). The Jacobian determinant $|\det \frac{\partial(x,y,z)}{\partial(\phi_1,\phi_2,\theta)}|$ will vanish on a set of points which we call R_{P_1} that include the surface of P_1 denoted as Σ_{P_1}. A priori, R_{P_1} might include points interior to P_1, but we will rule this out. We will show that $R_{P_1} = \Sigma_{P_2}$. Then, by inspection of the volume P_2 we can conclude that R_{P_1} can only be the surface Σ_{P_1}.

Proof that $R_{P_1} = \Sigma_{P_2}$

The Jacobian determinant of the transformation is

$$J = |4\cos\theta \sin\theta(\cos^2\phi_1 \sin^2\phi_1 - \cos^2\phi_2 \sin^2\phi_2)|. \tag{3.A.3}$$

First we show that the volume P_1 is unchanged under all permutations of $\epsilon_1, \ldots, \epsilon_4$. The permutations are generated by transpositions of two elements ϵ_i and ϵ_j. Transposition $(\epsilon_1\epsilon_2)(\epsilon_3)(\epsilon_4)$ will map (x, y, z) onto $(x, -z, -y)$, and similarly, the other transpositions interchange x, y and z and add minus signs. The Jacobian determinant is invariant under these transformations, thus the surface is a symmetric function of $\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4$. This implies that the volume itself is symmetric in $\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4$.

We will show that the Jacobian vanishes for points (x, y, z) iff $(x, y, z) \in \Sigma_{P_2}$, where Σ_{P_2} denotes the surface of P_2.

$R_{P_1} \supseteq \Sigma_{P_2}$ The Jacobian $(J = 0)$ vanishes iff one or more of the following hold

- $\cos\theta = 0 \Rightarrow \epsilon_1\epsilon_3 = \epsilon_2\epsilon_4$, or
- $\sin\theta = 0 \Rightarrow \epsilon_1\epsilon_2 = \epsilon_3\epsilon_4$, or
- $\cos^2\phi_1 \sin^2\phi_1 - \cos^2\phi_2 \sin^2\phi_2 = 0 \Rightarrow \epsilon_1\epsilon_4 = \epsilon_2\epsilon_3$.

These resulting equalities define points on Σ_{P_2}. Note that the surface is connected (see Ref. [63]), i.e. there is a way to reach every point on the surface from every starting point on the surface without leaving the surface and thus the volume P_1 is (edge) connected.

$\Sigma_{P_2} \supseteq R_{P_1}$ Take section S_1 (Eq.(3.A.2)) for which there are three parts of the surface C_1, C_2 , C_3 with $C_1 = \{\epsilon_1\epsilon_3 = \epsilon_2\epsilon_4, \epsilon_1\epsilon_2 \geq \epsilon_3\epsilon_4, \epsilon_1\epsilon_4 \geq \epsilon_2\epsilon_3\}$ and $C_2 = C_1(\epsilon_2 \leftrightarrow \epsilon_3)$ $C_3 = C_1(\epsilon_3 \leftrightarrow \epsilon_4)$. The Jacobian vanishes for $\epsilon_1\epsilon_3 = \epsilon_2\epsilon_4$. As C_2 and C_3 are obtained by permutations, the Jacobian will also vanish on these surfaces. Now by interchanging ϵ_1 with ϵ_2 we obtain S_2 etc..., but again because the Jacobian is invariant under this transposition, we know that the Jacobian will vanish also on the surface of S_2, S_3 and S_4.

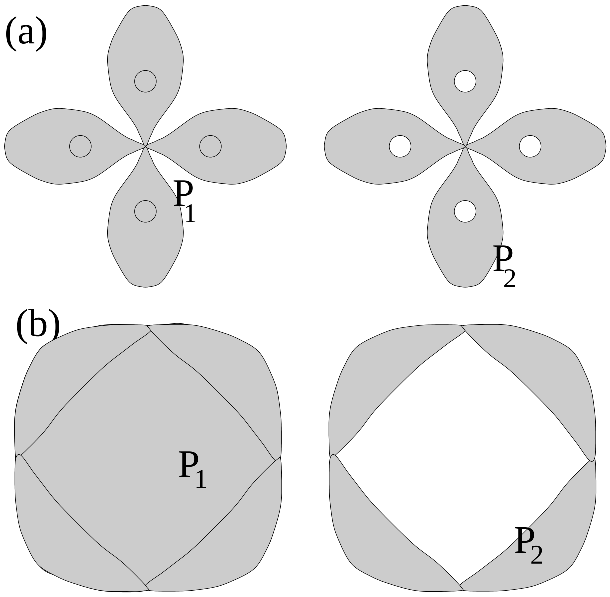

Figure 3.2: Schematic examples of volumes P_1 and P_2 that do not coincide even though $R_{P_1} = \Sigma_{P_2}$.

P_2 has no Holes or Non-contractible Surfaces

We must rule out the possibility that R_{P_1} includes points in the interior of P_1. In Fig. 3.2 a couple of example-volumes are given for which $R_{P_1} = \Sigma_{P_2}$ but $P_1 \neq P_2$. The volume P_2 consists of four sections. Any two sections intersect on a line. For volumes S_1 and S_2 this is the line given by $(\epsilon_1 = \epsilon_2, \epsilon_3 = \epsilon_4)$. Inspection of each section shows that it has no holes as in Fig. 3.2(a) [1]. By joining the sections it is possible that a non-contractible surface is created

[1] A rigorous proof could be established, say for section S_1, by showing that rays emanating from points on the line $\epsilon_2 = \epsilon_3 = \epsilon_4, 1/4 \leq \epsilon_1 \leq 1$ cross the surface of S_1 only once.

as in Fig. 3.2(b). Inspection of the volume also tells us that this is not the case for P_2. Thus the set of points R_{P_1} that coincides with Σ_{P_2} cannot contain points that lie inside the volume P_1. From this we can conclude that $\Sigma_{P_1} = \Sigma_{P_2}$.

Finally, there exists a point which is both inside P_1 and inside P_2, for example the point $\epsilon_1 = 5/8$, $\epsilon_2 = \epsilon_3 = \epsilon_4 = 1/8$, and P_1 is (edge)connected, therefore P_1 and P_2 coincide. $\square$

Chapter 4

On the Problem of Equilibration and the Computation of Correlation Functions on a Quantum Computer

4.1 The Limits of Classical Computation

The power of quantum computers has been demonstrated in several algorithms, of which the most striking have been Shor's factoring algorithm [8, 67] and Grover's search algorithm [14]. From the very start however, the quantum computer has also held the promise of being a simulator of physical systems. This is the content of the physical version of the Church-Turing thesis proposed by Deutsch [4] which says that every finitely realizable physical system can be perfectly simulated by a universal model computing machine operating by finite means. Thus we might expect that the universal quantum computer can be used to simulate any experiment that we could do on a real physical system. If such a simulation can be done efficiently (that is, without exponential slowdown), it is clear that this could be one of the major applications of a quantum computer [1]. This promise seems to have been only partly fulfilled until now; it has been shown by several researchers [32, 58] that a simulation of the unitary time evolution of a physical system that possesses some degree of locality (which realistic physical systems do) can be accomplished efficiently on a quantum computer. However, many quantities of interest that are determined by experiment, or by the use of classical simulation techniques, relate to open quantum systems, in particular to systems in thermal equilibrium. The thermal equilibrium (Gibbs) state (in the canonical ensemble) of a Hamiltonian H is given by

$$\rho_\beta = \sum_{m=1}^{N} \frac{e^{-\beta E_m}}{Z} |m\rangle\langle m|, \tag{4.1.1}$$

[1] Here we assume that the size of the system is not extremely large, for example it is not realistic to expect us to be able to simulate a system with 10^{23} degrees of freedom.

where $|m\rangle$ (E_m) are the eigenvectors (eigenvalues) of H. Z is the partition function

$$Z = \sum_{m=1}^{N} e^{-\beta E_m}, \tag{4.1.2}$$

and $\beta = \frac{1}{kT}$ where k is Boltzmann's constant and T the absolute temperature in degrees Kelvin. The physical systems that concern us here will have a finite dimensional Hilbert space $\mathcal{H}$ that can be decomposed as

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \ldots \otimes \mathcal{H}_n, \tag{4.1.3}$$

where each $\mathcal{H}_i$ represents a small, constant Hilbert space, typically associated with some (generalized) spin or other local degree of freedom. The Hamiltonian couples these local Hilbert spaces, for example in correspondence with a d-dimensional spatial lattice, so that there is only coupling between adjacent "spins" on this lattice. The quantities of interest, computed in experiment or in a classical computation, are of the form

$$\mathrm{Tr}\, O_1(t_1) O_2(t_2) \ldots O_k(t_k) \rho_\beta, \tag{4.1.4}$$

where the $O_i(t_i)$ are (possibly time-dependent) observables. Both for classical systems as well as for quantum systems, computational Monte Carlo methods have been developed to estimate correlation functions as in Eq. (4.1.4) [68, 69, 70]. The quantum Monte Carlo method for systems at finite temperature relies on a transformation introduced by Suzuki [71] that maps an initial quantum system on a d-dimensional lattice onto a $(d+1)$-dimensional classical system. This conversion then makes it possible to use classical computational sampling techniques to estimate correlation functions as in Eq. (4.1.4). There seem to be (at least) two situations when this approach runs into trouble and no good computational alternatives are available [70]: (1) the correlation functions depend explicitly on time t, and (2) the quantum system is of a fermionic nature. We will give a short explanation of why these problems are encountered.

The transformation from a classical to a quantum system is based on the generalized Trotter formula. Let $H = \sum_{i=1}^{k} H_i$ where each H_i is a Hamiltonian on a small constant Hilbert space. The Trotter formula reads

$$e^{\sigma H} = e^{\sigma(H_1 + H_2 + \ldots + H_k)} = \lim_{n \to \infty} \left(e^{\sigma H_1/n} e^{\sigma H_2/n} \ldots e^{\sigma H_k/n} \right)^n. \tag{4.1.5}$$

The partition function Eq. (4.1.2) (and similarly correlation functions as in Eq. (4.1.4)) can be rewritten, using the Trotter formula and the identity $\sum_m |a_{i,j}^m\rangle\langle a_{i,j}^m| = \mathbf{1}$, where the pair of indices (i,j) labels a choice of basis, as

$$Z = \mathrm{Tr}\, e^{-\beta H} = \sum_{\{a_{i,j}\}} p_{\{a_{i,j}\}}, \tag{4.1.6}$$

where $p_{\{a_{i,j}\}}$ is a distribution over the values of the collection of variables $\{a_{i,j}\}$ and j indexes the repetitions of the factors of Eq. (4.1.5) from 1 to n. If the distribution is nonnegative, we can write $p_{\{a_{i,j}\}} = e^{H_{eff}(\{a_{i,j}\})}$ where H_{eff} is now a classical Hamiltonian given by

$$H_{eff}(\{a_{i,j}\}) = \lim_{n\to\infty} \sum_{j=1}^{n} \sum_{i=1}^{k} \tilde{H}_i(a_{i,j}, a_{i+1,j}). \tag{4.1.7}$$

with $a_{k+1,j} = a_{1,j+1}$, $a_{k+1,n} = a_{1,1}$ and

$$\tilde{H}_i(a,b) = \log(\langle a|\exp(-\beta H_i/n)|b\rangle). \tag{4.1.8}$$

The distribution $p_{\{a_{i,j}\}}$ will only be nonnegative when the matrix elements $\langle a|\exp(-\beta H_i/n)|b\rangle$ are positive. Thus it is important to choose the right sets of basis states $|a_{ij}^m\rangle$ to make the conversion to a classical sampling problem with a nonnegative distribution. There are fermionic systems such as certain Hubbard models [70] in which it does not seem to be possible to choose such a good basis. For these systems it has turned out to be very hard to get good estimates of correlation functions by using classical Monte Carlo techniques. This problem is usually referred to as the "sign" problem.

When we are to compute time-dependent quantities, for example the function $f(it) = \mathrm{Tr}\, e^{iHt} O_1 e^{-iHt} O_2 \rho_\beta$, we need to use an imaginary time $\tau = it$ to perform the conversion of Eq. (4.1.5) to a classical system (we expand e^{iHt} with the Trotter formula). From the classical Monte Carlo sampling of the function $f(\tau)$ for real τ, we estimate $f(\tau)$ and then we could in principle analytically continue this function. However, we only have a finite number of samples of the function and each sample point has some inaccuracy. The errors that are introduced in estimating the Fourier components $\tilde{f}(\omega)$ from this data give rise to large fluctuations when we reconstruct $f(it)$ with the Laplace transform

$$f(it) = \int_{-\infty}^{\infty} d\omega\, e^{-\omega t} \tilde{f}(\omega), \tag{4.1.9}$$

resulting in a bad approximation for the time correlation function $f(it)$.

The relevance of estimating a simple time correlation function (an example of Eq. (4.1.4)) such as

$$\mathrm{Tr}\,[A_t, B_{t'}]\rho_\beta = \langle [A_t, B_{t'}]\rangle_s, \tag{4.1.10}$$

where A and B are some Hermitian Heisenberg operators of the system, cannot be overestimated. Let us recall the many contexts in which Eq. (4.1.10) is used in describing experimental properties of many-particle quantum systems [72]:

When $A = B = u$, where u is the displacement field of a crystal, (4.1.10) describes the phonon dynamics of solids as probed by inelastic neutron scattering. When A and B are the number-density operator, the dielectric susceptibility is represented; this correlation function describes a variety of other experiments, including x-ray photoemission and the so-called x-ray edge singularity. When we study the current-current response function, we obtain the electrical conductivity as described by the Kubo formula. (The density-density and current-current response functions are intimately related via the continuity equation.) Spin-dependent quantities are also of interest: with the spin-spin correlation function, information is obtained about the magnetic susceptibility, and thus the magnon dynamics of ferromagnets and antiferromagnets, the Kondo effect, and the magnetic-dipole channel in neutron scattering. And finally, if A and B involve anomalous pair amplitudes which involve Fermion operators like $a_{\downarrow}(k)a_{\uparrow}(-k)$, the presence and dynamics of a superconducting phase can be probed. In short, the dynamic pair correlation functions provide a window on many of the interesting quantities in experimental physics, and it would be highly desirable to have a method of obtaining estimates for these quantities by simulation on a quantum computer. In this chapter we develop an approach to tackle these problems on a quantum computer. We break the problem into two parts: First, we present an approach to prepare our quantum computer in the equilibrium state ρ_β of a given Hamiltonian (sections 4.2 and 4.3). We will give two alternative routes to prepare an equilibrium state. For the first quantum algorithm we can prove that in the limit of large time and space, the algorithm will successfully produce the equilibrium state as its output. In any realistic situation we are faced with finite resources in time and space. In the sections 4.2.7 and 4.2.8 we therefore present some numerical studies of the performance of the algorithm for small systems. In section 4.3 we present an alternative quantum equilibration algorithm that is based on eigenvalue estimation and the classical Metropolis algorithm. For this algorithm we prove as well that in the limit of large space and time equilibrium is achieved. In section 4.4 we describe a procedure for efficiently estimating quantities as in Eq. (4.1.4) given that the equilibrium state has been prepared. We will not attempt to prove that our algorithms run in polynomial time even for a restricted class of quantum systems given by an Hamiltonian H and/or for restricted ranges of β. The equilibration problem, in its full generality, is expected to be a computationally hard problem. Even classically there is a large class of systems that exhibit a feature called frustration, for which calculating the partition function Z as in Eq. (4.1.2) is a $P^\sharp$-complete problem [73]. Also, for these systems, deciding whether the energy of the ground state is lower than some constant K is an NP-complete problem [74]. The quantum problem has an added difficulty: We cannot assume that we know the eigenvectors (and eigenvalues) of the Hamiltonian of the system that we would like to equilibrate. There is no evidence (yet) that a quantum computer can exponentially outperform a classical computer in estimating the partition function for certain *classical* systems, which would enable us to sample efficiently from the classical Gibbs distribution [75].

The quantum algorithms that we present are hard to simulate on a classical computer. In both of our equilibration algorithms we use the fact that one can implement the unitary time evolution of a local Hamiltonian on n qubits in a polynomial number of steps in n on a quantum computer [32]. A direct simulation of this procedure on a classical computer would cost exponential (in n) space and time and is therefore unrealistic. As we will show in section 4.4, given a preparation of an equilibrium state, there exists an efficient polynomial time procedure on a quantum computer to calculate (time-dependent) correlation functions. As we discussed above, there is no general efficient classical algorithm with which one can estimate time-dependent correlation functions. Our quantum algorithm provides such an algorithm for a quantum computer. Lloyd and Abrams [56] have shown that the unitary simulation of a fermionic system such as the Hubbard model, either in first or second quantization, can be performed efficiently on a quantum computer. The quantum algorithms that we present will use this unitary evolution as a building block. Therefore these algorithms can be used to compute correlation functions for the Hubbard model on a quantum computer. This is a task for which we do not have a good classical algorithm, due to the "sign" problem, as we pointed out earlier.

We focus our efforts on quantum equilibration algorithms for Hamiltonians of which the eigenvalues and eigenvectors are not known beforehand. These are the Hamiltonians of, for example, Heisenberg models (in more than two dimensions), Hubbard models, t-J models, XY models, or many-electron Hamiltonians in quantum chemistry. On the other hand, knowing the eigenvectors and eigenvalues of a Hamiltonian, such as in the Ising model, is no guarantee that there exists an efficient (polynomial time) classical algorithm that produces the equilibrium distribution. The situation is similar for quantum algorithms; we do not know in what cases the equilibration algorithms presented in section 4.2 and 4.3 give rise to a polynomial time algorithm (see also Ref. [76] for quantum algorithms for Ising-type models).

The process of equilibration is also essential in the actual realization of a quantum computer. One of the assumptions underlying the construction of a quantum computer [38] is the ability to put a physical system initially into a known state (or a thermal equilibrium state in the NMR quantum computer [16]), the computational $|00\ldots0\rangle\langle00\ldots0|$ state. The way this is done in an experimental setup is to let this state be the ground state of a natural Hamiltonian and subsequently to cool to low temperature such that the probability of being in this ground state is some constant. This natural Hamiltonian must be sufficiently simple for this equilibration to be achievable efficiently and also be sufficiently weak or tunable not to disturb the computation later on.

Markov chains in the quantum domain

Our two quantum equilibration algorithms are the first examples of the use of quantum Markov chains on a quantum computer. The algorithms that we present are described as the repeated application of a $\mathbf{TCP}$ map $\mathcal{S}$ on some initial state. Such a $\mathbf{TCP}$ map can be viewed as a generalization of a Markov matrix to the quantum domain. A classical Markov chain corresponds to the following process (cf. [68]). Let $i = 1, \ldots, k$ be a set of states. We take time to be discrete variable taking the values $t = 0, 1, \ldots$. At some point in time $t = 0$, we start with a probability distribution $\{p_i \geq 0\}_{i=1}^k$ over the states i such that $\sum_i p_i = 1$. Through a stochastic process which we describe with a Markov matrix M this probability distribution is mapped onto a new probability distribution $\{p'_i\}_{i=1}^k$ at time $t = 1$, i.e.

$$p^T M = p'^T, \tag{4.1.11}$$

where p is the vector of probabilities at $t = 0$. A homogeneous Markov chain corresponds to a chain in which M is the same matrix during all timesteps. In the theory of Markov chains we study the properties of the matrix M determined by its eigenvalues and eigenvectors. Such a Markov process characterized by a matrix M can be viewed as a special kind of $\mathbf{TCP}$ map. The states $i = 1, \ldots, k$ now correspond to a set of orthonormal states $|i\rangle$: $\langle i|j\rangle = \delta_{ij}$. At $t = 0$ we start with a density matrix $\rho = \sum_i p_i |i\rangle\langle i|$. The $\mathbf{TCP}$ map corresponding to a classical Markov process maps

$$\mathcal{S}_M(\rho) = \rho', \tag{4.1.12}$$

where $\rho' = \sum_i p_i' |i\rangle\langle i|$. To give a full specification of $\mathcal{S}$ in terms of M we write

$$\begin{array}{l} \mathcal{S}_M(|i\rangle\langle i|) = \sum_{j=1}^k M_{ij} |j\rangle\langle j|, \\ \mathcal{S}_M(|i\rangle\langle j|) = 0, \quad i \neq j. \end{array} \tag{4.1.13}$$

In this classical chain the density matrices that result from this stochastic process are all diagonal in the same basis $\{|i\rangle\}_{i=1}^k$. For a general quantum Markov chain this will not be the case. In section 4.2.3 we establish several basic properties of $\mathbf{TCP}$ maps that can form the starting point for developing a theory of quantum Markov chains in a quantum computational setting.

4.2 Equilibration I

4.2.1 Introduction

The canonical ensemble is the ensemble of states $\{p_i, |\psi_i\rangle\}$, or a density matrix $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$, such that ρ has a given energy-expectation value

$$\mathrm{Tr}\, H\rho = \langle E\rangle. \tag{4.2.1}$$

The equilibrium state in this ensemble (Eq. (4.1.1)) can be obtained by maximizing the von Neumann entropy of ρ under this energy constraint (see Appendix 4.A). Another way in which the canonical ensemble is defined is by considering the possible states of a system that is in contact with an infinite heat bath at a certain temperature T. The total energy of system and bath is constant, but bath and system exchange energy, so that the system equilibrates. This directly suggests that the way to prepare the equilibrium state on a quantum computer is to mimic this process. In considering the computational complexity of such a procedure in a straightforward simulation without optimization or shortcuts, we will have to include the space and time cost of the bath, which may be large. Also, the intuitive picture of equilibration between a weakly coupled large bath and system does not tell us anything about the rate at which this equilibration occurs. Furthermore, the equilibration process assumes a bath that is already in its equilibrium state. Can we make the bath simple enough that this bath state can be prepared efficiently? In this section we study this process of equilibration. We present an algorithm and we derive expressions that completely characterize the equilibration process in an idealized case: the coupling between the bath and the system is very small, the bath is very large, and the time of interaction is large. In order to treat this problem analytically we develop a perturbation theory in the strength of the coupling between bath and system in section 4.2.4. In this perturbative regime we will see that the dynamics of our quantum Markov chain can be described by a classical Markov chain plus an additional 'dephasing' process. Only in the idealized regime, using the perturbation theory approach, are we able to show that the algorithm gives the equilibrium state as output. We then proceed by a numerical study of the algorithm in realistic cases where the bath is of finite dimension, the strength of the interaction is non-zero, and the interaction time is limited.

4.2.2 The Algorithm

Definition 1 *Equilibration algorithm I.*
Input-parameters:
-H_s, the Hamiltonian of a $N = 2^n$-dimensional quantum system.
-β , the inverse temperature.
-H_b, the Hamiltonian of a $K = 2^k$-dimensional "bath" quantum system.
-λH_{sb}, where H_{sb} is the NK-dimensional "bath-system" interaction Hamiltonian and λ is the parameter that measures the strength of the interaction between bath and system.
-t, the interaction time between bath and system.
-r, the number of times the bath is refreshed in the algorithm.
Define the total Hamiltonian of system and bath as

$$H = H_s \otimes \mathbf{1}_K + \mathbf{1}_N \otimes H_b + \lambda H_{sb}, \tag{4.2.2}$$

and the trace-preserving completely positive ($\mathbf{TCP}$) map $\mathcal{S}_{\lambda,t}$ as

$$\mathcal{S}_{\lambda,t}(\rho) \equiv \mathrm{Tr}_b\, e^{iHt} \rho \otimes \rho_{b,\beta}\, e^{-iHt}. \tag{4.2.3}$$

1. **Prepare system**. *We prepare the n qubits in the state $|00\ldots 0\rangle\langle 00\ldots 0|$.*

2. **Prepare bath**. *We prepare the k qubits of the bath in their equilibrium state $\rho_{b,\beta}$ of H_b.*

3. **Evolve** *system and bath for time t and discard bath, that is, perform the superoperator $\mathcal{S}_{\lambda,t}$ of Eq. (4.2.3).*

4. **Repeat** *steps 2 and 3 r times such that*

$$\| \mathcal{S}^{r+1}_{\lambda,t}(|00\ldots 0\rangle\langle 00\ldots 0|) - \mathcal{S}^{r}_{\lambda,t}(|00\ldots 0\rangle\langle 00\ldots 0|) \|_{tr} \leq \epsilon, \tag{4.2.4}$$

for all $r \geq r_0$ and ϵ is some accuracy. See Appendix 4.D for the definition of $\| \, . \, \|_{tr}$.

We put several constraints on H_s, H_b, and H_{sb}. We will use local Hilbert spaces as in Eq. (4.1.3) of dimension 2 (qubits). H_s must be a "local" Hamiltonian. We define a d-local Hamiltonian on n qubits as one that can be expressed as

$$H_s = \sum_{i=1}^{\mathrm{Poly}(n)} \mathbf{1}_{N/c} \otimes h_i, \tag{4.2.5}$$

where each h_i operates on a tensor product of several small qubit Hilbert spaces, whose total dimension is d. We will also assume that the eigenvalues of H_s are all distinct; the spectrum is non-degenerate. This will simplify the upcoming analysis. In order to treat Hamiltonians with degenerate spectra a change in the perturbation theory of Section 4.2.4 will have to be made. We expect however that with that change the main result of Section 4.2.5, namely successful equilibration in the idealized case, will still hold. H_{sb} has the linear coupling form

$$H_{sb} = S \otimes B, \tag{4.2.6}$$

where both $S \in B(\mathcal{H}_s)$ and $B \in B(\mathcal{H}_b)$ are local Hamiltonians. H_b is the Hamiltonian of a system of non-interacting qubits, i.e., it is a sum of single-qubit Hamiltonians:

$$H_b = \sum_{i=1}^{k} \mathbf{1}_{K/2} \otimes h_i. \tag{4.2.7}$$

The bath's equilibrium state factorizes into a tensor product of qubit equilibrium states associated with each h_i:

$$\rho_{b,\beta} = \rho^1_{b,\beta} \otimes \ldots \otimes \rho^k_{b,\beta}. \tag{4.2.8}$$

This enables us to prepare the bath (step 2) efficiently. Appendix 4.E shows that it will cost $2k$ elementary qubit operations to perform step 2. The locality of H_s, H_b, and H_{sb} is required in order to be able to simulate the unitary time evolution e^{iHt} in time $O(t^2/\delta)$ where δ is the accuracy with which the evolution is implemented [32, 77] (see also Appendix 4.B).

We also choose

$$\langle B \rangle_b \equiv \mathrm{Tr}\, B\rho_{b,\beta} = 0. \tag{4.2.9}$$

To understand the effect of a non-zero $\langle B \rangle_b$ we rewrite H as

$$H = (H_s + \lambda \langle B \rangle_b S) \otimes \mathbf{1}_K + \mathbf{1}_N \otimes H_b + \lambda S \otimes B', \tag{4.2.10}$$

with $B' = B - \langle B \rangle \mathbf{1}_K$ and thus $\langle B' \rangle_b = 0$. Choosing a non-zero $\langle B \rangle_b$ effectively corresponds to a change in the Hamiltonian of the system. The goal is for the bath to exchange energy with the system and not to disturb it otherwise. Choosing a non-zero $\langle B \rangle_b$ will give rise to an undesired shift in the energy spectrum of the system.

We now discuss the last step of the algorithm, step 4. When the superoperator $\mathcal{S}_{\lambda,t}$ has the equilibrium state $\rho_{s,\beta}$ as its unique fixed point, then Eq. (4.2.4) for all $r \geq r_0$ implies that

$$\| \, \mathcal{S}^r_{\lambda,t}(|00\ldots 0\rangle\langle 00\ldots 0|) - \rho_{s,\beta} \, \|_{tr} \leq C_N f(\epsilon, \kappa). \tag{4.2.11}$$

for all $r \geq r_0$, where κ is the absolute value of the second largest eigenvalue of $\mathcal{S}_{\lambda,t}$. C_N is some constant polynomial in N. The function $f(\epsilon, \kappa) \to 0$ for $\epsilon \to 0$ and $\kappa \neq 1$. The functional dependence of f on κ is such that when κ increases, –the equilibration slows down–, f increases. For sufficiently small ϵ the equilibration process will lead to successful convergence to the equilibrium state.

There does however not exist a straightforward implementation of step 4. The first problem is that we would have to check the closeness of the rth and the $(r+1)$th iteration of $\mathcal{S}_{\lambda,t}$ for all $r \geq r_0$. In practice this has to be replaced with choosing a finite set of iterations r for which the invariance of $\mathcal{S}^r_{\lambda,t}(|00\ldots0\rangle\langle00\ldots0|)$ is tested. This problem is also encountered in classical Monte Carlo simulations. The second problem, which is a purely quantum phenomenon, is that by measuring $\rho_r \equiv \mathcal{S}^r_{\lambda,t}(\rho)$ we might disturb ρ_r. Thus to compare ρ_r with ρ_{r+1} we would have to run $\mathcal{S}$ again for $r+1$ times. To assemble some statistics on the difference between ρ_r and ρ_{r+1} we have to run r iterations of $\mathcal{S}$ several times. These considerations about the verification of the convergence of the equilibration process are of course not special to the use of a quantum computer; they are the same as in the equilibration of a quantum physical system in an experimental setup. Furthermore, it would be an impractical task to try to measure all the matrix elements of ρ_r; namely ρ_r contains an exponential amount of data of which we can extract only a polynomial amount by measurement in polynomial time. The best way to proceed is the same as what one does in classical Monte Carlo simulations [70]. If the goal of the computation is to estimate $\mathrm{Tr}\, O\rho_{s,\beta}$ then one computes the datapoints

$$\mathcal{O}_r = \mathrm{Tr}\, O\rho_r, \tag{4.2.12}$$

until $|\mathcal{O}_r - \mathcal{O}_{r+1}| \leq \epsilon$ for a sufficiently large set of iterations $r \geq r_0$. The same procedure can be carried out when the goal of the equilibration is to compute a time-dependent correlation function such as Eq. (4.1.4).

In the remainder of this section we will analyse this algorithm. In section 4.2.3 we give some general properties of $\mathbf{TCP}$ maps. In section 4.2.4 we discuss the non-Hermitian perturbation theory that will be the basis of the analysis of $\mathcal{S}_{\lambda,t}$ in the idealized case. In section 4.2.5 we derive explicit expressions for the idealized case. The idealized case is the case obtained by taking the limits $k \to \infty$, $\lambda \to 0$, and $t \to \infty$, but $\lambda^2 t$ is constant [2]. Then we can show that in this idealized case the process has a unique fixed point which is the equilibrium state. Many parameters in the quantum algorithm that is given in Definition 1 are not yet fixed. These are the time t, the number of repetitions r, the size of the bath k, the Hamitonian of the bath H_b, the interaction Hamiltonian $S \otimes B$ and the interaction strength λ. In an actual realization of the quantum algorithm on a quantum computer with finite resources of time and space we should try to chose the optimal set of parameters. In sections 4.2.7 and 4.2.8 we present results from numerical simulations for small quantum systems that tell us how the algorithm depends on this set of parameters. In these sections the following questions will be addressed:
1. What is the influence of different choices for H_b, S and B (Eqs. (4.2.6), (4.2.7))?

[2]This limit is sometimes referred to as the Van Hove limit [78].

2. How do the parameters r, λ, and t required for successful equilibration depend on n generically? How does k, the number of bath qubits depend on n, the number of system qubits for successful equilibration? Are they polynomially related?

The dynamics of open quantum systems, like the system in our algorithm that interacts with a bath, are most often studied with the use of a generalized master equation [79]. The exact master equation in integral form describes the time evolution of $\rho(t) = \mathcal{S}_{\lambda,t}(\rho)$ of Eq. (4.2.3):

$$\rho(t) = e^{-i\mathcal{L}_s t}\rho(0) - \lambda^2 \int_0^t dt' \int_0^{t'} dt'' e^{-i\mathcal{L}_s(t-t')}\mathcal{M}(t',t'')\rho(t''). \tag{4.2.13}$$

where $\mathcal{L}$, the Liouvillian, is defined as

$$\mathcal{L}(\rho) = [H, \rho]. \tag{4.2.14}$$

so that $\mathcal{L}_s(\rho) = [H_s, \rho]$ etc. The operator $\mathcal{M}(t',t'')$ is the "memory kernel",

$$\mathcal{M}(t',t'') = \mathrm{Tr}_b\, \mathcal{L}_{sb}\, e^{-i(1-\rho_b \mathrm{Tr}_b)\mathcal{L}(t'-t'')}\mathcal{L}_{sb}\, \rho_b. \tag{4.2.15}$$

The form in which the master equation is most often used, however, is one in which two simplifying approximations are made: (1) the Born approximation. This relates to the weakness of the interaction parameter λ. (2) the Markov approximation. The process described by $\mathcal{S}_{\lambda,t}$ is Markovian if we can write

$$\mathcal{S}_{\lambda,t}(\mathcal{S}_{\lambda,s}(\rho)) = \mathcal{S}_{\lambda,t+s}(\rho), \tag{4.2.16}$$

for all $t \geq 0$ and $s \geq 0$. Note that the difference between the left- and the righthand side of this equation is the following. On the lefthand side the environment $\rho_{b,\beta}$ (see the definition of $\mathcal{S}_{\lambda,t}$ in Eq. (4.2.3)), is refreshed after time t, whereas on the righthand side, the environment is kept for the whole evolution time $t+s$.

Markovian behavior typically occurs when the rate at which the effect of the system on the bath is erased in the bath (in the sense of being spread throughout the bath) is much faster than the rate at which the system evolves; the system sees a "fresh" bath every time. In our algorithm this loss of correlations in the bath is enforced when after some time t the bath is replaced by a new bath (step 4). We would not be able to truly equilibrate a finite system with a finite-dimensional bath if we would not keep refreshing it. Since the global evolution of bath and system is unitary, eventually we will get back to the initial unentangled state and, after tracing over the bath, to the initial state of the system (a so-called Poincaré recurrence). Whether Markovian dynamics is justified will depend on the size of the bath, the strength of the interaction and the length of the interaction time. There are ways to make a simple but naive Markov approximation in Eq. (4.2.13) that lead to a master equation that fails to describe $\mathbf{TCP}$ dynamics [80, 81]. The form of the master equation that does incorporate both the approximations and yields a physical completely positive map is the master equation in Lindblad form [82]:

$$\frac{\partial \rho}{\partial t} = -i[H_s, \rho(t)] + L\rho(t), \tag{4.2.17}$$

where L [83, 81] can be expressed with a basis of operators F_i as

$$L\rho(t) = \frac{1}{2}\sum_{k,l=1}^{N^2-1} a_{kl}([F_k\rho(t), F_l^\dagger] + [F_k, \rho(t)F_l^\dagger]), \tag{4.2.18}$$

where a_{kl} is a positive semi-definite matrix. In a Lindblad equation describing the equilibration process, we expect L to depend on the system Hamiltonian H_s. The equilibrium state $\rho_{s,\beta}$ – if the algorithm is successful– should be a stationary state of the process, which implies that since $[H_s, \rho_{s,\beta}] = 0$, we must have that

$$L\rho_{s,\beta} = 0. \tag{4.2.19}$$

Davies [83, 84, 85] has demonstrated that a process described by $\mathcal{S}_{\lambda,t}$ where the bath is an infinite-dimensional quantum system (for example a quantum field) *does* equilibrate any quantum system in the limit where $\lambda \to 0, t \to \infty$, but $\lambda^2 t$ stays constant. By carefully taking a Born and Markov approximation, he derives a Lindblad equation of the form such that Eq. (4.2.19) is obeyed. We will perform a similar analysis here. The main point of difference is that we use a perturbative analysis of the dynamics which is only valid for small $\lambda^2 t$, but coincides in this regime with Davies' result. We furthermore obtain more explicit expressions for the dynamics in this limit.

One can write down the most general form of an operator L that obeys a quantum detailed balance condition [86], a stronger requirement than the stationarity of Eq. (4.2.19). Now, one might ask the following question: Could we implement the corresponding superoperator directly, without the use of a weakly coupled large bath, so as to save us time and space? We believe the answer is no, as L will depend on the eigenvectors and eigenvalues of H_s, which we do not know beforehand.

4.2.3 Some Useful Properties of TCP Maps

In this section, we study some essential properties of the superoperator $\mathcal{S}_{\lambda,t}$ defined as in Eq. (4.2.3). This superoperator is a $\mathbf{TCP}$ map

$$\mathcal{S}_{\lambda,t} : \; B(\mathcal{H}_N) \to B(\mathcal{H}_N), \tag{4.2.20}$$

where $B(\mathcal{H}_N)$ is the algebra of linear operators on a Hilbert space $\mathcal{H}_N$. The set $\mathbf{TCP}[N,N]$ is the set of $\mathbf{TCP}$ maps $\mathcal{S}\colon\ B(\mathcal{H}_N) \to B(\mathcal{H}_N)$. The elements of $B(\mathcal{H}_N)$ can be represented as $N \times N$ matrices. An alternative and convenient way to represent $B(\mathcal{H}_N)$ is as a N^2-dimensional complex vector space $\mathbf{C}^{N^2}$

$$I : \; \chi \in B(\mathcal{H}_N) \to (\chi)_{ij} \in \mathbf{C}^{N^2}. \tag{4.2.21}$$

This representation leads to a matrix representation of a $\mathbf{TCP}$ map $\mathcal{S}$ on $\mathbf{C}^{N^2}$. Let A_i be the operation elements of $\mathcal{S}$, i.e.

$$\chi' = \mathcal{S}(\chi) = \sum_i A_i \chi A_i^\dagger, \quad \sum_i A_i^\dagger A_i = \mathbf{1}_N. \tag{4.2.22}$$

Then

$$(\chi')_{mn} = (\mathcal{S}(\chi))_{mn} = \sum_i \sum_{k,l} (A_i)_{mk} (\chi)_{kl} (A_i^\dagger)_{ln} = \sum_{k,l} \mathcal{S}_{mn,kl} (\chi)_{kl}, \tag{4.2.23}$$

with

$$\mathcal{S}_{mn,kl} = \sum_i (A_i)_{mk} (A_i^\dagger)_{ln}. \tag{4.2.24}$$

One can then study the eigenvectors and eigenvalues of the matrix representation of a $\mathbf{TCP}$ map. First, we will give three useful properties of $\mathbf{TCP}$ maps that follow directly from their definition:

Property 1 *Let $B(\mathcal{H}_N)^+ \in B(\mathcal{H}_N)$ be the set of positive semi-definite matrices. Let $\mathcal{S} \in \mathbf{TCP}[N,N]$. Then*

$$\rho \in B(\mathcal{H}_N)^+ \Rightarrow \mathcal{S}(\rho) \in B(\mathcal{H}_N)^+, \tag{4.2.25}$$

as $\mathcal{S}$ is (completely) positive. Let χ be an eigenvector of $\mathcal{S}$ with eigenvalue μ, $\mathcal{S}(\chi) = \mu\chi$. We have

$$\mathrm{Tr}\,\chi \neq 0 \Rightarrow \mu = 1, \tag{4.2.26}$$

as $\mathcal{S}$ is trace-preserving. Let A_i be the operation elements in the decomposition of $\mathcal{S}$ as in Eq. (4.2.22). If χ is an eigenvector of $\mathcal{S}$ with eigenvalue μ, then $\chi^\dagger$ is also an eigenvector of $\mathcal{S}$ with eigenvalue μ^*. This follows from

$$(\mathcal{S}(\chi))^\dagger = \sum_i (A_i \chi A_i^\dagger)^\dagger = \mathcal{S}(\chi^\dagger). \tag{4.2.27}$$

Let $B_1(\mathcal{H}_N)^+$ be the set of positive semi-definite matrices on $\mathcal{H}_N$ that have trace 1, i.e. the density matrices. Thus Property 1 implies that if a *density matrix* ρ is an eigenvector of the superoperator, it must have eigenvalue 1, that is, it is a fixed point of the map. On the basis of the $\mathbf{TCP}$ property of a map $\mathcal{S}$, we can also show the following

Proposition 1 *Let $\mathcal{S} \in \mathbf{TCP}[N,N]$. All eigenvalues μ of $\mathcal{S}$ have $|\mu| \leq 1$.*

Proof (by contradiction): Assume χ is an eigenvector of $\mathcal{S}$ with eigenvalue $|\mu| > 1$. Note that Property 1 implies that χ has $\mathrm{Tr}\,\chi = 0$. If χ is Hermitian, μ will be real. As χ is traceless, it must have at least one negative eigenvalue. One can always find a density matrix ρ and a small enough ϵ such that $\rho' = \rho + \epsilon\chi$ is still a density matrix. Let $\mathcal{S}$ operate r times on this density matrix. For large enough r the result $\mathcal{S}^r(\rho + \epsilon\chi) = \mathcal{S}^r(\rho) + \epsilon\mu^r\chi$ will no longer be a positive semi-definite matrix: take the eigenvector $|\psi\rangle$ of χ corresponding to the lowest (negative) eigenvalue $\lambda_{\min}$. Then

$$\langle\psi|\mathcal{S}^r(\rho)|\psi\rangle + \epsilon\mu^r\langle\psi|\chi|\psi\rangle \leq 1 + \epsilon\mu^r\lambda_{\min}, \tag{4.2.28}$$

will become negative for large enough r. But Property 1 implies that $\mathcal{S}^r(\rho')$ is a density matrix, thus $|\mu|$ cannot be larger than 1. When χ is non-Hermitian, we reason similarly. One can find a density matrix ρ and a small enough ϵ such that $\rho' = \rho + \epsilon(\chi + \chi^\dagger)$ is a density matrix. Let $\mathcal{S}(\chi) = \mu\chi = |\mu|e^{i\phi}\chi$. Let $\lambda_{\mathrm{min,r}}$ be the smallest (and negative) eigenvalue of the traceless Hermitian matrix $e^{i\phi r}\chi + e^{-i\phi r}\chi^\dagger$ and let $|\psi\rangle$ be the corresponding eigenvector. Then

$$\begin{aligned}&\langle\psi|\mathcal{S}^r(\rho')|\psi\rangle = \\ &\qquad \langle\psi|\mathcal{S}^r(\rho)|\psi\rangle + \epsilon|\mu|^r\langle\psi|(e^{i\phi r}\chi + e^{-i\phi r}\chi^\dagger)|\psi\rangle \leq 1 + \epsilon|\mu|^r\lambda_{\mathrm{min,r}},\end{aligned} \tag{4.2.29}$$

will become negative for some large r ($\lambda_{\mathrm{min,r}}$ is a quasi-periodic function of r so it cannot be small for all large r). □

Another property about the existence of fixed points can be derived:

Proposition 2 *Let $\mathcal{S} \in \mathbf{TCP}[N, N]$. $\mathcal{S}$ has a fixed point (which is a density matrix).*

Proof The set of density matrices $B_1(\mathcal{H}_N)^+ \in B(\mathcal{H}_N)$ is convex and compact. $\mathcal{S}$ is a linear continuous map and $\mathcal{S}(\rho \in B_1(\mathcal{H}_N)^+) \in B_1(\mathcal{H}_N)^+$. Then the Markov-Kakutani Theorem V.10.6 of [87] applies. □

The existence of a fixed point does not by itself guarantee that the process described by $\mathcal{S}$ is "relaxing", that is $\lim_{r\to\infty}\mathcal{S}^r(\rho) = \rho_0$ for all density matrices ρ where ρ_0 is the fixed point. The existence of such a limit depends on whether the fixed point is unique. The following Proposition proves that when there is unique fixed point, relaxation will occur and the relaxation rate is determined by the second largest eigenvalue of $\mathcal{S}$ [88]:

Proposition 3 *Let $\rho_0 \in B_1(\mathcal{H}_N)^+$ be the unique fixed point of a $\mathbf{TCP}$ map $\mathcal{S}$. Let $\kappa = \max_{m|\mu_m\neq 1}|\mu_m|$, the absolute value of the second largest eigenvalue of $\mathcal{S}$. Then for all density matrices ρ we have*

$$\| \mathcal{S}^r(\rho) - \rho_0 \|_{tr} \leq C_N \mathrm{Poly}(r)\kappa^r. \tag{4.2.30}$$

where C_N is a constant depending on the dimension N of the system and $\mathrm{Poly}(r)$ denotes some polynomial in r. Thus for all density matrices ρ

$$\lim_{r\to\infty} \| \mathcal{S}^r(\rho) - \rho_0 \|_{tr} = 0. \tag{4.2.31}$$

Proof Let μ_i be the eigenvalues of $\mathcal{S}$. Let s be the number of distinct eigenvalues. We can bring any matrix $\mathcal{S}$ into Jordan form J by a similarity transformation M [90]:

$$\mathcal{S} = MJM^{-1}, \tag{4.2.32}$$

where

$$J = \sum_{i=1}^{s} (\mu_i P_i + N_i). \tag{4.2.33}$$

P_i are orthogonal projectors and N_i is a matrix of 1s above the diagonal in the ith block or N_i is the null matrix. When the eigenvalue μ_i is nondegenerate N_i is the null matrix. We therefore have $N_i N_j = 0$ for $i \neq j$ and $P_i N_j = 0$ for $i \neq j$. Call the unique largest eigenvalue $\mu_0 = 1$ and the corresponding projection P_0. As in Eq. (4.2.32) one can write

$$\mathcal{S}^r = MJ^r M^{-1}. \tag{4.2.34}$$

where J^r equals

$$J^r = \sum_{i=1}^{s} (\mu_i^r P_i + N_i') \tag{4.2.35}$$

where N_i' is a nilpotent matrix in the ith block. Note that N_0' is 0 as μ_0 is unique. Let $\mathcal{S}^0$ be MP_0M^{-1} or $\mathcal{S}^0(\rho) = \rho_0$. We use $\parallel A \parallel_{tr} \leq \sqrt{N} \parallel A \parallel_2$. Note that $\parallel A \parallel_2$ refers to the Euclidean norm of A represented as a vector. This follows from $(\sum_{i=1}^N |x_i|)^2 \leq N \sum_{i=1}^N |x_i|^2$ for complex numbers x_i. We have first of all

$$\parallel \mathcal{S}^r(\rho) - \rho_0 \parallel_{tr} \leq \sqrt{N} \parallel (\mathcal{S}^r - \mathcal{S}^0)(\rho) \parallel_2 . \tag{4.2.36}$$

This expression can be bounded with the use of the similarity transformation M to

$$\parallel \mathcal{S}^r(\rho) - \rho_0 \parallel_{tr} \leq \sqrt{N} \,|||M(\mathcal{S}^r - \mathcal{S}^0)M^{-1}|||_2 \leq C_{1,N} \,|||J^r - P_0|||_2 \tag{4.2.37}$$

where $|||.|||_2$ is defined in Appendix 4.D and we use $\parallel \rho \parallel_2 = \mathrm{Tr}\,\rho^2 \leq 1$ for density matrices. Using the expression for J^r, Eq. (4.2.35), we can also bound

$$|||J^r - P_0|||_2 \leq r^N C_{2,N} \kappa^r. \tag{4.2.38}$$

Combining Eq. (4.2.37) and Eq. (4.2.38) gives us the desired result Eq. (4.2.30). Eq. (4.2.31) then follows as $\kappa < 1$ by Proposition 1. If $\mathcal{S}$ is diagonalizable, the nilpotents N_i in expression Eq. (4.2.35) are not present. By going through the same steps, a bound as in Eq. (4.2.30) can be derived without the factor $\mathrm{Poly}(r)$. □

We refer the reader to [81] for discussions and references concerning the existence of a unique fixed point and other properties of relaxation for a process that is described by a Lindblad equation, Eq. (4.2.17).

Finally we give a result which relates members of $\mathbf{TCP}[N, N]$ to the stochastic matrices. A real matrix M is stochastic when the entries of its columns add up to 1, i.e. $\sum_i M_{ij} = 1$.

Proposition 4 *Let $\mathcal{S} \in \mathbf{TCP}[N,N]$. $\mathcal{S}_{mm,nn} \in \mathbf{R}^+$, and, $\forall n$, $\sum_m \mathcal{S}_{mm,nn} = 1$; that is, the elements $\mathcal{S}_{mm,nn}$ form an $N \times N$ stochastic matrix in the indices m and n. Also, $\forall n,k$, $n \neq k$, $\sum_m \mathcal{S}_{mm,nk} = 0$.*

Proof $\mathcal{S}_{mm,nn} \in \mathbf{R}^+$ follows directly from Eq. (4.2.24). For the rest, we impose the unit trace condition on Eq. (4.2.23):

$$1 = \sum_{m,k,l} \mathcal{S}_{mm,kl}\rho_{kl}. \tag{4.2.39}$$

This must be true for all density matrices represented by ρ. Taking $\rho_{kl} = \delta_{k,l}\delta_{k,k_0}$ gives the desired result

$$1 = \sum_{m} \mathcal{S}_{mm,k_0k_0}. \tag{4.2.40}$$

We now separate Eq. (4.2.39) into diagonal and off-diagonal parts, using the Hermiticity of the density matrix ρ:

$$\begin{aligned}1 = \sum_{m,k} \mathcal{S}_{mm,kk}\rho_{kk} + \sum_{m,k,l}^{k>l} (\mathcal{S}_{mm,kl} + \mathcal{S}_{mm,lk})\mathrm{Re}(\rho_{kl})\\ +i\sum_{m,k,l}^{k>l} (\mathcal{S}_{mm,kl} - \mathcal{S}_{mm,lk})\mathrm{Im}(\rho_{kl}).\end{aligned} \tag{4.2.41}$$

The first term of Eq. (4.2.41) is always 1, because of Eq. (4.2.40). If we require Eq. (4.2.41) when the off-diagonal terms in ρ are $\rho_{kl} = \delta_{k,k_0}\delta_{l,l_0}$ $(k > l)$, we obtain

$$\sum_m (\mathcal{S}_{mm,k_0l_0} + \mathcal{S}_{mm,l_0k_0}) = 0, \tag{4.2.42}$$

and setting the off-diagonal terms in ρ to $\rho_{kl} = i\delta_{k,k_0}\delta_{l,l_0}$ $(k > l)$ gives

$$\sum_m (\mathcal{S}_{mm,k_0l_0} - \mathcal{S}_{mm,l_0k_0}) = 0, \tag{4.2.43}$$

Adding these equations, we obtain the desired result

$$\sum_m \mathcal{S}_{mm,k_0l_0} = 0, \quad k_0 \neq l_0. \tag{4.2.44}$$

□

4.2.4 Perturbation Theory

In this section we develop a perturbation theory in the coupling λ for the superoperator $\mathcal{S}_{\lambda,t}$. The calculation will assume the diagonalizability of $\mathcal{S}_{\lambda,t}$. If all the eigenvalues of a matrix M are distinct, M is diagonalizable [90]. In many cases of interest for equilibration, this assumption for $\mathcal{S}_{\lambda,t}$ will be correct.

One can formally expand the superoperator $\mathcal{S}_{\lambda,t}$ as a power series in the coupling parameter λ,

$$\mathcal{S}_{\lambda,t} = \mathcal{S}_t^{(0)} + \lambda\mathcal{S}_t^{(1)} + \lambda^2\mathcal{S}_t^{(2)} + \lambda^3\mathcal{S}_t^{(3)} + \dots . \tag{4.2.45}$$

In section 4.2.5 we will explicitly calculate the expressions for these expansion operators. We will show (Eqs. (4.2.68)-(4.2.71)) that condition Eq. (4.2.9) implies that $\mathcal{S}_t^{(1)}$ is zero for all t. On the basis of this expansion, we will make a perturbative expansion of the eigenvalues and eigenvectors of $\mathcal{S}_{\lambda,t}$

$$\mu = \mu^{(0)} + \lambda\mu^{(1)} + \lambda^2\mu^{(2)} + \dots , \tag{4.2.46}$$

$$\chi = \chi^{(0)} + \lambda\chi^{(1)} + \lambda^2\chi^{(2)} + \dots . \tag{4.2.47}$$

Assuming that the perturbation expansion exists for this non-Hermitian operator, it will have the same structure as in the well established procedures familiar in quantum theory for bounded Hermitian operators (see textbooks on quantum mechanics such as [89] or [90] for a more mathematical background).

In the representation of Eq. (4.2.24) $\mathcal{S}_t^{(0)}$ reads

$$(\mathcal{S}_t^{(0)})_{mn,kl} = (U^t)_{mk}(U^{t^\dagger})_{ln}, \tag{4.2.48}$$

where $U = e^{iH_s}$. Unitarity of $\mathcal{S}_t^{(0)}$, as a *matrix* operator on vectors in $\mathbf{C}^{N^2}$, follows from

$$\sum_{k,l}(\mathcal{S}_t^{(0)})_{mn,kl}(\mathcal{S}_t^{(0)^\dagger})_{kl,ij} = \sum_{k,l}(U^t)_{mk}(U^{t^\dagger})_{ln}(U^t)_{jl}(U^{t^\dagger})_{ki} = \delta_{mi}\delta_{jn}. \tag{4.2.49}$$

The eigenbasis of $\mathcal{S}_t^{(0)}$ is formed by the set of matrices $|n\rangle\langle m|$ where $|n\rangle$ are the eigenvectors of H_s. These eigenvectors come with eigenvalues $\mu_{nm,t}^{(0)}$:

$$\{|n\rangle\langle m|, \mu_{nm,t}^{(0)} = e^{it(E_n - E_m)}\}_{n,m=1}^{N,N}, \tag{4.2.50}$$

where E_n are the eigenvalues of H_s. Thus all density matrices of the form $|n\rangle\langle n|$, and mixtures of these, have degenerate eigenvalues $\mu_{nn,t}^{(0)} = 1$. If the spectrum of H_s is non-degenerate (we assumed this in section 4.2.2), then all other eigenvectors $|n\rangle\langle m|$ for $n \neq m$ have non-degenerate eigenvalues. These eigenvectors $|n\rangle\langle m|$ form an orthonormal set with respect to the vector inner product on $\mathbf{C}^{N^2}$,

$$\mathrm{Tr}\,(|n\rangle\langle m|)^\dagger|k\rangle\langle l| = \delta_{nk}\delta_{ml}. \tag{4.2.51}$$

To carry out the perturbation theory, we switch to a ket notation for the density operators and a matrix notation for the superoperators. This will make it easier for us to perform the necessary manipulations of degenerate perturbation theory, in which the degenerate sector is isolated and a diagonalization performed within it.

We first organize the diagonal, degenerate part of this vector space to be indexed. To be specific, we introduce an orthogonal basis in this vector space such that

$$|\phi_i^{(0)}\rangle = |i\rangle\langle i|, \quad 1 \le i \le N, \tag{4.2.52}$$

$$|\phi_{i(m,n)}^{(0)}\rangle = |m\rangle\langle n|, \quad 1 \le m, n \le N, \quad m \ne n. \tag{4.2.53}$$

In the second equation the indexing i can be made consecutive by choosing

$$\begin{aligned} i(m,n) &= nN + m - \tfrac{1}{2}n(n+1), \quad m > n, \\ i(m,n) &= \tfrac{1}{2}N(N-1) + mN + n - \tfrac{1}{2}m(m+1), \quad n > m. \end{aligned} \tag{4.2.54}$$

This organizes this new vector space into a direct-sum form $\mathbf{C}^{N^2} = \mathbf{C}_D \oplus \mathbf{C}_{ND}$, where "$D$" and "$ND$" stand for diagonal and nondiagonal (or, degenerate and nondegenerate). $\mathbf{C}_D$ has dimension N and $\mathbf{C}_{ND}$ has dimension $N^2 - N$.

From the discussion above, we note that the degeneracy is lifted in lowest order by the second-order part of the superoperator $\mathcal{S}$ in the D sector, which we will denote $\mathcal{S}_{D,D}^{(2)}$. Assume that $\mathcal{S}_{D,D}^{(2)}$ is diagonalizable via the similarity transformation

$$M\mathcal{S}_{D,D}^{(2)}M^{-1} = \tilde{\mathcal{S}}_{D,D}^{(2)}, \tag{4.2.55}$$

where $\tilde{\mathcal{S}}_{D,D}^{(2)}$ is a diagonal matrix (the tilde will denote quantities expressed in the new basis $M_D \oplus \mathbf{1}_{ND}|\phi^{(0)}\rangle$, which is in general non-orthogonal). In this new basis, the degeneracy of the diagonal terms of $\mathcal{S}$ is lifted to second order in λ (the diagonal terms can be written to second order as $\mu_i = 1 + \lambda^2\tilde{\mathcal{S}}_{ii}^{(2)}$), and since the largest off-diagonal terms in the D sector are now third order, given by

$$\lambda^3 M\mathcal{S}_{D,D}^{(3)}M^{-1} = \lambda^3\tilde{\mathcal{S}}_{D,D}^{(3)}, \tag{4.2.56}$$

the condition for the successful application of non-degenerate perturbation theory is now satisfied, assuming that no additional, accidental degeneracy occurs. (The condition is satisfied from the start in the ND sector.) Its form is essentially no different from the conventional perturbation expansion [89] for Hermitian operators. This expansion for the eigenvalues is

$$\mu_i = \mu_i^{(0)} + \lambda^2\tilde{\mathcal{S}}_{ii}^{(2)} + \mathcal{O}(\lambda^3). \tag{4.2.57}$$

The form of this expansion is different depending on whether $i \in D$ or $i \in ND$, but only at $\mathcal{O}(\lambda^4)$ [3]. The perturbation expansions for the eigenvectors are

$$|\phi_i\rangle = |\tilde{\phi}_i^{(0)}\rangle + \lambda \sum_{j\in D,\ j\ne i} |\tilde{\phi}_j^{(0)}\rangle \frac{\tilde{\mathcal{S}}_{ji}^{(3)}}{\tilde{\mathcal{S}}_{ii}^{(2)} - \tilde{\mathcal{S}}_{jj}^{(2)}} + \mathcal{O}(\lambda^2), \quad i \in D, \tag{4.2.58}$$

$$|\phi_i\rangle = |\phi_i^{(0)}\rangle + \lambda^2 \sum_{j\ne i} |\tilde{\phi}_j^{(0)}\rangle \frac{\tilde{\mathcal{S}}_{ji}^{(2)}}{\mu_i^{(0)} - \mu_j^{(0)}} + \mathcal{O}(\lambda^3), \quad i \in ND. \tag{4.2.59}$$

[3] The notation $f(x) = \mathcal{O}(g(x))$ is equivalent with $f(1/y) = O(g(1/y))$.

This expansion indicates that there is no mixing between the D and ND sectors until second order in λ. This expansion strategy will be taken up again in the numerical simulations, Sec. 4.2.7 (Eq. (4.2.115)).

This perturbation analysis shows that the superoperator of Eq. (4.2.45) can be approximated by a simple one, for which the approximate eigenvectors are correct to zeroth order in λ, and the eigenvalues are correct to the next non-vanishing order (λ^2). In this approximation the D and ND sectors are completely decoupled. In the D sector the superoperator is written

$$(\mathcal{S}_{\lambda,t}(\rho))_{nn} \approx \sum_m P_{nm,t}\,\rho_{mm}, \tag{4.2.60}$$

$$P_{nm,t} = \delta_{nm} + \lambda^2(\mathcal{S}_t^{(2)})_{nn,mm}. \tag{4.2.61}$$

Note from Proposition 4 that $P_{nm,t}$ is exactly a stochastic matrix; therefore the approximate dynamics in the D sector is that of a classical Markov process. The approximate dynamics in the ND sector is diagonal in the eigenbasis:

$$(\mathcal{S}_{\lambda,t}(\rho))_{nm} \approx \mu_{nm,t}\,\rho_{nm} = \mu_{nm,t}^{(0)}\,\rho_{nm} + \lambda^2(\mathcal{S}_t^{(2)})_{nm,nm}\rho_{nm}, \quad n \neq m. \tag{4.2.62}$$

So, the full expression for the approximate superoperator is:

$$(\mathcal{S}_{\lambda,t})_{nm,kl} \approx P_{nk,t}\,\delta_{nm}\delta_{kl} + \mu_{nm,t}(1-\delta_{nm})\delta_{nk}\delta_{ml}. \tag{4.2.63}$$

The simplications of Eqs. (4.2.60) and (4.2.62) make it possible to answer questions about the uniqueness of the fixed point and, in principle, the mixing properties of a repeated application of $\mathcal{S}_{\lambda,t}$, using techniques from classical Markov chains [91]. The splitting in two sectors, each having its own relaxation times, is similar to the phenomenological description of a relaxation process by means of Bloch equations or the Redfield equation [92]. This description in terms of the longitudinal relaxation time T_1 (D sector) and transversal relaxation time T_2 (ND sector) is, for example, used in NMR [92].

Of course, the "smallness" of the operators $\lambda^2\mathcal{S}^{(2)}, \lambda^3\mathcal{S}^{(3)}, \ldots$ compared to $\mathcal{S}^{(0)}$ will determine how fast the perturbation series converges. We will calculate the eigenvectors of $\mathcal{S}_{\lambda,t}$ to zeroth order in λ and the eigenvalues to second order in λ. The stochastic matrix $P_{nm,t}$ is determined in this approximation. The justification of this approximation will be given when we explicitly determine the expressions for $\mathcal{S}_{\lambda,t}$ in section 4.2.5, where we set bounds on λ and t such that indeed λ^2 and higher order corrections are small within some norm (for example the $\|\,.\,\|_\diamond$ given in [93, 94, 67]).

4.2.5 Calculation of Expressions

Here we will calculate the elements of the superoperator described in the last section to lowest non-trivial order in λ (λ^2). Taking the second-order expression for P in Eq. (4.2.61) to second order, $Q_{nm,t}$ is defined by the expression

$$P_{nm,t} = \delta_{nm} + \lambda^2 Q_{nm,t}. \tag{4.2.64}$$

And taking μ of Eq. (4.2.62) and using Eq. (4.2.50), we define $\nu_{nm,t}$ by

$$\mu_{nm,t} = e^{it(E_n - E_m)}(1 + \lambda^2 \nu_{nm,t}). \tag{4.2.65}$$

In this section we will find expressions for $Q_{nm,t}$ and $\nu_{nm,t}$ and exhibit the regime in which they give a valid description of $\mathcal{S}_{\lambda,t}$. We also show that for a large enough bath, the equilibrium state is the fixed point of the map $\mathcal{S}_{\lambda,t}$. We discuss under what conditions this fixed point is unique.

We will use operators in the Heisenberg representation. We denote such operators (for example on the system) as

$$A_t = e^{iH_s t} A\, e^{-iH_s t}. \tag{4.2.66}$$

The total Liouvillian $\mathcal{L}$ is defined as

$$e^{-i\mathcal{L}t}(\rho \otimes \rho_{b,\beta}) = U^t(\rho \otimes \rho_{b,\beta}) U^{t\dagger}. \tag{4.2.67}$$

One can expand the operator $e^{-i\mathcal{L}t}$ in a perturbation series in λ [79], take a partial trace over the bath and identify the operators $\mathcal{S}_t^{(0)} = e^{-i\mathcal{L}_s t}$, $\mathcal{S}_t^{(1)}$ and $\mathcal{S}_t^{(2)}$ in Eq. (4.2.45):

$$\mathcal{S}_t^{(1)} = -i\,\mathrm{Tr}_b \int_0^t dt' e^{-i(\mathcal{L}_s+\mathcal{L}_b)(t-t')} \mathcal{L}_{sb}\, e^{-i(\mathcal{L}_s+\mathcal{L}_b)t'}, \tag{4.2.68}$$

and

$$\mathcal{S}_t^{(2)} = -\mathrm{Tr}_b \int_0^t dt' \int_0^{t'} dt''\, e^{-i(\mathcal{L}_s+\mathcal{L}_b)(t-t')} \mathcal{L}_{sb}\, e^{-i(\mathcal{L}_s+\mathcal{L}_b)(t'-t'')} \mathcal{L}_{sb}\, e^{-i(\mathcal{L}_s+\mathcal{L}_b)t''}. \tag{4.2.69}$$

First we consider $\mathcal{S}_t^{(1)}$. We use Eq. (4.2.67) and Eq. (4.2.14) to rewrite $\mathcal{S}_t^{(1)}$ acting on $\rho \otimes \rho_{b,\beta}$ as:

$$\begin{aligned}\mathcal{S}_t^{(1)}(\rho \otimes \rho_{b,\beta}) &= -i\lambda\,\mathrm{Tr}_b \int_0^t dt' \\ &e^{iH_s(t-t')} \otimes e^{iH_b(t-t')} [H_{sb}, \rho_{t'} \otimes \rho_{b,\beta_{t'}}]\, e^{-iH_s(t-t')} \otimes e^{-iH_b(t-t')},\end{aligned} \tag{4.2.70}$$

where $\rho_{t'}$ is the time-evolved (with H_s) ρ and $\rho_{b,\beta_{t'}}$ is the time-evolved (with H_b) $\rho_{b,\beta}$. The equilibrium state $\rho_{b,\beta}$ is invariant under unitary evolution with $e^{iH_b t'}$ and thus $\rho_{b,\beta_{t'}} = \rho_{b,\beta}$. We then use the cyclic permutation invariance of the trace and $H_{sb} = S \otimes B$ to rewrite equation (4.2.70) as a simpler sum of two terms

$$\begin{aligned}\mathcal{S}_t^{(1)}(\rho \otimes \rho_{b,\beta}) &= -i\lambda \int_0^t dt' \\ &\left[e^{iH_s(t-t')} S \rho_{t'} e^{-iH_s(t-t')} - e^{iH_s(t-t')} \rho_{t'} S e^{-iH_s(t-t')}\right] \mathrm{Tr}_b B \rho_{b,\beta}\end{aligned} \tag{4.2.71}$$

Then the condition Eq. (4.2.9) implies that $\mathcal{S}_t^{(1)}(\rho \otimes \rho_{b,\beta})$ is 0 for any ρ.

Let us consider the second order term. The expression for $\mathcal{S}_t^{(2)}$ reads

$$\mathcal{S}_t^{(2)} = -e^{-i\mathcal{L}_s t}\int_0^t dt' \int_0^{t'} dt'' \, [h(t'-t'')S_{-t'}S_{-t''}\rho \\ -h(t''-t')S_{-t'}\rho S_{-t''} - h(t'-t'')S_{-t''}\rho S_{-t'} + h(t''-t')\rho S_{-t''}S_{-t'}] \,, \quad (4.2.72)$$

where $h(t)$ is defined as $\langle BB_t\rangle_b$. We write

$$h(t) = \int_{-\infty}^{\infty} d\omega \; e^{it\omega}\tilde{h}(\omega). \quad (4.2.73)$$

Let S_{nm} be the matrix elements of the interaction S in this eigenbasis of H_s, $S_{nm} = \langle n|S|m\rangle$. Now we can find the expression for $Q_{nm,t} = (\mathcal{S}_t^{(2)})_{nn,mm}$. From Eq. (4.2.72) after integration over the variables t' and t'' and with the use of Eq. (4.2.73), we find:

$$Q_{nm,t} = 2\int_{-\infty}^{\infty} d\omega \; \tilde{h}(\omega) \left[\frac{|S_{mn}|^2(1-\cos t(\omega - E_n + E_m))}{(\omega - E_n + E_m)^2} \right. \\ \left. -\sum_l \frac{\delta_{nm}|S_{nl}|^2(1-\cos t(\omega - E_n + E_l))}{(\omega - E_n + E_l)^2}\right]. \quad (4.2.74)$$

For the "decay factor" $\nu_{nm,t}$ in the ND sector we find

$$\nu_{nm,t} = \int_{-\infty}^{\infty} d\omega \; \tilde{h}(\omega) \left[\frac{2S_{nn}S_{mm}(1-\cos t\omega)}{\omega^2} \right. \\ \left. -f(t,\omega,E_n) - f^*(t,\omega,E_m)\right], \quad (4.2.75)$$

with f^* the complex conjugate of f. The function f is given by

$$\text{Re } f(t,\omega,E_n) = \sum_l \frac{|S_{ln}|^2(1-\cos t(\omega - E_n + E_l))}{(\omega - E_n + E_l)^2}, \quad (4.2.76)$$

and

$$\text{Im } f(t,\omega,E_n) = \sum_l \frac{|S_{ln}|^2}{\omega - E_n + E_l}\left[1 - \frac{\sin t(\omega - E_n + E_l)}{t(\omega - E_n + E_l)}\right]. \quad (4.2.77)$$

We will now look at the idealized case, i.e., we take the limits (remember k is the number of qubits in the bath)

$$P_{nm}^{\lambda^2 t} \equiv \lim_{\substack{t\to\infty,\lambda\to 0 \\ \text{constant}\lambda^2 t}} \lim_{k\to\infty} P_{nm,t}, \\ \mu_{nm}^{\lambda^2 t} \equiv e^{it(E_n - E_m)} \lim_{\substack{t\to\infty,\lambda\to 0 \\ \text{constant}\lambda^2 t}} \lim_{k\to\infty} (1 + \lambda^2\nu_{nm,t}). \quad (4.2.78)$$

When the bath is infinitely large, it will have a continuous spectrum; $\tilde{h}(\omega)$ will be a smooth function. The rate of interaction vanishes, but as we take the limit $t\to\infty$, there is an effective non-zero interaction that is proportional to $\lambda^2 t$. Recall that

$$\delta(x) = \lim_{t\to\infty}\frac{1-\cos(tx)}{t\pi x^2}, \quad (4.2.79)$$

where $\delta(x)$ is the Dirac delta function, which is defined as $\int_{-\infty}^{\infty} dx\, \delta(x) = 1$ and, $\forall\, x \neq 0$, $\delta(x) = 0$. With the use of the δ function we find

$$P_{mn}^{\lambda^2 t} = \delta_{nm}[1 - \lambda^2 t 2\pi \sum_l |S_{nl}|^2 \tilde{h}(E_n - E_l)] + \lambda^2 t 2\pi |S_{mn}|^2 \tilde{h}(E_n - E_m), \tag{4.2.80}$$

and

$$\mu_{nm}^{\lambda^2 t} = e^{it(E_n - E_m)} \left[1 + \lambda^2 t 2\pi S_{nn} S_{mm} \tilde{h}(0) - \lambda^2 t \pi g(E_n) - \lambda^2 t \pi g^*(E_m)\right], \tag{4.2.81}$$

with

$$\text{Re } g(E_n) = \sum_l |S_{ln}|^2 \tilde{h}(E_n - E_l), \tag{4.2.82}$$

and

$$\text{Im } g(E_n) = \mathcal{P} \int_{-\infty}^{\infty} d\omega\, \tilde{h}(\omega) \sum_l \frac{|S_{ln}|^2}{\omega - E_n + E_l} \tag{4.2.83}$$

where $\mathcal{P}$ is the principal value of the integral. In order to see in what regime the perturbation theory is correct, we check whether the process described by Eq. (4.2.80) and Eq. (4.2.81) corresponds to that of a $\mathbf{TCP}$ map. First we verify Property 1 in Eq. (4.2.81); the eigenvalues of $|n\rangle\langle m|$ and $|m\rangle\langle n|$ are related by complex conjugation, or $\mu_{nm}^{\lambda^2 t^*} = \mu_{mn}^{\lambda^2 t}$. The trace-preserving property (also in Property 1) is also obeyed:

$$\sum_m P_{mn}^{\lambda^2 t} = 1. \tag{4.2.84}$$

Complete positivity of the map implies that $P_{mn}^{\lambda^2 t}$ must be a matrix of probabilities, that is we must have $P_{mn}^{\lambda^2 t} \geq 0$. Thus the first necessary condition for the validity of the perturbative approximation is

$$\text{Condition 1:} \quad \forall\, n:\ \lambda^2 t \ll \tfrac{1}{2\pi \sum_l |S_{ln}|^2 \tilde{h}(E_n - E_l)}. \tag{4.2.85}$$

Eq. (4.2.84) and Eq. (4.2.85) together ensure that $P_{mn}^{\lambda^2 t}$ is a stochastic matrix. Complete positivity also implies via Proposition 1 that $|\mu_{nm}^{\lambda^2 t}| \leq 1$. In order that $|1 + \lambda^2 t a| \leq 1$, where a is some complex number, we must have that Re $a \leq 0$ and $\lambda^2 t \leq 2/|\text{Re } a|$. This real part in Eq. (4.2.81) is indeed negative as $\tilde{h}(\omega)$ is positive, and we obtain a new condition:

$$\text{Condition 2: } \forall\, m, n:\ \lambda^2 t \ll \frac{1}{\pi| - S_{nn} S_{mm} \tilde{h}(0) + \frac{1}{2}\sum_l |S_{ln}|^2 \tilde{h}(E_n - E_l) + \frac{1}{2}\sum_l |S_{lm}|^2 \tilde{h}(E_m - E_l)|}. \tag{4.2.86}$$

Note that this condition is quite similar to the condition in Eq. (4.2.85).

It is not hard to see that the stochastic matrix $P_{mn}^{\lambda^2 t}$ obeys detailed balance for the equilibrium distribution:

$$P_{mn}^{\lambda^2 t} e^{-\beta E_n} = P_{nm}^{\lambda^2 t} e^{-\beta E_m}, \tag{4.2.87}$$

as the equilibrium condition of the bath implies that

$$\tilde{h}(-\omega) = e^{-\beta\omega}\tilde{h}(\omega). \tag{4.2.88}$$

Thus the equilibrium density matrix $\rho_{s,\beta}$ is a fixed point of the idealized equilibration process. To consider whether this fixed point is unique, we note the following: If a stochastic matrix M is such that all its matrix elements $M_{ij} > 0$, then M has a unique eigenvalue equal to 1 [68]. If Condition 1 is obeyed, we indeed have $P_{mn}^{\lambda^2 t} > 0$ and therefore the absolute value of the second largest eigenvalue (in the diagonal sector) is smaller than 1. For the off-diagonal sector, Condition 2 says that the largest eigenvalue in the off-diagonal sector is strictly smaller than 1 in absolute value. Thus under these conditions, with Proposition 3, we can conclude that the process converges to the equilibrium state. The expression of $P_{mn}^{\lambda^2 t}$ coincides with the derivation given by Davies [84] for small $\lambda^2 t$.

One can help to speed up the process in the off-diagonal sector by "dephasing"; that is, after having the system and the bath interact for some time t, we perform the operation

$$\mathcal{D}^v(\rho_s) = \frac{1}{v}\sum_{u=0}^{v-1} e^{iH_s u}\rho_s e^{-iH_s u}, \tag{4.2.89}$$

which can be implemented with the assistance of an extra register in the state $\frac{1}{\sqrt{v}}\sum_{u=0}^{v-1}|u\rangle$ which is used to condition the evolution $U = e^{iH_s u}$ and subsequently traced out. The dephasing has the effect of canceling off-diagonal terms in the eigenbasis of the system, i.e.

$$\lim_{v\to\infty}\mathcal{D}^v\left(\sum_{k,l=1}^{N}\alpha_{kl}|k\rangle\langle l|\right) = \sum_{k=1}^{N}\alpha_{kk}|k\rangle\langle k|. \tag{4.2.90}$$

A complete dephasing can in general not be achieved in polynomial time in n (see section 4.3), and thus must be understood as an extra aid but not a solution to the equilibration problem.

From the expressions for $P_{mn}^{\lambda^2 t}$ we can understand the physical picture of the interaction between bath and system. The system can make a transition from (eigen) level n to level m ($n \neq m$), i.e. $P_{mn}^{\lambda^2 t}$ is non-zero, when S_{mn} is non-zero and $\tilde{h}(E_n - E_m)$ is nonzero. The function $\tilde{h}$ that occurs in Eq. (4.2.80) can be expressed as

$$\tilde{h}(\Delta E) = \lim_{K\to\infty}\sum_{l,j}^{K}\delta((\Delta E) - (\omega_l - \omega_j))|B_{lj}|^2 e^{-\beta\omega_l}/Z. \tag{4.2.91}$$

Therefore in order that $\tilde{h}(\Delta E = E_n - E_m)$ is nonzero, there must be at least one matching energy difference in the bath, i.e. there is an l and an j such that $|\omega_l - \omega_j| = \Delta E$ *and* B_{lj} is non-zero. See Figure 4.1. Furthermore, the more such matching energy differences and corresponding nonzero matrix elements of B there are, the larger the probability $P_{mn}^{\lambda^2 t}$. This confirms the intuitive picture that one might have of equilibration. Note also the similarity with the Fermi Golden Rule [81, 95] that describes the transition probability from eigenlevel n to m in a unitary evolution that is perturbed by a time-dependent Hamiltonian.

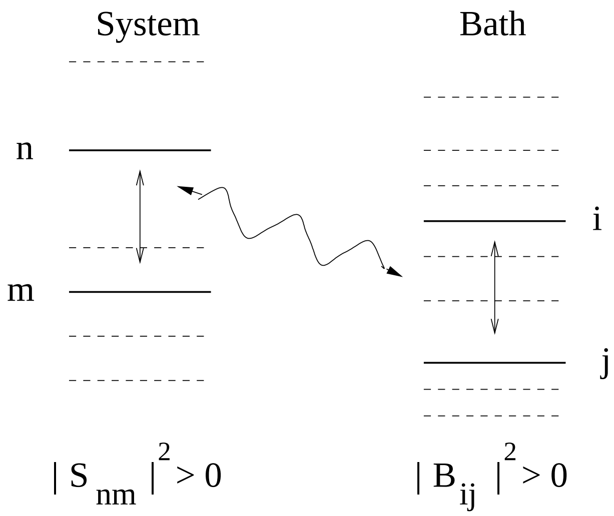

Figure 4.1: The process of energy exchange between bath and system that is important in the equilibration of the system.

For a finite-dimensional bath, we can express $h(t) \equiv \langle BB_t \rangle$ as

$$h(t) = \sum_{k,l} e^{it(\omega_k - \omega_l)} |B_{kl}|^2 e^{-\beta\omega_k}/Z_b, \tag{4.2.92}$$

where $B_{kl} = \langle k_b|B|l_b\rangle$ with $|l_b\rangle$ being the eigenstates of the bath Hamiltonian H_b and Z_b the partition function of the bath. Taking the limits $t \to \infty$ and $\lambda \to 0$ before letting the bath grow large leads to divergent expressions for $P_{mn}^{\lambda^2 t}$ and $\mu_{nm}^{\lambda^2 t}$, suggesting that the perturbation theory fails in this regime. This is not surprising, as the finiteness of the bath together with the limit $t \to \infty$ will lead to Poincaré recurrences (only the interaction cycle time is long due to $\lambda \to 0$).

The choice for an interaction of the form $S \otimes B$ between bath and system does not present a restriction in the process of equilibration in the limits of small coupling, large time and large bath; we have shown that equilibration is successful in these limits.

4.2.6 The Inverse Quantum Zeno Effect

In our numerical studies (sections 4.2.7 and 4.2.8) we have observed a phenomenon that one might call the inverse quantum Zeno effect. It is a way of mapping an arbitrary initial state onto the completely mixed state $\mathbf{1}_N$ by interacting repeatedly and strongly with the state for a very short time. Here we will give a theoretical analysis that explains this observation. Consider the weak coupling expansion $\mathcal{S}_{\lambda,t} = \mathcal{S}_t^{(0)} + \lambda^2 \mathcal{S}_t^{(2)} + \mathcal{O}(\lambda^3)$ with $\mathcal{S}_t^{(2)}$ given as in Eq. (4.2.69). We expand these operators around $t = 0$:

$$\mathcal{S}_{\lambda,t}(\rho) = \rho - it[H_s, \rho] + \frac{t^2\lambda^2}{2}([S\rho, S] + [S, \rho S])\,\langle B^2 \rangle_b + \mathcal{O}(t^2, \lambda^3 t^3). \tag{4.2.93}$$

In the limit $\lambda \to \infty$, but $t \to 0$, and *constant* $\lambda^2 t$, the higher order terms $\mathcal{O}(t^2, \lambda^3 t^3)$ will vanish. Thus we see that the fixed point of $\mathcal{S}_{\lambda,t}$ in this limit (assuming non-zero $\langle B^2 \rangle_b$) must obey

$$[H_s, \rho] = 0 \;\&\; [[S, \rho], S] = 0. \tag{4.2.94}$$

Notice that if we take the differential form of Eq. (4.2.93) and the prescribed limit, the equation is of the Lindblad form, Eq. (4.2.17). The state $\mathbf{1}_N$ certainly meets the requirements of Eq. (4.2.94), but is it unique? If S and H_s are such that they have no eigenspaces (except for the full space) in common, and both have a non-degenerate spectrum, we can show that $\mathbf{1}_N$ is the unique eigenvector. Eq. (4.2.94) requires that either $[S, \rho] = 0$ or $[S, \rho]$ is diagonal in the same basis as S. If $[S, \rho] = 0$ but also $[H_s, \rho] = 0$, then ρ can only be the state $\mathbf{1}_N$. What happens if $[S, \rho]$ is just diagonal in the same basis as S? Let $|n\rangle$ be an eigenvector of S with eigenvalue λ_n. We have for $n \neq m$

$$\langle n| [S, \rho] |m\rangle = 0. \tag{4.2.95}$$

Rewriting this expression gives

$$\forall n, m,\; n \neq m \;\; \langle n| \rho |m\rangle (\lambda_n - \lambda_m) = 0. \tag{4.2.96}$$

Now, because ρ is diagonal in the basis of H_s as $[H_s, \rho] = 0$ and H_s and S have no eigenvectors in common, there exist n and m such that $\langle n|\rho|m\rangle \neq 0$. But the eigenvalues of S were non-degenerate, thus we obtain a contradiction. □

When $\mathbf{1}_N$ is the unique eigenvector of this process, then, with the use of Proposition 3, the repeated application as in step 4 of the *Equilibration algorithm* I will eventually bring the system to the state $\mathbf{1}_N$.

We showed that for this "inverse quantum Zeno" effect to occur S and H_s have to be such that they have no partial eigenspace in common and both have a non-degenerate spectrum. If we assume that S and H_s are d-local with d larger or equal than 4, then this does not impose a very strong constraint on S and H_s; the effect will occur for a generic S and H_s.

4.2.7 Specifications of the Numerical Simulation

The main purpose of this study is to understand the effects of bath size and the choice of bath and interaction Hamiltonians for a specific system Hamiltonian. In Table 4.1 we list some of the choices that have been made in the numerical analysis. We have randomly generated the elementary Hamiltonians h_i that make up H_s, H_b and H_{sb}, Eqs. (4.2.5), (4.2.6), (4.2.7), with a measure $\mathcal{M}$. We choose the diagonal elements of each h_i uniformly in $[-a, a]$, where a is the sampling scale in Table 4.1. The absolute value of the above-the-diagonal elements of h_i are chosen uniformly in $[0, a]$ and its phase is chosen uniformly in $[0, 2\pi]$. The below-the-diagonal elements of h_i follow from hermiticity. This defines $\mathcal{M}$. Note that $\mathcal{M}$ is not a unitarily invariant measure.

We take the Hamiltonians S and B as sums of all possible local 2-qubit interactions ($d_s = 4$ in Table 4.1). For the Hamiltonian of the system H_s we also take a sum of all possible local

2-qubit interactions. Note that this includes a set of Hamiltonians that exhibit frustration, for which we don't expect equilibration to be particularly fast.

In section 4.2.5 we observed that matching energy differences between bath and system are an important ingredient in the equilibration of the system, which is consistent with the intuitive picture of equilibration that was sketched in section 4.2.1. However, as we do not know the eigenvalues of the system, we can only pick our bath so as to optimize the chance for matching level differences. The sampling scale of the bath $f(n, k, d_s, d_b)$ is determined by roughly optimizing these coincidences, $\Delta E_b = \Delta E_s$. This is done as follows.

Consider the density of states $p_s(E, a_s)$ of the system (the distribution of eigenvalues generated by the measure $\mathcal{M}$) and the density of states $p_b(E, a_b)$ of the bath. Here a_s is the sampling scale of the system which we set to 1 (see Table 4.1). The quantity $[\mathrm{Tr}\, H_s]_{\mathcal{M}}$ is the mean and $\frac{[\mathrm{Tr}\, H_s^2]_{\mathcal{M}}}{N}$ is the variance of the distribution $p_s(E, a_s)$. The choice for $\mathcal{M}$ ensures that the distributions are symmetric around $E = 0$:

$$[\mathrm{Tr}\, H_s]_{\mathcal{M}} = [\mathrm{Tr}\, H_b]_{\mathcal{M}} = 0. \tag{4.2.97}$$

To optimize for matching we choose the variances to be equal:

$$\frac{[\mathrm{Tr}\, H_s^2]_{\mathcal{M}}}{N} = \frac{[\mathrm{Tr}\, H_b^2]_{\mathcal{M}}}{K}. \tag{4.2.98}$$

For large K the bath distribution will be Gaussian (central limit theorem), whereas the system distribution will be similar to a Gaussian distribution for large N (see Fig. 4.2). Thus, setting the variances equal brings the distributions close together.

	H_s	H_b	S	B
dimension	$N = 2, .., 2^4$	$K = 2^2, .., 2^6$	N	K
locality	$d_s = 4$	$d_b = 2$	4	4
sampling scale a	1	$f(n, k, d_s, d_b)$	1	1

Table 4.1: Some settings in the numerical simulation.

Consider first $[\mathrm{Tr}\, H_b^2]_{\mathcal{M}}$. It is straightforward to calculate the variance of the eigenvalues of a qubit bath. Given a 2×2 Hermitian matrix m_{ij}, the eigenvalues $e_\pm = \frac{1}{2}(m_{11} + m_{22} \pm \sqrt{(m_{11} - m_{22})^2 + 4|m_{12}|^2})$ have the property

$$[e_\pm^2]_{\mathcal{M}} = \frac{1}{4a_b^3} \int_{-a_b}^{a_b} dm_{11} \int_{-a_b}^{a_b} dm_{22} \int_0^{a_b} d|m_{12}|\, e_\pm^2 = \frac{2a_b^2}{3}. \tag{4.2.99}$$

Let v_i be some $\pm$ pattern i of length k, corresponding to selecting e_+ or e_- for each qubit bath. Let E_{v_i} be an eigenvalue of the full bath, i.e., $E_{v_i} = \sum_{m=1}^{k} e_{v_i[m]}$ where $v_i[m]$ indicates that we select the mth bit in v_i. Then

$$\frac{[\mathrm{Tr}\, H_b^2]_{\mathcal{M}}}{K} = \frac{1}{K} \sum_{i=1}^{K} [E_{v_i}^2]_{\mathcal{M}} = \frac{2ka_b^2}{3}, \tag{4.2.100}$$

using $[e_{v_i[m]}e_{v_i[n]}]_{\mathcal{M}} = 0$ for $m \neq n$. We calculate $[\mathrm{Tr}\, H_s^2]_{\mathcal{M}} = \sum_{i,j}[|(H_s)_{ij}|^2]_{\mathcal{M}}$ for $n > 2$. We can write

$$\sum_{i,j}[|(H_s)_{ij}|^2]_{\mathcal{M}} = \sum_{i,j}\sum_{m=1}^{\binom{n}{2}}[|(h_m)_{ij}|^2]_{\mathcal{M}}, \qquad (4.2.101)$$

where h_m is the mth local interaction Hamiltonian. We have used $[(h_k^*)_{ij}(h_m)_{ij}]_{\mathcal{M}} = 0$ for $k \neq m$. Each row of h_m has only four non-zero entries as the dimension of the local Hamiltonians d_s was set to four. Using the fact that $[|(h_m)_{ij}|^2]_{\mathcal{M}} = \frac{1}{3}$ for all interaction terms m, we obtain

$$\frac{[\mathrm{Tr}\, H_s^2]_{\mathcal{M}}}{N} = \frac{4}{3}\begin{pmatrix} n \\ 2 \end{pmatrix}. \qquad (4.2.102)$$

For $n = 1$, we have $\frac{[\mathrm{Tr}\, H_s^2]_{\mathcal{M}}}{N} = \frac{2}{3}$. Comparing Eqs. (4.2.100) and (4.2.102) gives the expression for a_b:

$$a_b = f(n, k, 4, 2) = \sqrt{\frac{2}{k}\begin{pmatrix} n \\ 2 \end{pmatrix}}. \qquad (4.2.103)$$

For $n = 1$, $f(1, k, 4, 2) = \sqrt{1/k}$. Fig. 4.2 illustrates how this setting determines the density of states of bath and system.

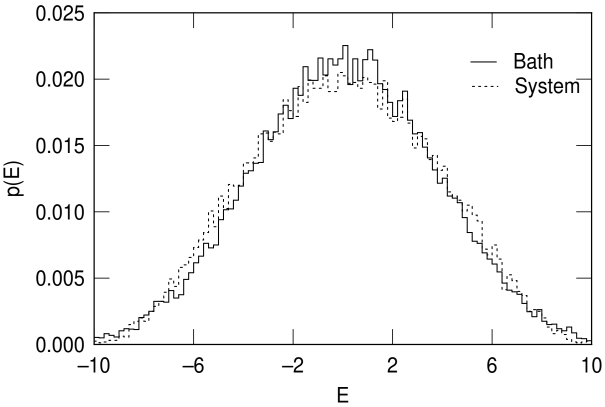

Figure 4.2: A histogram (500 samples) of the density of states (unnormalized) for $N = 32$ and $K = 64$ with sampling scale set as in Eq. (4.2.103).

The numerical work consists of a calculation of the fixed point of $\mathcal{S}_{\lambda,t}$ as a function of t for a fixed λ and the second largest eigenvalue for different baths and different systems and temperatures. We follow a numerical procedure based on perturbation theory (Section 4.2.4) to perform a stable numerical evaluation of these quantities. Also, we expect from the analysis in section 4.2.4 that the small coupling regime, the realm where the perturbation theory is valid, is the regime in which we find good equilibration. When the coupling between bath and system is too strong the bath does not just exchange energy with the system but higher order (in the coupling) effects will bias the dynamics of the system in a way that depends on the bath. We will derive an effective coupling parameter $c(t)$ that depends on λ, but also on t, the strength of the interaction Hamiltonian and the energy spectrum of the bath.

We can trust the answers from the numerical procedure only if we are in the regime in which perturbation theory is correct. This regime was heralded by the two conditions Eq. (4.2.85) and Eq. (4.2.86) in section 4.2. Whether these conditions are obeyed depends on the specific choices of H_s, H_b and S and B. We prefer to reformulate these conditions here such that they are obeyed for the average bath, system and interaction Hamiltonian obtained by sampling using $\mathcal{M}$ and the sampling scale. As the conditions are very similar, we take the first one, Eq. (4.2.85), and reformulate it as

$$c(t) \equiv \lambda^2 t\, 2\pi \frac{NK[S^2]_{\mathcal{M}}[B^2]_{\mathcal{M}}}{W_b} \leq 1. \tag{4.2.104}$$

where $[S^2]_{\mathcal{M}}$, the average matrix element, is defined as

$$[S^2]_{\mathcal{M}} = \frac{1}{N^2}\sum_{i,j}[|S_{ij}|^2]_{\mathcal{M}} = \frac{1}{N^2}[\mathrm{Tr}_s S^2]_{\mathcal{M}}, \tag{4.2.105}$$

and similarly for $[B^2]_{\mathcal{M}}$. W_b is the spectral width of the bath, i.e.,

$$W_b^2 = \frac{[\mathrm{Tr}\, H_b^2]_{\mathcal{M}}}{K}. \tag{4.2.106}$$

Here we indicate the approximations made in obtaining Eq. (4.2.104) from Condition 1 (Eq. (4.2.85)):

$$\lambda^2 t\; 2\pi \sum_l |S_{ln}|^2 \tilde{h}(E_n - E_l) \ll 1. \tag{4.2.107}$$

Using Eq. (4.2.92) and Eq. (4.2.73) we write the function $\tilde{h}$ as

$$\tilde{h}(E_n - E_l) = \sum_{k,m} \delta((E_n - E_l) - (\omega_k - \omega_m))|B_{km}|^2 e^{-\beta\omega_k}/Z. \tag{4.2.108}$$

We will approximate the matrix elements $|B_{km}|^2$ as constants and replace them by their average $[B^2]_{\mathcal{M}}$. Then we can use density-of-states arguments to approximate the m sum over the δ functions by the inverse of the average spacing between the δ functions; this spacing is given by W_b/K:

$$\sum_m \delta((E_n - E_l) - (\omega_k - \omega_m)) \approx \frac{K}{W_b}. \tag{4.2.109}$$

With these approximations, the partition-function sum over k in Eq. (4.2.108) becomes exactly one. So, Eq. (4.2.108) becomes

$$\tilde{h}(E_n - E_l) \approx \frac{K[B^2]_{\mathcal{M}}}{W_b}. \tag{4.2.110}$$

Now Eq. (4.2.107) is

$$\lambda^2 t \; 2\pi \frac{K[B^2]_{\mathcal{M}}}{W_b} \sum_l |S_{ln}|^2 \ll 1. \tag{4.2.111}$$

If we again approximate the matrix elements $|S_{ln}|^2$ as constants and replace them by their average $[S^2]_{\mathcal{M}}$, and note that the l sum in Eq. (4.2.111) has N terms, we obtain Eq. (4.2.104).

For the simulations that we have performed, we can find the values for $[S^2]_{\mathcal{M}}$ and $[B^2]_{\mathcal{M}}$ (note that these Hamiltonians have locality parameter $d = 4$, as does the system Hamiltonian H_s) and obtain the expression

$$c(t) = \lambda^2 t \, \frac{16\pi}{3\sqrt{3}} \begin{pmatrix} k \\ 2 \end{pmatrix} \sqrt{\begin{pmatrix} n \\ 2 \end{pmatrix}} \ll 1. \tag{4.2.112}$$

for $n > 1$ and $k > 1$. For a qubit system, $n = 1$, and $k > 1$ we obtain

$$c_1(t) \equiv \lambda^2 t \frac{8\pi\sqrt{2}}{3\sqrt{3}} \begin{pmatrix} k \\ 2 \end{pmatrix} \ll 1. \tag{4.2.113}$$

The quantity $c(t)$ in Eq. (4.2.104) will function as a rescaled time which depends on the strength of λ and the size of system and bath. In the regime where $c(t) \leq 1$ we expect a pertubative calculation of the eigenvectors and eigenvalues of the superoperator to be fairly accurate. The dimensionless parameter associated with the temperature is given by

$$\beta' = \beta W_s, \tag{4.2.114}$$

where W_s is the spectral width of the system, Eq. (4.2.106) ($W_s = W_b$). From here on, β will refer to this scaled dimensionless parameter. Instead of expanding the superoperator $\mathcal{S}$ in a series in λ as in Eq. (4.2.45), we write

$$\lambda^2 \bar{\mathcal{S}}_t^{(2)} \equiv \mathcal{S}_{\lambda,t} - \mathcal{S}_t^{(0)}, \tag{4.2.115}$$

where all higher order terms are grouped in $\bar{\mathcal{S}}_t^{(2)}$. The calculation of eigenvalues and eigenvectors then follows the analysis of Section 4.2.4. We find that the choice for the bath and the interaction Hamiltonian influences whether the equilibration will succeed or not. In Figs. 4.3 and 4.4 two extrema in dynamics are shown, each corresponding to a different choice for the system, bath, and interaction. In Fig. 4.3 the equilibration is successful, whereas in Fig. 4.4 the equilibration fails. The quantity rate$_D$ is defined as

$$\text{rate}_D = \frac{1 - \kappa_D}{\bar{c}(t)}, \tag{4.2.116}$$

where κ_D is the second largest eigenvalue in the diagonal sector and $\bar{c}(t)$ is the average coupling strength in the time interval that we consider, which is $c(t) \in [0, 0.3]$ here. Similarly, we define

$$\text{rate}_{ND} = \frac{1 - \kappa_{ND}}{\bar{c}(t)} \tag{4.2.117}$$

for the nondiagonal sector. The qualitative difference in the behavior in Fig. 4.3 and Fig. 4.4 depends on the three requirements that we also found to be of importance in the idealized case that was treated in section 4.2.5. These requirements were summarized in Fig. 4.1.

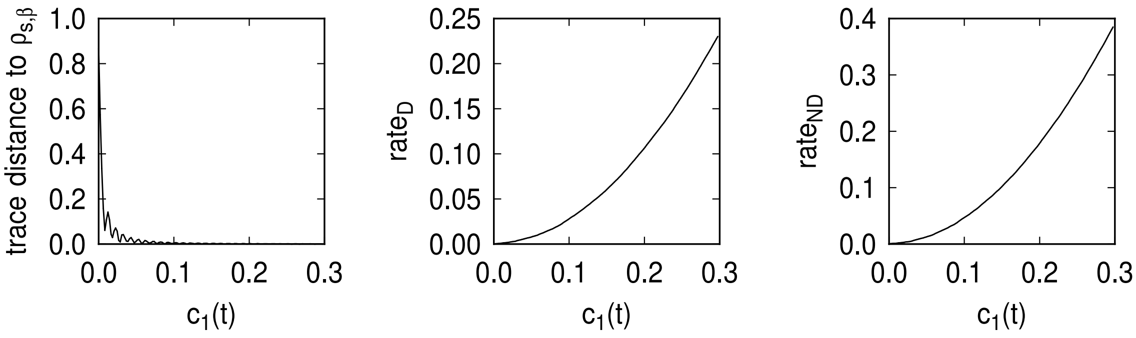

Figure 4.3: An example of successful equilibration for $n = 1$, $k = 3$ and $\beta = 3$.

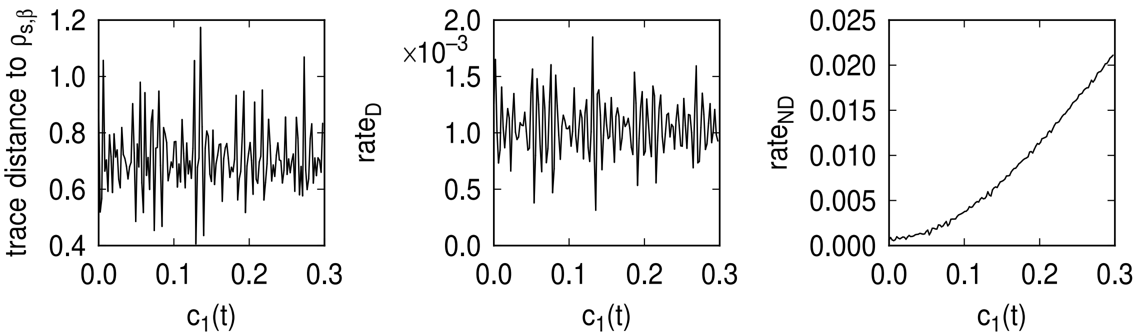

Figure 4.4: An example of an unsuccessful equilibration for $n = 1$, $k = 3$ and $\beta = 3$.

4.2.8 Numerical Results for Equilibration

We are interested in how well a randomly chosen bath and interaction equilibrate a system and how these averages are improved by choosing larger baths. As the mixing rates and the distance to the equilibrium state will in general be oscillating functions of the scaled time $c(t)$ (see Fig. 4.4) we will compute time averaged rates over a reasonable interval in $c(t)$,

$$[c(t_{init}) = 0, c(t_{end}) = 0.5], \tag{4.2.118}$$

such that we are in the realm where perturbation theory is valid, Eq. (4.2.104). We denoted these time averages (not to be confused with bath averages) as Av. rate$_D$ and Av. distance for the time averaged trace distance etc. In Fig. 4.5 we present histograms that show how, for a given fixed system *and* interaction, the equilibration process is different for a set of randomly chosen baths with fixed dimension. The insets show the distribution for the lowest bin. The vertical axis denotes the percentage of baths (the interval $[0\%, 100\%]$ is given as the interval $[0, 1]$) for a certain distance and rate. We observe that the diagonal rate distribution is very broad, and therefore the mean of the distribution is not a very good (or a very stable) measure of the generic behavior. Furthermore, we find that the rate in the diagonal sector is much worse than in the nondiagonal sector and thus is the dominant factor in setting the mixing time. This conforms to the pattern in many quantum systems, for example for nuclear spins as observed by NMR, for which T_1 is generically larger than T_2 [92].

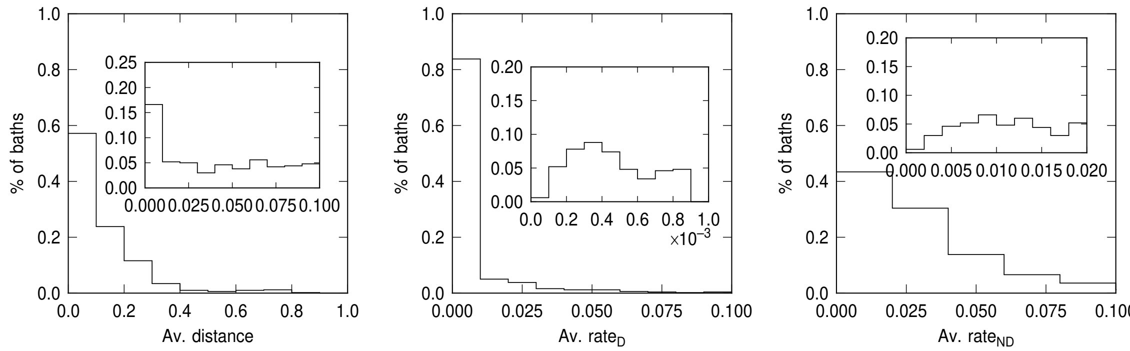

Figure 4.5: An example of the distribution of baths (500 samples) for $n = 2$ and $k = 3$ and $\beta = 2$.

To study the dependence on β and on the dimension of the bath versus the dimension of the system, we compute the following data. We pick a system Hamiltonian H_s of n qubits that has some well spread out spectrum. We set the dimension of the bath and then we randomly pick both the bath Hamiltonian and the interaction Hamiltonian. Means are denoted as $[.]_{\mathcal{M}_b}$. For the rates we look both at the mean and the median. The median is denoted as Median$_{\mathcal{M}_b}$, see Fig. 4.6. The results for $n = 1, 2, 3$ and 4 are shown in Figs. 4.6-4.9. We have given the median when the mean does not give a good representation of the distribution.

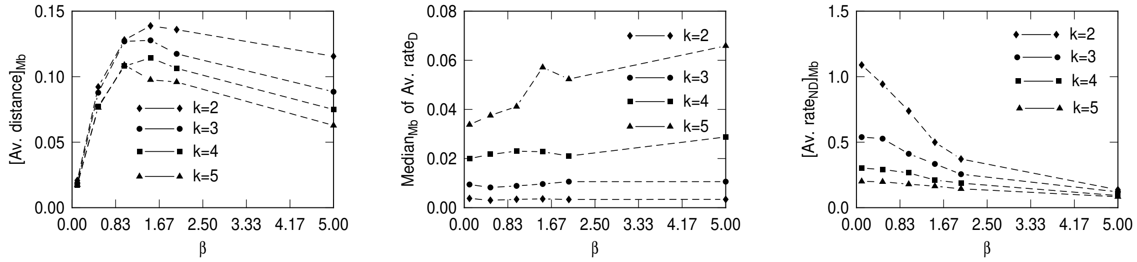

Figure 4.6: Means and median for n=1 (500 samples).

These data clearly indicate that larger baths improve the process of equilibration, both in the rates (D and ND) as well as in the closeness to the equilibrium state. The effects are the most pronounced at low temperature, where equilibration is in general harder as the system must relax to a single pure ground state. To understand the closeness scale, we show in Appendix 4.D how far apart two arbitrarily chosen density matrices are; this number lies around 1 for the dimensions that were considered. For these estimates, we see a trend towards approximations getting worse for larger system sizes for low temperature. The scaled rates [Av. rate$_D]_{\mathcal{M}_b}$ and [Av. rate$_{ND}]_{\mathcal{M}_b}$ seem to be fairly constant, thus we see behavior that

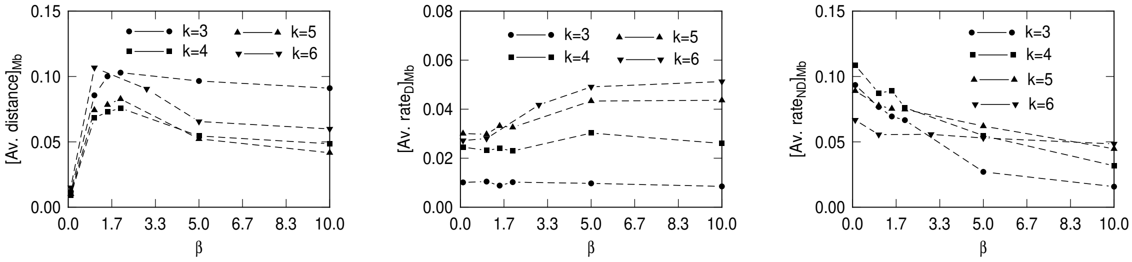

Figure 4.7: Means for n=2 (200-500 samples).

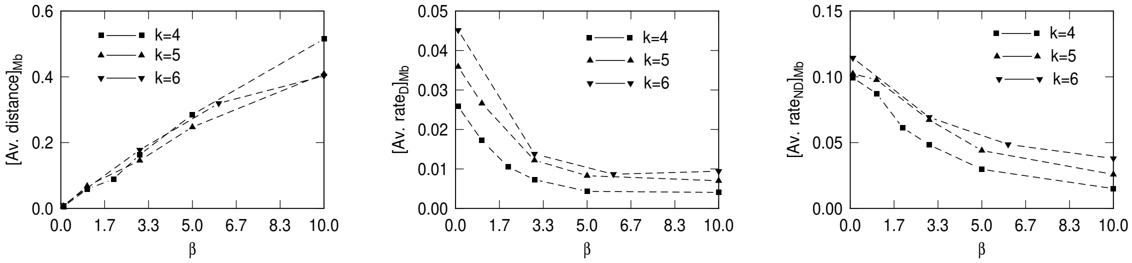

Figure 4.8: Means for n=3 (50-100 samples).

suggests that the rates are polynomially related to both system and bath number of qubits. We also observe that the nondiagonal rate (ND) is always higher than the diagonal rate (D). The data show a system Hamiltonian dependence, that is, the average equilibration for $n = 4$ seems to be more successful than for $n = 3$. We also observe that the difference between T_1 and T_2 becomes smaller with increasing β (lower temperature). Now we can try to give some answers to the question that were posed in section 4.2.2 (above Eq. (4.2.13). The parameters t and λ are grouped together in a single effective coupling $c(t)$. This effective coupling $c(t)$ should be small such that the perturbative approach is valid. How should one choose the value of parameter r, the number of iterations? This depends on the three quantities plotted in the Figures, the trace distance, the rate in the diagonal and the off-diagonal sector. An efficient equilibration would correspond to process in which, while using a bath that is polynomially related in size to the system, both the rates are polynomially related to the number of system qubits and the trace distance is a small constant.

Now, if we extrapolate the data to large systems we can see that if we pick a bath size (in number of qubits) that is polynomially related to the system size (note that the number of

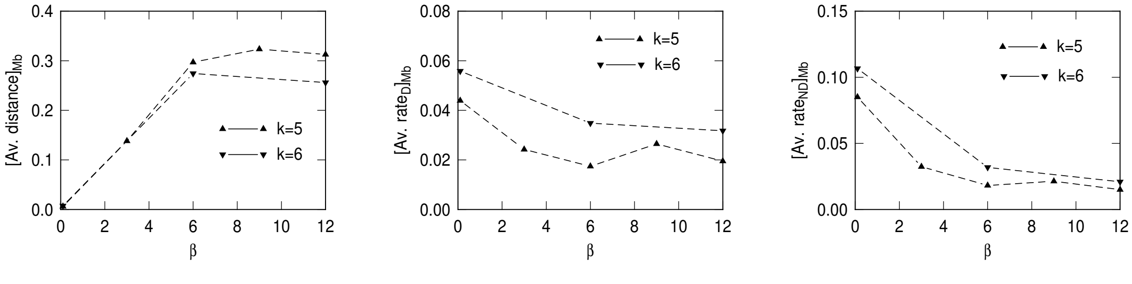

Figure 4.9: Means for n=4 (15-20 samples).

eigenvalues is then *exponentially* related), the rates of relaxation are polynomially related to the system size (in qubits), in fact we find that the scaled rates are constant. This behavior we do not find for the scaling of the trace distance. When we choose a polynomial relation between system and bath size, the data suggest that the relaxed state could be still fairly far away from the true equilibrium state for large system sizes.

In choosing the Hamiltonians H_b, S and B one should try to optimize for energy matching, Fig. 4.1. The numerical data show that choosing a large bath is beneficial for equilibration. The constraint that the bath consists of a set of uncoupled qubits does not seem to impose a serious restriction on the equilibration process. We believe however that when H_b corresponds to a set of uncoupled qubits *and* the Hamiltonian B does *not* couple these qubits, i.e. $B = \sum_i B_i$ where each B_i acts on a single qubit, then the equilibration process might be somewhat impaired. The reason is that the pairs of energy levels in the bath for which $|B_{ij}|^2$ is nonzero are then restricted to the energy levels for each qubit separately. The number of matching energy level pairs for a n-qubit bath is thus n. For a general interaction term B it will be $O(2^n)$. This can lead to quite different dynamics.

4.3 Equilibration II

We present an alternative to the algorithm in section 4.2. This algorithm relies on the technique for the estimation of eigenvalues, originally given in Ref. [67] (see Refs. [96, 97]). This eigenvalue estimation routine has also been used as a building block in quantum algorithms in Refs. [77] and [98].

Let H_s be a d-local Hamiltonian with non-degenerate eigenvalues as in section 4.2.

Definition 2 *Equilibration algorithm II.*

1. **Initialize** *the system in the (infinite temperature) completely mixed state $\mathbf{1}_N$. Also add one m-qubit register set to $|00\ldots0\rangle\langle 00\ldots0|$. This last register will be used to compute an m-bit estimate of an eigenvalue.*

2. **Compute eigenvalues** *with the use of the Fourier transform and* **dephase** *in the computational (eigenvalue) basis. Let U be the eigenvalue computation routine, i.e.*

$$U|n\rangle \otimes \underbrace{|00\ldots0\rangle}_{m} = |n\rangle \otimes \underbrace{|(00\ldots0) \oplus s_n\rangle}_{m}, \tag{4.3.1}$$

where $|s_n\rangle$ is an m-bit estimate of the eigenvalue E_n defined by $H|n\rangle = E_n|n\rangle$. The dephasing is a simple superoperator $\mathcal{D}$ on the eigenvalue register that operates as

$$\mathcal{D}(|s_i\rangle\langle s_i|) = |s_i\rangle\langle s_i|, \quad \mathcal{D}(|s_i\rangle\langle s_j|) = 0, \quad i \neq j. \tag{4.3.2}$$

The total transformation maps

$$\mathcal{D}(U(\mathbf{1}_N \otimes |00\ldots0\rangle\langle 00\ldots0|)U^\dagger) = \sum_{n=0}^{N-1}\sum_{s_n=0}^{2^m-1} p(n, s_n)|n\rangle\langle n| \otimes |s_n\rangle\langle s_n|, \tag{4.3.3}$$

where $p(n, s_n)$ is a probability distribution, peaked at $\frac{s_n}{2^m} \sim E_n$ for large m (see Appendix 4.C).

3. **Prepare** *an additional N-dimensional quantum system, the bath, also in $\mathbf{1}_N$. Add a m-qubit register and one qubit register set to $|00\ldots0\rangle\langle 00\ldots0|$.*

4. **Compute eigenvalues** *of the bath as for the system in step 2.*

5. **Interact** *system and bath according the following rule $\mathcal{R}$ ("partial swap"):*

$$U_{\mathcal{R}}|n,m\rangle|0\rangle|s,t\rangle = \begin{cases} |m,n\rangle|0\rangle|s,t\rangle & \text{if } t < s, \\ (p_{st}^{\beta/2}|m,n\rangle|0\rangle + \sqrt{1-p_{st}^{\beta}}|n,m\rangle|1\rangle)|s,t\rangle & \text{if } t \geq s, \end{cases} \tag{4.3.4}$$

where $p_{st}^{\beta} = e^{-\beta(t-s)}$. Here $|n\rangle$ and $|s\rangle$ are the registers of the system and $|m\rangle$ and $|t\rangle$ are the registers of the bath.

6. **Trace** *over the single-qubit register, all bath registers, and the eigenvalue register of the system. The system will be in some state*

$$\rho_s = \sum_n \alpha_n |n\rangle\langle n|. \tag{4.3.5}$$

The steps 2-6 define a $\mathbf{TCP}$ map $\mathcal{S}$ with $\mathcal{S}(\mathbf{1}_N) = \rho_s$.

7. **Repeat** *steps 2-6 r times such that*

$$\| \mathcal{S}^{r+1}(|00\ldots0\rangle\langle 00\ldots0|) - \mathcal{S}^{r}(|00\ldots0\rangle\langle 00\ldots0|) \|_{tr} \leq \epsilon, \tag{4.3.6}$$

for all $r \geq r_0$ and ϵ is some accuracy.

The advantage of this algorithm is its simplicity and its similarity to a classical algorithm; we create a Markov chain in the eigenbasis of the system. The disadvantage of the algorithm is that it is very likely to be slow; the computation of the eigenvalues to high accuracy with the use of the Fourier transform can take exponential time in the number of qubits of the system for an arbitrary d-local Hamiltonian H. This routine has to be performed twice, for system and bath, in each round of the chain. First, let us show that in the case when the eigenvalues are computed exactly in steps 2 and 4, i.e, $p(n, s_n) = \delta_{E_n, s_n/2^m}/N$ the Markov chain equilibrates the system. Recall [96] that the routines of steps 2 and 4 compute rescaled eigenvalues

$$E_n' = f_1 E_n + f_2, \tag{4.3.7}$$

with f_1 and f_2 depending on the maximum and minimum eigenvalue (of which we assume that we can find an estimate) such that $E_n' \in [0, 1)$. In the following we will drop these primes. The chain that is created can be represented as

$$\sum_n \alpha_n^{(k)} |n\rangle\langle n|, \tag{4.3.8}$$

where $\alpha_m^{(k)} = \sum_n \alpha_n^{(k-1)} P_{n\to m}$. We have

$$P_{n\to m} = \begin{cases} \frac{1}{N} & \text{if } E_m < E_n, \\ \frac{1}{N}(1 + \sum_{E_k \geq E_n}(1 - p_{nk}^\beta)) & \text{if } E_m = E_n, \\ \frac{1}{N} p_{nm}^\beta & \text{if } E_m > E_n, \end{cases} \tag{4.3.9}$$

where $p_{nk}^\beta = p_{E_n,E_k}^\beta$. Note that $\sum_m P_{n\to m} = 1$ as required. The equilibrium state Eq. (4.1.1) obeys the detailed balance condition:

$$\forall n, m \;\; P_{n\to m} e^{-\beta E_n} = P_{m\to n} e^{-\beta E_m}. \tag{4.3.10}$$

All the matrix elements of the Markov matrix $P_{n\to m}$ are nonzero. Therefore the chain will have a unique fixed point which is equal to the equilibrium state due to detailed balance. Thus for all probability distributions α_n we have

$$\lim_{k\to\infty} \sum_n \alpha_n P_{n\to m}^{(k)} = \frac{e^{-\beta E_m}}{Z}. \tag{4.3.11}$$

Notice that it is not hard to prepare the initial states of system and bath. One way to make the completely mixed state $\mathbf{1}_N$ is to make a maximally entangled state $\frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} |i\rangle|i\rangle$ and trace over the second register. This takes $O(n)$ steps. The partial swap in step 5 can be implemented with $\mathrm{Poly}(n, m)$ elementary steps. The dephasing in step 2 is introduced to keep the form of the algorithm clean, but it does not affect its output. This dephasing is implemented by measuring the eigenvalue register in the computational basis and discarding its answer. When using an m-bit eigenvalue register the joint probability $p(n, s_n)$ in the first round (after step 2) is equal to

$$p(n, s_n) = \frac{1}{N} \left| \frac{1}{2^m} \sum_{t=0}^{2^m-1} e^{2\pi i t(E_n - s_n/2^m)} \right|^2. \tag{4.3.12}$$

When $p(n, s_n)$ is not a delta function on the eigenvalue, the Markov chain will still be in the eigenbasis of the system; It will be a concatenation of chains; the transition probability of this new chain is

$$P'_{n\to m} = \sum_{s,t} p(s \,|\, n) P_{s\to t}\, p(m \,|\, t), \tag{4.3.13}$$

where $p(s|n)$ is a conditional probability, defined by $p(n, s) = p(s|n)p(n)$, and $P_{s\to t}$ is the exact chain (when $p(s|n) = \delta_{E_n, s_n/2^m}$). Note that $\sum_s p(s|n) = 1$ and $\sum_m p(m|t) = 1$, so that $P'_{n\to m}$ is a stochastic matrix. Let us make a few remarks about the behavior of such an approximate equilibration process. If this new Markov chain is close to the exact Markov chain, we can bound the deviation from the exact fixed point with perturbation theory [99]. Let

$$P'_{n\to m} = P_{n\to m} + E_{nm}, \tag{4.3.14}$$

where E_{nm} is a deviation matrix defined by Eq. (4.3.13). Let $\rho_\Delta = \rho'_{s,\beta} - \rho_{s,\beta}$ where $\rho'_{s,\beta}$ is the fixed point of the Markov chain P'_{nm}. Assume that P is diagonalizable. Let Y be the matrix defined as

$$Y = (\mathbf{1} - P + P^{(\infty)})^{-1} - P^{(\infty)}, \tag{4.3.15}$$

where $P^{(\infty)}$ is the infinite iteration of P. We can write $P^{(\infty)} = \mathrm{diag}(1, 0, \dots, 0)$ in the basis where the stationary state $\rho_{s,\beta}$ is an eigenvector. In this basis, with diagonalizability, P is of the form $\mathrm{diag}(1, \lambda_2, \dots, \lambda_N)$. We can then write

$$Y = \mathrm{diag}(0, \frac{1}{1-\kappa}, \dots, \frac{1}{1-\lambda_N}), \tag{4.3.16}$$

where κ is the absolute value of the second largest eigenvalue. For later use we note that the norm $|||Y|||_2 = \frac{1}{|1-\kappa|}$. It is possible to write the deviation ρ_Δ in terms of Y and E:

$$\rho_\Delta = (\mathbf{1} - YE)^{-1} Y E \rho_{s,\beta} \tag{4.3.17}$$

when E is small enough such that $\mathbf{1} - YE$ is invertible. This expression can be derived from $P^{(\infty)}\rho_\Delta = 0$, which follows from the uniqueness of the stationary state ρ. We now use

$$\| \rho_\Delta \|_{tr} \leq \sqrt{N} \| \rho_\Delta \|_2, \tag{4.3.18}$$

as in Proposition 3. Then using the expression for Y, Eq. (4.3.17) and Eq. (4.3.18) (see also below Eq. (4.2.37)) we can bound

$$\begin{aligned} \| \rho_\Delta \|_{tr} &\leq C_N \, \mathrm{Tr}\, \rho_{s,\beta}^2 \left(1 - \frac{|||E|||_2}{|1-\kappa|}\right)^{-1} \frac{|||E|||_2}{|1-\kappa|} \\ &\leq C_N \left(1 - \frac{|||E|||_2}{|1-\kappa|}\right)^{-1} \frac{|||E|||_2}{|1-\kappa|}. \end{aligned} \tag{4.3.19}$$

Thus the size of the correction ρ_Δ will be determined by the strength of the perturbation $|||E|||_2$ and the rate of convergence of the original Markov chain P.

For a general H_s, the computation of an m-bit approximation of the eigenvalues can cost an exponential (in m) number of elementary gates. As there are 2^n eigenvalues, knowing the $m = \log \mathrm{Poly}(n)$ bits of the values of E_n still leaves groups of $\frac{2^n}{\mathrm{Poly}(n)}$ eigenvalues indistinguishable. Thus only in very special cases, if the gates $U_s^{2^m}$ can be implemented with a polynomial (in m) number of elementary steps (as in Shor's factoring algorithm [8]) is it possible to compute the eigenvalues to high accuracy efficiently.

We have demonstrated a way to set up a Markov chain on a quantum computer that will converge to the equilibrium state for long enough time. For special Hamiltonians, there might be more efficient ways to tune and modify this kind of algorithm. The rule $\mathcal{R}$ might be chosen to depend on other features of the eigenstates $|n\rangle$ and $|m\rangle$ as in the classical Metropolis algorithm where transitions are made between states that are related by local spin flips. There might be Hamiltonians for which the calculation of an eigenvalue, given the eigenvector, is efficient.

4.4 (Time-dependent) Observables

Given that we have prepared n qubits in the equilibrium state corresponding to a certain Hamiltonian H_s, we can then proceed by experimenting and measuring. The simplest measurement that we could try to perform is the estimation of the expectation value of a d-local (Hermitian) observable O:

$$\langle O \rangle_s = \mathrm{Tr}\, \rho_{s,\beta}\, O. \tag{4.4.1}$$

As O is local, we write $O = \sum_{i=1}^{\mathrm{Poly}(n)} O_i$ where each operator O_i operates on a Hilbert space of constant dimension d. We can calculate the eigenvectors and eigenvalues of each O_i rapidly on a (possibly) classical computer, which takes $\mathrm{Poly}(n, c)$ operations. If O_i has eigenvalues μ_i that are both smaller as well as larger than zero, we define O_i^+ as

$$O_i^+ = \frac{O_i + |\min_k \mu_k| \mathbf{1}}{\max_k \mu_k + |\min_k \mu_k|} \tag{4.4.2}$$

such that O_i^+ is positive semi-definite and has eigenvalues smaller than or equal than 1. If O_i has only positive or zero eigenvalues, we just "normalize" the operator by dividing by $\max_k \mu_k$, and similarly if O_i has only negative eigenvalues. Let I_i be a positive operator valued measurement (POVM [62]) with operation elements $A_{1,i}$ and $A_{2,i}$ and corresponding outcomes 1 and 2 such that

$$\begin{aligned} E_{1,i} &= A_{1,i}^\dagger A_{1,i} = O_i^+, \\ E_{2,i} &= A_{2,i}^\dagger A_{2,i} = \mathbf{1} - O_i^+. \end{aligned} \tag{4.4.3}$$

This measurement will give outcome 1 with probability

$$p_{1,i} = \mathrm{Tr}\, O_i^+ \rho \ \text{ etc.} \tag{4.4.4}$$

The operators $A_{1,i}$ and $A_{2,i}$ are given by

$$A_{1,i} = U_{o_i} \, (\mathrm{diag}_{O_i^+})^{1/2} \, U_{o_i}^\dagger \quad \text{and } A_{2,i} = U_{o_i} \, (\mathbf{1} - \mathrm{diag}_{O_i^+})^{1/2} \, U_{o_i}^\dagger, \tag{4.4.5}$$

where $\mathrm{diag}_{O_i^+}$ is the diagonal form of O_i^+ and U_{o_i} the diagonalizing matrix of O_i^+. We summarize these results in a Proposition:

Proposition 5 *The estimation of* $\mathrm{Tr}\, \rho\, O$ *where* O *is a* d*-local observable with precision* δ *and error probability* ϵ *and* $\rho \in B_1(\mathcal{H}_N)^+$ *(*$N = 2^n$*) takes* $T\, O(\frac{\ln \epsilon^{-1}}{\delta^2}) \mathrm{Poly}(n, c)$ *operations where* T *is the time to prepare the state* ρ*.*

Proof All commuting observables O_i can be measured once with a single preparation of ρ. To estimate a probability p with precision δ and error probability ϵ we need $O(\frac{\ln \epsilon^{-1}}{\delta^2})$ samples [52]. □.

More interesting is an algorithm to estimate time-dependent expectation values. Let O_1 and O_2 be two d-local observables. We consider how to estimate a time-dependent quantity (identical to Eq. (4.1.10))

$$\mathrm{Tr}\,\rho_\beta\,[O_1, O_{2t}], \tag{4.4.6}$$

where O_{2t} is in the Heisenberg representation. Notice that O_{2t}, the time-evolved operator, will for general t *not* be local. Thus we cannot use Proposition 5. The way these quantities come about in linear response theory [72] provides the key for how to estimate them on a quantum computer. One considers a system that is perturbed at some initial time $t = 0$: its time evolution is generated by the perturbed Hamiltonian $H_s + \lambda O_1(t)$ $(O_1(t < 0) = 0)$ and λ is small. After time t we consider the response of the system to the perturbation by measuring another observable O_2. Notice that with Proposition 5, it is simple to perform this experiment. Linear response means that we take into account corrections of order λ, but no higher order, in the estimation of

$$\delta\langle O_2\rangle_s = \mathrm{Tr}\,O_2\rho_t - \mathrm{Tr}\,O_2\rho_\beta, \tag{4.4.7}$$

where ρ_t is the time-evolved system density matrix. This first-order correction takes the form [95]

$$\delta\langle O_2\rangle_s \approx i\lambda \int_0^t dt'\ \mathrm{Tr}\,\rho_\beta\,[O_1(t'), O_{2t-t'}]. \tag{4.4.8}$$

If the disturbance $O_1(t) = O_1\delta(t = 0)$ we find on the right hand side the correlation function of Eq. (4.4.6). The quantity of Eq. (4.4.6) is interesting, because it can be used to compute the simplest reponse of the system, the linear response of Eq. (4.4.8), which we can directly estimate on our quantum computer, provided that both O_1 and O_2 are local. But we are of course not restricted to a linear response regime: λ is a parameter that we can tune freely. A sequence of measurements could determine higher response functions that will involve quantities such as

$$\langle O_{1t_1} O_{2t_2} O_{3t_3} \ldots O_{kt_k}\rangle_s. \tag{4.4.9}$$

4.5 Conclusion

It seems that by asking the question of how fast real quantum systems equilibrate, we have opened a Pandora's box of hard-to-answer questions. If there are many simple quantum systems in nature that equilibrate slowly (that is, *not* in polynomial time and space) by any dynamics that does not require extensive preknowledge of the system, then it would be unreasonable to ask our quantum computer to perform this task efficiently. By relaxation in polynomial time we mean the following: in polynomial time in n and $1/\epsilon$ we obtain a state that is within ϵ trace distance of the equilibrium state where ϵ is a small constant. It might be the case that leaving aside the classical phenomenon of frustration, relaxation does *not* take place in polynomial time. The idea here is that for a quantum system, the eigenbasis is not known beforehand, but must be singled out on the basis of an estimation of the eigenvalues, which is generically a hard problem.

This however is not in contradiction with physical and experimental reality as we know it, as the quantities that are measured in an experimental setup usually involve operators on a small number of qubits; these are the experiments that can be done efficiently (in polynomial time) and thus do not necessarily probe the system's complete state. For example, the outcomes of the set of measurements $\sigma_{i_1} \otimes \sigma_{i_2} \otimes \ldots \sigma_{i_n}$ where σ_{i_j} is one of the Pauli matrices or $\mathbf{1}$, completely determines the state, but there are 4^n measurements in this set. In an experimental setup, we might randomly select a polynomial subset of them and there is some small chance of order $\frac{\mathrm{Poly}(n)}{4^n}$ that these are the measurements that distinguish the equilibrated state from the present state in the lab that is supposed to approximate it. The estimates of time-dependent correlations could possibly be more sensitive to distance from equilibrium, as these involve time-evolved, non-local operations. The numerical study suggests that product baths whose size is polynomially related to the system can function as adequate baths in the sense of providing relaxation in polynomial time. The relaxed state could still be a rather rough approximation to the true equilibrium state, but, as we argued above, it might be a good starting point for subsequent measurements.

We have taken the bath to be part of the (cost of) the quantum computer. In any experimental setup, there is a natural bath that is used to equilibrate and cool the quantum computer. Can we use this bath for a computational problem such as equilibration? Consider for example the NMR quantum computer [16] where computation takes place at room temperature. In the regime in which the heat bath has a non-Markovian character it has been shown to be possible to alter the Hamiltonian of the system and the coupling to the bath dynamically (see Ref. [100], but also standard books on NMR [92]). These techniques could make it possible to simulate the time-evolution of a "designer" Hamiltonian and also to equilibrate the system to the equilibrium state of this designer Hamiltonian.

Finally, we have taken the first steps in developing a theory of quantum Markov chains for quantum computational purposes. It will be interesting to bring this theory to the next level. For example, we can try to define the notions of irreducibility and ergodicity for these quantum Markov chains. One of the essential question is then, can we find quantum Markov chains that are rapidly mixing for interesting computational problems for which no good classical algorithms are available. Such a problem could for example be a problem of equilibration for a specific Hamiltonian H. In this study we have laid the ground work for this future research.

4.A Gibbs State is the Equilibrium State

We prove that the equilibrium state as defined in Eq. (4.1.1) is the state that has maximum entropy on the "energy shell", that is, given a value for $\langle E \rangle = \mathrm{Tr}\, H\rho$ (see Ref. [101] for example for an alternative proof). Let $\rho = \sum_{m,n=1}^{N} r_{mn}|n\rangle\langle m| \in B_1(\mathcal{H}_N)^+$ and $|n\rangle$ are the eigenvectors of H. Then we can observe that

$$\langle E\rangle = \sum_m r_{mm}E_m, \tag{4.A.1}$$

does not depend on the values of r_{mn} for $n \neq m$.

We will need the following lemma

Lemma 2 *A density matrix* $\rho = \sum_{m,n=1}^{N} r_{mn}|n\rangle\langle m|$ *with fixed diagonal elements* r_{mm} *has maximum entropy when* $r_{mn} = 0$ *for* $n \neq m$.

Proof Ky Fan's inequality [101] for Hermitian matrices reads

$$\sum_{i=1}^{m} \mu_i \geq \sum_{i=1}^{m} \langle i|\,\rho'\,|i\rangle, \tag{4.A.2}$$

for all $m \leq N$. Here $\mu_1 \geq \mu_2 \geq \ldots \geq \mu_N$ are the eigenvalues of ρ' and $\{|i\rangle\}_{i=1}^{N}$ is an orthogonal set of vectors. Let $\{|i\rangle\}_{i=1}^{N}$ be the eigenvectors of ρ and let ρ' be a density matrix whose diagonal elements are identical to ρ in the basis $\{|i\rangle\}_{i=1}^{N}$. This implies that

$$\sum_{i=1}^{N} \mu_i \geq \sum_{i=1}^{N} \lambda_i, \tag{4.A.3}$$

with λ_i (μ_i) a decreasing series of eigenvalues of ρ (ρ'). Given the majorization of Eq. (4.A.3) it follows that for a concave function as the von Neumann entropy [101, 102]

$$S(\rho) \geq S(\rho'). \tag{4.A.4}$$

$\square$.

With this lemma we find that the equilibrium density matrix ρ will be diagonal in the eigenbasis of H, $\rho = \sum_{m=1}^{N} r_{mm}|m\rangle\langle m|$. In order to find the density matrix ρ with maximum entropy we have to optimize over the values r_{mm}. In Ref. [103] a derivation for this optimization is given. One maximizes the von Neumann entropy of ρ, which is now the classical Shannon entropy associated with r_{mm}, under the energy constraint:

$$H = -\sum_{m=1}^{N} r_{mm}\log r_{mm} - \beta\left(\sum_m r_{mm}E_m - \langle E\rangle\right) - \alpha\left(\sum_m r_{mm} - 1\right), \tag{4.A.5}$$

where β and α are Lagrange multipliers. This leads to the solution $r_{mm} = \frac{e^{-\beta E_m}}{Z}$ with $Z = \sum_m e^{-\beta E_m}$. $\square$

4.B Implementing a Local Hamiltonian Evolution

This result was originally derived in Ref. [32] and [77]. Let $U^t = e^{iHt}$ be a time-evolution of a 2^n dimensional d-local quantum system (n qubits). An d-dim gate is a gate on a d-dimensional

Hilbert space. We estimate how many of these d-dim gates, where $d = 2^k$ with some k, are needed to simulate U^t with accuracy δ, i.e.

$$\| U^t - U^t_{\text{approx}} \| \leq \delta. \tag{4.B.1}$$

where $\| \, . \, \|$ is defined in Appendix 4.D. We can write

$$U^t = e^{iH\epsilon t/\epsilon}. \tag{4.B.2}$$

Using a Baker-Campbell-Hausdorff expansion we approximate

$$e^{iH\epsilon} = e^{iH_1\epsilon} \ldots e^{iH_k\epsilon} [1 + \mathcal{O}(\epsilon^2)\text{Poly}(n)X], \tag{4.B.3}$$

where X is some operator with $\| X \| \leq c_0$ with a constant c_0. Each operator $e^{iH_i\epsilon}$ can be implemented with a constant number of d-dim gates. Thus for the full simulation we have

$$\| U^t - U^t_{\text{approx}} \| \leq t\mathcal{O}(\epsilon)\text{Poly}(n) \equiv \delta. \tag{4.B.4}$$

To achieve an accuracy δ, the number of d-dim gates is therefore proportional to $t/\epsilon = O(t^2/\delta)\text{Poly}(n)$.

Interestingly, the simulation is proportional to t^2 and not to t. This is due to the fact that an elementary operation that lasts for a small time ϵ must be accounted for as at least one elementary gate or elementary time step, although the action of this gate is very short. This is the cost that we have to pay for breaking up the evolution in tiny steps.

4.C Eigenvalue Estimation

An alternative view on Shor's factoring algorithm in terms of eigenvalue estimation was found by Kitaev [67]. In [96] Kitaev's eigenvalue estimation algorithm was analyzed. We give the analysis in Ref. [96]. Let $E_{max} = \max_n E_n$ and $E_{min} = \min_n E_n$. Rescale the energy eigenvalues such that

$$E'_n = f_1 E_n + f_2 \in [0, 1), \tag{4.C.1}$$

with f_1 and f_2 depending on estimates of E_{min} and E_{max}. Given is an eigenvector of H: $H|\psi_n\rangle = E'_n|\psi_n\rangle$. To compute the first m bits of $E'_n = a_n/2^m + \delta_n$ with $|\delta_n| \leq 1/2^{m+1}$ we do the following steps

1. **Prepare** a "time"-register and eigenstate as

$$|\psi_n\rangle \otimes \frac{1}{\sqrt{2^m}} \sum_{t=0}^{2^m-1} |t\rangle. \tag{4.C.2}$$

2. **Evolve the eigenvector in time** using the rescaled Hamiltonian H' of the system:

$$U_C|\psi_n\rangle|t\rangle = e^{2\pi i H' t}|\psi_n\rangle|t\rangle = e^{2\pi i E'_n t}|\psi_n\rangle|t\rangle. \tag{4.C.3}$$

3. **Fourier** transform the "time-register":

$$U_F|t\rangle = \frac{1}{\sqrt{2^m}} \sum_{s=0}^{2^m-1} e^{-2\pi its/2^m}|s\rangle. \tag{4.C.4}$$

This sequence of transformations will result in the following state:

$$\sum_{s=0}^{2^m-1}\left(\frac{1}{2^m}\sum_{t=0}^{2^m-1} e^{2\pi it(a_n-s)/2^m}e^{2\pi i\delta_n t}\right)|\psi_n\rangle|s\rangle. \tag{4.C.5}$$

In a normal eigenvalue estimation algorithm, the next step is to measure the Fourier transformed time-register. In Ref. [96] it was estimated that the probability to measure state $|s = a\rangle$ was larger than $\frac{4}{\pi^2}$, independent of m. The number of operations in the algorithm is a set of controlled $U^t \equiv e^{2\pi iH't}$ where $t = 0, \ldots, 2^m - 1$ and $O(m^2)$ other operations. The implementation of an operation such as U^{2^m} can take an exponential number (in m) of elementary operations for arbitary local Hamiltonians. The probability of succes for the same m-bit estimate of the eigenvalue can be made arbitary large by running the algorithm for a longer time.

4.D Norms

In this Appendix we give the definitions of several norms and inner products. The inner product between vectors in $\mathbf{C}^{N^2}$ can be represented on $B(\mathcal{H}_N)$ as

$$\langle\chi_1|\chi_2\rangle = \mathrm{Tr}\,\chi_1^\dagger\chi_2. \tag{4.D.1}$$

The trace norm [93, 94] is defined as

$$\parallel A \parallel_{tr} = \mathrm{Tr}\sqrt{A^\dagger A}. \tag{4.D.2}$$

What makes this norm attractive is that it captures a measurable closeness of two density matrices ρ_1 and ρ_2 [94]:

$$\parallel \rho_1 - \rho_2 \parallel_{tr} = \max_A \sum_j |P_1^A(j) - P_2^A(j)|, \tag{4.D.3}$$

where P_1^A and P_2^A are the probability distributions over outcomes j that are obtained by measuring observable A on ρ_1 and ρ_2. The matrix norm $|||.\,|||_2$ is defined as

$$|||A|||_2 = \max_{x:\|x\|_2=1} \parallel Ax \parallel_2 . \tag{4.D.4}$$

where $\parallel . \parallel_2$ is the Euclidean norm on $\mathbf{C}^{N^2}$: $\sqrt{\langle v|v\rangle}$ for $|v\rangle \in \mathbf{C}^{N^2}$. We have

$$\parallel Ax \parallel_2 \leq |||A|||_2 \parallel x \parallel_2 . \tag{4.D.5}$$

The conventional operator norm is defined as

$$\| U \| = \sup_{|\xi\rangle \neq 0} \frac{\| U|\xi\rangle \|}{\| |\xi\rangle \|}, \tag{4.D.6}$$

In order to aid in the interpretation of the numerical results of section 4.2.7, we present some numerical estimates for the average $\| \, . \, \|_{tr}$ distance of two randomly chosen density matrices. We first have to choose a measure over $B_1(\mathcal{H}_N)^+$. All density matrices can be written as $\rho = \sum_i \lambda_i \rho_{ii}$ with $\sum_{i=1}^N \lambda_i = 1$. The eigenvalues λ_k lie on a $(N-1)$- dimensional simplex S in $\mathbf{R}^N$. We use the Euclidean metric $\| \, . \, \|_2$ induced on the simplex. The Haar measure on the group of unitary matrices $U(N)$ induces a uniform measure on the set of projectors $\{\rho_{ii}\}_{i=1}^{N^2}$. Together this defines a measure $\mathcal{M}_{B_1^+}$ [104]. Within this measure, one can express the average distance between two density matrices ρ_1 and ρ_2, using the unitary invariance of $\| \, . \, \|_{tr}$, as

$$\begin{aligned}[\| \rho_1 - \rho_2 \|_{tr}]_{\mathcal{M}_{B_1^+}} &= \frac{1}{\mathrm{Vol}(S)^2 V(U(N))} \int dU \int_0^1 d\lambda_1 \dots d\lambda_k \int_0^1 d\mu_1 \dots d\mu_k \\ &\delta\left(\sum_i \lambda_i - 1\right) \delta\left(\sum_i \mu_i - 1\right) \mathrm{Tr}\, |\sum_j \lambda_j \rho_{jj} - U \sum_j \mu_j \rho_{jj} U^\dagger |. \end{aligned} \tag{4.D.7}$$

The values obtained by a numerical calculation of Eq. (4.D.7) are tabulated in Table 4.2.

dim	mean	standard error= $\sqrt{\mathrm{var}/(n-1)}$, $n = 1000$
4	0.90388	0.00740588
8	0.96190	0.00514057
16	1.00294	0.00341226
32	1.01452	0.00220363
64	1.02617	0.00132233

Table 4.2: The average distance between two randomly selected density matrices.

4.E Preparation of the Bath

To prepare the state

$$\rho_{b,\beta} = \rho_{b,\beta}^1 \otimes \dots \otimes \rho_{b,\beta}^k, \tag{4.E.1}$$

given $H_b = \sum_{i=1}^k \mathbf{1}_{K/2} \otimes h_i$, we first calculate the eigenvalues and eigenvectors of each qubit Hamiltonian h_i. We prepare the state

$$\Pi_{i=1}^k \left(e^{-\beta e_{i,0}} |0\rangle\langle 0| + e^{-\beta e_{i,1}} |1\rangle\langle 1|\right) / Z_i. \tag{4.E.2}$$

with $\{e_{0,i}, e_{1,i}\}$ the eigenvalues of qubit Hamiltonian h_i. This can be done by changing an initial state $|0\rangle\langle 0|$ with probability $e^{-\beta e_{i,1}}/Z_i$ into state $|1\rangle\langle 1|$ for each i. We then rotate each qubit to its eigenbasis $\{|b_{i0}\rangle, |b_{i1}\rangle\}$:

$$\otimes_{i=1}^{k} U_{b_i} = \otimes_{i=1}^{k}(|b_{i0}\rangle\langle 0| + |b_{i1}\rangle\langle 1|). \tag{4.E.3}$$

In total we perform $2k$ elementary qubit operations plus some constant classical overhead.

Chapter 5

Product Bases, Local Distinguishability and Bound Entanglement

5.1 Introduction

In this chapter we study fundamental properties of quantum mechanical states, operations and measurements. In section 5.2 we review the notions of entanglement, distillation of entanglement and its relation to positive linear maps. In section 5.3 we establish a relation between Bell inequalities and the separability criterion and show under what restrictions they are equivalent. In section 5.4 we present results that relate local distinguishability of sets of product states to bound entanglement. Central in this construction is the notion of an unextendible product basis, of which we will give many examples. We prove that uncompletable product bases cannot be distinguished by a finite number of local operations and classical communication. These uncompletable product bases form new examples of the phenomenon of nonlocality without entanglement. In section 5.5 we present a new family of indecomposable positive linear maps. In the following sections we use the notation $n \otimes m$ or $\mathcal{H}_n \otimes \mathcal{H}_m$ to denote the tensorproduct between a n-dimensional Hilbert space and a m-dimensional Hilbert space.

5.2 Quantum Entanglement

The study of entanglement is essential for the understanding of quantum mechanics and the use of quantum mechanics in computation and information processing tasks. Erwin Schrödinger introduced the notion of entanglement and he was the first to understand its fundamental importance; in Ref. [105] we find

> "*When two systems, of which we know the states by their respective representatives, enter into temporary physical interaction due to known forces between them, and when after a time of mutual influence the systems separate again, then they can no longer be described in the same way as before, viz. by endowing each*

of them with a representative of its own. I would not call that one but rather the characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought. By the interaction the two representatives (or ψ-functions) have become entangled."

The simplest form of entanglement is the entanglement that we find in bipartite pure states. Let $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$. We call the state $|\psi\rangle$ entangled iff $|\psi\rangle$ cannot be written as a product of pure states:

$$|\psi\rangle \neq |\psi_a\rangle \otimes |\psi_b\rangle, \tag{5.2.1}$$

where $|\psi_a\rangle \in \mathcal{H}_A$ and $|\psi_b\rangle \in \mathcal{H}_B$. Equivalently, when we express the pure state $|\psi\rangle$ as a density matrix $|\psi\rangle\langle\psi|$, the density matrix is entangled iff it cannot be written as

$$|\psi\rangle\langle\psi| \neq |\psi_a\rangle\langle\psi_a| \otimes |\psi_b\rangle\langle\psi_b|. \tag{5.2.2}$$

The famous example of a bipartite entangled state in $2 \otimes 2$ is the Einstein-Podolsky-Rosen (EPR) singlet state [106, 107]

$$|\Psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle). \tag{5.2.3}$$

We are not only concerned with pure states, but also with mixed states, represented by positive semidefinite Hermitian matrices ρ with $\mathrm{Tr}\,\rho = 1$, the density matrices. Let us give the definition of entanglement for a bipartite density matrix ρ:

Definition 3 *Let ρ be a density matrix on a finite-dimensional Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$. A state $|\psi\rangle$ of the form $|\psi^A\rangle \otimes |\psi^B\rangle$ is a (pure) product state in $\mathcal{H}_A \otimes \mathcal{H}_B$. The density matrix ρ is entangled iff ρ cannot be written as a convex combination of pure product states, i.e. there does not exist an ensemble $\{p_i \geq 0, |\psi_i^A \otimes \psi_i^B\rangle\}$ such that*

$$\rho = \sum_i p_i \, |\psi_i^A\rangle\langle\psi_i^A| \otimes |\psi_i^B\rangle\langle\psi_i^B|. \tag{5.2.4}$$

When ρ is not entangled ρ is called separable.

One would also like to classify entanglement in multipartite systems. A famous example of a tripartite pure entangled state is the Greenberger-Horne-Zeilinger (GHZ) state:

$$|\mathrm{GHZ}\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle). \tag{5.2.5}$$

The state cannot be written as a product of three single party states, nor as a product of an entangled two party and a single party state. The example illustrates that in a full characterization of the entanglement of multipartite systems we will have specify between which parties the entanglement occurs. When we look at multipartite density matrices, this can lead to surprising results. In section 5.4.3 we will give an example (Example 3) of a tripartite density matrix which cannot be written as a convex combination of pure product states for all three parties. When viewed as a bipartite density matrix on $\mathcal{H}_{AB} \otimes \mathcal{H}_C$ or $\mathcal{H}_A \otimes \mathcal{H}_{BC}$ or $\mathcal{H}_{AC} \otimes \mathcal{H}_B$, it can be shown that the density matrix is separable. This is impossible if the density matrix is a pure state, but apparently allowed when we consider general density matrices.

5.2.1 Quantification of Entanglement

It is important to have a measure of entanglement that quantifies 'how much entanglement' a state contains. Here we will only consider a measure of entanglement for bipartite states. For multipartite states the measure of entanglement has to take into account between which subsystems the entanglement occurs. It is an open problem how to define a measure of multipartite entanglement that gives a complete description of the various kinds of entanglement that are present in the state.

Any measure of entanglement $E(\rho)$ where ρ is a density matrix on $\mathcal{H}_A \otimes \mathcal{H}_B$ must have the following four natural properties [37, 108]:

1. $E(\rho) \geq 0$ for all density matrices ρ and $E(\rho) = 0$ when ρ is a separable density matrix.

2. $E(\rho)$ is invariant under local unitary transformations, that is, unitary transformations of the form $U = U_A \otimes U_B$.

3. The entanglement $E(\rho)$ cannot increase under local operations and classical communication, that is

$$E(\mathcal{S}(\rho)) \leq E(\rho), \tag{5.2.6}$$

   where $\mathcal{S}$ is a superoperator that can be implemented with local quantum operations of the two parties A and B and an unlimited amount of classical communication between them.

4. The entanglement $E(\rho)$ is a convex function of ρ, i.e.

$$E(\rho = \sum_i p_i \rho_i) \leq \sum_i p_i \, E(\rho_i). \tag{5.2.7}$$

For pure states $|\psi\rangle$ the conventional measure that obeys the four requirements is the entropy of entanglement. It is defined as

$$E(|\psi\rangle\langle\psi|) = S(\mathrm{Tr}_A |\psi\rangle\langle\psi|) = S(\mathrm{Tr}_B |\psi\rangle\langle\psi|), \tag{5.2.8}$$

where S is the von Neumann entropy:

$$S(\rho) = -\mathrm{Tr}\, \rho \log \rho. \tag{5.2.9}$$

With this measure the EPR singlet in Eq. (5.2.3) has an entanglement of 1 bit, which is the maximum for a state in $2 \otimes 2$. For mixed states several entanglement measures have been proposed. One favorite measure that was introduced in Ref. [37] that obeys all the requirements is the *entanglement of formation*. The entanglement of formation for bipartite mixed states is more complicated than for pure states as the decomposition of a mixed state into a convex combination of pure states is not unique. Let ρ be a bipartite density matrix and let $\mathcal{E}_\rho = \{p_i \geq 0, |\psi_i\rangle\}$ be an ensemble into which ρ can be decomposed:

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|. \tag{5.2.10}$$

The entanglement of formation of ρ is defined as

$$E(\rho) = \min_{\mathcal{E}_\rho} \sum_i p_i E(|\psi_i\rangle\langle\psi_i|). \tag{5.2.11}$$

The entanglement of formation equals the minimal average amount of pure state entanglement that is needed to build the density matrix ρ. The minimization in Eq. (5.2.11) makes an analytical computation of the entanglement of formation of mixed states a nontrivial task. Only in $2 \otimes 2$ has the problem of determining the entanglement of formation of any density matrix been completely solved by Wootters [109].

One may require that a measure of entanglement has the additional property of additivity, which does not follow from properties 1-4. An entanglement measure E for bipartite states is additive when for any two density matrices ρ_1 on $\mathcal{H}_{A_1} \otimes \mathcal{H}_{B_1}$ and ρ_2 on $\mathcal{H}_{A_2} \otimes \mathcal{H}_{B_2}$ and $\rho = \rho_1 \otimes \rho_2$ on $\mathcal{H}_A \otimes \mathcal{H}_B$ where $\mathcal{H}_A = \mathcal{H}_{A_1} \otimes \mathcal{H}_{A_2}$ and $\mathcal{H}_B = \mathcal{H}_{B_1} \otimes \mathcal{H}_{B_2}$, the following holds:

$$E(\rho_1 \otimes \rho_2) = E(\rho_1) + E(\rho_2). \tag{5.2.12}$$

The entanglement of formation is certainly subadditive

$$E(\rho_1 \otimes \rho_2) \leq E(\rho_1) + E(\rho_2), \tag{5.2.13}$$

where the equality holds when one uses the optimal individual ensembles $\mathcal{E}_{\rho_1}$ and $\mathcal{E}_{\rho_2}$ in the decomposition of $\rho_1 \otimes \rho_2$. The entanglement for pure states can be shown to be additive, but it has not yet been proved that the entanglement of formation is additive for all density matrices. If the entanglement of formation were not additive then this would mean that the entanglement costs for making $\rho_1 \otimes \rho_2$ would be strictly less than the entanglement costs for making ρ_1 and ρ_2 separately. It is possible that the entanglement of formation obeys only the requirement of partial additivity, that is, for all $n = 1, 2, \ldots$ $E(\rho^{\otimes n}) = nE(\rho)$ for mixed states ρ [108].

5.2.2 Distillation of Quantum Entanglement

The sharing of quantum entanglement between two or more parties is a resource that for many quantum information processing tasks is more powerful than the sharing of classically correlated states. As mentioned in Chap. 1 section 1.2.3, an EPR state can be used to send quantum information via teleportation. The protocols that employ quantum communication to solve a classical communication complexity problem (sec. 1.2.2) can be replaced by protocols that start from sharing a set of entangled states, which are then used to teleport the quantum data. If in these protocols the two parties start out with a mixed entangled state, they will have to "purify" this state to a pure entangled state before using it in some protocol. This procedure is called distillation [36]. The allowed set of quantum operations in the distillation procedure is restricted to the class of superoperators that is implementable by local quantum operations (LO) and classical communication (CC). Let us give the definition of distillable entanglement:

Definition 4 *[37, 110] The* distillable entanglement *of a bipartite density matrix ρ on $\mathcal{H}_A \otimes \mathcal{H}_B$ with an unlimited amount of local operations and an unlimited amount of classical communication (LO+CC) is the maximum number $D(\rho)$ such that there exists a sequence of LO+CC* **TCP** *maps $\mathcal{S}_i$*

$$\mathcal{S}_i \colon B((\mathcal{H}_A \otimes \mathcal{H}_B)^{n_i}) \to B(\mathcal{K}_i \otimes \mathcal{K}_i), \tag{5.2.14}$$

with $n_i \to \infty$,

$$\frac{1}{n_i} \log \dim \mathcal{K}_i \to D(\rho), \tag{5.2.15}$$

and fidelity with respect to a maximally entangled state

$$\langle \Phi^+ | \, \mathcal{S}_i(\rho^{\otimes n_i}) \, | \Phi^+ \rangle \to 1, \tag{5.2.16}$$

where

$$\rho^{\otimes n_i} = \overbrace{\rho \otimes \ldots \otimes \rho}^{n_i}, \tag{5.2.17}$$

and

$$|\Phi^+\rangle = \frac{1}{\sqrt{\dim \mathcal{K}_i}} \sum_{j=1}^{\mathcal{K}_i} |jj\rangle. \tag{5.2.18}$$

A density matrix ρ is called distillable if we can distill a non-zero amount of maximally entangled states from an arbitrary number of copies of ρ.

In words, this definition says that a density matrix ρ is distillable by LO+CC if, when given a large number of copies of the density matrix ρ, there is a LO+CC procedure that maps these copies onto a set of states in a (smaller) Hilbert space $\mathcal{K}_i \otimes \mathcal{K}_i$ such that these remaining (distilled) states have a high fidelity with respect to a maximally entangled state–for example the state $|\Phi^+\rangle$–in $\mathcal{K}_i \otimes \mathcal{K}_i$. Note that we call a density matrix ρ distillable when *some* pure state entanglement can be distilled from it. If only a constant number of maximally entangled states can be distilled from an infinite number of copies of ρ, then $D(\rho) = 0$. We call such a density matrix distillable though. From property 3, Eq. (5.2.6), it follows that $D(\rho) \leq E(\rho)$; if $D(\rho)$ would exceed $E(\rho)$, we would have increased $E(\rho)$ by LO+CC.

It has been shown [111] that any entangled density matrix on $2 \otimes 2$ is distillable, in fact it was found that $D(\rho) > 0$ for all entangled density matrices on $2 \otimes 2$. In higher dimensions the problem of distillation has turned out to be complicated by the richer structure of the manifold of entangled states and their relation to positive linear maps.

5.2.3 Positive Linear Maps

The problem of deciding whether a bipartite density matrix ρ on $\mathcal{H}_A \otimes \mathcal{H}_B$ is entangled can be quite hard. It has been shown by the Horodeckis [112] that there exist an intimate connection between the classification of entangled density matrices and the theory of positive linear maps.

Let $\mathcal{S}\colon\ B(\mathcal{H}_n) \to B(\mathcal{H}_m)$ be a linear map. $\mathcal{S}$ is positive when $\mathcal{S}\colon\ B(\mathcal{H}_n)^+ \to B(\mathcal{H}_m)^+$, where $B(\mathcal{H}_n)^+$ denotes the set of positive semidefinite matrices on $\mathcal{H}_n$. Let id_k be the identity map on $B(\mathcal{H}_k)$. We define the map $\mathrm{id}_k \otimes \mathcal{S}\colon\ B(\mathcal{H}_k \otimes \mathcal{H}_n) \to B(\mathcal{H}_k \otimes \mathcal{H}_m)$ for $k = 1, 2, \ldots$ by

$$(\mathrm{id}_k \otimes \mathcal{S}) \left(\sum_i \sigma_i \otimes \tau_i \right) = \sum_i \sigma_i \otimes \mathcal{S}(\tau_i), \tag{5.2.19}$$

where $\sigma_i \in B(\mathcal{H}_k)$ and $\tau_i \in B(\mathcal{H}_n)$. The map $\mathcal{S}$ is k-positive when $\mathrm{id}_k \otimes \mathcal{S}$ is positive. The map $\mathcal{S}$ is completely positive when $\mathcal{S}$ is k-positive for all $k = 1, 2, \ldots$. Following Lindblad [113], the set of physical operations on a density matrix $\rho \in B(\mathcal{H}_n)^+$ is given by the set of completely positive trace-preserving maps $\mathcal{S}\colon\ B(\mathcal{H}_n) \to B(\mathcal{H}_m)$. Similarly as k-positive, one can define a k-copositive map. Let $T\colon B(\mathcal{H}_n) \to B(\mathcal{H}_n)$ be defined as matrix transposition in a chosen basis for $\mathcal{H}_n$, i.e.

$$(T(A))_{ij} = A_{ji}, \tag{5.2.20}$$

on a matrix $A \in B(\mathcal{H}_n)$. The map $\mathcal{S}$ is k-copositive when $\mathrm{id}_k \otimes (\mathcal{S} \circ T)$ is positive. A positive linear map $\mathcal{S}\colon\ B(\mathcal{H}_n) \to B(\mathcal{H}_m)$ is decomposable if it can be written as

$$\mathcal{S} = \mathcal{S}_1 + \mathcal{S}_2 \circ T, \tag{5.2.21}$$

where $\mathcal{S}_1\colon\ B(\mathcal{H}_n) \to B(\mathcal{H}_m)$ and $\mathcal{S}_2\colon\ B(\mathcal{H}_n) \to B(\mathcal{H}_m)$ are completely positive maps. It has been shown by Woronowicz [114] that all positive linear maps $\mathcal{S}\colon\ B(\mathcal{H}_2) \to B(\mathcal{H}_2)$ and $\mathcal{S}\colon\ B(\mathcal{H}_2) \to B(\mathcal{H}_3)$ are decomposable.

In Ref. [115] Peres made the observation that every separable density matrix $\rho \in B(\mathcal{H}_A \otimes \mathcal{H}_B)$ remains positive semidefinite under partial transposition of ρ, $(\mathrm{id}_A \otimes T_B)(\rho)$. He conjectured that this would not only be a necessary but also a sufficient condition for separability. His conjecture turned out to be true for density matrices on $2 \otimes 2$ and $2 \otimes 3$.

The following theorem by the Horodeckis [112] formulates a necessary and sufficient condition for a density matrix ρ to be entangled:

Theorem 1 (Horodecki) *A density matrix ρ on $\mathcal{H}_A \otimes \mathcal{H}_B$ is entangled iff there exists a positive linear map $\mathcal{S}\colon\ B(\mathcal{H}_B) \to B(\mathcal{H}_A)$ such that*

$$(\mathrm{id}_A \otimes \mathcal{S})\,(\rho), \tag{5.2.22}$$

is not positive semidefinite. Here id_A denotes the identity map on $B(\mathcal{H}_A)$.

Remark An equivalent statement as Theorem 1 holds for positive linear maps $\mathcal{S}\colon\ B(\mathcal{H}_A) \to B(\mathcal{H}_B)$ and the positive semidefiniteness of $(\mathcal{S} \otimes \mathrm{id}_B)(\rho)$.

The consequences of Theorem 1 and Woronowicz' result is that a bipartite density matrix ρ on $\mathcal{H}_2 \otimes \mathcal{H}_2$ and $\mathcal{H}_2 \otimes \mathcal{H}_3$ is entangled iff $(\mathrm{id}_A \otimes [\mathcal{S}_1 + \mathcal{S}_2 \circ T])\,(\rho)$ is not positive semidefinite for some $\mathcal{S}_1$ and $\mathcal{S}_2$. As $\mathcal{S}_1$ and $\mathcal{S}_2$ are completely positive maps this is equivalent to testing whether the requirement that $(\mathrm{id}_A \otimes T)\,(\rho)$ is not positive semidefinite is satisfied.

In the following sections we will sometimes refer to a density matrix having the NPT-property, which means that the density matrix is not positive semidefinite under partial transposition, or a density matrix having the PPT-property.

The partial transposition map is a powerful tool in characterizing entanglement, even in high dimensional Hilbert spaces. In Ref. [116] it was shown that if a bipartite density matrix ρ has the PPT-property, the density matrix *cannot* be distilled (see definition 4). It is not known whether the converse it true; all density matrices that have the NPT-property are distillable. There are indications that this might not be the case.

In Ref. [117] P. Horodecki found the first examples of density matrices on $\mathcal{H}_2 \otimes \mathcal{H}_4$ and $\mathcal{H}_3 \otimes \mathcal{H}_3$ that are provably entangled, but remain positive semidefinite under the partial transposition map. These states which are not distillable are called *bound entangled* states. In sec. 5.4 we will present many new examples of bound entangled states and show how their construction is intimately connected with LO+CC distinguishability of sets of orthogonal product states. In section 5.5 we show how this new class of bound entangled states gives rise to a new family of indecomposable positive linear maps.

5.3 Bell Inequalities and the Separability Criterion

We will start by reproducing a lemma of [112]. This lemma expresses a necessary and sufficient condition for separability of a bipartite density matrix:

Lemma 3 *[112] A density matrix $\rho \in B(\mathcal{H}_A \otimes \mathcal{H}_B)^+$ is entangled iff there exists a Hermitian operator $\mathrm{H} \in B(\mathcal{H}_A \otimes \mathcal{H}_B)$ with the properties:*

$$\mathrm{Tr}\,\mathrm{H}\,\rho < 0 \quad \text{and} \quad \mathrm{Tr}\,\mathrm{H}\,\sigma \geq 0, \tag{5.3.1}$$

for all separable density matrices $\sigma \in B(\mathcal{H}_A \otimes \mathcal{H}_B)^+$.

The lemma follows from basic theorems in convex analysis [118]. The proof invokes the existence of a separating hyperplane between the closed convex set of separable density matrices on $\mathcal{H}_A \otimes \mathcal{H}_B$ and a point, the entangled density matrix ρ, that does not belong to it. This separating hyperplane is characterized by the vector H that is normal to it; the hyperplane is the set of density matrices τ such that $\mathrm{Tr}\,\mathrm{H}\,\tau = 0$.

From a physics point of view, the Hermitian operator H is the observable that would reveal the entanglement of a density matrix ρ. We will call H an entanglement witness. The lemma tells us that there exists such an observable H for any entangled bipartite density matrix. Thus, if one can prove that there exists no such observable for a density matrix ρ, it follows that ρ must be separable.

We now turn to the formulation of Bell inequalities. The question of whether quantum mechanics provides a complete description of reality underlies the formulation of Bell's original inequality [119]. The issue is whether the results of measurements can be described by assuming the existence of a classical local hidden variable. The variable is hidden as its value cannot necessarily be measured directly; the average outcome of any measurement is a statistical average over different values that this hidden variable can take. The locality of this variable is required by the locality of classical physics [1]. Bell demonstrated that for the state in Eq. (5.2.3), the EPR singlet state, there exists a set of local measurements performed by two parties, Alice and Bob, whose outcomes *cannot* be described by any local hidden variable theory. The first experimental verification of his result with independently chosen measurements for Alice and Bob was carried out by Alain Aspect [120]. Since Bell's result, much attention has been devoted to finding stronger "Bell inequalities ", that is, inequalities that demonstrate the nonlocal character of other entangled states, pure *and* mixed. It has been found that any bipartite pure entangled state violates some Bell inequality [62]. The situation for mixed states is less clear. Multiple copies of bipartite mixed states that can be distilled (see definition 4) will violate a Bell inequality. The distillability makes it possible to map these states onto pure entangled states after which a pure-state Bell-inequality test will reveal their nonlocal character. But there are many entangled states, such as the ones that we will introduce and discuss in section 5.4.3 for which it is not known whether they violate a Bell inequality.

Interestingly, the general formulation of Bell inequalities [121, 122, 123] has great similarity with the separability criterion of Lemma 3 and there exists a relation between the two.

The general formulation of Bell inequalities comes about in the following way. We will consider only bipartite states here, but the formulation also holds for multipartite states. Let $\mathcal{M}_1^A, \ldots \mathcal{M}_{n_A}^A$ be a set of possible measurements for Alice and $\mathcal{M}_1^B, \ldots \mathcal{M}_{n_B}^B$ be a set of measurements for Bob. For simplicity let us consider measurements in which each outcome corresponds to a single operation element (see section 3.2, Eq. (3.2.6)). The analysis is completely analogous for measurements with more than one operation element per outcome. Thus each measurement is characterized by its operation elements corresponding to its possible outcomes. We write for the ith Alice measurement with k outcomes,

$$\mathcal{M}_i^A = (A_{i,1}, A_{i,2}, \ldots, A_{i,k(i)}), \quad \textstyle\sum_{m=1}^{k(i)} A_{i,m}^\dagger A_{i,m} = \mathbf{1}, \tag{5.3.2}$$

[1] No information can travel faster than the speed of light.

and similarly for the jth measurement of Bob,

$$\mathcal{M}_j^B = (B_{j,1}, B_{j,2}, \ldots, B_{j,l(j)}), \quad \sum_{m=1}^{l(j)} B_{j,m}^\dagger B_{j,m} = \mathbf{1}. \tag{5.3.3}$$

Let $\vec{P}$ be a vector of probabilities of outcomes of measurements by Alice and Bob on a quantum state ρ. The vector $\vec{P}$ has three parts denoted with the components $(P_{A:i|k,B:j|l}, P_{A:i|k}, P_{B:j|l})$. For example, when Alice has two measurements with two outcomes each and Bob has one measurement with three outcomes, $\vec{P}$ will be a 12+4+3 component vector with its components equal to

$$\begin{aligned} P_{A:i|k,B:j|l} &= \mathrm{Tr}\, E_{i,k}^A \otimes E_{j,l}^B\, \rho, \\ P_{A:i|k} &= \mathrm{Tr}\, E_{i,k}^A \otimes \mathbf{1}\, \rho, \\ P_{B:j|l} &= \mathrm{Tr}\, \mathbf{1} \otimes E_{j,l}^B\, \rho, \end{aligned} \tag{5.3.4}$$

with $E_{i,k}^A = A_{i,k}^\dagger A_{i,k}$ for $i = 1, 2$, $k = 1, 2$ and $E_{j,l}^B = B_{j,l}^\dagger B_{j,l}$ for $j = 1$, $l = 1, 2, 3$. We call $\vec{P}$ the event vector.

Let λ be a local hidden variable. We choose λ such that when λ takes a specific value, each measurement outcome is made either impossible or made to occur with probability 1. In other words, given a value of λ a probability of either 0 or 1 is assigned to Alice's outcomes and similarly for Bob. Then we choose λ to take as many values as are needed to produce all possible patterns of 0s and 1s, all Boolean vectors. These outcome patterns are denoted as Boolean vectors $\vec{B}_\lambda^A$ and $\vec{B}_\lambda^B$. For example, when Alice has three measurements each with two outcomes there will be 2^6 vectors $\vec{B}_\lambda^A \in \{0,1\}^6$. The vector $\vec{B}_\lambda^A$ has ofcourse the same number of entries as Alice's part of the event vector $\vec{P}_A$ and similarly for Bob. The locality constraint comes in by requiring that the vector of joint probabilities $\vec{B}_\lambda^{AB}$ is a product vector, i.e. $\vec{B}_\lambda^{AB} = \vec{B}_\lambda^A \otimes \vec{B}_\lambda^B$. The total vector is denoted as $\vec{B}_\lambda = (\vec{B}_\lambda^{AB}, \vec{B}_\lambda^A, \vec{B}_\lambda^B)$. An example will serve to elucidate the idea. When, as before, Alice has two measurements each with two outcomes and Bob has one measurement with three outcomes, an example of the vector $\vec{B}_\lambda$ is

$$\vec{B}_\lambda = [(\overbrace{1,0,}^{\mathcal{M}_1^A}\overbrace{0,1}^{\mathcal{M}_2^A}) \otimes (\overbrace{0,1,0}^{\mathcal{M}_1^B}), (1,0,0,1), (0,1,0)]. \tag{5.3.5}$$

We denote the vector $\vec{B}_{\lambda=\lambda_1}$, when λ takes the value λ_1 as $\vec{B}_{\lambda_1}$. Any local hidden variable theory can be represented as a vector $\vec{V}$:

$$\vec{V} = \sum_i p_i \left(\vec{P}_i^A \otimes \vec{P}_i^B, \vec{P}_i^A, \vec{P}_i^B \right), \tag{5.3.6}$$

with $p_i \geq 0$ and $\vec{P}_i^A$ and $\vec{P}_i^B$ are vectors of (positive) probabilities. These vectors are convex combinations of the vectors $\vec{B}_{\lambda_1}, \ldots, \vec{B}_{\lambda_N}$, where N is such that $\vec{B}_\lambda^A$ and $\vec{B}_\lambda^B$ are all possible Boolean vectors (see Ref. [123]):

$$\vec{V} = \sum_i q_i \vec{B}_{\lambda_i}, \tag{5.3.7}$$

with $q_i \geq 0$. Thus we see that the set of local hidden variable theories forms a convex cone $L_{LHV(\mathcal{M})}$. The label $\mathcal{M}$ is a reminder that the cone depends on the chosen measurements for Alice or Bob, in particular the number of them and the number of outcomes for each of them. The vectors $\vec{B}_{\lambda_i}$ are the extremal rays [121] of $L_{LHV(\mathcal{M})}$. The question then of whether the probabilities of the outcomes of the chosen set of measurements on a density matrix ρ can be reproduced by a local hidden variable theory, is equivalent to the question whether or not

$$\vec{P} \in L_{LHV(\mathcal{M})}. \tag{5.3.8}$$

It is not hard to see that all separable *pure* states have event vectors $\vec{P} \in L_{LHV(\mathcal{M})}$ as the event vector $\vec{P}$ for a separable pure state has a product structure $\vec{P} = (\vec{P}_A \otimes \vec{P}_B, \vec{P}_A, \vec{P}_B)$. It follows that all separable states have event vectors in $L_{LHV(\mathcal{M})}$, as they are convex combinations of separable pure states. What about the entangled states? We can use the Minkowski-Farkas lemma for convex sets in $\mathbf{R}^n$ [118]. The lemma implies that $\vec{P} \notin L_{LHV(\mathcal{M})}$ iff there exists a vector $\vec{F}$ such that

$$\vec{F} \cdot \vec{P} < 0 \quad \text{and} \;\; \forall \; \lambda_i \;\left[\vec{F} \cdot \vec{B}_{\lambda_i} \geq 0\right]. \tag{5.3.9}$$

The equation $\forall \; \lambda_i \;\left[\vec{F} \cdot \vec{B}_{\lambda_i} \geq 0\right]$ is a Bell inequality. The equation $\vec{F} \cdot \vec{P} < 0$ corresponds to the violation of a Bell inequality. Thus, finding a set of measurements and exhibiting the vector $\vec{F}$ with the properties of Eq. (5.3.9) is equivalent to finding a violation of a Bell inequality. If one can prove that for a density matrix ρ no such sets of inequalities of the form Eq. (5.3.9) for all possible measurement schemes can be found, then it follows that ρ can be described by a local hidden variable theory. This concludes our discussion of the literature on the general formulation of Bell inequalities.

There is a nice correspondence between Eq. (5.3.9) and Lemma 3, captured in the following construction: Given a (Farkas) vector $\vec{F}$ of Eq. (5.3.9) and a set of measurements $\mathcal{M}$ for a bipartite entangled state ρ, one can construct an entanglement witness for ρ as in Lemma 3. Denote the components of the Farkas vector $\vec{F}$ as $(F_{A:i|k,B:j|l}, F_{A:i|k}, F_{B:j|l})$. Then

$$\mathrm{H} = \sum_{i,k,j,l} F_{A:i|k,B:j|l} E^A_{i,k} \otimes E^B_{j,l} + \sum_{i,k} F_{A:i|k} E^A_{i,k} \otimes \mathbf{1} + \sum_{j,l} F_{B:j|l} \mathbf{1} \otimes E^B_{j,l}, \tag{5.3.10}$$

where $E^A_{i,k} = A^\dagger_{i,k} A_{i,k}$, and $A_{i,k}$ are the operation elements of the ith measurement with outcome k for Alice and similarly for Bob. With this construction $\vec{F} \cdot \vec{P} = \mathrm{Tr}\,\mathrm{H}\,\rho$. Also, one has $\mathrm{Tr}\,\mathrm{H}\,\sigma \geq 0$ for any separable density matrix σ as $\vec{P}_\sigma \in L_{LHV(\mathcal{M})}$ for all separable density matrices σ. Thus a violation of a Bell inequality for a bipartite density matrix ρ can be reformulated as a entanglement witness H for ρ. One may ask whether this relation holds in the opposite direction: Given an entanglement witness H for a bipartite density matrix ρ, does there exist a decomposition of H into a set of measurements and a vector $\vec{F}$ as in Eq. (5.3.10), that leads to a violation of a Bell inequality for ρ. The answer to this question seems to be negative for certain mixed states [124]. The reason for the discrepancy between the inequalities of Lemma 3 and Eq. (5.3.9) is that the hidden variable cone $L_{LHV(\mathcal{M})}$ contains more than just the separable states; it can also contain vectors which do *not* correspond to probabilities of outcomes of measurements on a quantum mechanical system. If quantum mechanics is correct then we will never find these sets of outcomes. An example of such an unphysical vector is the following. Let Alice perform two possible measurements on a two-dimensional system. Her first measurement $\mathcal{M}_1^A$ is a projection in the $\{|0\rangle, |1\rangle\}$ basis and her second measurement $\mathcal{M}_2^A$ is a projection in the $\{\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\}$ basis. The hidden variable cone $L_{LHV(\mathcal{M})}$ will contain vectors such as

$$\vec{B}_\lambda = [(\overbrace{1,0,}^{\mathcal{M}_1^A}\overbrace{0,1}^{\mathcal{M}_2^A}) \otimes (\ldots), (1,0,0,1), (\ldots)]. \tag{5.3.11}$$

This vector $\vec{B}_\lambda$ which assigns a probability 1 to outcome $|0\rangle$ *and* a probability 1 to outcome $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ cannot describe the outcome of these measurements on any quantum mechanical state ρ.

These unphysical vectors play an important role in the construction of hidden variable theories for entangled states: their importance is emphasized by the following observation. If we restrict the cone $L_{LHV(\mathcal{M})}$ to contain only vectors that are consistent with quantum mechanics, then we can prove that there exists a "violation of a Bell inequality" for *any* entangled state. By this we mean the following; We demand that all vectors in the set $L_{LHV(\mathcal{M})}$ correspond to sets of outcomes that can be obtained by measurements on a quantum mechanical system in $\mathcal{H}_A \otimes \mathcal{H}_B$. Here $\mathcal{H}_A \otimes \mathcal{H}_B$ is the Hilbert space on which the density matrix that we would like to describe with a restricted local hidden variable theory is defined. We can call this restricted local hidden variable theory a local quantum mechanical hidden variable theory. One can prove that in this restricted scenario, there will always be a set of measurements under which ρ reveals its nonlocality and its entanglement:

Theorem 2 *Let ρ be a bipartite density matrix on $\mathcal{H}_A \otimes \mathcal{H}_B$. The density matrix ρ is separable iff there exists a restricted local hidden variable theory of ρ.*

Proof The idea of the proof is the following. All vectors in the restricted local hidden variable theory now correspond to outcomes of measurements on a quantum mechanical system. We chose a set of measurements that completely determines a quantum state in a given Hilbert space. Then there is a 1-1 correspondence between vectors of measurement outcomes and quantum states. Then we show that all vectors in the restricted local hidden variable set correspond to measurement outcomes of separable states. Therefore measurement outcomes from entangled states do not lie in the set described by a restricted local hidden variable theory.

We write the density matrix ρ as

$$\rho = \sum_{i,j} \mu_{ij} \sigma_i \otimes \tau_j + \sum_i \mu_i^A \sigma_i \otimes \mathbf{1} + \sum_j \mu_j^B \mathbf{1} \otimes \tau_j, \tag{5.3.12}$$

where the Hermitian matrices $\{\sigma_i \otimes \tau_j\}_{i=1,j=1}^{d_A^2-1,d_B^2-1}$, $\{\sigma_i \otimes \mathbf{1}\}_{i=1}^{d_A^2-1}$, $\{\mathbf{1} \otimes \tau_j\}_{j=1}^{d_B^2-1}$ with $d_A = \dim \mathcal{H}_A$ etc., form a basis for the Hermitian operators on $\mathcal{H}_A \otimes \mathcal{H}_B$. Let $|w_{i,k}^A\rangle$ be the eigenvectors of the matrix σ_i and $|w^B_{j,l}\rangle$ be the eigenvectors of τ_j. The projector onto the state $|w^A_{i,k}\rangle$ is denoted as $\pi_{w^A_{i,k}}$ and similarly, the projector onto the state $|w^B_{j,l}\rangle$ is denoted as $\pi_{w^B_{j,l}}$.

Alice and Bob choose a set of measurements such that the probabilities of outcomes of these measurements are given by

$$\begin{aligned} &\mathrm{Tr}\,\pi_{w^A_{i,k}} \otimes \pi_{w^B_{j,l}}\,\rho = p_{i,k,j,l},\\ &\mathrm{Tr}\,\pi_{w^A_{i,k}} \otimes \mathbf{1}\,\rho = p^A_{i,k},\\ &\mathrm{Tr}\,\mathbf{1} \otimes \pi_{w^B_{j,l}}\,\rho = p^B_{j,l}\ , \end{aligned} \tag{5.3.13}$$

for all i, k, j and l. In order to construct these POVMs they may get outcomes whose probability is given by other expressions than Eq. (5.3.13). What is important is that they, if they would carry out these measurements repeatedly on ρ (a single measurement on each copy of ρ), would be able to determine the probabilities $(p_{i,k,j,l}, p^A_{i,k}, p^B_{j,l})$. Then they can uniquely infer from these probabilities the state ρ. We call this set of measurements $\mathcal{M}_c$, a complete set of measurements. Let $L^r_{LHV(\mathcal{M}_c)}$ be the convex set of restricted local hidden variable theories [2]. We first consider which density matrices ρ can be described by restricted local hidden variable vectors of the form $(\vec{P}_A \otimes \vec{P}_B, \vec{P}_A, \vec{P}_B)$, where $\vec{P}_A$ $(\vec{P}_B)$ is a vector of probabilities $p^A_{i,k}$ $(p^B_{j,l})$. The density matrix $\rho = \rho_A \otimes \rho_B$ where $\rho_A = \mathrm{Tr}_B\rho$ and $\rho_B = \mathrm{Tr}_A\rho$ is a solution of the equations

$$\begin{aligned} &\mathrm{Tr}\,\pi_{w^A_{i,k}} \otimes \pi_{w^A_{j,l}}\,\rho = p^A_{i,k}\,p^B_{j,l},\\ &\mathrm{Tr}\,\pi_{w^A_{i,k}} \otimes \mathbf{1}\,\rho = p^A_{i,k},\\ &\mathrm{Tr}\,\mathbf{1} \otimes \pi_{w^B_{j,l}}\,\rho = p^B_{j,l}\ , \end{aligned} \tag{5.3.14}$$

for all i, k, j and l, since

$$\mathrm{Tr}\,\pi_{w^A_{i,k}} \otimes \pi_{w^B_{j,l}}\,\rho = \mathrm{Tr}\,\pi_{w^A_{i,k}} \otimes \pi_{w^B_{j,l}}(\rho_A \otimes \rho_B), \tag{5.3.15}$$

As the set of measurements completely determines the density matrix ρ it follows that the solution $\rho = \rho_A \otimes \rho_B$ is the only solution of Eq. (5.3.14) for all i, k, j and l . Therefore all the restricted local variable vectors of the form $(\vec{P}_A \otimes \vec{P}_B, \vec{P}_A, \vec{P}_B)$ correspond to separable states. If follows that any convex combination of the restricted local hidden variable vectors $\vec{V} = \sum_i p_i\,(\vec{P}^i_A \otimes \vec{P}^i_B, \vec{P}^i_A, \vec{P}^i_B)$ corresponds to a separable state. As the map from the vectors $\vec{P}$ to states ρ is 1-1, this is the only density matrix that corresponds to $\vec{V}$. Thus we can conclude that no vector in the convex set $L^r_{LHV(\mathcal{M}_c)}$ corresponds to an entangled state. On the other hand the outcome vector of any separable density matrix lies in $L^r_{LHV(\mathcal{M}_c)}$ by the argument given below Eq. (5.3.8). This completes the proof. □

We are now ready to clarify the relation between the separability criterion, Lemma 3, and Bell inequalities. Theorem 2 shows that $L^r_{LHV(\mathcal{M}_c)}$ only contains outcome vectors of separable

[2]Note that $L^r_{LHV(\mathcal{M}_c)}$ is a set and not a cone, as $\vec{V} \in L^r_{LHV(\mathcal{M}_c)}$ does not imply that $\lambda\vec{V} \in L^r_{LHV(\mathcal{M}_c)}$ with $\lambda > 0$, as we now require that all vectors in $\vec{V}$ correspond to probabilities of outcomes of measurements on a quantum mechanical system.

states. We decompose the entanglement witness H in terms of the vectors $|w^A_{i,k}\rangle$ and $|w^B_{j,l}\rangle$, given by $\mathcal{M}_c$:

$$\mathrm{H} = \sum_{i,k,j,l} F_{A:i|k,B:j|l} \pi_{w^A_{i,k}} \otimes \pi_{w^B_{j,l}} + \sum_{i,k} F_{A:i|k} \pi_{w^A_{i,k}} \otimes \mathbf{1} + \sum_{j,l} F_{B:j|l} \mathbf{1} \otimes \pi_{w^B_{j,l}}. \tag{5.3.16}$$

This is always possible as the set $\{\sigma_i \otimes \tau_j\}_{i=1,j=1}^{d_A^2-1,d_B^2-1}$,$\{\sigma_i \otimes \mathbf{1}\}_{i=1}^{d_A^2-1}$, $\{\mathbf{1} \otimes \tau_j\}_{j=1}^{d_B^2-1}$ forms a basis for the Hermitian operators on $\mathcal{H}_A \otimes \mathcal{H}_B$. The coefficients $(F_{A:i|k,B:j|l}, F_{A:i|k}, F_{B:j|l})$ are real and are identified with the components of the vector $\vec{F}$. We then have an equivalence between the separability criterion and a "violation of a Bell inequality" with restricted local hidden variables:

$$\begin{array}{rcl} \vec{F} \cdot \vec{P} & = & \mathrm{Tr}\, \mathrm{H}\, \rho, \\ & \text{and} & \\ \forall \vec{V} \in L^r_{LHV(\mathcal{M}_c)},\ \vec{F} \cdot \vec{V} \geq 0 & \Leftrightarrow & \forall \text{ separable } \sigma, \mathrm{Tr}\, \mathrm{H}\, \sigma \geq 0. \end{array} \tag{5.3.17}$$

To conclude, we have been able to show that there is an equivalence relation between the separability criterion and a weak form of Bell inequality, namely one that assumes that the variables take a restricted set of values, consistent with quantum mechanics. The analysis as presented does not resolve the question whether all entangled states violate a Bell inequality in the strong sense, one where the variable can take 'unphysical' values.

5.4 Product Bases, Local Distinguishability and Bound Entanglement

5.4.1 Nonlocality without Entanglement

The EPR singlet, Eq. (5.2.3), or any other pure entangled state, is a prime demonstration of the nonlocality of quantum mechanics. Its entanglement is an asset in protocols such a teleportation and its intrinsic nonlocal character is demonstrated by its violation of a Bell inequality. It is tempting to think that only entangled states, pure or mixed, exhibit some form of nonlocality. Reality however is more subtle than this. In Ref. [125] it was demonstrated that there exists a form of quantum nonlocality that does not need entanglement. The authors of [125] presented a set of nine *orthogonal product* states in a bipartite $3 \otimes 3$ Hilbert space:

$$\begin{array}{ll} |v_0\rangle = \frac{1}{\sqrt{2}}|0\rangle|0-1\rangle, & |v_1\rangle = \frac{1}{\sqrt{2}}|0\rangle|0+1\rangle, \\ |v_3\rangle = \frac{1}{\sqrt{2}}|2\rangle|1-2\rangle, & |v_4\rangle = \frac{1}{\sqrt{2}}|2\rangle|1+2\rangle, \\ |v_5\rangle = \frac{1}{\sqrt{2}}|0-1\rangle|2\rangle, & |v_6\rangle = \frac{1}{\sqrt{2}}|0+1\rangle|2\rangle, \\ |v_7\rangle = \frac{1}{\sqrt{2}}|1-2\rangle|0\rangle, & |v_8\rangle = \frac{1}{\sqrt{2}}|1+2\rangle|0\rangle, \end{array} \tag{5.4.1}$$

where $\frac{1}{\sqrt{2}}|0-1\rangle$ denotes $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ etc. Here and further in the text tensorproducts $\otimes$ are sometimes omitted; the state $|\psi_a, \psi_b\rangle$ is equivalent with $|\psi_a\rangle|\psi_b\rangle$ or $|\psi_a\rangle \otimes |\psi_b\rangle$. Two parties, Alice and Bob, are given a single copy of one of these nine states, but they are not told which one they are given. Their task is to determine which one of the nine states they are given by performing local measurements on the state and communicating classically to each other about the outcomes. As these states are mutually orthogonal they *can* be distinguished when a measurement is done on the joint system of Alice and Bob. It was shown however that it is not possible for the two parties to find out with certainty which state they were given even if they could use an unlimited amount of classical communication and could perform an unlimited number of local measurements and other computational operations. It is not even possible to get the right answer with arbitrary small probability of error. What is important is that these states, being orthogonal product states, do not exhibit any entanglement at all. They form an example of a phenomenon that one could call nonlocality without entanglement.

There could be an interesting use of such states, that surpasses anything that can be done in a strictly classical world. These states could be used in a protocol of *secret sharing*. Consider the following situation. Alice and Bob are given a secret by a third authorized party Charlie. The idea of the secret sharing is that Alice and Bob are not able to determine the secret alone. For example the American government (the authorized party) lets two employees at Los Alamos National Laboratory share the secret of new weapon. One of the employees is malevolent and the other one can be trusted to keep his part of the secret. The malevolent employee needs information of the trusted employee in order to determine the secret. The trusted employee refuses to reveal the information that he/she has to the malevolent employee. In this scenario if both the employees are malevolent, it is impossible to keep the secret safe, if we do not assume any restrictions on the computational resources and the cunning of the employees.

In a quantum world it might be possible that two malevolent parties are able to keep a secret if their communication is restricted to the transmission of classical messages. Such a quantum scenario that uses entanglement has been investigated in Ref. [126]. Here we propose a protocol that does *not* use entanglement between the sharing parties. The idea would be the following. The secret is encoded as a word in an alphabet with nine letters, the nine states. We know that the two parties are not able to distinguish between one of the nine states exactly. However, if they would be allowed to send each other quantum data, they are able to uncover the secret. In order for the protocol to be absolutely safe, one will have to be able to show that the two parties will obtain less than a certain small amount of information about the secret for any attack that they can carry out. The establishment of such a protocol and the proof of its safety is a question of current research.

5.4.2 Unextendible Product Bases

We introduce a new concept, that of an unextendible product basis:

Definition 5 *Consider a multipartite quantum system $\mathcal{H} = \bigotimes_{i=1}^{m} \mathcal{H}_i$ with m parties of respective dimension $d_i, i = 1, ..., m$. A (partial orthogonal)* product basis (PB) *is a set S of pure orthogonal product states spanning a proper subspace $\mathcal{H}_S$ of $\mathcal{H}$. An* unextendible product

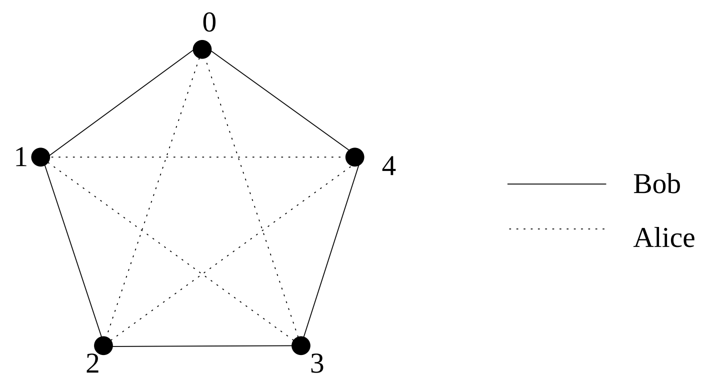

Figure 5.1: The UPB **Pent** *and* **Tiles** represented as two "color" graphs.

basis (UPB) *is a* PB *whose complementary subspace* $\mathcal{H}_S^\perp$ *contains no product state.*

Here are two examples of a UPB on $3 \otimes 3$ (two qutrits):

Example 1: Consider five vectors in real three-dimensional space forming the apex of a regular pentagonal pyramid, the height h of the pyramid being chosen such that nonadjacent apex vectors are orthogonal. The vectors are

$$\vec{v}_i = N(\cos\frac{2\pi i}{5}, \sin\frac{2\pi i}{5}, h), \quad i = 0, \ldots, 4, \tag{5.4.2}$$

with $h = \frac{1}{2}\sqrt{1+\sqrt{5}}$ and $N = 2/\sqrt{5+\sqrt{5}}$. The following five states in $3 \otimes 3$ form the UPB **Pent**

$$\vec{p}_i = \vec{v}_i \otimes \vec{v}_{2i \bmod 5}, \quad i = 0, \ldots, 4. \tag{5.4.3}$$

Example 2: The following five states on $3 \otimes 3$ form the UPB **Tiles**

$$\begin{array}{ll} |v_0\rangle = \frac{1}{\sqrt{2}}|0\rangle|0-1\rangle, & |v_2\rangle = \frac{1}{\sqrt{2}}|2\rangle|1-2\rangle, \\ |v_1\rangle = \frac{1}{\sqrt{2}}|0-1\rangle|2\rangle, & |v_3\rangle = \frac{1}{\sqrt{2}}|1-2\rangle|0\rangle, \\ \multicolumn{2}{c}{|v_4\rangle = (1/3)|0+1+2\rangle|0+1+2\rangle.} \end{array} \tag{5.4.4}$$

The first four states are the interlocking tiles of [125], Eq. (5.4.1), and the fifth state works as a "stopper" to force the unextendibility. (In fact, the sets **Pent** and **Tiles** are both members of a single six-parameter family of UPBs [129])

The orthogonality relations between the members of the set, both for **Pent** as well as **Tiles**, are depicted in Fig. 5.1. The states are given as vertices in the graph. When two vertices are connected by, say, a Bob-edge, it means that the states are orthogonal on Bob's side and similar for Alice.

In both these examples one can observe that any subset of three vectors on either side spans the full three-dimensional space. This implies that there cannot be a product vector orthogonal to these states and thus both these PBs are UPBs.

We can formalize the way in which these two examples were constructed to give a necessary and sufficient condition for a PB on a multipartite system to be a UPB:

Lemma 4 *Let P be a partition of a* PB *S into disjoint sets $S = S_1 \cup S_2 \cup \ldots S_m$ where S_i is a set of states associated with the i^{th} party. Let π_j be the projector onto the j^{th} state in S. Let $\rho_i^P = \sum_{j \in S_i} \mathrm{Tr}_{(\bigotimes_{k \neq i} \mathcal{H}_k)} \pi_j$. The set S forms a* UPB *on $\bigotimes_{i=1}^m \mathcal{H}_i$ iff for all partitions P at least one ρ_i^P has full rank, equal to d_i.*

Proof If for all partitions P there is a local ρ_i^P with full rank, then it is not possible to add a new product state to S that is orthogonal to all the members in S. If a set of states S forms a UPB, but there exists a partition P for which all ρ_i^P have less than full rank, then we arrive at a contradiction, as we can add a product state to the UPB in the following way. One takes the partition P that gives rise to the matrices ρ_i^P that all have less than full rank d_i. Then a new state can be added that is orthogonal to the states in S_1 on $\mathcal{H}_1$, the states in S_2 on $\mathcal{H}_2$ etc., such that this new state is orthogonal to all the members in S. □

In principle Lemma 4 can be used recursively to explore whether a set of states S can be completed to a full basis, but it is not known whether there exist an efficient algorithm that performs this task.

The lemma provides a simple lower bound on the number of states k in a UPB :

$$k \geq \sum_i (d_i - 1) + 1. \tag{5.4.5}$$

If k is equal to $\sum_i (d_i - 1)$ or smaller then one can partition S into sets of size $|S_i| \leq d_i - 1$. This partition has the property that $\mathrm{Rank}(\rho_i^P) < d_i$ for all i and therefore the set S cannot be a UPB.

5.4.3 Bound Entanglement

Bipartite UPBs lead directly to the construction of density matrices that have bound entanglement. These bipartite bound entangled states are positive semidefinite under partial transposition (PPT), although they are entangled. The PPT-property implies that no pure state entanglement can be distilled from these states.

Proposition 6 *Let S be a bipartite unextendible product basis $\{|\alpha_i\rangle \otimes |\beta_i\rangle\}_{i=1}^{|S|}$ in $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. We define a density matrix ρ_S as*

$$\rho_S = \frac{1}{\dim \mathcal{H} - |S|} \left(\mathbf{1}_{AB} - \sum_{i=1}^{|S|} |\alpha_i\rangle\langle\alpha_i| \otimes |\beta_i\rangle\langle\beta_i| \right), \tag{5.4.6}$$

where $\mathbf{1}_{AB}$ is the identity on $\mathcal{H}$. The density matrix ρ_S is entangled. Furthermore, the matrix $(\mathrm{id}_A \otimes [\mathcal{S}_1 + T \circ \mathcal{S}_2])(\rho_S)$ for all completely positive maps $\mathcal{S}_1$ and $\mathcal{S}_2$, is positive semidefinite.

Proof The density matrix ρ_S is proportional to the projector on the complementary subspace $\mathcal{H}_S^\perp$. As S is unextendible $\mathcal{H}_S^\perp$ contains no product states. Therefore the density matrix is entangled. It is not hard to see that $(\mathrm{id}_A \otimes T)(\rho_S)$ is positive semidefinite. It has been proved in Ref. [116] that when $(\mathrm{id}_A \otimes T)(\rho_S)$ is positive semidefinite that $(\mathrm{id}_A \otimes T \circ \mathcal{S}_2)(\rho_S)$ where $\mathcal{S}_2$ is any completely positive map, is also positive semidefinite. Therefore $(\mathrm{id}_A \otimes [\mathcal{S}_1 + T \circ \mathcal{S}_2])(\rho_S)$ is also positive semidefinite. □

We now give an example of a tripartite UPB :

Example 3: Consider a set **Shifts** of orthogonal product states between three parties A, B, and C:

$$\{|0,1,+\rangle, |1,+,0\rangle, |+,0,1\rangle, |-,-,-\rangle\}, \tag{5.4.7}$$

with $\pm = |0 \pm 1\rangle$ (unnormalized). There is no product state that is orthogonal to these four states, as any subset of two states spans the full 2-dimensional space on one side. The complementary state constructed as in Eq. (5.4.6) has the curious property that it is 2-way separable, i.e., the entanglement between every split into two parties is zero. This refutes a conjecture that was made in Ref. [130]. To show that, for example, the entanglement between A and BC is zero, one writes the BC parts of the states in Eq. (5.4.7) as $a = |1,+\rangle$, $b = |+,0\rangle$, $c = |0,1\rangle$ and $d = |-,-\rangle$. Note that $\{a,b\}$ are orthogonal to $\{c,d\}$. Consider the vectors $a^\perp$ and $b^\perp$ in the $\mathrm{Span}(a,b)$ and the vectors $c^\perp$ and $d^\perp$ in the $\mathrm{Span}(c,d)$. Now, one can complete the original set of vectors to a full product basis between A and BC with the states $\{|0,a^\perp\rangle, |1,b^\perp\rangle, |+,c^\perp\rangle, |-,d^\perp\rangle\}$. By the symmetry of the states, this is also true for the other splits AB versus C and AC versus B. This implies that the density matrix $\rho_{\mathbf{Shifts}}$ constructed as in Eq. (5.4.6) of the UPB **Shifts** has multipartite bound entanglement. If the entanglement was distillable, then it would be possible to make entanglement over some bipartite split, say, A and BC. This is in contradiction with the fact that these states are 2-way separable and the entanglement cannot be created by local operations and classical communication.

This argument can be generalized to any multipartite UPB, even though the partial transposition criterion cannot be applied directly to a multipartite state. Let S be a multipartite UPB. The density matrix ρ_S derived as in Eq. (5.4.6) from S has bound entanglement. This follows from the fact that for any bipartite split on the multipartite system into a system 1 and a system 2 by grouping of the parties, the matrix $(\mathrm{id}_1 \otimes T_2)(\rho_S)$ is positive semidefinite . This implies that ρ_S considered as a density matrix on system 1 and system 2 is either separable or has bound entanglement. Then it follows that any global entanglement can never be distilled, as this distillation would create free entanglement over some bipartite split. Any entanglement in the density matrix ρ_S of a multipartite UPB S must therefore be bound.

General UPBs

It is possible to generalize the first three examples of UPBs to higher dimensions and more parties. We list some of these generalizations:

- **GenShifts**, a UPB on $\bigotimes_{i=1}^{2k-1}\mathcal{H}_2$ with $2k$ members. The first state is $|0,\dots 0,0\rangle$. The second is

$$|1,\psi_1,\psi_2,\dots,\psi_{k-1},\psi_{k-1}^{\perp},\dots,\psi_2^{\perp},\psi_1^{\perp}\rangle. \tag{5.4.8}$$

The states $|\psi_i\rangle$ and $|\psi_j\rangle$ for all $i\neq j$ are neither orthogonal nor identical. Also, $|\psi_i\rangle$ is neither orthogonal nor identical to the state $|0\rangle$ for all i. The other states in the UPB are obtained by (cyclic) right shifting the second state, i.e. the third state is

$$|\psi_1^{\perp},1,\psi_1,\psi_2,\dots,\psi_{k-1},\psi_{k-1}^{\perp},\dots,\psi_2^{\perp}\rangle. \tag{5.4.9}$$

These states are all orthogonal in the following way. The state $|0,\dots,0,0\rangle$ is special and it is orthogonal to all the other states as they all have a $|1\rangle$ for some party. Leaving this special state aside, all states are orthogonal to the next state, their first right-shifted state, by the orthogonality of $|\psi_{k-1}\rangle$ and $|\psi_{k-1}^{\perp}\rangle$. All states are orthogonal to the 2nd right-shifted state by the orthogonality of $|\psi_1\rangle$ and $|\psi_1^{\perp}\rangle$. The 3rd rightshifted state is made orthogonal with $|\psi_{k-2}\rangle$ and $|\psi_{k-2}^{\perp}\rangle$. We can continue this until the last $(2k-2)$th right-shifted state and we are done.

As there are no states repeated on one side in the UPB all sets of two states span a two-dimensional space and Lemma 4 implies that the set is a UPB .

- **GenTiles**, a bipartite PB on $n\otimes n$ where n is even. These states have a tile structure which in the case of $6\otimes 6$ is shown in Fig. 5.2. A tile represents one or more states. For example, the tile in the upperleft corner of Fig. 5.2 represents 2 states each of which is of the form

$$(\alpha_0|0\rangle+\alpha_1|1\rangle+\alpha_2|2\rangle)\otimes|0\rangle. \tag{5.4.10}$$

The general construction is the following. One labels a set of n orthonormal states as $|0\rangle,\dots,|n-1\rangle$. One takes $n(n/2-1)$ states of the form $|k\rangle\otimes|\omega_{m,k+1}\rangle$, $m=1,\dots,n/2-1$, and $k=0,\dots,n-1$ where $|\omega_{m,k}\rangle$ is defined as

$$|\omega_{m,k}\rangle=\sum_{j=0}^{n/2-1}\omega^{jm}|j+k \bmod n\rangle, \tag{5.4.11}$$

where $\omega=e^{i4\pi/n}$ and thus $\omega^{jm}=e^{i4\pi jm/n}$. Note that $\langle\omega_{m,k}|\omega_{n,k}\rangle$ is proportional to δ_{mn}. Similarly, one takes $|\omega_{m,k}\rangle\otimes|k\rangle$, $m=1,\dots,n/2-1$, and $k=0,\dots,n-1$. Then one adds the "stopper" which is the state $\sum_i|i\rangle\otimes\sum_j|j\rangle$. Note that the set has n^2-2n+1 states, which is much more than the minimum of Eq. (5.4.5). This construction can be proved to be a UPB in $4\otimes 4$ and $6\otimes 6$ by exhaustive checking of all partitions. This procedure runs into problems for arbitrary high dimension, but one may conjecture that

Conjecture 1 *The set of states* **GenTiles** *forms a* UPB *on $n\otimes n$ for all even $n\geq 4$.*

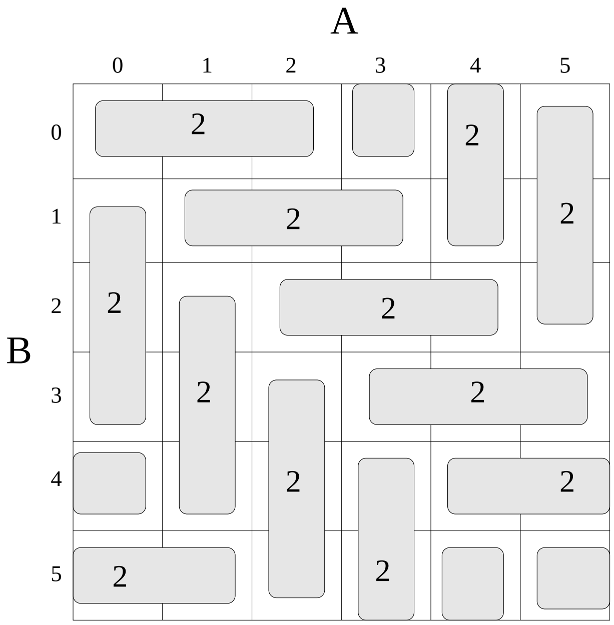

Figure 5.2: Tile structure of the bipartite $6 \otimes 6$ UPB.

- Tensor powers of UPBs. The following theorem holds:

Theorem 3 *Given two bipartite* UPBs *S_1 and S_2 with members $|\psi_i^1\rangle = |\alpha_i^1\rangle \otimes |\beta_i^1\rangle$, $i = 1, \ldots, l_1$ on $n_1 \otimes m_1$ and members $|\psi_i^2\rangle = |\alpha_i^2\rangle \otimes |\beta_i^2\rangle$, $i = 1, \ldots, l_2$ on $n_2 \otimes m_2$ respectively. The* PB *$\{|\psi_i^1\rangle \otimes |\psi_j^2\rangle\}_{i,j=1}^{l_1,l_2}$ is a bipartite* UPB *on $n_1 n_2 \otimes m_1 m_2$.*

Proof Assume the contrary, i.e. there is a product state that is orthogonal to this new ensemble which we call PB^2. The idea is to show that this leads to a contradiction and thus PB^2 is a UPB. Note first that for any UPB a partition P into a set with 0 states for Bob and all states for Alice give rise to a ρ_A^P (see Lemma 4) that has full rank; the states on Alice's side together must span the entire Hilbert space of Alice. Also note that if one takes a tensor product of two UPBs this partition in which all states are assigned to Alice still leads to a ρ_A^P that has full rank on Alice's side as $\mathrm{Rank}(\rho_1 \otimes \rho_2) = \mathrm{Rank}(\rho_1)\mathrm{Rank}(\rho_2)$.

The set PB^2 has $l_1 l_2$ members. The new hypothetical product state to be added to the set has to be orthogonal to each member either on Bob's side or on Alice's side, or on both sides. One can represent such an orthogonality pattern as a rectangle of size l_1 (number of columns) by l_2 (number of rows) filled with the letters A and B, depending on how the new state is orthogonal to a member of the PB^2, see Fig. 5.3. When this hypothetical state is orthogonal on both sides, we are free to choose an A or B in the corresponding square. Consider a row of this rectangle. The pattern of As and the Bs can be viewed as a partition of the S_1 UPB. For example in the partition of S_1 corresponding to the first row in Fig. 5.3 Alice gets the states $|\alpha_1^1\rangle$ and $|\alpha_3^1\rangle$ and Bob gets $|\beta_2^1\rangle$ and $|\beta_4^1\rangle, \ldots, |\beta_7^1\rangle$. Since S_1 is a UPB, either Alice's states span the full Hilbert space of dimension n_1 or Bob's states span the full Hilbert space of dimension m_1. Assume that Bob's states span the full Hilbert space of dimension m_1. Then ofcourse the states $|\beta_1^1\rangle$ and $|\beta_3^1\rangle$ lie in the space spanned by $|\beta_2^1\rangle$ and $|\beta_4^1\rangle, \ldots, |\beta_7^1\rangle$. Thus any hypothetical product state that is orthogonal to all states $|\beta_2^1\rangle \otimes |\psi_1^2\rangle$, $|\beta_4^1\rangle \otimes |\psi_1^2\rangle$, ..., $|\beta_7^1\rangle \otimes |\psi_1^2\rangle$, is also orthogonal to $|\beta_1^1\rangle \otimes |\psi_1^2\rangle$ and $|\beta_3^1\rangle \otimes |\psi_1^2\rangle$. Thus filling the whole row with Bs is a possible way to make the new state orthogonal. This argument can be applied to every row and every column, making them either all As or all Bs. This will eventually lead to a rectangle with only As or with only Bs. This implies however that PB^2 does not have full rank on either Alice's or Bob's side which is in contradiction with the original sets forming UPBs . □

L_1

L_2

A	B	A	B	B	B	B
A	B	B	A	A	A	A
B	B	A	A	A	B	B
A	A	B	B	A	B	A
B	B	A	A	B	A	B

Figure 5.3: The As and Bs denote on what side a hypothetical product state is orthogonal to the members of PB^2.

The theorem has the consequence that arbitrary tensor powers of bipartite UPBs are again UPBs . The theorem holds for multipartite states as well, where patterns of As and Bs are replaced by As, Bs, Cs etc.

- A generalization of the UPB **Pent** to $3 \otimes 3 \otimes 3$. Define the following states

$$\vec{v}_i = N(\cos\frac{2\pi i}{7}, \sin\frac{2\pi i}{7}, h), \quad i = 0, \ldots, 6, \tag{5.4.12}$$

with $h = \sqrt{-\cos\frac{4\pi}{7}}$ and $N = 1/\sqrt{1 + |\cos\frac{4\pi}{7}|}$. The following seven states in $3 \otimes 3 \otimes 3$ form the UPB **Sept**

$$\vec{p}_i = \vec{v}_i \otimes \vec{v}_{2i \bmod 7} \otimes \vec{v}_{3i \bmod 7}, \quad i = 0, \ldots, 6. \tag{5.4.13}$$

The orthogonality of these vectors $\vec{p}_i$ is shown in Fig. 5.4. To prove that these states form a UPB , we must show that any subset of three of them on one of the three sides (Lemma 4) spans the full three-dimensional space. As the vectors $\vec{v}_i$ form the apex of a regular septagonal pyramid, there is no subset of three of them that lies in a two-dimensional plane. It is not known whether the complementary state $\rho_{\mathbf{Sept}}$ has bipartite bound entanglement or whether it is a separable state over a bipartite split.

This construction can be extended to $3^{\otimes n}$; we have n parties and $p = 2n + 1$ states where p is a prime number. Thus one can have $(n, p) = (2, 5), (3, 7), (5, 11)$ etc. The states in the polygonal pyramid with p vertices are defined as

$$\vec{v}_i = N_p(\cos\frac{2\pi i}{p}, \sin\frac{2\pi i}{p}, h_p), \quad i = 0, \ldots, 2n. \tag{5.4.14}$$

In **Sept** and **Pent**, h_p was chosen such that nonadjacent vertices were orthogonal. For higher primes p one has to make a choice dependent on p. In order for the vectors to the vertex i and to the vertex $m + i$ to be made orthogonal by lifting the vectors out of the plane of the polygon, we must have

$$\frac{\pi}{2} \leq \frac{2\pi m}{p}(\leq \pi), \tag{5.4.15}$$

i.e. the angle between the vectors in the plane must be larger than 90 degrees. One can always find such an m given a p, for example, for $p = 7$, $m = 2$ or 3. With the choice of m one fixes h_p and N_p as

$$h_p = \sqrt{-\cos\tfrac{2\pi m}{p}}, \quad N_p = 1/\sqrt{1 + |\cos\tfrac{2\pi m}{p}|}. \tag{5.4.16}$$

Finally, the UPB is

$$\vec{p}_i = \vec{v}_i \otimes \vec{v}_{2i \bmod p} \otimes \ldots \otimes \vec{v}_{ni \bmod p}, \quad i = 0, \ldots, 2n. \tag{5.4.17}$$

The primality of p ensures that there are no states repeated on one side: if $ki \mod p = kj \mod p$ for some integers $i \neq j$ and some integer $k = 1, \ldots, 2n$ then this would imply that p is divisible. Orthogonality is also ensured by primality. As in Fig. 5.4 there will be a party for whom next neighbor states are orthogonal, there will be a party for whom every second neighbor states is orthogonal, etc. up to the nth neighbor. This implies that all vertices in the orthogonality graph are mutually connected (orthogonal). From basic three-dimensional geometry it follows that any set of three vectors has full rank when $h_p \neq 0$ and thus these generalized sets form UPBs .

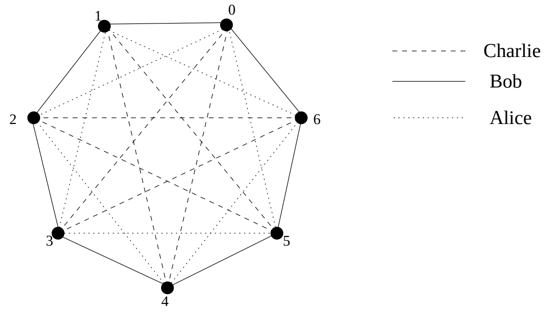

Figure 5.4: The **Sept** UPB on $3 \otimes 3 \otimes 3$.

- We would like to mention a conjecture by Peter Shor on the existence of a UPB based on quadratic residues. These are sets of orthogonal product states on $n \otimes n$ where n is such that $2n-1$ is a prime p of the form $4m+1$. Thus we can have $(m,p,n) = (1,5,3),(3,13,7)$ etc. The sets contains $p = 2n-1$ members, the minimal number for a UPB (Eq. (5.4.5)). Let $\mathbf{Z}_p^*$ be $\mathbf{Z}_p \setminus \{0\}$. Let Q_p be a group of quadratic residues, that is, elements $q \in \mathbf{Z}_p^*$ such that

$$q = x^2 \mod p, \tag{5.4.18}$$

for an integer x. The set Q_p is a group under multiplication. The order of the group is $\frac{p-1}{2}$. The following properties hold: when $q_1 \in Q_p$ and $q_2 \notin Q_p$, a quadratic nonresidue, then $q_1 q_2 \notin Q_p$. Also, if $q_1 \notin Q_p$ and $q_2 \notin Q_p$, then $q_1 q_2 \in Q_p$ [127]. The states of the UPB are

$$|Q(a)\rangle \otimes |Q(xa)\rangle \quad \text{for } a \in \mathbf{Z}_p, \quad x \in \mathbf{Z}_p^*, x \notin Q_p, \tag{5.4.19}$$

where

$$|Q(a)\rangle = (N,0,\ldots,0) + \sum_{q \in Q_p} e^{2\pi i q a/p} \hat{e}_q, \tag{5.4.20}$$

where N is a normalization constant to be fixed for orthogonality and $\hat{e_q}$ are unit vectors of the form $(0,1,0,\ldots,0)$, $(0,0,1,0,\ldots,0)$ etc. The dimension n of the Hilbert space is $\frac{p+1}{2}$, one more than the order of Q_p. One can prove that these vectors can be made orthogonal by an appropriate choice of N:

$$\begin{aligned} &\langle Q(a)|Q(b)\rangle\langle Q(xa)|Q(xb)\rangle = \\ &\qquad (|N|^2 + \sum_{q \in Q_p} e^{2\pi i q(b-a)/p})(|N|^2 + \sum_{q \in Q_p} e^{2\pi i q x(b-a)/p}). \end{aligned} \tag{5.4.21}$$

One uses the properties of Q_p to find that for $(b-a) \neq 0$:

$$\sum_{q \in Q_p} e^{2\pi i q(b-a)/p} + \sum_{q \in Q_p} e^{2\pi i q x(b-a)/p} = \sum_{z \in \mathbf{Z}_p^*} e^{2\pi i z(b-a)/p} = -1. \tag{5.4.22}$$

Thus the orthogonality relation of Eq. (5.4.21) for $b \neq a$ is of the form

$$(|N|^2 + s)(|N|^2 - 1 - s) = 0, \tag{5.4.23}$$

where

$$s = \sum_{q \in Q_p} e^{2\pi i q(b-a)/p}. \tag{5.4.24}$$

Note that s can take two values depending on whether $b-a$ is a quadratic residue or a quadratic nonresidue. In order to show that s is real, one considers s^* in which one sums over $-q$. As $q \in Q_p$ and $-1 \in Q_p$ when p is of the form $4m+1$ (see Theorem 82,[127]), we have that $-q \in Q_p$. Therefore $s = s^*$. Thus for all values that s can take, Eq. (5.4.23) has a solution for N.

Conjecture 2 (Shor) *The states given in Eq. (5.4.19) and Eq. (5.4.20) on $n \otimes n$ with $2n-1$ a prime of the form $4m+1$ with the appropriate value of N determined by the solution of Eq. (5.4.23) form a* UPB.

The proof will require the application of Lemma 4, that is, one must show that any set of n states on either side spans the full n-dimensional Hilbert space. This conjecture has been proved for $p=5$, $p=13$ and $p=17$. These sets form a generalization of the **Pent** UPB that was presented in section 5.4.2 and Figure 5.1. Drawn as graphs as in Fig. 5.1, they are regular polygons, with a prime number p (of the form $4m+1$) of vertices. The elements of the quadratic residue group Q_p correspond to the periodicity of the vectors that are orthogonal on one side. For example, when $p=13$, one has quadratic residues $1, 3, 4, 9, 10$ and 12. Thus on, say, Alice's side, every vertex is connected to its first neighbor (1), every vertex is connected with the 3rd neighbor (3) etc. On Bob's side the orthogonality pattern follows from the quadratic nonresidues.

5.4.4 Global versus Local Rank

The construction of bound entangled states based on UPBs suggests that bound entangled density matrices only come with a large rank. The idea is that when a basis is nearly complete, it is always possible to extend the basis and therefore our construction fails. The following theorem captures this observation and is relevant for any kind of bipartite bound entangled state with PPT. The theorem was conjectured by the author and proved together with P. Horodecki [128].

Theorem 4 *Let ρ be a bipartite density matrix on $\mathcal{H}_A \otimes \mathcal{H}_B$. Define $R_A = \mathrm{Rank}(\mathrm{Tr}_A \rho)$ and similarly R_B. Let R be the rank of ρ itself. If*

$$\max(R_A, R_B) > R, \tag{5.4.25}$$

then ρ is distillable.

Proof First of all, one can observe that when $\max(R_A, R_B) > R$ the state has to be entangled. Any separable state can be written as a convex combination of a set of product states $\{|\psi_i, \phi_i\rangle\}$. The number of linearly independent states $|\psi_i\rangle$ which determines R_A is a lower bound on the number of linearly independent states $|\psi_i, \phi_i\rangle$ which determines R, and similarly for R_B.

Without loss of generality let R_A be the largest local rank. Let $\rho_A = \mathrm{Tr}_A \rho$ in its diagonal form be $\mathrm{diag}(\lambda_1, \ldots, \lambda_{R_A}, 0, \ldots, 0)$. One can apply a local filter [131] on Alice's side to the state ρ

$$\rho_W = \frac{(W \otimes \mathbf{1})\, \rho\, (W^\dagger \otimes \mathbf{1})}{\mathrm{Tr}\, (W \otimes \mathbf{1})\, \rho\, (W^\dagger \otimes \mathbf{1})}, \tag{5.4.26}$$

where $W = \mathrm{diag}(1/\sqrt{\lambda_1}, \ldots, 1/\sqrt{\lambda_{R_A}}, 0, \ldots, 0)$ in the same basis as ρ_A. The filtering corresponds to the performance of a POVM measurement by Alice. The operation elements of her POVM measurements are cW and $\sqrt{\mathbf{1} - |c|^2 W^\dagger W}$ where c is chosen such that $\mathbf{1} - |c|^2 W^\dagger W$ has eigenvalues in the interval $[0,1]$. Then with probability $p_W = \mathrm{Tr}\,(cW \otimes \mathbf{1})\,\rho\,(c^* W^\dagger \otimes \mathbf{1})$ the state ρ_W is obtained and with probability $1 - p_W$ the filtering fails. Note that it is not a problem to have a certain probability of failure in a distillation protocol as one will have an arbitrary number of copies of the state. The reduced density matrix $\rho_{A,W}$ of the filtered state ρ_W has the same rank and its eigenvalues are equal to $\frac{1}{R_A}$ or 0. The rank of ρ can only decrease or stay the same by filtering. From this it follows that for any eigenvalue $\lambda_{\rho_{A,W}}$

$$\lambda_{\rho_{A,W}} = \frac{1}{R_A} < \frac{1}{R} \leq \lambda_{\rho_W}^{max}, \tag{5.4.27}$$

where $\lambda_{\rho_W}^{max}$ is the largest eigenvalue of ρ_W. Now we invoke a theorem in Ref. [131] which says that any bipartite density matrix ρ for which there exists a pure state $|\psi\rangle$ such that

$$\langle\psi|\,(\rho_A \otimes \mathbf{1}) - \rho\,|\psi\rangle < 0, \tag{5.4.28}$$

is distillable. Take $|\psi\rangle$ to be the eigenvector of ρ_W with maximum eigenvalue and it follows from Eq. (5.4.27) and Eq. (5.4.28) that ρ_W and therefore ρ is distillable. This completes the proof. □

The consequence of this theorem is that there exists no bipartite bound entangled state on any $\mathcal{H}_A \otimes \mathcal{H}_B$ that has rank 2. The reason is that when the maximum local rank of a bipartite rank 2 density matrix ρ exceeds 2, Theorem 4 implies that the state is distillable. On the other hand if both local ranks of ρ are smaller than or equal to 2, then the density matrix ρ effectively has support only on a $2 \otimes 2$ subspace. But it is known that all entangled density matrices on $2 \otimes 2$ are distillable [111].

It also follows that any bipartite PB S on $n \otimes m$ with a number of states $k = nm - 2$ is extendible. By construction the complementary state ρ_S has the PPT-property. However, ρ_S has rank 2 and there do not exist bound entangled states with rank 2. Therefore ρ_S must be separable and it follows that S is extendible. One can carry the argument one step further. After adding the new product state to the set S, we can ask whether one can find the last product state of the basis. Again, that state, which is a pure state must have the PPT-property. It is not hard to show that all entangled pure state have the NPT-property and therefore this last basis state must be a product state. Hence we have shown that any bipartite PB S in $\mathcal{H}$ which has $\dim \mathcal{H} - 2$ states is not only extendible but also *completable*.

5.4.5 Local Distinguishability and Uncompletable Product Bases

In the preceding sections we showed how UPBs give rise to entangled states that cannot be distilled. It turns out that this is not the only interesting property that these sets of states have. One can ask whether the members of a UPB are distinguishable by Local quantum Operations and Classical Communication (LO+CC). The situation is the same as we described in section

5.4.1. We will consider sets of product states which are mutually orthogonal, such as the UPBs. This implies that these states are distinguishable when arbitrary quantum measurements are allowed. When the set of states is given by $\{|\psi_j\rangle\}_{j=1}^{|S|}$ then a projection measurement with projectors $\{\pi_j = |\psi_j\rangle\langle\psi_j|\}_{j=1}^{|S|}$ and $\pi_{S^\perp} = \mathbf{1} - \sum_j \pi_j$ would distinguish the states in the set S. The question is whether measurements that exactly distinguish the set of states can be implemented with local operations and classical communication only. Let us assume that two parties Alice and Bob are given one of the five states in the **Pent** set and they have to determine by LO+CC which one they have. It is not hard to see that for a set such as **Pent** straightforward attempts at finding an appropriate series of measurements are bound to fail; the way in which the states are made orthogonal, partially on Alice's side, partially on Bob's side seems to preclude the existence of a perfect measurement. The parties seem to end up with disturbing the states by measuring them and this disturbance then results in a set of non-orthogonal product states that can no longer be distinguished.

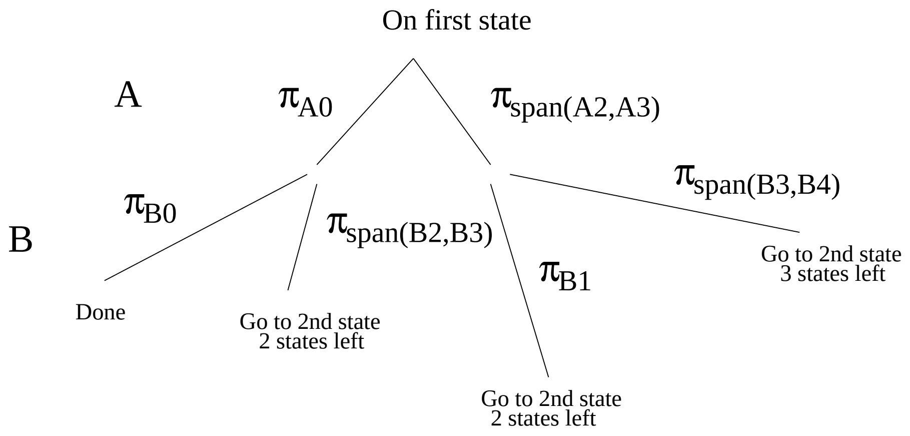

Figure 5.5: Measurement tree for two copies of the **Pent** ensemble.

In order to understand what kinds of measurement protocols are possible, we consider the situation in which the two parties are given *two* copies of the same state of the **Pent** set. We can then show that there is a LO+CC that reliably identifies the states. This measurement procedure is presented in Fig. 5.5. Each level of the tree corresponds to a measurement by either Alice or Bob. After each measurement round Alice and Bob communicate classically to discuss the results. In this protocol only (incomplete) von Neumann measurements (see Chap. 3, sec. 3.2) are performed and they are denoted by their operation elements which are projectors. The states of the $\mathbf{Pent}$ set themselves are denoted as $A0, \ldots, A4$ for Alice's part of the **Pent** states and $B0, \ldots, B4$ for Bob's part in correspondence with $i = 0, \ldots, 4$ in Eq. (5.4.3). Thus in the first round Alice's measurement has two outcomes, associated with the projector on her $|0\rangle$ state and the projector on the span of her $|2\rangle$ and $|3\rangle$ states, which obey the relation

$$\pi_0 + \pi_{\mathrm{Span}(A3,A2)} = \mathbf{1}. \tag{5.4.29}$$

As one can see in Fig. 5.5 by a two-round protocol on the first state, Alice and Bob have reduced the number of states to be distinguished to at most three. Now it is not hard to see that three orthogonal product states can always be distinguished by von Neumann measurements (see also sec. 5.4.8).

One can associate with the leafs of such a measurement tree the series of projections that resulted in the series of measurement outcomes. For example, in the tree of Fig. 5.5, the leaf all the way to the left can be associated with the projector:

$$\pi_{A0} \otimes \pi_{B0}. \tag{5.4.30}$$

These projectors at the ends of the tree are called 'leaf-projectors'. Aside from local von Neumann measurements a party can also perform local POVM measurements. But by Neumark's theorem [62] any POVM measurement on a Hilbert space $\mathcal{H}$ can be viewed as an (incomplete) von Neumann measurement in an extended higher dimensional Hilbert space $\mathcal{H}_{ext}$. When the number of POVM outcomes is finite, this extended Hilbert space is finite. Note that the conversion of a local POVM to a local von Neumann measurement corresponds to a local extension of the Hilbert space, i.e. for a bipartite system a Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ is extended to $\mathcal{H}_{ext} = (\mathcal{H}_A \oplus \mathcal{H}'_A) \otimes (\mathcal{H}_B \oplus \mathcal{H}'_B)$ where $\mathcal{H}'_A$ and $\mathcal{H}'_B$ are the extensions of the local Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$. In the following analysis the local measurements are restricted to have a *finite* number of outcomes. Furthermore we require that the number of rounds in the entire protocol is finite. Thus one can say that we restrict ourselves to finite resources in time and space. Every local measurement with a finite number of outcomes can be decomposed into a series of local measurements with two outcomes only with the understanding that subsequent levels of the measurement tree can correspond to actions of the same party.

A special class of local measurements are the measurements that we call *dissections*. A dissection measurement is one in which the set of states S is split into two sets 1 and 2. The states themselves are unchanged by the measurement, but the outcome of the measurement tells us whether the state that one is given was in set 1 or in set 2. More general measurement schemes can project the states in S onto other states that might or might not be orthogonal.

In the last section we briefly mentioned the notion of a completion of a set of orthogonal product states. Let us now give the definition of an uncompletable set of orthogonal product states.

Definition 6 *Consider a multipartite quantum system $\mathcal{H} = \bigotimes_{i=1}^{m} \mathcal{H}_i$ with m parties of respective dimension $d_i, i = 1, ..., m$. An* uncompletable product basis *in $\mathcal{H}$ is a* PB *that cannot be completed with orthogonal product states to a full orthonormal product basis for $\mathcal{H}$.*

Remark The uncompletable product basis is defined with respect to $\mathcal{H}$. An orthogonal product basis in $\mathcal{H}$ could be uncompletable in $\mathcal{H}$, but completable to a full product basis for $\mathcal{H}_{ext}$, when the set is embedded in $\mathcal{H}_{ext}$. In section 5.4.6 we will give an example of such a set.

The following theorem captures an essential connection between completability and exact local distinguishability:

Theorem 5 *Given a set S of orthogonal product states* (PB) *on $\mathcal{H} = \bigotimes_{i=1}^{m} \mathcal{H}_i$ with $\mathrm{dim}\mathcal{H}_i = d_i, i = 1, ..., m$. If the set S is exactly distinguishable with finite resources by local incomplete von Neumann measurements on $\mathcal{H}_{ext}$,–which can be $\mathcal{H}$ itself or any finite local extension of $\mathcal{H}$–, and classical communication, then the set S is completable to an orthogonal basis for $\mathcal{H}_{ext}$.*

Proof One shows how the measurement protocol leads directly to a way to complete the set S. At some stage of their protocol, the parties (1) may have been able to eliminate members of the original set of states S and (2) they may have mapped, by performing their von Neumann measurements, the remaining set of orthogonal states into a new set of *orthogonal* states S'. Note that the remaining states have to be orthogonal, otherwise the measurement could never be exact. Determining which member they have in this new set uniquely determines with which state of S they started with. At this stage, party i_0 performs a von Neumann measurement. Let $\mathcal{K}$ be the Hilbert space in which the remaining states are known to lie (including the local extensions that are needed in order for any subsequent measurement to be described as a von Neumann measurement). The measurement of party i_0 is given by a decomposition of the Hilbert space $\mathcal{K} = \mathcal{K}_{else} \otimes \mathcal{K}_{i_0}$ with $\mathcal{K}_{else} = \bigotimes_{j \neq i_0} \mathcal{K}_j$, into 2 orthogonal subspaces, $\mathcal{K}_{else} \otimes \pi_1 \mathcal{K}_{i_0}$ and $\mathcal{K}_{else} \otimes \pi_2 \mathcal{K}_{i_0}$.

If a state in S' lies in one of these subspaces, it will be unchanged by the measurement. If a state $|\alpha\rangle \otimes |\beta\rangle$, where $|\alpha\rangle \in \mathcal{K}_{else}$, is not contained in one of the subspaces, it will be projected onto one of the states $\{|\alpha\rangle \otimes \pi_1 |\beta\rangle, |\alpha\rangle \otimes \pi_2 |\beta\rangle\}$. Let S" be this new projected set of states, containing both the unchanged states in S' as well as the possible projections of the states in S'. If one of the subspaces $\mathcal{K}_{else} \otimes \pi_1 \mathcal{K}_{i_0}$ or $\mathcal{K}_{else} \otimes \pi_2 \mathcal{K}_{i_0}$ does not contain a member of S', it can be 'completed' directly; one can freely choose a product basis for this space. For a subspace that does contain members of S", let us assume that it can be completed with product states orthogonal to members of S". In this way one has completed S" on the full Hilbert space $\mathcal{K}$ since the two orthogonal-subspace completions are orthogonal sets and they are a decomposition of $\mathcal{K}$. However, one has now completed the set S" rather than the set S'. Fortunately, one can replace the projected states $|\alpha\rangle \otimes \pi_1 |\beta\rangle, |\alpha\rangle \otimes \pi_2 |\beta\rangle$ by the original state $|\alpha\rangle \otimes |\beta\rangle$ and 1 orthogonal state by making a linear combination of $|\alpha\rangle \otimes \pi_1 |\beta\rangle$ and $|\alpha\rangle \otimes \pi_2 |\beta\rangle$ orthogonal to $|\alpha\rangle \otimes |\beta\rangle$. These two states are orthogonal to all other states as each $|\alpha\rangle \otimes \pi_i |\beta\rangle$ was already orthogonal to all other states. Thus at each round of measurement, a completion of the set of states S' is achieved assuming a completion of the subspaces determined by the measurement.

The tree of nested subspaces will always lead to a subspace that contains only a single state of the set, as the measurement protocol was able to tell the states in S apart exactly. But such a subspace containing only one state can easily be completed and thus, by induction, we have proved that the original set S can be completed. □

Before discussing the consequences of this theorem, we will show how one can strengthen the result to include measurements that have an arbitrary small probability of error. First, one has to define what it means for a set of states to be distinguishable with some probability of error:

Definition 7 *Given a set S of orthogonal product states (*PB*) on* $\mathcal{H} = \bigotimes_{i=1}^m \mathcal{H}_i$ *with* $\mathrm{dim}\mathcal{H}_i = d_i, i = 1, ..., m$. *Let* $\mathcal{M}$ *be a local incomplete von Neumann measurement protocol on a finite-dimensional Hilbert space* $\mathcal{H}_{ext}$ *which can be a local extension of* $\mathcal{H}$*, that includes classical communication between the local parties. Let* $\mathcal{D}(\mathcal{M})$ *be a decision scheme that associates each leaf of the measurement tree of* $\mathcal{M}$ *with a state of the set S, meaning that upon the outcomes of leaf* j*, we decide that the associated state* i *is the state that we were given of the set S. A set S is* ϵ*-distinguishable if there exists an* $\mathcal{M}$ *and a* $\mathcal{D}(\mathcal{M})$ *such that*

$$P_{suc,\mathcal{M}} = \min_{i \in S} \sum_{j \,|\, j \to i} \mathrm{Prob}(\pi_j | i) \geq 1 - \epsilon, \tag{5.4.31}$$

where $\pi_1, \ldots, \pi_k$ *are the leaf-projectors of the measurement tree of* $\mathcal{M}$. $\mathrm{Prob}(\pi_j|i)$ *is the probability that given the state* i *we obtain the measurement outcomes of leaf* π_j. *The sum over the leaf-projectors is constrained to leaf-projectors that lead to deciding for state* i*, which is indicated by* $j \to i$.

In words this definition says that the set S is ϵ-distinguishable if the probability of deciding correctly for a state in S is greater than or equal to $1 - \epsilon$ for any state that the parties are given from S. This definition makes it possible to state the following lemma:

Lemma 5 *Given a set S of orthogonal product states (PB) on* $\mathcal{H} = \bigotimes_{i=1}^m \mathcal{H}_i$ *with* $\mathrm{dim}\mathcal{H}_i = d_i, i = 1, ..., m$. *If the set S is* ϵ*-distinguishable for all* $\epsilon > 0$ *then S is exactly distinguishable.*

Proof As one is restricted to using finite resources, one can set the total number of levels in the binary measurement tree to a certain large number L and set the dimensions of the local extensions $\dim \mathcal{H}_{i,ext} = d_{i,ext}$. All measurement trees (plus decision schemes) that have at most L levels and correspond to local extensions of which the dimensions are upper bounded by $\dim \mathcal{H}_{i,ext} = d_{i,ext}$ can then by characterized by five sets of variables:
(1) the structure of the tree $\mathcal{T}$, i.e. the distribution of the length of its branches,
(2) an assignment $\mathcal{A}$ of levels to the various parties A, B, C etc.,
(3) the dimension $\mathcal{D}im$ in which the von Neumann measurement takes place for each node of the tree,
(4a) the rank $\mathcal{R}$ and the number of the projectors pertaining to each node of the tree,
(4b) and the projectors $\mathcal{P}$ themselves pertaining to each node of the tree,
(5) a decision scheme $\mathcal{D}$ that infers from measurement outcomes –the leafs of the tree– decisions about what the original state was.

Consider the function $P_{suc,\mathcal{M}}$ as in Eq. (5.4.31). The domain of this function is the set $(\mathcal{T}, \mathcal{A}, \mathcal{D}im, \mathcal{R}, \mathcal{P}, \mathcal{D})$. The set of trees $\mathcal{T}$, assignments $\mathcal{A}$, decisions $\mathcal{D}$, dimensions $\mathcal{D}im$ and the set of ranks $\mathcal{R}$ of the projectors are all discrete sets with a finite number of elements.

Consider a measurement at a single node. We fix the number of projectors, the dimension of the Hilbert space and the rank of the projectors at this node. Let (π_1, π_2) be a set of projectors at this node. Then another set (π_1', π_2') can be obtained by unitary transformations $U_i \pi_i = \pi_i'$. This implies that the set (π_1, π_2) is a compact set, as the set of unitary transformations in a finite-dimensional Hilbert space is a compact set. The function $P_{suc,\mathcal{M}}$ is continuous on this compact set. The entire domain of the function $P_{suc,\mathcal{M}}$ is the union of a finite number of compact sets. Then, if there exists measurements and decision schemes such that $P_{suc,\mathcal{M}}$ is larger or equal than $1-\epsilon$ for all ϵ, there also exist a scheme for which $P_{suc,\mathcal{M}} = 1$. This measurement corresponds to exactly distinguishing the members of S. □

Corollary 1 *Given a set S of orthogonal product states (*PB*) on $\mathcal{H} = \bigotimes_{i=1}^m \mathcal{H}_i$ with $\dim \mathcal{H}_i = d_i, i = 1, ..., m$. If this set S is ϵ-distinguishable for all $\epsilon > 0$, then S is exactly completable on $\mathcal{H}_{ext}$, a locally extended Hilbert space or $\mathcal{H}$ itself.*

This follows from Theorem 5 and Lemma 5.

Let us give a final theorem, that relates the question of completability to the property of entanglement:

Theorem 6 *Given a set S of orthogonal product states (*PB*) on $\mathcal{H} = \bigotimes_{i=1}^m \mathcal{H}_i$ with $\dim \mathcal{H}_i = d_i, i = 1, ..., m$. If the set S is ϵ-distinguishable for all $\epsilon > 0$, then the complementary density matrix ρ_S, Eq. (5.4.6), is separable.*

Proof If the exact measurement is a von Neumann measurement on $\mathcal{H}$, then by Corollary 1 the set S can be completed with product states to a basis for $\mathcal{H}$. Thus, the density matrix ρ_S which is the uniform mixture of these product states that complete S, is separable. If the exact measurement is a von Neumann measurement on $\mathcal{H}_{ext}$, then the density matrix $\rho_{S,ext}$ is separable. One can obtain ρ_S on $\mathcal{H}$ by local projections from $\mathcal{H}_{ext}$ onto $\mathcal{H}$ and therefore $\rho_S = P_{\mathcal{H}} \rho_{S,ext} P_{\mathcal{H}}$ is separable as well. □

This theorem implies that a multipartite UPB S is not distinguishable with arbitrary small probability of error by LO+CC, using finite resources, since the density matrix ρ_S is entangled. The UPBs are new illustrations of the phenomenon of nonlocality without entanglement. The strength of the result as compared to the results in Ref. [125], is that the indistinguishability has been proved for any UPB, whereas in Ref. [125] only the set of states of Eq. (5.4.1) were shown not to be distinguishable with arbitrary small probability of error. The weakness of this result is that we have restricted the set of measurements to ones that can be performed using finite resources. This was necessary as a POVM measurement with an infinite number of outcomes describes a von Neumann measurement in an infinite dimensional Hilbert space. It is not clear how to extend the notions of completability on an infinite dimensional Hilbert space. Also, it is unclear whether a result such as Lemma 5 would hold for measurements that could use infinite resources.

5.4.6 Local Extensions and Deficits of Product States

In the last section care has been taken to include POVM measurements in a local measurement scheme. But do there exist PBs that are exactly distinguishable by POVMs but not by von Neumann measurements in the original Hilbert space? It turns out that there are such sets. Here is an example of such a set on $3\otimes 4$, the set **PO**. Consider the states $\vec{v}_j\otimes\vec{w}_j,\quad j=0,\ldots,4$ with $\vec{v}_j$ the states of the **Pent** UPB as in Eq. (5.4.2) and $\vec{w}_j$ defined as

$$\begin{aligned}\vec{w}_j \;&= \sqrt{2/\sqrt{5}}(\sqrt{\cos(\pi/5)}\cos(2j\pi/5),\sqrt{\cos(\pi/5)}\sin(2j\pi/5),\\ &\quad\sqrt{\cos(2\pi/5)}\cos(4j\pi/5),\sqrt{\cos(2\pi/5)}\sin(4j\pi/5)).\end{aligned} \tag{5.4.32}$$

Note that $\vec{w}_j^T\vec{w}_{j+1}=0$ (addition mod 5). The orthogonality graph of these states is the same as for **Pent**, Figure 5.1. One can show that this set, albeit extendible on $3\otimes 4$, is not *completable*: One can at most add three vectors: $\vec{v}_0\otimes(\vec{w}_0,\vec{w}_1,\vec{w}_4)^\perp$, $\vec{v}_3\otimes(\vec{w}_2,\vec{w}_3,\vec{w}_4)^\perp$ and $(\vec{v}_0,\vec{v}_3)^\perp\otimes(\vec{w}_1,\vec{w}_2,\vec{w}_4)^\perp$.

The POVM measurement that is performed by Bob on the four-dimensional side has five projector elements, each projecting onto a vector

$$\vec{u}_j=\frac{1}{\sqrt{2}}(-\sin(2j\pi/5),\cos(2j\pi/5),-\sin(4j\pi/5),\cos(4j\pi/5)), \tag{5.4.33}$$

with $j=0,\ldots,4$. Note that $\vec{u}_0$ is orthogonal to vectors $\vec{w}_0,\vec{w}_2$ and $\vec{w}_3$, or, in general, $\vec{u}_i$ is orthogonal to $\vec{w}_i,\vec{w}_{i+2},\vec{w}_{i+3}$ (addition mod 5). This means that upon Bob's POVM measurement outcome, three vectors are excluded from the set; then the remaining two vectors on Alice's side, $\vec{v}_{i+1}$ and $\vec{v}_{i+4}$, are orthogonal and can thus be distinguished.

Since the set is distinguishable by a POVM, it is completable. The completion of this set in $3\otimes 5$ is particularly simple. Bob's Hilbert space is extended to a five-dimensional space. The POVM measurement can be extended to a projection measurement in this five-dimensional space with orthogonal projections onto the states $\vec{x}_i=(\vec{u}_i,0)+\frac{1}{2}(0,0,0,0,1)$. Then a completion of the set in $3\otimes 5$ are the following ten states:

$$\begin{array}{ll}(\vec{v}_1,\vec{v}_4)^\perp\otimes\vec{x}_0, & \vec{v}_0\otimes(\vec{w}_0^\perp\in\mathrm{Span}(\vec{x}_4,\vec{x}_1)),\\ (\vec{v}_0,\vec{v}_2)^\perp\otimes\vec{x}_1, & \vec{v}_1\otimes(\vec{w}_1^\perp\in\mathrm{Span}(\vec{x}_0,\vec{x}_2)),\\ (\vec{v}_1,\vec{v}_3)^\perp\otimes\vec{x}_2, & \vec{v}_2\otimes(\vec{w}_2^\perp\in\mathrm{Span}(\vec{x}_1,\vec{x}_3)),\\ (\vec{v}_2,\vec{v}_4)^\perp\otimes\vec{x}_3, & \vec{v}_3\otimes(\vec{w}_3^\perp\in\mathrm{Span}(\vec{x}_2,\vec{x}_4)),\\ (\vec{v}_0,\vec{v}_3)^\perp\otimes\vec{x}_4, & \vec{v}_4\otimes(\vec{w}_4^\perp\in\mathrm{Span}(\vec{x}_3,\vec{x}_0)).\end{array} \tag{5.4.34}$$

There is another interesting feature of this set. Let's take the set **PO** and add one of the product states, say the vector,

$$\vec{v}_0\otimes(\vec{w}_0,\vec{w}_1,\vec{w}_4)^\perp, \tag{5.4.35}$$

to make it a six-state ensemble **PO**$^+$. The complementary density matrix $\rho_{\mathbf{PO}^+}$, Eq. (5.4.6), has rank $12-6=6$. Is $\rho_{\mathbf{PO}^+}$ a separable density matrix? We can enumerate the product states that are orthogonal to the members of $\mathbf{PO}^+$, but are not necessarily mutually orthogonal:

$$\begin{aligned} \vec{v}_3 \otimes (\vec{w}_2, \vec{w}_3, \vec{w}_4)^\perp, \\ \vec{v}_2 \otimes (\vec{w}_1, \vec{w}_2, \vec{w}_3)^\perp, \\ (\vec{v}_0, \vec{v}_3)^\perp \otimes (\vec{w}_1, \vec{w}_2, \vec{w}_4)^\perp, \\ (\vec{v}_0, \vec{v}_2)^\perp \otimes (\vec{w}_1, \vec{w}_3, \vec{w}_4)^\perp. \end{aligned} \tag{5.4.36}$$

This means that the space on which $\rho_{\mathbf{PO}^+}$ has support contains only four product states, whereas $\rho_{\mathbf{PO}^+}$ has rank 6. Therefore $\rho_{\mathbf{PO}^+}$ must be entangled. The entanglement of $\rho_{\mathbf{PO}^+}$ is bound by construction.

We have constructed a new bound entangled state whose range is not without product states but has a product state *deficit*. As any UPB set, the set $\mathbf{PO}^+$ is not locally distinguishable with finite means. This construction works in a very general way:

Lemma 6 *Given a set S of orthogonal product states (*PB*) on $\mathcal{H} = \bigotimes_{i=1}^m \mathcal{H}_i$ with $\dim \mathcal{H}_i = d_i, i = 1, ..., m$. If the set S is not completable to a full basis for $\mathcal{H}$, but is completable in some $\mathcal{H}_{ext}$, then there always exist a set of orthogonal product states in $\mathcal{H}_S^\perp$ such that when we add these states to S to make the ensemble S^+, the complementary density matrix ρ_{S^+} is bound entangled.*

Proof Assume that there does not exist such an augmented ensemble S^+. This will lead to a contradiction. The fact that S is completable in $\mathcal{H}_{ext}$ makes it possible to add at least one product state to S, see Theorem 6. Let us call this state $|\psi_1\rangle$. The density matrix ρ_{S_1} complementary to the ensemble $S_1 = S \cup \{|\psi_1\rangle\}$ is either entangled or separable. If it is entangled, then we have found the desired augmentation of S. If it is separable, then there is at least one state, let us call it $|\psi_2\rangle$, in the range of ρ_{S_1} which is a product state. Note that $|\psi_2\rangle$ is orthogonal to $|\psi_1\rangle$. Then we can augment ρ_{S_1} with this new orthogonal product state $|\psi_2\rangle$ to make the set ρ_{S_2}. Consider its complementary density matrix ρ_{S_2}. Repeat the arguments as before. If we find that all density matrices $\rho_{S_1}, \rho_{S_2}, \ldots, \rho_{S_{\dim \mathcal{H} - |S|}}$ are separable, then we have found a completion of the original set S with *orthogonal* product states. This leads to a contradiction, because S is not completable in $\mathcal{H}$. □

Remark While the lemma shows that there exists a set of orthogonal product states that when added to the set S leads to a bound entangled state, the proof also suggests a simple way to find this set. The bound entangled state that we find by this procedure can be a bound entangled state corresponding to a UPB, that is the augmented set S^+ is a UPB, or it can be an entangled state, based on a set such as $\mathbf{PO}^+$, which is supported on a subspace which has a product state deficit.

5.4.7 Rank and the Optimal Decomposition of a Density Matrix

Some of the uncompletable sets of product states exhibit additional interesting properties. It was shown by Uhlmann [132] that every bipartite density matrix ρ admits an optimal decomposition, that is, a decomposition that achieves the entanglement of formation $E(\rho)$, Eq.

(5.2.11), with at least $\mathrm{Rank}(\rho)$ and at most $\mathrm{Rank}(\rho)^2$ different pure states. No examples of density matrices for which more than $\mathrm{Rank}(\rho)$ states are needed to construct the optimal decomposition were previously known. By numerical minimization it was found that the state complementary to the **Tiles** UPB , as introduced in section 5.4.2, has an entanglement of formation of 0.213726 bits. This complementary state $\rho_{\mathbf{Tiles}}$ has rank 4. However, it was found that the optimal ensemble of pure states consists of five pure states. A similar result was found for the **Pent** UPB which has an entanglement of 0.232635 bits, made by mixing together five pure states. Thus both these density matrices are numerical examples of states for which more states are needed in the optimal decomposition than the rank of the state.

The following result exhibits a whole class of *separable* states which have this peculiar property.

Theorem 7 *Let* $\{|\alpha_i\rangle \otimes |\beta_i\rangle\}_{i=1}^{|S|}$ *be a* PB *S in* $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$*. If S is uncompletable in* $\mathcal{H}$*, but S is completable in some local extension* $\mathcal{H}_{ext}$ *of* $\mathcal{H}$*, then* ρ_S *has the property that the number of different pure states in the optimal decomposition of* ρ_S *exceeds the rank of* ρ_S*.*

Proof Since the set of states S is completable in a local extension of $\mathcal{H}$, the state ρ_S is separable by Theorem 6. However we know that ρ_S was uncompletable in $\mathcal{H}$, therefore ρ_S cannot be decomposed with an ensemble of orthogonal product states. Any optimal decomposition of ρ_S has to use nonorthogonal product states. The von Neumann entropy of ρ_S is equal to $S(\rho_S) = \log \mathrm{Rank}(\rho_S)$ as ρ_S is the identity matrix on a space of dimension $\mathrm{Rank}(\rho_S)$. In order to achieve this entropy the optimal decomposition of ρ_S has to use more than $\mathrm{Rank}(\rho_S)$ product states, as any density matrix ρ which is a mixture of only n non-orthogonal states has entropy strictly less than $\log n$ bits. $\square$

The only example so far is the set **PO** in $3 \otimes 4$, which was distinguishable by a set of orthogonal projectors in $3 \otimes 5$, but could not be completed in $3 \otimes 4$. The complementary density matrix $\rho_{\mathbf{PO}}$ on $3 \otimes 4$ has rank seven, but the separable decomposition consists of ten non-orthogonal states. These ten non-orthogonal product states are obtained by projecting the orthogonal states of the completion, Eq. (5.4.34) back into the $3 \otimes 4$ Hilbert space. It is not known whether there exists a separable decomposition with more than seven but with less than ten states.

5.4.8 Restrictions

The method to create bound entangled states from UPBs is not always successful. In particular one can show that

Theorem 8 *Any set of orthogonal product states* $\{|\alpha_i\rangle \otimes |\beta_i\rangle\}_{i=1}^{k}$ *in* $2 \otimes n$ *for any* $n \geq 2$ *is distinguishable by local measurements and classical communication and therefore completable to a full product basis for* $2 \otimes n$*.*

Proof The measurement is a three round protocol. Let Alice be associated with the two-dimensional side and Bob with the n-dimensional side. Alice divides S in subsets P_i in the following way:

$$P_i = \{|\alpha_i\rangle \otimes |\beta_i^1\rangle, |\alpha_i\rangle \otimes |\beta_i^2\rangle, \ldots, |\alpha_i^\perp\rangle \otimes |\beta_i^{l_i}\rangle\}, \tag{5.4.37}$$

i.e., Alice's part of the states in set P_i is either $|\alpha_i\rangle$ or $|\alpha_i^\perp\rangle$. The states $|\alpha_i\rangle$ and $|\alpha_j\rangle$ for $i \neq j$ are neither orthogonal nor identical. When P_i contains a set of states $\{|\alpha_i\rangle \otimes |\beta_i^1\rangle, \ldots, |\alpha_i\rangle \otimes |\beta_i^k\rangle\}$ for some $k > 1$ then due to the orthogonality of the states, we must have that $\langle\beta_i^j|\beta_i^m\rangle = \delta_{jm}$ for $j, m = 1, \ldots, k$. The same is true for P_i containing a set of states in which $|\alpha_i^\perp\rangle$ is repeated. Furthermore, all the members of the set P_i have to be orthogonal to all the members of a set P_j for $i \neq j$ on Bob's side as they are never orthogonal on Alice's side. The measurement goes as follows. Bob performs a measurement of which the operation elements are projectors π_i on the subspace spanned by his side of the states in each P_i. These projectors have the property that $\pi_i\pi_j = 0$ for $i \neq j$. The outcome tells Bob in which set P_i the original state lies. After Bob sends this information, the label i, to Alice, she does a measurement that distinguishes $|\alpha_i\rangle$ from $|\alpha_i^\perp\rangle$. Then Bob is left to finish the protocol by distinguishing between states that repeat on Alice's side, for example the states $|\alpha_1\rangle \otimes |\beta_1^1\rangle$ and $|\alpha_1\rangle \otimes |\beta_1^2\rangle$. He can distinguish between these states, because they are mutually orthogonal on his side. Theorem 5 then implies that S is completable. □

Local Dissectibility as a Graph Problem

One can express local dissections on a set of orthogonal product states (PB) as operations on the orthogonality graph of the PB. Consider for example the sets in Fig. 5.7. The dimension of Alice's and Bob's Hilbert space is larger than or equal to four, otherwise these patterns would not be possible. In case (a) it is not hard to see how Alice and Bob would go about measuring the set. State 2 is orthogonal to all the others on Alice's side. Thus Alice can distinguish the sets (2) and (134) by measuring with the local projectors π_{α_2} and $\pi_{\alpha_2}^\perp$. If she finds (2) the protocol is finished. If she finds (134) Bob continues the measurement. The states 1, 3 and 4 form a clique –a graph in which all the vertices are connected– on Bob's side. Therefore a measurement that uses the projectors $\pi_{\beta_1}, \pi_{\beta_3}$ and π_{β_4} distinguishes perfectly between 1, 3 and 4. These measurements are all dissection measurements, that is, they split the total set of states into two or more subsets in which each state occurs only once. In this case Alice first makes the dissection into the sets (2) and (134), after which Bob dissects (134) as (1)(3)(4). Let us translate such a dissection measurement in graph language. A complete bipartite graph is a graph G in which the set of vertices V can be split in two sets V_1 and V_2 such that every vertex in V_1 is connected to every vertex in V_2 by an edge and vertices of V_1 and V_2 are not directly connected amongst each other. The vertices of the orthogonality graph G of a m-partite PB S in $\mathcal{H}$ represent the members of S. Recall that $\mathcal{H} = \mathcal{H}_1 \otimes \ldots \otimes \mathcal{H}_m$. The edges are colored with m different colors such that if two members of S are orthogonal on $\mathcal{H}_i$ color i is used for the edge. One can have multiply colored edges between vertices. To describe whether a multipartite PB is dissectible we will need the notion of a complete bipartite graph of a single color in which vertices in V_1 (or V_2) can be connected amongst each other. We call this kind of graph an (over)complete bipartite graph of a single color [3]. Fig. 5.6 shows an example of such a graph; some of the vertices in V_1 are mutually connected by Bob's edges.

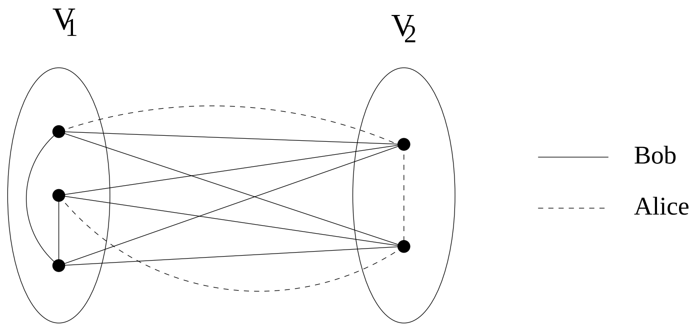

Figure 5.6: An example of an (over)complete bipartite graph of Bob's color.

One can write down the following translation of a series of dissection measurements into a decomposition of the graph:

Proposition 7 *Let S be a multipartite PB on* $\mathcal{H} = \bigotimes_{i=1}^{m} \mathcal{H}_i$ *with* $\dim \mathcal{H}_i = d_i, i = 1, ..., m$ *represented by an orthogonality graph* G. *Then S is* exactly distinguishable by local dissections and classical communication *iff there exists a decomposition of the graph* G *of S into a hierarchical tree of (over)complete bipartite graphs of a single color such that the leafs of the tree correspond to single vertices.*

Proof At each round of the dissection measurement the set of vertices V of the graph G is cut in two sets V_1 and V_2 such that each state represented by a vertex in V_1 is orthogonal to each state represented by a vertex in V_2 for the ith party. In graph language: G is an (over)complete bipartite graph of color i. Let G_1 be the graph with vertices V_1 and the edges connecting vertices in V_1 and similarly for G_2. In the next round of measurement each of the graphs G_1 and G_2 is cut according to the bipartition in an (over)complete bipartite graph of a possibly different color. This process is repeated until the resulting graphs consist a single vertex. Then we conclude that all the states in S have been distinguished. □

Let us consider Fig. 5.7(b). There is no (over)complete bipartite graph of a single color. However we can perform a more general von Neumann measurement. Alice can measure with the projector π_{α_1,α_2} and the projector $\pi^{\perp}_{\alpha_1,\alpha_2}$ where $\pi_{\alpha_1,\alpha_2} = \mathbf{1} - \pi^{\perp}_{\alpha_1,\alpha_2}$. The projector $\pi^{\perp}_{\alpha_1,\alpha_2}$ projects onto states that are orthogonal to $|\alpha_1\rangle$ and $|\alpha_2\rangle$. This will distinguish the sets (12) and (4), but it will project state 3 on either $\pi_{\alpha_1,\alpha_2}|\alpha_3\rangle$ or $\pi^{\perp}_{\alpha_1,\alpha_2}|\alpha_3\rangle$. Thus one can say that we cut the set into subsets (123) and (34). When (34) is found, Bob can finish the protocol directly. When Alice gets (123), she now can distinguish between (1) and (23) since they form a complete bipartite graph. What about the final distinction between 2 and 3? These

[3]There does not appear to exist a standard terminology for this kind of graph.

states started out as orthogonal on Alice's side, then with Alice's first measurement $|\alpha_3\rangle$ was mapped onto $\pi_{\alpha_1,\alpha_2}|\alpha_3\rangle$. This projected state is however still orthogonal to $|\alpha_2\rangle$, as $|\alpha_2\rangle$ did not have a component outside the space spanned by $|\alpha_1\rangle$ and $|\alpha_2\rangle$ and thus

$$0 = \langle\alpha_2|\alpha_3\rangle = \langle\alpha_2|\pi_{\alpha_1,\alpha_2}|\alpha_3\rangle. \tag{5.4.38}$$

Here we can notice a more general rule. A local 2-outcome von Neumann measurement is called orthogonality preserving on a set of (multipartite) orthogonal product states S if after measurement the states in S are still mutually orthogonal. Suppose we can find an (over)complete bipartite graph of a single color that includes all the vertices *but 1* in the orthogonality graph of a set of (multipartite) orthogonal product states S. Then one can show that there exists a local von Neumann measurement that is orthogonality preserving. The party associated with the color of the (over)complete bipartite graph does the measurement and splits the set of states S minus one state $|p\rangle$ into the sets S_1 and S_2. Let π be the projector associated with the set S_1 and $\mathbf{1}-\pi$ be the projector associated with the set S_2. The state $|p\rangle$ is projected onto $\pi|p\rangle$ and $(\mathbf{1}-\pi)|p\rangle$. However, $\pi|p\rangle$ is orthogonal to all the members of S_1 since the states in S_1 have support only on the subspace $\pi\mathcal{H}$, and also orthogonal to S_2 since they only have support on the subspace $(\mathbf{1}-\pi)\mathcal{H}$. The same argument holds for the state $(\mathbf{1}-\pi)|p\rangle$. Unfortunately, there are von Neumann measurements that are orthogonality preserving which are not expressible in terms of graph language only. With these tools, one can easily show that bipartite PBs of 2, 3 and 4 orthogonal product states in any dimension (consistent with the number of states) are always completable. One shows that these sets are locally distinguishable by considering their graphs, then one invokes Theorem 5 to conclude their completability.

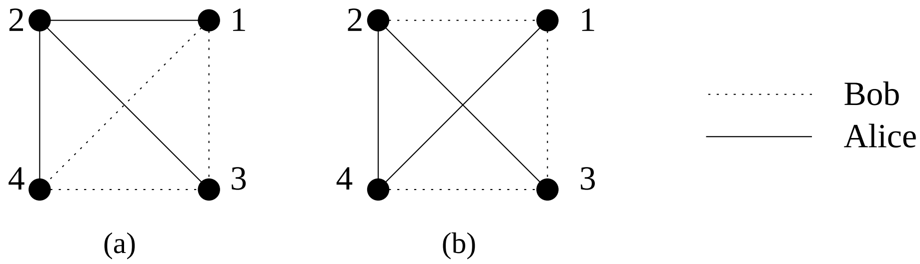

Figure 5.7: Examples of PBs, represented as two color graphs, that are distinguishable by LO+CC.

5.4.9 Transfer of Indistinguishable Product States

Consider a set of bipartite orthogonal product states S that are not exactly distinguishable by local operations and classical communication. Say Alice holds half of the unknown product state out of the ensemble S. She wants to send this state to Charlie by means of classical communication and local measurements in her lab. Let Bob, the holder of the other half, be fully cooperative in this scheme, and that he, as is Alice, is restricted to using local measurements and classical communication. We will show that when the states in S cannot be locally distinguished, Alice will not be able to transfer her state exactly to Charlie. In a scenario where Alice and Bob share some classically correlated information, it is always possible for Alice to transmit this correlation to a third party, even when Alice and Bob are ignorant about the information that they share. In the quantum scenario we ask whether it is possible for Alice to transfer her half of an unknown product state which is correlated with Bob to a third party, possibly in a way that does not reveal to Alice which state she transferred. If Alice and Bob could measure what state they have, then it is clear that the information about the state can be transferred from Alice to Charlie. They convert their correlated quantum state to a piece of classical information. However, when Alice and Bob cannot tell with absolute certainty what state they have, even though it is a product state, it is not possible to transfer Alice's part to Charlie. The reason is the following. If it could be achieved by means of classical communication and local measurements only, then all the classical information that Alice sends to Charlie might as well be sent to Bob. Thus if Charlie can reconstruct half of the state with this information, then so can Bob. Now Bob, possessing both parts of the product state, does a local projection measurement on the orthogonal states in S. Since the states are orthogonal this measurement will uniquely determine which state in S he shared with Alice. This is in contradiction with the premise that the states in S are not exactly distinguishable by LO+CC. What this construction illustrates is that even though no entanglement is involved, the quantum states that the two parties share exhibit some essential nonlocal quantum features.

5.4.10 The Use of Separable Superoperators

We have shown in the last sections that there exist sets of orthogonal product states that are not locally distinguishable. In this section we address the question of what kind of measurement does distinguish them. We are interested in finding measurements that need the least amount of resources in terms of entanglement between the two or more parties.

Let us introduce a class of quantum operations that are close relatives of operations that can be implemented by local quantum operations and classical communication, the *separable superoperators and measurements*:

Definition 8 *[133] Let $\mathcal{H} = \bigotimes_{i=1}^n \mathcal{H}_i$. Let $\mathcal{H}' = \bigotimes_{i=1}^n \mathcal{H}'_i$. A* $\mathbf{TCP}$ *map $\mathcal{S}\colon\ B(\mathcal{H}) \to B(\mathcal{H}')$ is separable iff one can write the action of $\mathcal{S}$ on an arbitrary density matrix $\rho \in B(\mathcal{H})$ as*

$$\mathcal{S}(\rho) = \sum_i A_{1,i} \otimes A_{2,i} \otimes \ldots \otimes A_{n,i}\, \rho\, A_{1,i}^\dagger \otimes A_{2,i}^\dagger \otimes \ldots \otimes A_{n,i}^\dagger, \tag{5.4.39}$$

where $A_{k,i}$ is a $\dim\mathcal{H}'_i \times \dim\mathcal{H}_i$ matrix and

$$\sum_i A_{1,i}^\dagger A_{1,i} \otimes A_{2,i}^\dagger A_{2,i} \otimes \ldots \otimes A_{n,i}^\dagger A_{n,i} = \mathbf{1}. \tag{5.4.40}$$

Similarly, a quantum measurement (Chap. 3, sec. 3.2) on a multipartite Hilbert space is separable iff for each outcome m, the operation elements A_i^m for all i are of a separable form:

$$A_i^m = A_{1,i}^m \otimes A_{2,i}^m \otimes \ldots \otimes A_{n,i}^m. \tag{5.4.41}$$

The results of [125] show that separable superoperators are not equivalent to local quantum operations and classical communication. There is a separable measurement for the nine states of Eq. (5.4.1); it is the measurement whose operation elements are the projectors onto the nine states. The nine states are not locally distinguishable by LO+CC.

The following theorem gives a sufficient condition under which a bipartite set of orthogonal product states is distinguishable with the use of separable measurements. Unfortunately, it is not known what entanglement resources are needed to implement such separable measurements. They do however form a rather restricted class of operations. It is for example not possible to use them to create entanglement where previously none existed.

Theorem 9 *Let S be a set of bipartite orthogonal product states in $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ with k members. If S has the property that it is completable in $\mathcal{H}$ or local extensions of $\mathcal{H}$ ($\mathcal{H}_{ext}$) when any single member is removed from S, then the members of S are distinguishable by means of a separable measurement.*

Proof Denote by $\{\pi_i\}_{i=1}^k$ the orthogonal rank 1 product projectors onto the states in S. Let S_i be the set S without a state i. As S_i is completable, the projectors

$$\pi_{S_i^\perp} = \mathbf{1} - \sum_{m \neq i} \pi_m, \tag{5.4.42}$$

for all $i = 1, \ldots, k$, are separable. Note that $\pi_{S_i^\perp} = \pi_{S_i^\perp}^\dagger$. The projectors $\pi_{S_i^\perp}$ and π_i for $i = 1, \ldots, k$ with the right coefficients sum up to $\mathbf{1}$:

$$\frac{1}{k}\sum_{i=1}^{k} \pi_{S_i^\perp}^\dagger \pi_{S_i^\perp} + \frac{k-1}{k}\sum_{i=1}^{k} \pi_i^\dagger \pi_i = \mathbf{1}, \tag{5.4.43}$$

using $\pi^2 = \pi$ for projectors. As the projectors $\pi_{S_i^\perp}$ are separable, one can decompose them into a set of N_i rank 1 product projectors, $\pi_{(S_i^\perp, m_i)}$ labeled by an index $m_i = 1, \ldots, N_i$. Note that one can choose mutually orthogonal projectors (for a given i) $\pi_{(S_i^\perp, m_i)}$ when S_i is completable in the given Hilbert space $\mathcal{H}$. When S_i is completable only in a local extension of $\mathcal{H}$, these projectors will be non-orthogonal. In both cases the set of product projectors

$$\left\{ \frac{1}{\sqrt{k}} \pi_{(S_i^\perp, m_i)}, \sqrt{\frac{k-1}{k}} \pi_i \right\}_{i=1, m_i=1}^{k, N_i}, \tag{5.4.44}$$

are the operation elements of a separable measurement. This measurement projects onto states in S or onto separable states that are orthogonal to all but one state in S. With a slight modification of this measurement one can construct a measurement which distinguishes the states in S locally. Formally one replaces the projectors of Eq. (5.4.44) by

$$\begin{aligned} \pi_i = |\alpha_i, \beta_i\rangle\langle\alpha_i, \beta_i| \to |i_A, i_B\rangle\langle\alpha_i, \beta_i|, \\ \pi_{(S_i^{\perp}, m_i)} = |\delta_{i,m_i}, \gamma_{i,m_i}\rangle\langle\delta_{i,m_i}, \gamma_{i,m_i}| \to |i', m_{iA}, i', m_{iB}\rangle\langle\delta_{i,m_i}, \gamma_{i,m_i}|, \end{aligned} \tag{5.4.45}$$

such that the set of states $|i_A\rangle$, $|i', m_{i_A}\rangle$ is an orthonormal set and the same for B. This modification leaves Eq. (5.4.43) unchanged, so that this new set of operation elements again corresponds to a (separable) measurement. Upon this measurement, however, Alice and Bob both get a classical record of the outcome. If they perform this measurement on states in S, their outcomes will uniquely determine which state in S they were given. □

Both the **Pent** UPB as well as the **Tiles** UPB are examples of sets that are completable in $3 \otimes 3$ when anyone state in the set is omitted. These sets are thus distinguishable by a separable measurement.

5.5 A Family of Indecomposable Positive Linear Maps

We introduce a new family of indecomposable positive linear maps based on entangled states. Central to our construction is the notion of an unextendible product basis. The construction lets us create and conjecture indecomposable positive linear maps in matrix algebras of arbitrary high dimension.

5.5.1 Introduction

One of the central problems in the emergent field of quantum information theory [37] is the classification and characterization of the entanglement of quantum states. Entangled quantum states have been shown to be valuable resources in (quantum) communication and computation protocols. In this context it has been shown [112] that there exists a strong connection between the classification of the entanglement of quantum states and the structure of positive linear maps. Very little is known about the structure of positive linear maps even on low dimensional matrix algebras, in particular the structure of indecomposable positive linear maps. We denote the $n \times n$ matrix algebra as $M_n(\mathbf{C})$. The first example of an indecomposable positive linear map in $M_3(\mathbf{C})$ was found by Choi [134]. There have been only a couple of other examples of indecomposable positive linear maps (see Ref. [135] for some recent literature); they seem to be hard to find and no general construction method is available. In this section we make use of the connection with quantum states to develop a method to create indecomposable positive linear maps on matrix algebras $M_n(\mathbf{C})$ for any $n > 2$. This construction exhibits some of the structure of positive linear maps which is present in almost any dimension. In section 5.5.2 we present the general construction. In section 5.5.3 we present two examples and discuss various open problems.

5.5.2 Unextendible Product Bases and Indecomposable Maps

The more complicated structure of the positive linear maps in higher dimensional matrix algebras, namely the existence of indecomposable positive maps (see the introductory section 5.2.3) is reflected in the existence of entangled density matrices ρ on $\mathcal{H}_A \otimes \mathcal{H}_B$ for which $(\mathrm{id}_A \otimes T)(\rho)$ is positive semidefinite.

In Ref. [136] a method was discovered to construct entangled density matrices ρ with positive semidefinite $(\mathrm{id}_A \otimes T)(\rho)$ in various dimensions $\dim \mathcal{H}_A > 2$ and $\dim \mathcal{H}_B > 2$. The construction was based on the notion of an unextendible product basis.

We will present our results relating these density matrices obtained from the construction in Proposition 6 to indecomposable positive linear maps. We will need the definition of a maximally entangled pure state:

Definition 9 *Let $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. Let $|\psi\rangle$ be a normalized state ($\langle\psi|\psi\rangle = 1$) in $\mathcal{H}$ and*

$$\rho_{A,\psi} = \mathrm{Tr}_B \, |\psi\rangle\langle\psi|, \tag{5.5.1}$$

where Tr_B indicates that the trace is taken with respect to Hilbert space $\mathcal{H}_B$ only. The state $|\psi\rangle \in \mathcal{H}$ is maximally entangled if

$$S(\rho_{A,\psi}) = -\mathrm{Tr}\, \rho_{A,\psi} \log_2 \rho_{A,\psi} = \log_2 \min(\dim \mathcal{H}_A, \dim \mathcal{H}_B) \tag{5.5.2}$$

The function $S(\rho_{A,\psi})$ is the von Neumann entropy of the density matrix $\rho_{A,\psi}$.

Remarks For pure states $|\psi\rangle$ the von Neumann entropy of $\rho_{A,\psi}$ is always less than or equal to $d \equiv \log_2 \min(\dim \mathcal{H}_A, \dim \mathcal{H}_B)$. For maximally entangled states we will have $\rho_{A,\psi} = \mathrm{diag}(1/d, \ldots, 1/d, 0, \ldots, 0)$ so that the maximum von Neumann entropy, Eq. (5.5.2), is achieved. When $\dim \mathcal{H}_A = \dim \mathcal{H}_B$ one can always make an orthonormal basis for $\mathcal{H}$ with maximally entangled states [137]. For pure states $|\psi\rangle$ we always have $S(\rho_{A,\psi}) = S(\rho_{B,\psi})$.

The following lemma bounds the innerproduct between a maximally entangled state and any product state.

Lemma 7 *Let $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. Let $|\Phi\rangle \in \mathcal{H}$ be a maximally entangled state. Let $d = \min(\dim \mathcal{H}_A, \dim \mathcal{H}_B)$. For all product states $|\phi_A\rangle \otimes |\phi_B\rangle$ of norm 1 we have*

$$|\langle\Phi|\phi_A\rangle \otimes |\phi_B\rangle|^2 \leq \frac{1}{d}. \tag{5.5.3}$$

Proof We write the maximally entangled state $|\Phi\rangle$ in the Schmidt polar form [138] as

$$|\Phi\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^{d} |a_i\rangle \otimes |b_i\rangle, \tag{5.5.4}$$

where $\langle a_i|a_j\rangle = \delta_{ij}$ and $\langle b_i|b_j\rangle = \delta_{ij}$. Thus we can write

$$|\langle\Phi|\phi_A\rangle\otimes|\phi_B\rangle|^2 = \frac{1}{d}\left|\sum_{i=1}^{d}\langle\phi_A|a_i\rangle\langle\phi_B|b_i\rangle\right|^2 \leq \frac{1}{d}, \tag{5.5.5}$$

using the Schwarz inequality and $\sum_{i=1}^{d}|\langle\phi_A|a_i\rangle|^2 \leq 1$ and $\sum_{i=1}^{d}|\langle\phi_B|b_i\rangle|^2 \leq 1$. □

We will also need the following lemma:

Lemma 8 *Let S be an unextendible product basis* $\{|\alpha_i\rangle\otimes|\beta_i\rangle\}_{i=1}^{|S|}$ *in* $\mathcal{H} = \mathcal{H}_A\otimes\mathcal{H}_B$. *Let*

$$f(|\phi_A\rangle,|\phi_B\rangle) = \sum_{i=1}^{|S|}|\langle\phi_A|\alpha_i\rangle|^2|\langle\phi_B|\beta_i\rangle|^2. \tag{5.5.6}$$

The minimum of f *over all pure states* $|\phi_A\rangle\in\mathcal{H}_A$ *and* $|\phi_B\rangle\in\mathcal{H}_B$ *exists and is strictly larger than 0.*

Proof The set of all pure product states $|\phi_A\rangle\otimes|\phi_B\rangle$ on $\mathcal{H}$ is a compact set. The function f is a continuous function on this set. Therefore, if there exists a set of states $|\phi_A\rangle\otimes|\phi_B\rangle$ for which f is arbitrary small then there would also exist a pair $|\phi'_A\rangle\otimes|\phi'_B\rangle$ for which $f=0$. This contradicts the fact that S is an unextendible product basis. □

The following two theorems contain the main result of this section.

Theorem 10 *Let S be an unextendible product basis* $\{|\alpha_i\rangle\otimes|\beta_i\rangle\}_{i=1}^{|S|}$ *in* $\mathcal{H} = \mathcal{H}_A\otimes\mathcal{H}_B$. *Let* ρ *be the density matrix*

$$\rho = \frac{1}{\dim\mathcal{H}-|S|}\left(\mathbf{1}_{AB}-\sum_{i=1}^{|S|}|\alpha_i\rangle\langle\alpha_i|\otimes|\beta_i\rangle\langle\beta_i|\right), \tag{5.5.7}$$

Let $d=\min(\dim\mathcal{H}_A,\dim\mathcal{H}_B)$. *Let* $|\Phi\rangle$ *be a maximally entangled state on* $d\otimes d$ *such that*

$$\langle\Phi|\,\rho\,|\Phi\rangle > 0, \tag{5.5.8}$$

and

$$\epsilon = \min_{|\phi_A\rangle\otimes|\phi_B\rangle}\sum_{i=1}^{|S|}|\langle\phi_A|\alpha_i\rangle|^2|\langle\phi_B|\beta_i\rangle|^2, \tag{5.5.9}$$

where the minimum is taken over all pure states $|\phi_A\rangle\in\mathcal{H}_A$ *and* $|\phi_B\rangle\in\mathcal{H}_B$. *Let* H *be a Hermitian operator given by*

$$\mathrm{H} = \sum_{i=1}^{|S|}|\alpha_i\rangle\langle\alpha_i|\otimes|\beta_i\rangle\langle\beta_i| - d\epsilon|\Phi\rangle\langle\Phi|. \tag{5.5.10}$$

For every unextendible product basis S there is a maximally entangled state $|\Phi\rangle$ *such that Eq. (5.5.8) holds. Moreover* H *has the following properties:*

$$\mathrm{Tr}\,\mathrm{H}\,\rho < 0, \tag{5.5.11}$$

and for all product states $|\phi_A\rangle \otimes |\phi_B\rangle \in \mathcal{H}$,

$$\mathrm{Tr}\,\mathrm{H}|\phi_A\rangle\langle\phi_A| \otimes |\phi_B\rangle\langle\phi_B| \geq 0. \tag{5.5.12}$$

Proof Eq. (5.5.12) follows from the definition of ϵ, Eq. (5.5.9), and Lemma 7. Consider Eq. (5.5.11). Since the density matrix ρ is proportional to the projector on $\mathcal{H}_S^{\perp}$, one has

$$\mathrm{Tr}\,\mathrm{H}\,\rho = -d\epsilon\,\langle\Phi|\,\rho\,|\Phi\rangle, \tag{5.5.13}$$

which is strictly smaller than zero by Lemma 8 and the choice of the maximally entangled state, Eq. (5.5.8). When $\dim\mathcal{H}_A = \dim\mathcal{H}_B$ there exists a basis of maximally entangled states and thus there will be a basis vector $|\Phi\rangle$ for which $\langle\Phi|\,\rho\,|\Phi\rangle$ is nonzero. In case, say, $\dim\mathcal{H}_A > \dim\mathcal{H}_B = d$, there exists a basis of maximally entangled states for every subspace $\mathcal{H}' = \mathcal{H}'_A \otimes \mathcal{H}_B$ with $\mathcal{H}'_A \subset \mathcal{H}_A$ and $\dim\mathcal{H}'_A = d$. Therefore there will be a maximally entangled state $|\Phi\rangle$ such that $\langle\Phi|\,\rho\,|\Phi\rangle$ is nonzero. This completes the proof. □

Remark Note that H is the entanglement witness that was introduced in Lemma 3, section 5.3.

Theorem 11 *Let S be an unextendible product basis $\{|\alpha_i\rangle \otimes |\beta_i\rangle\}_{i=1}^{|S|}$ in $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. Let H be defined as in Theorem 10, Eq. (5.5.10). Choose an orthonormal basis $\{|i\rangle\}_{i=1}^{\dim\mathcal{H}_A}$ for $\mathcal{H}_A$. Let $\mathcal{S}\colon\ B(\mathcal{H}_A) \to B(\mathcal{H}_B)$ be a linear map defined by*

$$\mathcal{S}\,(|i\rangle\langle j|) = \langle i|\,\mathrm{H}\,|j\rangle. \tag{5.5.14}$$

Then $\mathcal{S}$ is positive but not completely positive. Moreover, $\mathcal{S}$ is indecomposable.

Proof The relation between $\mathcal{S}$ and H, Eq. (5.5.14), follows from the isomorphism between Hermitian operators on $\mathcal{H}_A \otimes \mathcal{H}_B$ with the property of Eq. (5.5.12) and linear positive maps, see Refs. [112, 139]. In particular, iff a Hermitian H operator on $\mathcal{H}_A \otimes \mathcal{H}_B$ has the property of Eq. (5.5.12) then the linear map $\mathcal{R}\colon\ B(\mathcal{H}_A) \to B(\mathcal{H}_B)$ defined by

$$\mathrm{H} = \sum_{i,j} (|i\rangle\langle j|)^{\dagger} \otimes \mathcal{R}(|i\rangle\langle j|), \tag{5.5.15}$$

is positive for any choice of the orthonormal basis $\{|i\rangle\}_{i=1}^{\dim\mathcal{H}_A}$. The map $\mathcal{S} = \mathcal{R} \circ T$, where T is matrix transposition in the basis $\{|i\rangle\}_{i=1}^{\dim\mathcal{H}_A}$, in Eq. (5.5.14) is then positive as well.

Using the density matrix ρ that is derived from the unextendible product basis, Eq. (5.5.7), we will show that $\mathcal{S}$ is not completely positive. At the same time we will prove that the assumption that $\mathcal{S}$ is decomposable leads to a contradiction. We can rewrite Eq. (5.5.14) as

$$\mathrm{H} = (\mathrm{id}_A \otimes \mathcal{S})(|\Phi^+\rangle\langle\Phi^+|), \tag{5.5.16}$$

where $|\Phi^+\rangle$ is equal to the (unnormalized) maximally entangled state $\sum_{i=1}^{\dim \mathcal{H}_A} |i\rangle \otimes |i\rangle$. Let $\mathcal{S}^\dagger\colon\ B(\mathcal{H}_B) \to B(\mathcal{H}_A)$ be the Hermitian conjugate of the map $\mathcal{S}$. We use the definition of $\mathcal{S}^\dagger$

$$\mathrm{Tr}\,\mathcal{S}^\dagger(A^\dagger)\,B = \mathrm{Tr}A^\dagger\,\mathcal{S}(B), \tag{5.5.17}$$

and Eq. (5.5.16) to derive that Eq. (5.5.11) can be rewritten as

$$\mathrm{Tr}\,\mathrm{H}\,\rho = \langle\Phi^+|\left(\mathrm{id}_A \otimes \mathcal{S}^\dagger\right)(\rho)|\Phi^+\rangle < 0. \tag{5.5.18}$$

Thus $\mathcal{S}^\dagger$ cannot be completely positive and therefore $\mathcal{S}$ itself is not completely positive. If $\mathcal{S}$ were decomposable, then $\mathcal{S}^\dagger$ would be of the form $\mathcal{S}_1 + T \circ \mathcal{S}_2$ where $\mathcal{S}_1$ and $\mathcal{S}_2$ are completely positive maps. The density matrix ρ is positive semidefinite under any positive linear map of the form $\mathcal{S}_1 + T \circ \mathcal{S}_2$ by Proposition 6. This is in contradiction with Eq. (5.5.18) and therefore $\mathcal{S}$ cannot be decomposable. □

5.5.3 Examples and Discussion

As we have shown the structure of unextendible product bases carries over to indecomposable positive linear maps. The results on unextendible product bases that we presented in sections 5.4.2-5.4.3 give us many examples of these indecomposable positive linear maps. In this section we will take two examples of unextendible product bases and demonstrate the construction of Theorem 10 and Theorem 11.

Example 1: Consider the **Pent** UPB, Eq. (5.4.3). Let $\rho_{\mathbf{Pent}}$ be the bound entangled state derived from **Pent** as in Eq. (5.4.6), Proposition 6. We choose a maximally entangled state $|\Phi\rangle$, here named $|\Phi^+\rangle$,

$$|\Phi^+\rangle = \frac{1}{\sqrt{3}}(|11\rangle + |22\rangle + |33\rangle). \tag{5.5.19}$$

One can easily compute that

$$\langle\Phi^+|\,\rho_{\mathbf{Pent}}\,|\Phi^+\rangle = \frac{1}{4}\left(1 - \frac{7+\sqrt{5}}{3(3+\sqrt{5})}\right) > 0. \tag{5.5.20}$$

The map $\mathcal{S}$ as defined in Eq. (5.5.14) Theorem 11, follows directly:

$$\mathcal{S}(|i\rangle\langle j|) = \sum_{k=0}^{4}\langle i|v_k\rangle\langle v_k|j\rangle|v_{2k \bmod 5}\rangle\langle v_{2k \bmod 5}| - \epsilon|i\rangle\langle j|. \tag{5.5.21}$$

A positive linear map $\mathcal{S}\colon\ B(\mathcal{H}_n) \to B(\mathcal{H}_m)$ is called unital if $\mathcal{S}(\mathbf{1}_n) = \mathbf{1}_m$. We will demonstrate that $\mathcal{S}$ is not unital. One can write

$$\mathcal{S}(\mathbf{1}_A) = \mathrm{Tr}_A\,\mathrm{H} = \sum_{k=0}^{4}|v_{2k \bmod 5}\rangle\langle v_{2k \bmod 5}| - 3\epsilon\;\mathrm{Tr}_A|\Phi^+\rangle\langle\Phi^+|, \tag{5.5.22}$$

which in turn is equal to

$$\mathcal{S}(\mathbf{1}_A) = \mathrm{diag}\left(\frac{10}{5+\sqrt{5}}, \frac{10}{5+\sqrt{5}}, \sqrt{5}\right) - \epsilon\,\mathbf{1}_B. \tag{5.5.23}$$

A numerical approximation of ϵ as defined in Eq. (5.5.9) Theorem 10, gives the value

$$\epsilon \approx 0.037911, \tag{5.5.24}$$

but we don't know whether this is the minimum of the function in Eq. (5.5.9).

The next example is based on a more general unextendible product bases that was presented in Ref. [129].

Example 2: The states of S, **Tiles3n** in $\mathcal{H}_3 \otimes \mathcal{H}_n$ are

$$|F_k^0\rangle = \frac{1}{\sqrt{n-2}}|0\rangle \otimes (|1\rangle + \sum_{l=3}^{n-1} \omega^{k(l-2)}|l\rangle), \quad 1 \le k \le n-3, \tag{5.5.25}$$

$$|F_k^1\rangle = \frac{1}{\sqrt{n-2}}|1\rangle \otimes (|2\rangle + \sum_{l=3}^{n-1} \omega^{k(l-2)}|l\rangle), \quad 1 \le k \le n-3, \tag{5.5.26}$$

$$|F_k^2\rangle = \frac{1}{\sqrt{n-2}}|2\rangle \otimes (|0\rangle + \sum_{l=3}^{n-1} \omega^{k(l-2)}|l\rangle), \quad 1 \le k \le n-3, \tag{5.5.27}$$

$$|\psi_3\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \otimes |0\rangle, \tag{5.5.28}$$

$$|\psi_4\rangle = \frac{1}{\sqrt{2}}(|1\rangle - |2\rangle) \otimes |1\rangle, \tag{5.5.29}$$

$$|\psi_5\rangle = \frac{1}{\sqrt{2}}(|2\rangle - |0\rangle) \otimes |2\rangle, \tag{5.5.30}$$

$$|\psi_6\rangle = \frac{1}{\sqrt{3n}}\sum_{i=0}^{2}\sum_{j=0}^{n-1} |i\rangle \otimes |j\rangle, \tag{5.5.31}$$

and we have $\omega = \exp(2\pi i/(n-2))$. Here the states $\{|k\rangle\}_{k=0}^{n-1}$ form an orthonormal basis. In total there are $3n-5$ states in the basis. We choose a maximally entangled state, again we take $|\Phi^+\rangle$, Eq. (5.5.19). One can show that

$$\langle\Phi^+|\,\rho_{\mathbf{Tiles3n}}\,|\Phi^+\rangle = \frac{1}{5}\left(\frac{1}{2} - \frac{1}{3n}\right) > 0. \tag{5.5.32}$$

The map $\mathcal{S}\colon\ B(\mathcal{H}_3) \to B(\mathcal{H}_n)$ is given as

$$\mathcal{S}(|i\rangle\langle j|) = \sum_{k=1}^{n-3}\sum_{p=0}^{2}\langle i|F_k^p\rangle\langle F_k^p|j\rangle + \sum_{p=3}^{6}\langle i|\psi_p\rangle\langle\psi_p|j\rangle - \epsilon\,|i\rangle\langle j|. \tag{5.5.33}$$

The following questions concerning the positive linear maps that were introduced in this paper are left open.

1. Is $\mathcal{S}$ always non unital? We conjecture it is. As we showed, see Eq. (5.5.22), the answer to this question depends on whether

$$\sum_{i=1}^{|S|} |\beta_i\rangle\langle\beta_i| = c\mathbf{1}_B, \tag{5.5.34}$$

where the set of states $\{|\beta_i\rangle\}_{i=1}^{|S|}$ are one side of the unextendible product basis and c is some constant. The states $|\beta_i\rangle$ will span $\mathcal{H}_B$ but they will not be all orthogonal, nor all non-orthogonal.

2. It was shown in Theorem 11 that the new indecomposable positive linear maps $\mathcal{S}\colon\ B(\mathcal{H}_m) \rightarrow B(\mathcal{H}_n)$ are not m-positive, as they are not completely positive. Are these maps $\mathcal{S}$ k-positive for some k with $1 < k < m$? The answer to this question will require on a better understanding of the structure of unextendible product bases.

3. In Ref. [136] (see section 5.4.6) a single example was given of a entangled density matrix on $\mathcal{H}_3 \otimes \mathcal{H}_4$ which was positive semidefinite under the map $\mathrm{id}_3 \otimes T$. The density matrix was based not on an unextendible product basis, but a 'barely' completable product basis S. We did show that the Hilbert space $\mathcal{H}_S^{\perp}$ had a product state *deficit*, i.e. the number of product states in $\mathcal{H}_S^{\perp}$ was less than $\dim \mathcal{H}_S^{\perp}$. It is open question how to generalize this example and whether these kinds of density matrices will give rise to more general family of indecomposable positive linear maps.

5.6 Discussion

Both the bound entangled states and the unextendible product bases from which we derived the bound entangled states exhibit a form of local irreversibility. A bound entangled state can be viewed as the result of a local entropy-increasing process on a pure entangled state. It is the nature of a bound entangled state that this process cannot be reversed locally; we cannot distill pure entanglement out of a bound entangled state. The states in an unextendible product basis are states that can be prepared locally. The uniform ensemble of the UPB states can be viewed as the result of an entropy-increasing process in which the local preparers forget which state was prepared. We have shown that this process cannot be reversed locally, as the preparers are not able to confidently distinguish the members of the ensemble.

The question what minimal nonlocal means are needed to undo both kinds of processes is not answered by this work. Another question that is related to this work, is the question of the use of bound entangled states. Bound entangled states are not helpful in the transmission of quantum data via a teleportation protocol; it has been shown [140] that any (attempt at) teleportation that is done with the use of bound entangled states can also be done without the use of entanglement. It is possible that bound entangled states are instrumental in the implementation of separable superoperators and measurements.

Bibliography

[1] R. P. Feynman, Foundations of Physics **16**, 507 (1986).

[2] R. Landauer, IBM J. Res. Develop. **5**, 183 (1961).

[3] C. H. Bennett, IBM J. Res. Develop. **17**, 525 (1973).

[4] D. Deutsch, Proc. Roy. Soc. Lond. A **400**, 97 (1985).

[5] D. Deutsch and R. Jozsa, Proc. Roy. Soc. Lond. A **439**, 553 (1992).

[6] S. Wiesner, SIGACT News **15**: 1, 78 (1983)

[7] C. H. Bennett and G. Brassard, *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing* (IEEE Press, 1984), 175.

[8] P. W. Shor, *Proceedings of the 35th Annual Symposium on the Foundations of Computer Science* (IEEE Press, 1994), 124.

[9] W. K. Wootters and W. H. Zurek, Nature **299**, 802 (1982).

[10] D. Mayers, "Unconditional Security in Quantum Cryptography", quant-ph/9802025.

[11] David DiVincenzo, private communication.

[12] A. Muller, H. Zbinden and N. Gisin, Europhys. Lett. **33**, 335 (1996).

[13] W. T. Buttler, R. J. Hughes, P. G. Kwiat, S. K. Lamoreaux, G. G. Luther, G. L. Morgan, J. E. Nordholt, C. G. Peterson and C. M. Simmons, Phys. Rev. Lett. **81**, 3283 (1998).

[14] L. K. Grover, *Proceedings of the 28th Annual ACM Symposium on Theory of Computing* (ACM, 1996), 212.

[15] J. I. Cirac and P. Zoller, Phys. Rev. Lett. **74**, 4091 (1995).

[16] N. Gershenfeld and I. L. Chuang, Science **275**, 350 (1997); D. Cory, A. Fahmy and T. Havel, Proc. Nat. Acad. Sci. **94**, 1634 (1997).

[17] D. Loss and D. P. DiVincenzo, Phys. Rev. A **57**, 120 (1998).

[18] I. L. Chuang, N. Gershenfeld and M. Kubinec, Phys. Rev. Lett. **80**, 3408, (1998); J. A. Jones and M. Mosca, J. Chem. Phys. **109**, 1648 (1998).

[19] Q. A. Turchette, C. S. Wood, B. E. King, C. J. Myatt, D. Leibfried, W. M. Itano, C. Monroe and D. J. Wineland, Phys. Rev. Lett. **81**, 3631 (1998).

[20] D. G. Cory, M. D. Price, W. Maas, E. Knill, R. Laflamme, W. H. Zurek, T. F. Havel and S. S. Somaroo, Phys Rev. Lett. **81**, 2152 (1998); D. Leung, L. V. Vandersypen, X. Zhou, M. Sherwood, N. Yannoni, M. Kubinec and I. Chuang, "Experimental Realization of A Two Bit Phase Damping Quantum Code", quant-ph/9811068.

[21] B. Schumacher, Phys. Rev. A **51**, 2738 (1995).

[22] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres and W. K. Wootters, Phys. Rev. Lett. **70**, 1895 (1993).

[23] P. W. Shor, Phys. Rev. A **52**, 2493 (1995).

[24] D. Bouwmeester, J. -W. Pan, K. Mattle, M. Eibl, H. Weinfurter and A. Zeilinger, Nature **390**, 575 (1997).

[25] D. Boschi, S. Branca, F. De Martini, L. Hardy and S. Popescu, Phys. Rev. Lett. **80**, 1121 (1998).

[26] M. Nielsen, E. Knill and R. Laflamme, Nature **395**, 5 (1998).

[27] A. Furusawa, J. L. Sorensen, S. L. Braunstein, C. A. Fuchs, H. J. Kimble and E. S. Polzik, Science **282**, 706 (1998).

[28] H. Buhrman, R. Cleve and A. Wigderson, *Proceedings of the 30th Annual ACM Symposium on the Theory of Computing* (ACM Press, 1998), 63.

[29] A. Ambainis, L. Schulman, A. Ta-Shma, U. Vazirani and A. Wigderson, *Proceedings of the 39th Annual Symposium on the Foundations of Computer Science* (IEEE Press, 1998), 342.

[30] C. Fuchs, Phys. Rev. Lett. **79**, 1162 (1997).

[31] C. H. Bennett, P. W. Shor, J. A. Smolin and A. V. Thapliyal, "Entanglement-Assisted Classical Capacity of Noisy Quantum Channels", quant-ph/9904023.

[32] S. Lloyd, Science **273**, 1073 (1996).

[33] S. S. Somaroo, C. H. Tseng, T. F. Havel, R. Laflamme and D. G. Cory, Phys. Rev. Lett. **82**, 5381 (1999).

[34] A. Holevo, IEEE Trans. Inf. Theory **44**, 269 (1998).

[35] A. Holevo, Prob. Inf. Transm. **15**, 247 (1979).

[36] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin and W. K. Wootters, Phys. Rev. Lett. **76**, 722 (1996).

[37] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Phys. Rev. A **54**, 3824 (1996).

[38] D. P. DiVincenzo, "Topics in Quantum Computers", in *Mesoscopic Electron Transport*, Vol. 345 of NATO Advanced Study Institute, Series E: Applied Sciences, eds. L. Sohn, L. Kouwenhoven, and G. Schoen (Kluwer, Dordrecht, 1997), 657.

[39] A. Steane, Appl. Phys. B**64**, 623 (1997).

[40] Private communication, Herman te Riele, CWI.

[41] M. Boyer, G. Brassard, P. Høyer, and A. Tapp, Fortschr. Phys. **46** (4-5), 493 (1998).

[42] C. H. Bennett, E. Bernstein, G. Brassard and U. Vazirani, SIAM Journal on Computing, **26**(5), 1510-1523 (1997).

[43] G. Brassard and P. Høyer, *Proceedings of the Fifth Israeli Symposium on Theory of Computing and Systems (ISTCS'97)* (IEEE Computer Science Press, 1997), 12.

[44] G. Brassard, P. Høyer, and A. Tapp, *Proceedings of the 25th ICALP* (Springer-Verlag, 1998), 820.

[45] L. Grover, *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing* (ACM Press, 1998), 53.

[46] E. Bernstein and U. Vazirani, *Proceedings of the 25th Annual ACM Symposium on Theory of Computing* (ACM Press, 1993), 11.

[47] E. Bernstein and U. Vazirani, SIAM Journal on Computing **26**(5), 1411-1473 (1997).

[48] cf. Martin Aigner, *Combinatorial Search* (John Wiley & Sons, London, 1988).

[49] A. S. Holevo, Problemy Peredachi Informatsii **9**, 3 (1973). This paper has appeared in English translation in Problems of Information Transmission **9**, 177 (1973).

[50] A. Barenco, C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. A. Smolin and H. Weinfurter, Phys. Rev. A **52**, 3457 (1995).

[51] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley Series in Telecommunications, London, 1991).

[52] W. Feller, *An introduction to Probability Theory and Its Applications*, Vol. I (John Wiley & Sons, London, 1957), page 33.

[53] R. Beals, H. Buhrman, R. Cleve, M. Mosca, and R. de Wolf, *Proceedings of the 39th IEEE Symposium on Foundations of Computer Science* (IEEE Press, 1998), 351.

[54] E. Farhi, J. Goldstone, S. Gutmann and M. Sipser, Phys. Rev. Lett. **81**, 5442 (1998).

[55] Y. Ozhigov, "Quantum Computer Can Not Speed Up Iterated Applications of a Black Box", quant-ph/9712051.

[56] D. S. Abrams and S. Lloyd, Phys. Rev. Lett. **79**, 2586 (1997).

[57] M. Brune, E. Hagley, J. Dreyer, X. Maitre, A. Maali, C. Wunderlich, J. M. Raimond and S. Haroche, Phys. Rev. Lett. **77**, 4887 (1996).

[58] C. Zalka, Proc. Roy. Soc. Lond. A **454**, 313 (1998); S. Wiesner, "Simulations of Many-Body Quantum Systems by a Quantum Computer", quant-ph/9603028; B. M. Boghosian and W. Taylor, Physica D **120** (1998) 30.

[59] B. Schumacher, Phys. Rev. A **54**, 2614 (1996).

[60] I. L. Chuang and M.A. Nielsen, J. Mod. Opt. **44**, 2455 (1997).

[61] M.-D. Choi, Lin. Alg. and Its Appl. **10**, 285 (1975).

[62] A. Peres, *Quantum Theory: Concepts and Methods* (Kluwer, Dordrecht, 1993).

[63] http://stout.physics.ucla.edu/~smolin/tetrahedron

[64] D. Cox, J. Little, and D. O'Shea, *Ideals, Varieties, and Algorithms* (Springer-Verlag, New York, 1992).

[65] W. Bosma, J. J. Cannon, and C. Playoust, Journal of Symbolic Computation **24**, 235 (1997).

[66] cf. E. Knill "Non-binary unitary error bases and quantum codes", quant-ph/9608048, and D. I. Fivel, Phys. Rev. Lett. **74**, 835 (1995).

[67] A. Kitaev, Russian Math. Surveys **52**: 6, 1191 (1997); A. Kitaev, "Quantum measurements and the abelian stabilizer Problem", quant-ph/9511026.

[68] J. M. Hammersley and D. C. Handscomb, *Monte Carlo Methods* (John Wiley & Sons, London, 1964).

[69] M. Suzuki ed., *Quantum Monte Carlo Methods in Equilibrium and Nonequilibrium Systems*, Springer Series in Solid-state Sciences (Springer-Verlag, Berlin, 1986).

[70] H. De Raedt and W. von der Linden, "Quantum Lattice Problems" in *The Monte Carlo Method in condensed matter physics*, Topics in Applied Physics Vol. 71 (Springer-Verlag, Berlin, 1991).

[71] M. Suzuki, Prog. Theor. Phys. **56**, 1454 (1976).

[72] S. Doniach and E. H. Sondheimer, *Green's functions for solid state physicists* (Benjamin-Cummings, Reading MA, 1974).

[73] M. Jerrum and A. Sinclair, SIAM Journal of Computation **22**(5), 1087 (1993).

[74] F. Barahona, J. Physics. A **15**, 3241 (1982).

[75] In [76] quantum algorithms were presented to sample from the equilibrium distribution of arbitrary Ising spin glass models. The performance of these algorithms does not necessarily provide an exponential speedup over the best classical algorithms (see [73]) for the hard instances though.

[76] D. Lidar and O. Biham, Phys. Rev. E **56**, 3661 (1997).

[77] D. Abrams and S. Lloyd, "A quantum algorithm providing exponential speed increase for finding eigenvalues and eigenvectors", quant-ph/9807070.

[78] L. van Hove, Physica **23**, 441 (1957).

[79] E. Fick, G. Sauermann and W. D. Brewer, *Quantum Statistics of Dynamic Processes*, Springer Series in Solid-State Sciences Vol. 86 (Springer-Verlag, Berlin, 1990).

[80] M. Celio and D. Loss, Physica A **158**, 769-783 (1989).

[81] R. Alicki and K. Lendi, *Quantum Dynamical Semigroups and Applications*, Lec. Notes in Physics (Springer-Verlag, Berlin, 1987).

[82] G. Lindblad, Comm. Math. Phys. **48**, 119-130 (1976).

[83] E. B. Davies, *Quantum Theory of Open Systems* (Academic, New York, 1976).

[84] E. B. Davies, Comm. Math. Phys. **39**, 91-110 (1974).

[85] E. B. Davies, Math. Annalen **219**, 147-158 (1976).

[86] R. Alicki, Rep. Math. Phys. **10**, 249 (1976).

[87] N. Dunford and J. T. Schwartz, *Linear Operators*, Part I (Interscience, New York, 1958).

[88] R. F. Werner, private communication.

[89] A. Messiah, *Quantum Mechanics*, Vol. II, Chapter XVI (J. Wiley & Sons, London, 1976).

[90] R. A. Horn and C. R. Johnson, *Matrix Analysis* (Cambridge University Press, Cambridge, 1985).

[91] cf. M. Jerrum and A. Sinclair, *Approximation Algorithms for NP-hard Problems*, ed. by D. S. Hochbaum (PWS Publishing, Boston, 1996), 482.

[92] A. Abragam, *The principles of Nuclear Magnetism* (Oxford University Press, Oxford, 1961).

[93] S. Richter and R. F. Werner, J. Stat. Phys. **82**, 963 (1996).

[94] D. Aharonov, A. Kitaev, N. Nissan, *Proceedings of the 30th Annual ACM Symposium on Theory of Computation* (ACM, 1997), 20.

[95] A. L. Fetter and J. D. Walecka, *Quantum Theory of Many-Particle Systems* (McGraw-Hill Book Company, New York, 1971).

[96] R. Cleve, A. Ekert, C. Macchiavello and M. Mosca, Proc. Roy. Soc. Lond. A **454**, 339 (1998).

[97] R. Jozsa, Proc. R. Soc. Lond. A **454**, 323 (1998).

[98] D. A. Lidar and H. Wang, Phys. Rev. E **59**, 2429 (1999).

[99] cf. G. W. Stewart and Ji-guang Sun, *Matrix Perturbation Theory*, Computer science and scientific computing (Academic, San Diego, 1990); P. J. Schweitzer, Journal of Appl. Prob. **5**, 401 (1968); M. Haviv and L. van der Heyden, Adv. Appl. Prob. **16**, 804 (1984).

[100] L. Viola, E. Knill and S. Lloyd, Phys. Rev. Lett. **82**, 2417 (1999).

[101] A. Wehrl, Rev. Mod. Phys. **50** 221 (1978).

[102] T. Ando, Lin. Alg. and Its Appl. **199**, 17 (1994).

[103] M. Plischke and B. Bergersen, *Equilibrium Statistical Physics* (World Scientific, Singapore, 1994).

[104] K. Zyczkowski, P. Horodecki, A. Sanpera and M. Lewenstein, Phys. Rev. A **58**, 883 (1998).

[105] E. Schrödinger, Proc. Cambridge Phil. Soc. **31**, 555 (1935).

[106] A. Einstein, B. Podolsky, N. Rosen, Phys. Rev **47**, 777 (1935).

[107] D. Bohm, *Quantum Theory* (Prentice Hall, New York, 1951), page 614.

[108] M. Horodecki, P. Horodecki and R. Horodecki, "Limits for entanglement measures", quant-ph/9908065.

[109] W. K. Wootters, Phys. Rev. Lett. **80**, 2245 (1998).

[110] E. Rains, "An Improved Bound on Distillable Entanglement", quant-ph/9809082.

[111] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. Lett. **78**, 574 (1997).

[112] M. Horodecki, P. Horodecki and R. Horodecki, Phys. Lett. A **223**, 1 (1996).

[113] G. Lindblad, Comm. Math. Phys. **40**, 147 (1975).

[114] S. L. Woronowicz, Rep. Math. Phys. **10**, 165 (1976).

[115] A. Peres, Phys. Rev. Lett. **77**, 1413 (1996).

[116] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. Lett. **80**, 5239 (1998).

[117] P. Horodecki, Phys. Lett. A **232**, 333 (1997).

[118] R. T. Rockafellar, *Convex Analysis* (Princeton University Press, Princeton, 1970).

[119] J. S. Bell, Physics **1**, 195 (1964).

[120] A. Aspect, J. Dalibard and G. Roger, Phys. Rev. Lett. **49**, 1804 (1982).

[121] A. Peres, "All the Bell Inequalities", quant-ph/9807017.

[122] A. Garg and N. D. Mermin, Foundations of Physics **14**, 1 (1984).

[123] I. Pitowsky, Mathematical Programming **50**, 395 (1991).

[124] R. F. Werner, "Quantum States with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model", Phys. Rev. A **40**, 4277 (1989).

[125] C. H. Bennett, D. P. DiVincenzo, C. A. Fuchs, T. Mor, E. Rains, P. W. Shor, J. A. Smolin, and W. K. Wootters, Phys. Rev. A **59**, 1070 (1999).

[126] M. Hillery, V. Buzek and A. Berthiaume, Phys. Rev. A **59**, 1829 (1999).

[127] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, Oxford University Press (1979).

[128] Private communication, P. Horodecki.

[129] D. P. DiVincenzo, T. Mor, P. W. Shor, J. A. Smolin and B. M. Terhal, "Unextendible Product Bases, Uncompletable Product Bases and Bound Entanglement", quant-ph/9908070.

[130] G. Brassard and T. Mor, "Multi-particle entanglement via 2-particle entanglement", presented at NASA QCQC'98 conference.

[131] M. Horodecki and P. Horodecki, Phys. Rev. A **59**, 4206 (1999).

[132] A. Uhlmann, Open Sys. & Information Dyn. **5**, 209 (1998).

[133] E. Rains, "A Rigorous Treatment of Distillable Entanglement", quant-ph/9809078.

[134] M. -D. Choi, Lin. Alg. and Its Appl. **12**, 95 (1975).

[135] H. Osaka, Lin. Alg. and Its Appl. **186**, 45 (1993).

[136] C. H. Bennett, D. P. DiVincenzo, T. Mor, P. W. Shor, J. A. Smolin and B. M. Terhal, Phys. Rev. Lett. **82**, 5385 (1999).

[137] cf. D. I. Fivel, Phys. Rev. Lett. **74**, 835 (1995).

[138] L. Hughston, R. Jozsa and W. Wootters, Phys. Lett. A **183**, 14 (1993).

[139] A. Jamiołkowski, Rep. Math. Phys. **3**, 275 (1972).

[140] M. Horodecki, P. Horodecki, and R. Horodecki, "General Teleportation Channel, Singlet Fraction and Quasi-distillation", quant-ph/9807091.

Samenvatting

Het onderzoeksgebied quantum informatie verwerking ligt op het grensvlak tussen de natuurkunde en de informatica. Het informatica aspect komt tot uitdrukking in het feit dat we bij dit onderzoek geïnteresseerd zijn in computers en informatie verwerkende processen en communicatie. Het fysica aspect is erin gelegen dat we voor het uitvoeren van deze computationele processen meerwaarde proberen te vinden in het gebruik van quantum mechanische fenomenen.

De fysische technologie die aan de basis staat van de huidige generatie computers is sinds de jaren '60 niet veel veranderd. De ontwikkelingen in de halfgeleider technologie hebben ons in staat gesteld steeds kleinere en snellere rekenchips en logisch geheugen te bouwen. Deze trend gaat redelijk gelijk op met de voortuitgang in technieken van magnetische opslag en retrieval, die ervoor zorgen dat op een harde schijf van heden ten dage gigabytes kunnen worden opgeslagen waar dat nog één megabyte betrof 10 jaar geleden.

Als we gaan kijken hoe een chip er in detail uitziet, dan vinden we dat het een dun stukje gedoteerd silicium is met daarop geëtst een netwerk van transistoren. De transistoren worden gebruikt om onder andere logische poorten te maken. De meest elementaire poort die men met een transistor kan bouwen is een NOT poort. Het is een apparaatje dat één bit als input en één bit als output heeft. Als het input bit een 1 is, is de output een 0 en vice versa. In de transistor wordt dit effect bewerkstelligd doordat als de input een 1 is, een klein voltage op de 'gate' van de transistor wordt gezet. De transistor geleidt dan stroom door de gedoteerde laagjes silicium en de output is een 0. In afwezigheid van een voltage, wat de input 0 representeert, is de transistor niet geleidend en is de output een 1. De transistor geleidt niet als de mobiele ladingsdragers in het silicium, de elektronen of juist de afwezigheid daarvan, een 'potentiaal landschap' zien dat hen er niet toe aanzet om zich te verplaatsen. U kan de ladingsdragers vergelijken met water wat nu eenmaal niet graag bergopwaarts stroomt. Om de aard van dit potentiaal landschap te begrijpen, hebben we uiteindelijk de taal van de quantum mechanica nodig.

Als we de huidige trend van de miniaturisatie van chips doorzetten naar de toekomst, dan zullen we rond het jaar 2020 misschien met bouwstenen werken van atomaire (10^{-10} meter) grootte. Op deze schaal is de quantum mechanica onontbeerlijk in de beschrijving van hoe de materie zich gedraagt. Het opschalen van de huidige computer architectuur naar dit regime wordt, voor zover het technologisch mogelijk zou zijn, nog ernstig beperkt door het probleem van oververhitting van een dergelijke 'nanochip'. Het ligt dan voor de hand om te vragen of we de capaciteit van de materie niet beter gebruiken als we quantum toestanden als onze elementaire bouwstenen gebruiken in plaats van bits en bytes. We hebben het dan over een quantum computer. Waar een klassiek bit slechts twee waardes aan kan nemen, een 1 of een 0, heeft een quantum bit een hele twee dimensionale complexe vectorruimte tot zijn beschikking, waarin de klassieke 1 en 0 slechts twee loodrecht staande vectoren zijn. Essentiële eigenschappen van quantum toestanden zoals interferentie, entanglement en superpositie staan aan de basis van dit nieuwe computer model.

In de afgelopen 15 jaar heeft men voor diverse toepassingen gevonden dat het gebruik van quantum bits grote voordelen kan opleveren. Deze voordelen zijn gevonden in de vorm van de snelheid waarmee op een quantum computer een probleem kan worden opgelost. Het quantum algoritme van Peter Shor is hiervan het grote voorbeeld; met zijn algoritme zou men op een quantum computer een groot getal in priemfactoren kunnen ontbinden in een tijdsbestek van minuten, waar 300 PCs twee maanden voor nodig hebben. Een ander voorbeeld ligt in de cryptografie. Met behulp van quantum toestanden is het mogelijk een geheime cryptografische sleutel aan te maken tussen twee partijen ondanks dat er tussen deze twee partijen een onveilig kanaal bestaat, waarop luistervinken berichten kunnen onderscheppen en afluisteren.

Het probleem van de quantum bits is echter dat ze veel meer dan klassieke bits een grote mate van fragiliteit kennen. Als wij ze willen gebruiken om mee te rekenen, moeten we ze enerzijds hermetisch afsluiten van oncontroleerbare invloeden van buitenaf. Anderzijds moet de afsluiting van de buitenwereld niet totaal zijn; we willen de quantum bits ook kunnen manipuleren en meten. Het is een reusachtige uitdaging voor de gemeenschap van experimentele fysici om quantum systemen te bouwen die aan dergelijke eisen voldoen. Op het moment houdt men zich pas met de allereerste bits van een quantum processor bezig.

Er zijn twee hoofdlijnen aanwezig in dit proefschrift. Enerzijds bestuderen we de kracht van een quantum computer. Anderzijds worden fundamentele eigenschappen van quantum toestanden in kaart gebracht.

De grote vraag over de vooralsnog 'papieren quantum computer' is welke problemen hij sneller kan oplossen dan een klassieke (gewone) computer en voor welke problemen hij geen voordeel biedt. Deze vraag is niet eenvoudig; zelfs voor klassieke computers is het voor een grote klasse van problemen, de zogeheten NP-complete problemen, niet bewezen maar wel algemeen aanvaard dat een klassieke computer ze niet efficiënt kan oplossen. In Hoofdstuk 2 van dit proefschrift geven we drie soorten problemen. Het eerste probleem is het bepalen van het gemiddelde van een functie en we laten zien dat een redelijke, maar geen exponentieel grote, tijdswinst behaald kan worden op een quantum computer. De tweede groep problemen waarvoor de quantum computer een aanzienlijk voordeel biedt, zijn problemen van informatie extractie met als voorbeeld het munt-weeg probleem. Stel er is een verzameling van n munten, waarvan er een onbekend aantal vals zijn. Deze valse munten kenmerken zich erdoor dat zij in hun gewicht verschillen van de echte munten; van beide kennen we echter het gewicht. Alle echte munten hebben een bepaald gewicht E en alle valse munten hebben een bepaald gewicht V. Door deelverzamelingen van munten te wegen, bijvoorbeeld iedere munt afzonderlijk, kunnen we bepalen welke munten vals en welke munten echt zijn. Welke wegingen zullen wij uitvoeren en hoeveel wegingen hebben we nodig? Als we de wegingen klassiek uitvoeren, dan kan men bewijzen dat tenminste $\frac{n}{\log(n+1)}$ wegingen nodig zijn. Als we de wegingen met behulp van een quantum computer mogen uitvoeren, dan kan men laten zien dat slechts één weging voldoende is. Als laatste geven we een probleem waarvoor geen winst behaald kan worden door het gebruik van een quantum computer.

In de Hoofdstukken 3 en 4 doen we een uitgebreide studie naar het simuleren van fysische systemen op een quantum computer. Stelt U zich het volgende experiment voor. We hebben een pannetje met water dat zich op een temperatuur bevindt van 80 graden Celsius. In het pannetje plaatsen we een bekertje koude melk uit de koelkast. Na verloop van tijd, als we het water op constante temperatuur houden, zal de melk dezelfde temperatuur als het water aannemen, 80 graden Celsius. We zeggen dan dat we de melk 'au bain Marie' hebben opgewarmd of in de natuurkunde dat de melk in thermisch evenwicht is gekomen met het water. Dit experiment, of een versie ervan waarbij het water, melk, pannetje en bekertje quantum mechanische systemen zijn, is een experiment dat men zou willen simuleren op een quantum computer. Het einddoel van het experiment is een bekertje melk van 80 graden Celsius. In Hoofdstuk 4 beschrijven we twee algoritmes die een dergelijke proces nabootsen. We laten zien dat een quantum computer een natuurlijke omgeving is om deze simulatie van quantum systemen uit te voeren; voor zover we weten kan een klassieke computer deze algoritmes niet op efficiente wijze uitvoeren.

Nadat ons bekertje melk in thermisch evenwicht is gekomen met de pan water, zouden we misschien het volgende experiment kunnen uitvoeren. Op een bepaald tijdstip laten we een kleine druppel honing in de melk vallen. Een tijdje later halen we het bekertje uit de pan en meten we hoe stroperig de melk is. Dit vergelijken we met de situatie waarin we geen honing toevoegen aan de melk. Ook voor een dergelijk soort experiment, waarin de honing etc. weer quantum mechanische systemen voorstellen, beschrijven we in Hoofdstuk 4 hoe het op een quantum computer kan worden nagebootst.

In het laatste hoofdstuk presenteren we diverse resultaten die behaald zijn bij het in kaart brengen van fundamentele eigenschappen van quantum mechanische systemen. Een van de cruciale eigenschappen van quantum systemen, waarin ze zich onderscheiden van klassieke systemen, is hun mogelijkheid tot 'entanglement' oftewel *verstrengeling* [4]. Verstrengeling kan aanleiding geven tot een vorm van niet-lokaliteit. Stel wij hebben twee lichtdeeltjes, fotonen. Een foton kenmerkt zich onder meer door zijn polarisatie. Denkt U bijvoorbeeld aan een polarisatie filter in een zonnebril dat er voor zorgt dat veel van het verstrooide zonlicht niet wordt doorgelaten. Het is mogelijk om een paar fotonen te maken welker polarisatie toestand verstrengeld is. Alhoewel de twee fotonen zich op grote afstand van elkaar kunnen bevinden, bestaat er een extra sterke correlatie tussen hen die niet altijd te beschrijven is op

[4] In het oorspronkelijke Duitse artikel spreekt Erwin Schrödinger van 'verschränken' dat door het Van Dale woordenboek Duits-Nederlands vertaald wordt als kruislings over elkaar leggen, kruislings verbinden. Een Nederlandse vertaling van 'verschränken' als verstrengelen correspondeert echter beter met de fysische betekenis van de term en blijft dichter bij de Engelse vertaling, 'entangle'.

een klassieke lokale wijze. Met lokaal bedoelen we hier dat de variabelen in een klassieke wijze van beschrijven zodanig worden gekozen dat er geen schending van causaliteit plaats vindt; informatie kan zich niet sneller dan met de snelheid van het licht, ongeveer 300.000 km/s in vacuüm, verplaatsen. De Bell ongelijkheden die in Hoofdstuk 5 worden besproken, vormen een uitdrukking van de niet-lokaliteit van sommige verstrengelde quantum toestanden. In Hoofdstuk 5 wordt een nieuwe relatie gelegd tussen een criterium dat bepaalt of een toestand verstrengeld is en een Bell ongelijkheid.

Verstrengeling is een belangrijk onderwerp in het bestuderen van het gebruik van quantum informatie. Laten we aannemen dat de verstrengelde fotonen in bezit zijn van twee personen, die doorgaans Alice en Bob worden genoemd; Alice krijgt foton 1, dat wil zeggen, ze kan metingen en andere operaties op foton 1 verrichten en Bob krijgt foton 2. Men kan laten zien dat het bezit van Alice en Bob van paren van verstrengelde fotonen hen in staat stelt elkaar informatie, klassieke of quantum, toe te sturen op een efficientere wijze dan mogelijk zou zijn geweest als zij niet een dergelijk paar hadden bezeten. Als zodanig vormen verstrengelde quantum mechanische toestanden een nuttige bron die men graag wil kwantificeren. Niet alle verstrengelde quantum mechanische toestanden zijn echter gelijkwaardig geschapen. Het is daarom van groot belang hiërarchieën te ontwikkelen van de verschillende *soorten* verstrengelde toestanden waarvan de zwakste soorten niet converteerbaar zijn tot sterkere verstrengelde toestanden via lokale operaties van Alice en Bob en klassieke communicatie tussen Alice en Bob.

In Hoofdstuk 5 wordt een speciale klasse van verstrengelde toestanden geconstrueerd die niet converteerbaar zijn tot de meest krachtige bron van verstrengelde toestanden. Deze toestanden worden ook wel gebonden verstrengelde toestanden genoemd; hun verstrengeling zit als het ware opgesloten en kan niet voor goed gebruik worden vrij gemaakt. De gebonden verstrengelde toestanden kunnen worden gezien als het eindprodukt van een lokaal proces op de meeste krachtige soort van verstrengelde toestanden. Hun gebondenheid impliceert echter dat dit proces niet lokaal ongedaan kan worden gemaakt. Bij de constructie van deze toestanden komt niet-lokale onomkeerbaarheid ook op een andere wijze aan de orde. Stel er zijn twee personen, Alice en Bob, die met behulp van lokale operaties één uit een set van quantum toestanden bouwen. Na de constructie vergeten zij tijdelijk welke uit de set van toestanden zij hadden geprepareerd; deze informatie is nu verloren gegaan. Tenslotte trachten zij met behulp van lokale operaties en klassieke communicatie, bijvoorbeeld via de telefoon, erachter te komen welke toestand zij hadden gemaakt. Wij bewijzen dat de twee personen niet in staat zijn dit lokaal uit te voeren. Echter als aan Alice en Bob de volledige vrijheid was gegeven en zij *samen* quantum operaties zouden kunnen uitvoeren, dan kan men laten zien dat zij deze taak wel kunnen volbrengen. In deze zin kan men spreken van een lokaal proces dat lokaal *niet* omkeerbaar, maar globaal wel omkeerbaar is.

Dankwoord

Welke van de vele plannen, die men wel eens maakt over de toekomst, uiteindelijk worden gerealiseerd is een samenspel van willekeur, wilskracht, talent en hulp en stimulans van anderen: de collega's, begeleiders, vrienden en familie. Op deze plaats zou ik willen aangeven hoe deze laatsten mij geholpen hebben gedurende mijn promotie en hen hiervoor willen bedanken.

Beginnen moet ik dan met Remko Scha, die mij rond mijn doctoraal examen, erop attendeerde dat Paul Vitányi erg enthousiast was over een nieuw onderzoeksgebied, quantum computation. Nadat ik contact had gelegd met Paul, was het de vraag hoe wij een promotie aanstelling voor mij konden creeëren aan de Universiteit van Amsterdam. Ik had ook inmiddels contact gelegd met Bernard Nienhuis, die zich nu bereid verklaarde in zee te gaan met dit nieuwe project. Ik wil Bernard ervoor bedanken dat hij vertrouwen heeft geschonken aan deze nieuwe onderzoeksrichting. Ook wil ik Bernard bedanken voor de gesprekken en de begeleiding die plaatsvonden vooral in het eerste anderhalf jaar, waarin hij mij probeerde op een goed spoor te zetten. Paul wil ik graag bedanken voor de vrijheid die hij mij gelaten heeft in mijn onderzoek en de keuze om een groot deel van mijn promotie tijd in de Verenigde Staten door te brengen.

Mijn kamergenoten, collega's en mede-AIOs Harry Buhrman, Peter Grünwald, Alec Maassen van den Brink, Nathalie Muller, John Tromp en Ronald de Wolf bedank ik graag voor discussie en soms afleiding en plezier.

I would like to thank John Smolin for the many things that he has done for me when I first came to the United States (for example, teaching me how to drive). I would like to thank him as well for the enjoyable collaborations that resulted in papers that we wrote together. From Charlie Bennett I learned that research is play and play is a lot of fun. I would like to thank him for making it possible to stay at IBM for an extended period of time and for creating a very stimulating and creative atmosphere in the IBM group. I would like to thank my other co-authors Howard Barnum, Ike Chuang, Markus Grassl, Tal Mor, Peter Shor and Ashish Thapliyal for joint research.

It has been a great pleasure to work with David DiVincenzo; he is a wonderful collaborator. More importantly he has become my partner during my stay at IBM. I am grateful for the way he has enriched my life and the love that we share.