



universität
wien

DISSERTATION

Titel der Dissertation

Quantum Markov processes and
applications in many-body systems

Verfasser

Paul Kristan Temme

angestrebter akademischer Grad

Doktor der Naturwissenschaften (Dr. rer. nat.)

Studienkennzahl laut Studienblatt: A 091 411

Dissertationsgebiet laut Studienblatt: Physik

Betreuer: Univ.-Prof. Dr. Frank Verstraete

Wien, im Dezember 2010

Für Pilar, meine Frau

“Frauen sind die Holzwolle der Glaskiste des Lebens.”

– Kurt Tucholsky

Contents

Zusammenfassung	7
Abstract	9
Acknowledgments	11
Introduction	13
1 Preliminaries for classical and quantum Markov processes	19
1.1 Formal setting and notation	21
1.2 Classical Markov chains	23
1.2.1 Spectral properties and ergodicity	24
1.2.2 Example: The Metropolis algorithm	26
1.3 The master equation	28
1.4 Quantum Markov chains	31
1.4.1 Completely positive maps	32
1.4.2 Perron Frobenius and irreducibility	36
1.5 The Lindblad equation	40
1.6 Matrix product states	45
2 Mixing time analysis of quantum Markov chains	51
2.1 Mixing time in Markov chains	53
2.2 The quantum χ^2 -divergence	54
2.2.1 Monotone Riemannian metrics and generalized relative entropies.	56
2.2.2 Properties of the quantum χ^2 -divergence	59
2.3 Mixing time bounds and contraction of the χ^2 -divergence under tcp-maps	60
2.3.1 Mixing time bounds	60
2.3.2 Contraction coefficients	63
2.4 Quantum detailed balance	66
2.5 Quantum Cheeger's inequality	69
2.5.1 Example: Conductance bound for unital qubit channels	72

3	Quantum Metropolis sampling	73
3.1	Summary of the algorithm	75
3.2	Implementation	80
3.3	Description of the quantum Metropolis algorithm	83
3.4	Running time of the quantum Metropolis algorithm	87
3.4.1	The rejection procedure	88
3.4.2	Mixing time scaling for an example system	91
3.5	The completely positive map	93
3.6	Fixed point of the ideal chain	95
3.7	The influence of imperfections to the fixed point	99
3.7.1	Error bounds and realistic phase estimation	100
3.8	An experimental implementation	106
4	Stochastic matrix product states	109
4.1	Quantum states and correlation measures	111
4.2	Entropy cost and stochastic matrix products states	113
4.3	Normal form and bounds on the entropy cost	114
4.4	Application to the asymmetric exclusion process	118
5	Stochastic exclusion processes versus coherent transport	125
5.1	Exclusion processes	127
5.2	Formulation of the quantum master equation	128
5.3	The symmetric exclusion process	130
5.4	The asymmetric exclusion process	133
5.4.1	The Derrida algebra	133
5.4.2	Dynamical MPS approach to open quantum systems	136
	Conclusions and Outlook	141
	Bibliography	146
	Lebenslauf	157

Zusammenfassung

Gegenstand der vorliegenden Arbeit ist die Untersuchung von quantenmechanischen und klassischen Markov-Prozessen und deren Anwendung im Bereich der stark korrelierten Vielteilchensysteme. Unter einem Markov-Prozess versteht man eine spezielle Art eines stochastischen Prozesses, dessen weitere dynamische Entwicklung unabhängig ist von der Vorgeschichte seiner Entwicklung und nur von der derzeitigen Konfiguration abhängt. Die Anwendung von Markov-Prozessen im Bereich der statistischen Mechanik von klassischen Vielteilchensystemen hat eine lange Geschichte. Markov-Prozesse dienen nicht nur der Beschreibung der Dynamik von stochastischen Systemen, sondern liefern vielmehr auch eine sehr praktische Methode, mit deren Hilfe grundlegende Eigenschaften komplexer Vielteilchenprobleme in Form eines probabilistischen Algorithmus berechnet werden können. Ziel dieser Arbeit ist es das Verhalten von quantenmechanischen Markov Prozessen, dies sind Markov-Prozesse, welchen ein quantenmechanischer Konfigurationsraum zu Grunde liegt, zu untersuchen und mit deren Hilfe komplexe Vielteilchensysteme besser zu verstehen. Darüber hinaus formulieren wir einen Quantenalgorithmus, mit dessen Hilfe es möglich ist, die thermischen- und Grundzustandseigenschaften von quantenmechanischen Vielteilchensystemen zu berechnen. Nachdem wir eine kurze Einführung in das Feld der quantenmechanischen Markov-Prozesse gegeben haben, untersuchen wir deren Konvergenzeigenschaften. Wir finden Schranken für die Konvergenzraten der quantenmechanischen Prozesse, basierend auf einer Verallgemeinerung von geometrischen Schranken, welche für klassische Prozesse gefunden wurden. Wir verallgemeinern ein Abstandsmaß, die χ^2 -Divergenz für nicht kommutative Wahrscheinlichkeitsräume, welches unseren Untersuchungen zu Grunde liegt. Diese Divergenz ermöglicht auch eine Verallgemeinerung der detaillierten Balance für quantenmechanische Prozesse. Danach konstruieren wir den Quantenalgorithmus, der als natürliche Verallgemeinerung des Metropolisalgorithmus für quantenmechanische Hamiltonoperatoren verstanden werden kann. Wir beabsichtigen damit zu zeigen, dass ein Quantencomputer in der Lage ist, als universeller Quantensimulator zu fungieren, welcher nicht nur die Dynamik eines Quantensystems beschreiben kann, sondern auch den Zugang zu statischen Berechnungen ermöglicht. Danach untersuchen wir die Korrelationseigenschaften von klassischen Nichtgleichgewichtszuständen mit Methoden der Quanteninformationstheorie. Wir konstruieren eine Klasse von Matrix-Produkt-Zuständen, deren Korrelationen anhand von klassischen Markov-Prozessen verstanden werden können. Schließlich untersuchen wir die Transporteigenschaften eines stationären Nichtgleichgewichtszustandes. Die dynamische Gleichung ist so konstruiert, dass der Transport je nach Parameterwahl entweder hauptsächlich stochastisch oder hauptsächlich kohärent stattfindet. Wir können somit die unterschiedlichen Formen des Transports innerhalb eines Modells miteinander vergleichen.

Abstract

This thesis is concerned with the investigation of quantum as well as classical Markov processes and their application in the field of strongly correlated many-body systems. A Markov process is a special kind of stochastic process, which is determined by an evolution that is independent of its history and only depends on the current state of the system. The application of Markov processes has a long history in the field of statistical mechanics and classical many-body theory. Not only are Markov processes used to describe the dynamics of stochastic systems, but they predominantly also serve as a practical method that allows for the computation of fundamental properties of complex many-body systems by means of probabilistic algorithms. The aim of this thesis is to investigate the properties of quantum Markov processes, i.e. Markov processes taking place in a quantum mechanical state space, and to gain a better insight into complex many-body systems by means thereof. Moreover, we formulate a novel quantum algorithm which allows for the computation of the thermal and ground states of quantum many-body systems. After a brief introduction to quantum Markov processes we turn to an investigation of their convergence properties. We find bounds on the convergence rate of the quantum process by generalizing geometric bounds found for classical processes. We generalize a distance measure that serves as the basis for our investigations, the χ^2 -divergence, to non-commuting probability spaces. This divergence allows for a convenient generalization of the detailed balance condition to quantum processes. We then devise the quantum algorithm that can be seen as the natural generalization of the ubiquitous Metropolis algorithm to simulate quantum many-body Hamiltonians. By this we intend to provide further evidence, that a quantum computer can serve as a fully-fledged quantum simulator, which is not only capable of describing the dynamical evolution of quantum systems, but also gives access to the computation of their static properties. After this, we turn to an investigation of classical non-equilibrium steady states with methods derived from quantum information theory. We construct a special class of matrix product states that exhibit correlations which can best be understood in terms of classical Markov processes. Finally, we investigate the transport properties of non-equilibrium steady states. The dynamical equations are constructed in such a manner that they allow for both stochastic as well as coherent transport in the same formal framework. It is therefore possible to compare different forms of transport within the same model.

Acknowledgments

Foremost, I would like to take this opportunity to thank my supervisor Frank Verstraete. After graduating from university, I found it quite challenging to find and choose the “right” place to pursue my graduate studies. I am very glad to say that I have found it here. I am sincerely grateful for his excellent supervision during the course of this thesis, for his uncompromising open-door policy and his dedication and enthusiasm he has put into every project I was fortunate to work on with him. He not only gave me good direction, but also created an environment to pursue ideas freely.

I gratefully acknowledge the generous support by the doctoral program Complex Quantum Systems (W1210) of the FWF and the ERC grant QUERG. Moreover, I would like to thank Markus Arndt for his support and the effort he has put into bringing the doctoral program to life, which I have had the honor and pleasure of profiting from so greatly.

Furthermore, I am very grateful for the many discussions and collaborations that have emerged with so many interesting people during my time as a graduate student. I would like to thank everybody I have had a chance to collaborate with and the honor to learn from in the past three years. I would especially like to thank Michael Wolf, not only for the intriguing projects on which I have had the fortune to collaborating with him, but more importantly for the many discussions I have learned from so very much; Michael Kastoryano for the fruitful collaboration and the delightful time we spent discussing also non-physics related issues during my visit in Copenhagen. Moreover, I would like to thank my coauthors Mary Beth Ruskai, Tobias Osborne, Karl Gerd Vollbrecht and David Poulin, it has been a pleasure working with them. Furthermore, I would like to thank Eddie Farhi for his generous hospitality and the chance to learn from him during the very delightful visit to his group.

It is a pleasure to thank the people here at the University of Vienna and in the doctoral program, who have made the time spent during the course of this thesis so very pleasant and memorable. I would like to thank Bogdan Pirvu and Johannes Wilms for being such enjoyable colleagues and for the countless heated discussions we had. Furthermore, I would like to thank Borivoje Dakic, Sven Ramelow, Stefan Nimmrichter and Enrique Rico for the nice time on as well as off campus.

And last but not least I am thankful to my family for the enduring support and encouragement, their calm and their wisdom. Looking back at these past three years as a graduate student I feel great joy, because I was fortunate enough to share them with my wife.

Introduction

The success story of the exact sciences is based to a large extent on the fact that its methodology allows us to break down many complex mechanisms to a set of simple rules, which are more accessible to us. The question that arises, however, is whether boundaries are set to this endeavor or whether it will always be possible to express an innately complex issue in a more simple manner. This undertaking appears to be quite paradox to some extent. On one hand, we seek a description as simple as possible, on another, we want a solution which is capable of reproducing the richness nature provides. Thus if we do not look for a solution that restricts itself to explaining only individual aspects of the problem, we have to find a description that on its own is complex enough to encompass the richness, which naturally occurs in the system. It seems that the complexity is somehow conserved and at a certain point we will have to pay the price for the large amount of descriptive power. Let us try to be more concrete. Since the early days of quantum mechanics, it has been clear that there is a fundamental difficulty in studying quantum many-body systems. The size of the configuration space, i.e. Hilbert space, of a collection of particles grows exponentially with the number of particles. Even though we might know the equations which govern the evolution of a particular many-body system, the sheer size of the configuration space renders an attempt of exactly solving these equations in all generality futile. This observation is best expressed in a quote by Dirac [1]:

“The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble. It therefore becomes desirable that approximate practical methods of applying quantum mechanics should be developed, which can lead to an explanation of the main features of complex atomic systems without too much computation.”

Many of the important breakthroughs in quantum physics during the 20th century have resulted from efforts to address this problem, leading to fundamental theoretical and numerical methods to approximate solutions of the many-body Schrödinger equation. One of the most prominent approaches to finding approximate solutions is perturbation theory [2]. The assumption that enters into this approximation is that the system does not deviate too strongly from a solution which is already known. Even though this approach has given incredibly remarkable results, it is nevertheless limited to providing solutions which are similar in their features to

those of the known system. Another example where this becomes evident is the application of variational states [2]. Here, one constructs a family of states and one seeks to find the best approximation within this family. Hence, one already made an assumption about the form of the solution one seeks to find. A significant amount of the research effort has been devoted to constructing ample families of variational states as universal as possible and capable of capturing the most significant aspects of the solution. The field of quantum information theory has made fecund contributions to this enterprise by studying the correlation properties of quantum states as such [3, 4, 5, 6].

However, exceedingly complex quantum many-body systems exist, for which thus far no definite solutions have been found by means of classical approaches. Notable examples include high- T_c superconductors, electronic structure in large molecules, and quark confinement in quantum chromodynamics. We are thus confronted with constructing increasingly complex simulation methods for the increasingly complex physical systems we seek to understand.

Does this problem only pertain to complex quantum systems, or does it also occur in classical systems? The problem of the configuration-space explosion is not unique to quantum mechanics. The task of simulating interacting classical particles is challenging for the same reason. It was only with the advent of computers in the 1950's, that a systematic approach of simulating classical many-body systems was made possible. In their seminal paper [7] Metropolis *et al.* devised a general method to calculate the properties of any substance comprising individual molecules with classical statistics. This landmark paper is a cornerstone in the simulation of interacting systems and has had a huge influence on a wide variety of fields (see e.g. [8, 9, 10]). The Metropolis method can also be used to simulate certain quantum systems by a "quantum-to-classical map" [11]. Unfortunately, this quantum Monte Carlo method is only scalable when the mapping conserves the positivity of the statistical weights, and fails in the case of fermionic systems due to the infamous sign-problem.

The Metropolis algorithm is a prime example of a probabilistic algorithm based on the concept of Markov chains. This concept was introduced in 1906 by the Russian mathematician Alexey Markov, when he was investigating a special class of stochastic processes that have the defining property of being memoryless. A Markov process is a stochastic process which is independent of the history of the stochastic evolution and only depends on the current configuration. Such processes find a large variety of applications throughout different disciplines such as computer science, statistics, engineering and in particular physics. In fact, Markov processes are not only a useful tool to simulate physical systems, as is the case for the Metropolis algorithm, but also serve as a tool for describing the dynamics of stochastic processes that occur in nature. It is their striking simplicity which makes them so ubiquitous.

We are convinced that the construction of a quantum computer is indeed a viable option to tackle the problems that pertain to the simulation of complex quantum many-body systems, much in the same fashion as the advent of classical computers has lead to a systematic approach

for simulating complex classical systems. If we look back, we see that the original motivation for building a quantum computer came from Feynman to achieve this very task [12]. The seminal work of Lloyd [13] demonstrated that a quantum computer can reproduce the dynamical evolution of any quantum many-body system. With this work, we hope to contribute to the establishment of the idea of the quantum computer as a universal quantum simulator, by showing that a quantum computer can also give access to the computation of thermal and ground state properties of quantum many-body systems.

The central topic of this thesis is the investigation of quantum Markov processes and the application thereof to many-body systems. A quantum Markov process can be seen as the natural generalization of a memoryless stochastic process to non-commuting probability spaces, i.e. quantum state spaces, and is described by either a continuous-time Lindblad equation or by the subsequent application of completely positive trace preserving maps. These processes are often used to model open quantum systems or driven non-equilibrium systems. Recently, work was put forward that showed that the concept of quantum Markov processes is as powerful as universal quantum computation itself [14]. The objective of this thesis is twofold. On one hand we use Markov processes to better understand the correlations present in strongly correlated one-dimensional systems and we want to use them as a tool to investigate driven non-equilibrium systems. On the other hand we use them to propose a quantum algorithm to simulate the static properties of quantum many-body systems. A large part of our investigation of quantum Markov processes relates to their convergence properties. We assume, that the reader is familiar with the basic concepts of quantum information theory and in particular with the circuit formulation of quantum computation [15]. We do not assume a prior knowledge of quantum Markov processes, as they will be introduced in the following. The thesis is organized as follows:

Outline and Summary of the results

Chapter 1: This chapter is devoted to an introduction to classical as well as quantum Markov processes. Several important facts about Markov processes are reviewed and we provide some of the fundamental theorems we will frequently be making use of. We start by introducing the reader to the formal setting of this thesis. We then proceed to introduce classical time-discrete Markov processes, which are also known as Markov chains. We only consider time-homogenous processes throughout this thesis. Several fundamental theorems that pertain to classical chains, such as the Perron-Frobenius theorem, are given. Then the Metropolis algorithm, which we seek to quantize later, is introduced and we discuss the form of the stochastic transition matrix of the algorithm. We then proceed to discuss time-continuous Markov processes and show that the Markov property immediately gives rise to a semi-group structure. From this we derive the classical master equation for probability distributions. We then turn

to quantum Markov processes. The elementary building block of quantum Markov chains, the trace-preserving completely positive map, is introduced and several lemmata that provide representations of such maps are given. We provide generalizations of Perron-Frobenius results in the quantum setting. We then turn to time continuous quantum master equations and prove the most general form of the generator, i.e. the Lindblad map, for such processes. The starting point of the derivation will be that we assume that the corresponding trace preserving completely positive map has a continuous semi-group structure. This chapter is concluded with an introduction to matrix product states. These are states that support a limited amount of entanglement and approximate states of one-dimensional systems very well. We highlight their connection to quantum Markov chains and emphasize how the correlations these states exhibit can be understood in terms of the Markov chain framework.

Chapter 2: We give mixing time bounds for quantum Markov processes. The mixing time is the number of times a stochastic map has to be applied to an arbitrary initial state to be close to the fixed point distribution of the Markov chain. The “closeness” is determined here in terms of the trace-distance between the fixed point and the current state of the chain. We generalize results that pertain to classical Markov chains to the quantum setting. The derivation of the classical results relies on a distance measure called the χ^2 -divergence, which serves as an upper bound to the total variational distance. We start our analysis by defining a non-commuting generalization of the χ^2 -divergence on quantum state spaces. This quantum divergence is intimately related to so-called monotone Riemannian metrics. We show, that the quantum χ^2 -divergence also gives an upper bound to the trace-distance and derive a mixing time bound for the application of any primitive quantum stochastic map. This upper bound is determined by the singular values of a map that is similar to the actual trace preserving completely positive map. We then investigate the contractive behavior of this divergence under the application of quantum stochastic maps. Moreover, the χ^2 -divergence gives rise to a convenient way of defining a generalization of the classical detailed balance condition for completely positive maps. This condition will prove very useful in the analysis of the fixed point structure of our quantum Markov chain based algorithm. We conclude this section by deriving a quantum generalization of the classical conductance bound for unital maps. The conductance bound yields a way of bounding the singular values of any unital map by a geometric constant, which can be seen as a generalization of the well known Cheeger’s constant.

Chapter 3: In this chapter we construct a quantum algorithm, which allows one to prepare the Gibbs state of a quantum many-body Hamiltonian. The algorithm can be seen as a quantization of the classical Metropolis algorithm. The fixed point of the corresponding quantum Markov chain is ensured to be the Gibbs state of the Hamiltonian. We begin this chapter with a brief summary of the algorithm by making some simplifying assumptions to facilitate the presentation. Then we turn to a presentation of the basic building block of the quantum algorithm,

the quantum phase estimation procedure for the many-body Hamiltonian we want to simulate. We assume that the reader is familiar with the formalism of circuit-based quantum computation [15]. Thereafter, we proceed to a more elaborate description of the algorithm that explains the individual steps in greater detail. In the following section we turn to a discussion on the general runtime of the algorithm. We discuss the runtime of the algorithm based on mixing time arguments derived in the previous chapter. The mixing time of the algorithm crucially depends on the problem Hamiltonian. It is impossible to prove a polynomial runtime of the algorithm for an arbitrary Hamiltonian in all generality. In fact, it is even expected that some classes of Hamiltonians, which have been proven to be QMA-complete, will give rise to a mixing time that is exponential in the system size. However, we provide a simple spin Hamiltonian for which the mixing time is estimated to scale linearly in the number of spins. We then proceed by explicitly constructing the completely positive map of the chain and show that for a specific set of Metropolis updates the unique fixed point of the map is the Gibbs state of the many-body Hamiltonian. To show this, we will make use of the quantum detailed balance condition and the chain's ergodicity properties. We then give an error bound on the deviation from the idealized fixed point of the quantum Markov chain, which is due to the fact that we discuss the implementation of the algorithm on a quantum computer with finite resources. We conclude the chapter with presenting an experimental implementation of this algorithm for a two qubit example system that is accessible with today's technology.

Chapter 4: There are striking similarities between classical non-equilibrium steady states, i.e. fixed points of a classical multi-particle master equation, and the ground states of quantum many-body Hamiltonians. It turns out that matrix product states play an important role in both fields. We devise a program to characterize the classical correlations present in non-equilibrium steady states and introduce a special class of matrix product states, which we call stochastic matrix product states. These are geared towards approximating multi-partite classical probability distributions that support only a limited amount of correlations. We introduce a new correlation measure we call the entropy cost and show that it is an upper bound to the mutual information. Furthermore, we show that a multi-partite probability distribution can be well approximated by a stochastic matrix product state, if the entropy cost is low. A normal form of these states is derived that establishes a connection to classical Markov chains, which can be seen as the classical analog of the connection quantum matrix product states have to quantum Markov chains. This decomposition also gives rise to a set of so-called source probabilities, which can be seen as the stochastic analog of Schmidt coefficients. We apply these concepts to a driven stochastic system called the asymmetric exclusion process. We estimate the mutual information of the non-equilibrium steady state numerically for large chain sizes and show that the steady state obeys an area law, which is only corrected by logarithmic contributions when the system is critical.

Chapter 5: We formulate a quantum master equation that can be seen to generate the classical non-equilibrium dynamics of one-dimensional exclusion processes, when no Hamiltonian evolution is present. The classical dynamics of the stochastic exclusion processes is mediated by the Lindblad operators in the quantum master equation, which usually account for dissipation in quantum systems. Since we incorporated the classical dynamics in this more general framework, we can contrast its behavior to that of the quantum transport mediated by a simple Hamiltonian. We investigate two types of exclusion process known as the fully symmetric exclusion process and the totally asymmetric exclusion process. We can compute the particle density and the current density in the steady state for the symmetric exclusion process with coherent evolution exactly. We observe, that the stochastic transport properties are not modified significantly in the presence of coherent transport and the transport remains diffusive. Only when the stochastic contribution to the transport vanishes the transport properties change and we have ballistic transport. We then turn to the totally asymmetric exclusion process. The steady state of this process cannot be computed exactly in the presence of coherent evolution. We therefore turn to a numerical simulation of the steady state properties. For the asymmetric exclusion process we find evidence for the change in transport behavior only when the state is already correlated classically. Otherwise the steady state solution remains unchanged in the presence of coherent evolution.

Chapter 1

Preliminaries for classical and quantum Markov processes

Synopsis:

We will use this chapter to introduce the reader to the basic concepts in classical and quantum Markov chains. These concepts will repeatedly be made use of during the course of this thesis. The chapter is kept rather formal because it states some of the central definitions and theorems for Markov chains we will make use of throughout this thesis. First we will fix the formal framework of this thesis and then turn to the discussion of classical Markov chains and their transition matrices. We will discuss their spectral and convergence properties and introduce the ubiquitous Metropolis algorithm. We then turn to the discussion of time - continuous Markov processes that can be described by a classical rate equation which is commonly known as the master equation. The reason for introducing classical processes is twofold. We will be making use of classical processes a number of times in this thesis, for instance when we introduce the concept of stochastic matrix product states, but it is also useful to introduce the concept of Markov processes for probability spaces before we turn to generalizing them to non-commuting, i.e. quantum, probability spaces. We then discuss the concepts of completely positive maps on non-commuting probability spaces and their spectral properties. These completely positive maps can be seen as the natural generalization of Markov processes to quantum state spaces. It is their spectral and structural similarity to the classical transition matrices that allows for this connection. With the general framework of quantum stochastic maps at hand we will discuss the structure of the dynamical semi group that models continuous time quantum Markov processes and derive the Lindblad master equation for quantum states. Finally, we will close this chapter with a discussion on matrix product states and highlight their connection to completely positive maps.

1.1 Formal setting and notation

In this section we briefly introduce the notation and some elementary facts we will frequently be making use of in the course of this thesis. A good general introduction to the field of quantum information theory can be found for instance in [15]. Throughout this thesis we will only consider finite dimensional classical stochastic or quantum stochastic systems. A good reference for the formal preliminaries and the Matrix analysis tools we need can be found in [16, 17].

The classical stochastic systems will be described by a finite state space Ω that contains the elementary events, which we use to define a probability space. We denote random variables that take values in Ω by capital letters such as X, Y, \dots . The probability of an event H is denoted by $\mathbb{P}(H)$. The probability space will be taken to be $L^1(\Omega)$, i.e all states which are normalized with respect to the one - norm. We will describe the probability distributions either by $p \in L^1(\Omega)$ or by vectors $|p\rangle \in L^1(\Omega)$, with $|p\rangle = \sum_{i \in \Omega} p_i |i\rangle$ in the standard Dirac notation, where the $p_i \geq 0$ are required to be non-negative, and $\sum_i p_i = 1$. The space L^1 is normed with respect to the one-norm that can be expressed for an arbitrary finite dimensional vector $|a\rangle$ as $\| |a\rangle \|_1 = \sum_i |a_i|$.

The quantum mechanical systems will be described by a finite dimensional Hilbert space \mathcal{H} over the complex numbers. Hence, the Hilbert spaces we will be dealing with are typically just $\mathcal{H} = \mathbb{C}^d$. As is common in the literature we will work in units where $\hbar = 1$. The algebra of the observables are all bounded operators on this Hilbert space and are commonly denoted by $\mathcal{B}(\mathcal{H})$ in the literature. Since the Hilbert space we consider is typically just \mathbb{C}^d our observables are described by the matrix algebra \mathcal{M}_d of $d \times d$ - dimensional complex matrices over the Hilbert space \mathcal{H} , for which we also often write $\mathcal{M}(\mathcal{H})$. This space is equipped with a norm that can be seen as the nature quantum generalization of the one-norm, the so called trace - norm $\|A\|_{\text{tr}} = \text{tr} \left[\sqrt{A^\dagger A} \right] = \sum_i \sigma_i(A)$, where the $\sigma_i(A)$ denote the singular values of the matrix $A \in \mathcal{M}_d$. Note, that \mathcal{M}_d itself turns into a Hilbert space when equipped with the Hilbert - Schmidt scalar product $\langle A|B \rangle_{HS} = \text{tr} [A^\dagger B]$, for $A, B \in \mathcal{M}_d$. It is therefore possible to choose a complete orthonormal Basis of \mathcal{M}_d , which we will denote by $\{F_i\}_{i=1 \dots d^2}$, such as for instance the canonical product basis $\{|i\rangle \langle j|\}_{i,j=1, \dots, d}$. Choosing a fixed basis immediately gives rise to an isomorphism $\mathcal{M}_d \simeq \mathbb{C}^{d^2}$. The states are density matrices $\rho \in \mathcal{S}_d$, where $\mathcal{S}_d = \{\rho \in \mathcal{M}_d | \rho = \rho^\dagger, \rho \geq 0, \text{tr}[\rho] = 1\}$, acting on $\mathcal{H} = \mathbb{C}^d$. The set of pure states given by projectors $|\psi\rangle \langle \psi|$ on \mathbb{C}^d is denoted by \mathcal{S}_d^1 , while the set of positive definite states, i.e. states $\rho \in \mathcal{S}_d$ of full rank, is denoted by \mathcal{S}_d^+ .

Since we are dealing with many - body systems, we will need to consider state spaces of composite systems. The composition of two quantum systems is described by the tensor product of the individual spaces $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. Multipartite or many-body systems are therefore described by spaces that are of the form $\mathcal{H} = \mathcal{H}_1 \otimes \dots \mathcal{H}_N$, i.e. have an endowed tensor product structure. It is this tensor product structure, which gives rise to a particularly striking

phenomenon of quantum mechanics called entanglement. At this point, we would like to introduce a set of formal properties of entangled states, which will be very useful throughout this thesis.

Formal tricks with entanglement If a bipartite state $|\psi\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$ cannot be written as a product state $|\psi\rangle \neq |v\rangle \otimes |w\rangle$, it is defined as being entangled. Any bipartite state that is written with respect to some arbitrary local bases of \mathcal{H}_1 and \mathcal{H}_2 as $|\psi\rangle = \sum_{ab} M_{ab} |a\rangle \otimes |b\rangle$ can be brought into the so called Schmidt form

$$|\psi\rangle = \sum_{i=1}^d \sigma_i |i_1\rangle \otimes |i_2\rangle, \quad (1.1)$$

for some orthogonal bases $\{|i_1\rangle\}$ and $\{|i_2\rangle\}$, known as Schmidt basis, with coefficients $\sigma_i > 0$ for which $\sum_{i=1}^d \sigma_i^2 = 1$. These coefficients are known as Schmidt coefficients. Here, the sum is taken up to $d = \min(\dim(\mathcal{H}_1), \dim(\mathcal{H}_2))$, the minimal dimension of either of the two Hilbert spaces. This decomposition can be constructed by making use of the ubiquitous singular value decomposition [16, 17], which states that any matrix M can be decomposed as $M = U\Sigma V^\dagger$, where $\Sigma > 0$ is diagonal and U, V are isometries. A state $|\Omega\rangle$ is called maximally entangled, if all its Schmidt coefficients are equal $\sigma_i = 1/\sqrt{d}$. We write in the Schmidt basis

$$|\Omega\rangle = \frac{1}{\sqrt{d}} \sum_{k=1}^d |kk\rangle. \quad (1.2)$$

This definition of a maximally entangled state is robust with respect to transformations of the form $|\psi\rangle = (U_1 \otimes U_2) |\Omega\rangle$, since local unitaries only change the local basis in each Hilbert space, but leave the Schmidt coefficients $\sigma_i = 1/\sqrt{d}$ invariant. We therefore refer to (1.2) as the maximally entangled state, from which any other state that is also maximally entangled can be obtained by a local unitary transformation. In fact, it is possible to construct an entire basis of maximally entangled states by choosing an orthogonal unitary basis $\{U_i\}_{i=1\dots d^2}$ of \mathcal{M}_d and by writing $|\psi_i\rangle = (\mathbb{1} \otimes U_i) |\Omega\rangle$. This basis is orthogonal, because we have for any two matrices $A, B \in \mathcal{M}_d$ the general correspondence

$$\langle \Omega | (A \otimes B) | \Omega \rangle = \frac{1}{d} \text{tr} [AB^T] \quad \text{and} \quad (\mathbb{1} \otimes A^T) | \Omega \rangle = (A \otimes \mathbb{1}) | \Omega \rangle, \quad (1.3)$$

where the transpose is to be taken with respect to the Schmidt basis of $|\Omega\rangle$. Furthermore, we observe that any state $|\Psi\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$ can be written as

$$|\Psi\rangle = (\mathbb{1} \otimes R) |\Omega\rangle, \quad (1.4)$$

for some general matrix $R \in \mathcal{M}_d$. This matrix can be related to $|\psi\rangle = \sum_{ij} c_{ij} |i\rangle \otimes |j\rangle$ via $R = \sum_{ij} \sqrt{d} c_{ij} |j\rangle \langle i|$ when $|\psi\rangle$ is expressed in terms of the c_{ij} with respect to the Schmidt basis of $|\Omega\rangle$.

1.2 Classical Markov chains

Before we turn to the discussion of quantum Markov chains, let us first revise classical Markov chains and their transition matrices. A more exhaustive and excellent discussion of Markov chains and non-negative matrices can be found in [18, 19, 17]. In this section we will only present the basics of Markov chains, so we can better understand the generalization to non-commuting probability spaces, i.e. quantum Markov chains. We will furthermore need to make use of classical transition matrices in chapter 4, where we discuss a certain class of multipartite probability distributions that can best be understood in terms of stochastic transition matrices. We only consider finite Markov chains. A finite Markov chain can be seen as stochastic process that traverses the elements of a finite set Ω , known as the state space, in the following manner: When we are at some point $i \in \Omega$, the next position $j \in \Omega$ is chosen at random with respect to a fixed transition probability P_{ji} . To be more precise:

Definition 1. *A sequence of random variable (X_0, X_1, \dots) is called a **Markov chain** with finite state space Ω and transition matrix P , if for all $i, j \in \Omega$ and all $n \geq 1$, and all events $H_{n-1} = \bigcap_{k=0}^{n-1} \{X_k = i_k\}$ (sequence of random variables) with non-vanishing $\mathbb{P}(H_{n-1} \cap \{X_n = j\}) > 0$, we have that the conditional probability is given by*

$$\mathbb{P}(X_{n+1} = j | H_{n-1} \cap \{X_n = i\}) = \mathbb{P}(X_{n+1} = j | X_n = i) = P_{ji}. \quad (1.5)$$

Furthermore, the matrix P is called the **stochastic matrix**, or transition matrix of the Markov chain. All elements of $[P_{ji}]_{j,i=1 \dots |\Omega|}$ are real and non-negative. Furthermore, the matrix obeys

$$\sum_{j \in \Omega} P_{ji} = 1 \quad \text{for all } i \in \Omega. \quad (1.6)$$

The condition 1.5 is often called Markov property, which essentially means, that the next state along the chain of random variables is independent of the history of the chain and only depends on the current position in the state space Ω . This is the central property of all Markovian evolutions and the reason why it suffices to describe every transition on the state space by the stochastic matrix P . Note, that we are considering only time homogeneous Markov chains here, i.e. we assume that at each step the same transition rule applies. In principle one could consider an alternative scenario, where the transition matrix P changes for each step, so P itself depends on n , but we will not consider this case here.

We can view the stochastic matrix P as a dynamical evolution law for the probability distribution

$$\mu_i(n) = \mathbb{P}(X_n = i) \geq 0 \quad \text{with} \quad \sum_{i \in \Omega} \mu_i(n) = 1. \quad (1.7)$$

We will often write the probability μ_i as a vector in Dirac notation as $|\mu\rangle = \sum_{i \in \Omega} \mu_i |i\rangle$, where we use for $\{|i\rangle\}$ the canonical basis indexed by Ω . The new probability $|\mu(n)\rangle$ is

obtained from $|\mu(n-1)\rangle$ by application of the stochastic matrix P via

$$|\mu(n)\rangle = P|\mu(n-1)\rangle. \quad (1.8)$$

Note, that we use a slightly different notation here than is commonly used in the mathematical literature. We use the notation common in the physics literature, where the stochastic matrix acts from the left on the probability distribution, that is $\mu_j(n) = \sum_i P_{ji}\mu_i(n-1)$ as opposed to $\mu_i(n) = \sum_j \mu_j(n-1)P_{ji}$. We can now see that the condition (1.6) amounts to the conservation of probability, so that $\sum_i \mu_i(n) = 1$ if $|\mu(n-1)\rangle$ was normalized. Suppose we start in some initial configuration $|\mu(0)\rangle$, then all other probabilities along the Markov chain are determined only by the transition Matrix P so that

$$|\mu(n)\rangle = P^n |\mu(0)\rangle. \quad (1.9)$$

We can now ask what happens as $n \rightarrow \infty$. Does the state converge somehow? If so, what does the limiting distribution, mostly called fixed point distribution

$$|\sigma\rangle = \lim_{n \rightarrow \infty} P^n |\mu(0)\rangle, \quad (1.10)$$

look like? The traditional theory of Markov chains is precisely concerned with convergence statements of this type. The general questions are, for instance, whether or not such a fixed point exists and whether this fixed point is unique. How long does it take the Markov chain to reach such a fixed point? These questions are of importance in several fields, not only in physics, but also in other areas ranging from computer science to finance, from communication science to biology. The physical motivation for investigating the convergence of such Markov chains often stems from questions concerning the equilibration of statistical mechanical systems. Here one is not only interested in the equilibration to a thermal equilibrium, which is described for instance in the canonical ensemble by the Gibbs-distribution, but also to what is often referred to as non-equilibrium steady states [20, 21]. Such processes, if they describe the physical process of equilibration, are usually continuous in time and we will discuss the corresponding formal framework in the next section. Here, we will focus on some of the characteristic properties of stochastic transition matrices.

1.2.1 Spectral properties and ergodicity

Stochastic matrices are matrices that have only positive elements and whose columns sum to one. This information alone already ensures several nice spectral properties [17, 18]. Note, that the eigenvalues of a stochastic matrix P will usually be complex, since P is generally not self-adjointed. But we will see, that all the eigenvalues λ of P are confined to lie in the unit disk $|\lambda| \leq 1$ of the complex plain.

Lemma 2. *Let P be a stochastic matrix, as defined in (1) and $\lambda_i(P) \in \mathbb{C}$ the eigenvalues of P , then we have*

$$|\lambda_i(P)| \leq 1. \quad (1.11)$$

Furthermore, there exists at least one $\lambda_i(P) = 1$.

PROOF: Assume, that $|v^i\rangle$ is an eigenvector of P that belongs to the eigenvalue λ_i . We see that $\|P|v^i\rangle\|_1 = |\lambda_i| \| |v^i\rangle \|_1$ in the one norm. On the other hand, we have that $\|P|v^i\rangle\|_1 \leq \|P\|_1 \| |v^i\rangle \|_1$, i.e. the matrix norm $\|P\|_1 = \max_j \sum_i |P_{ij}|$ is natural with respect to the one-norm. Thus it holds that $|\lambda_i| \leq \|P\|_1$. Since $\|P\|_1 = \max_j (\sum_i P_{ij}) = 1$ due to (1.6) we are ensured that $|\lambda_i| \leq 1$ for all i . We furthermore know that there is an eigenvalue $\lambda = 1$, since there exists the left eigenvector $\langle I | = \sum_i \langle i |$ for which due to (1.6) $\langle I | P = \langle I |$. \square

This lemma ensures that there exists at least one stationary eigenvector of the stochastic matrix, namely the right eigenvector to the eigenvalue $\lambda(P) = 1$. We already know the left eigenvector corresponding to this eigenvalue. It is simply given by $\langle I | = \sum_i \langle i |$, i.e. the vector that has unity in each entry. Furthermore, the lemma ensures that in the limit of $n \rightarrow \infty$ applications $P^n |\mu^0\rangle$ only the eigenvectors that have eigenvalues of unit magnitude are relevant. However, it neither ensures that there is a unique fixed point independent of the starting configuration $|\mu^0\rangle$, nor does it ensure that the Markov chain converges at all. Consider for example the stochastic matrix

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (1.12)$$

This matrix clearly has the two eigenvalues $\lambda_0 = 1$ and $\lambda_1 = -1$. This Markov chain, however, never converges to a fixed point distribution, as it just swaps the probabilities $\mu_0 |0\rangle + \mu_1 |1\rangle \rightarrow \mu_1 |0\rangle + \mu_0 |1\rangle$. To ensure convergence and the uniqueness of the fixed point we have to revoke an additional constraint the stochastic matrices have to satisfy. There is a whole field in stochastic mathematics that deals with the classification of Markov chains and their ergodicity properties [18, 19, 22]. The central Theorem is the Perron-Frobenius Theorem that ensures the convergence properties for element-wise strictly positive matrices. The Perron-Frobenius Theorem pertains to a set of matrices which are called strictly positive matrices, i.e. matrices that have only non-vanishing positive entries. For this Theorem it is not required that the probability preservation condition (1.6) holds and the considered matrices are therefore slightly more general. Since the proof is rather lengthy and is not immediately relevant to this thesis, we will skip it here and just state the theorem. The proof can be found in [17].

Theorem 3 (Perron - Frobenius). *Let A denote a matrix that is **element-wise strictly positive**, i.e. $A_{ij} > 0$ for all $i, j \in \Omega$ and let $\rho(A) = \sup \{|\lambda| \mid \lambda \in \mathbb{C}, A|x\rangle = \lambda|x\rangle\}$ denote the spectral radius of A , then the following holds:*

1. *The spectral radius is strictly positive $\rho(A) > 0$.*
2. *Furthermore $\rho(A)$ is an eigenvalue of A .*
3. *There is a $|x\rangle \in \mathbb{R}^d$ with $x_i > 0$ and $A|x\rangle = \rho(A)|x\rangle$.*

4. $\rho(A)$ is an algebraically (and therefore also geometrically) simple eigenvalue of A
5. Every eigenvalue $\lambda \neq \rho(A)$ of A is strictly smaller than the spectral radius, i.e. $|\lambda| < \rho(A)$. Hence, $\rho(A)$ is the unique eigenvalue of this modulus
6. We have the unique convergence $\lim_{n \rightarrow \infty} [\frac{1}{\rho(A)} A]^n = |x\rangle \langle y|$, where $|x\rangle$ and $\langle y|$ are the right and left eigenvector corresponding to $\rho(A)$.

The requirement that P has to be element-wise strictly positive is rather strong and will not be met in general. To the contrary, most physically motivated Markov chains will have a large number of vanishing transition elements. So for all practical purposes one needs to have a simpler criterion that ensures the desired convergence properties. To this end, one considers a type of stochastic matrix, that is referred to as being *primitive*. These matrices are defined so they have the desired properties.

Definition 4 (Primitive maps). *We say that a stochastic matrix P is called **primitive**, if it has only one eigenvalue $\lambda(P)$ of magnitude $|\lambda(P)| = 1$ and a fixed point $|\sigma\rangle$ that is strictly positive on the total state space Ω .*

In practice one always faces the problem of testing a given stochastic matrix for primitivity. One ideally would hope to be able to do so without an explicit diagonalization of the matrix. The following criterion of primitivity gives rise to useful criteria for verification.

Lemma 5 (Condition for primitive maps). *The stochastic matrix P , with $P_{ij} \geq 0$ is primitive if and only if there exists some natural number $m \geq 1$ so that*

$$[P^m]_{ij} > 0 \tag{1.13}$$

for all pairs $i, j \in \Omega$.

PROOF: Given the Perron - Frobenius Theorem 3 the proof is straight forward. Since P is element-wise positive and we have that $[P^m]_{i,j} > 0$, we can apply the points 4 and 5 of Theorem 3 directly to P^m . Conversely, if P is taken to be primitive by Definition (4), then $\lim_{m \rightarrow \infty} P^m = |\sigma\rangle \langle I| \equiv L$. Since P is stochastic, it has the left eigenvector $\langle I|$. Primitivity furthermore requires $\sigma_i > 0$ for all i and hence $L_{ij} > 0$. Hence, there must exist an $m \geq 1$ so that (1.13) holds. \square

It is relatively easy to convince yourself, that the stochastic matrix in the above example (1.12) does not satisfy this condition.

1.2.2 Example: The Metropolis algorithm

Let us consider an example of a classical Markov chain that will be made use of extensively throughout this thesis. The prime example for a classical Markov chain is the Metropolis

algorithm introduced by Metropolis, the Rosenbluth's and the Teller's in [7]. The algorithm was first proposed to compute averages with respect to the thermal Gibbs distribution of the form

$$\langle x \rangle = \frac{\sum_i x_i e^{-\beta E_i}}{\sum_i e^{-\beta E_i}}. \quad (1.14)$$

It is evident that the explicit calculation of this sum will become intractable for a classical computer if the state space is too large. Consider for example the simple case of a one-dimensional spin chain of N spins that can assume the values $s_i \in \{+1, -1\}$. Counting the number of different spin configurations we realize that the total state space has $|\Omega| = 2^N$. So the number of summands in (1.14), and by that the complexity, increases exponentially in the number of spins. A direct computation of the average therefore becomes intractable for larger system sizes. This problem is tackled by turning the computation of (1.14) into a sampling problem. Rather than computing the full average, one just draws random configurations from the probability distribution $\sigma_i = 1/Z e^{-\beta E_i}$ and computes the empirical average. Due to the law of large numbers, we are assured that after taking M samples the statistical error scales as $\mathcal{O}(1/\sqrt{M})$, cf. [23]. The challenge is therefore to find a method that allows one to sample directly from the probability distribution σ_i . This is precisely what the Metropolis algorithm accomplishes. The Metropolis algorithm simulates, so to speak, the random walk that underlies a stochastic map which has the distribution σ_i as its unique fixed point. That is, the Metropolis algorithm can be seen as a set of rules a random walker which traverses the different configurations $i \in \Omega$ (e.g. in our example different spin states $\uparrow\uparrow\downarrow \dots \downarrow\downarrow$) has to obey as to generate an evolution that corresponds to the desired stochastic map. These rules can be cast into following randomized algorithm:

- 0 Initialization:** We randomly pick a single configuration $i \in \Omega$ and assign it to the starting position of the random walker $X_0 = i$.
- 1 Update:** The random walker is in the position $X_n = j$. From this position we propose a new configuration according to some stochastic transition matrix c which we require to be symmetric $c^T = c$. Hence, with probability c_{ij} we propose a new configuration i .
- 2 Accept/reject:** We have to decide, whether we accept the proposed update or keep the old state. We accept the new configuration with the probability

$$a = \min\left(1, \frac{\sigma_i}{\sigma_j}\right).$$

That is, with probability a we have $X_{n+1} = i$ and with probability $1 - a$ we keep $X_{n+1} = j$. We return to **1**.

The number of steps the fictive random walker has to take, i.e. the number of times the Metropolis rule has to be applied, depends on the convergence time, also called mixing time,

of the underlying stochastic map P . The stochastic map P that is generated by this algorithm is of the form

$$P_{ij} = \begin{cases} \text{if } i \neq j : & \min\left(1, \frac{\sigma_i}{\sigma_j}\right) c_{ij} \\ \text{else } i = j : & c_{ij} + \sum_k c_{ik} \left(1 - \min\left(1, \frac{\sigma_i}{\sigma_k}\right)\right) \end{cases}. \quad (1.15)$$

We know that $|\sigma\rangle$ is a fixed point distribution of this P , because the stochastic map, together with the probabilities σ_i , obeys the so called detailed balance condition, i.e. $P_{ij}\sigma_j = P_{ji}\sigma_i$, as can be checked easily. This condition ensures that $\sum_j P_{ij}\sigma_j = (\sum_j P_{ji})\sigma_i = \sigma_i$. In fact, the entire random walk was set up as to ensure the detailed balance condition. This condition, however, does not ensure that the σ_i constitute the only fixed point of the Markov chain. As we have seen in the previous discussion, pathological degeneracies can actually occur. One has to verify whether the generated map P_{ij} is primitive, c.f. Definition 4 to ensure uniqueness of the fixed point. This of course strongly depends on the updates c_{ij} . One is in general at liberty to choose the update rules in a Metropolis simulation. Note, however, that the convergence rate of the algorithm as well as the primitivity resp. ergodicity strongly depends on the choice of updates. It is therefore quite a refined art to choose good updates that meet the requirements. The overall runtime of the algorithm, i.e. the number of Metropolis steps, is known as the mixing time. This ‘time’ n characterizes the error we make when sampling from the distribution after n applications in the one norm, i.e. $\|P^n |\mu_0\rangle - |\sigma\rangle\|_1 \leq \epsilon_{mix}$. The entire complexity and runtime of the algorithm is measured in terms of the mixing time, see chapter 2 for a more in-depth discussion. There certainly are Hamiltonians for which the runtime of the algorithm scales exponential in the system size $n_{mix} \sim 2^N$, because we can encode problem instances in these Hamiltonians that correspond to NP-complete problems [24], such as for example spin glasses.

1.3 The master equation

Thus far we have only considered Markov processes that are time discret. The physically more realistic scenario, however, is when the evolution is continuous in time. In this chapter we will introduce Markov processes for which time is a continuous parameter $t \in [0, \infty)$, even though our processes will continue to take their values in some finite state space Ω . A good introduction can be found in [25, 26]. In complete analogy to (1), we define a continuous time Markov process as:

Definition 6 (Continuous time). *A stochastic process $\{X(t)|t \geq 0\}$ is called a **continuous time Markov process** if the conditional probability obeys*

$$\mathbb{P}(X(t) = i | X(s) = j, X(t_{n-1}) = i_{n-1}, \dots, X(t_{n-1}) = i_{n-1}) = \mathbb{P}(X(t) = i | X(s) = j), \quad (1.16)$$

for all non - decreasing sequences $0 \leq t_1 \dots \leq t_{n-1} \leq s \leq t$ of times and all $i_1, \dots, i_n \in \Omega$.
Moreover, we will call the process **time homogeneous**, if

$$\mathbb{P}(X(t) = i | X(s) = j) = \mathbb{P}(X(t-s) = i | X(0) = j). \quad (1.17)$$

For such a process we define the transition matrix

$$P_{ij}(t) = \mathbb{P}(X(t) = i | X(0) = j), \quad (1.18)$$

which is by construction a stochastic matrix $P(t)$ for all times t .

A homogeneous time continuous Markov process is completely described by the transition functions $P_{ij}(t)$. The transition functions can be understood as the probability to jump from state j to state i at some given time t . We will always require that $P(t)$ is continuous as a function of t and that the derivatives of the transition functions exist. In fact, it can be shown that all continuous Markov processes [25] are differentiable and we actually don't need to require this explicitly. However, since the proof is rather lengthy, we will not present it here. Furthermore it is easy to see that at $t = 0$ we have that $P(0) = \mathbb{1}$, since $P_{ii}(0) = \mathbb{P}(X(0) = i | X(0) = i) = 1$ and $P_{ij}(0) = \mathbb{P}(X(0) = i | X(0) = j) = 0$ for $i \neq j$. Note, that $P(t)$ can be seen as the probabilistic analogue of the quantum mechanical Feynman propagator. It is important to point out, that the Markov condition (1.16) immediately gives rise to a semi-group structure for the transition matrix $P(t)$. We will later on, when we consider quantum stochastic processes, make this the defining criterion for a continuous time quantum stochastic Markov process.

Lemma 7 (Chapman-Kolmogorov). *If $\{X(t) | t \geq 0\}$ is a homogeneous time continuous Markov process on a state space Ω with transition function $P_{ij}(t)$, then for any $t, s \geq 0$, we have that*

$$P_{ij}(t+s) = \sum_{k \in \Omega} P_{ik}(t) P_{kj}(s). \quad (1.19)$$

This can alternatively be written as

$$P(t+s) = P(t)P(s). \quad (1.20)$$

PROOF: The proof is straight forward; by making use of the Markov property (1.16), and since we are dealing with time-homogeneous processes, we have that

$$\begin{aligned} P_{ij}(t+s) &= \mathbb{P}(X(t+s) = i | X(0) = j) = \\ &= \sum_k \mathbb{P}(X(t+s) = i | X(s) = k, X(0) = j) \mathbb{P}(X(s) = k | X(0) = j) = \\ &= \sum_k \mathbb{P}(X(t+s) = i | X(s) = k) \mathbb{P}(X(s) = k | X(0) = j) = \sum_k P_{ik}(t) P_{kj}(s) \end{aligned} \quad (1.21)$$

□

This multiplication rule only gives rise to a semi-group, because $P(t)$ does in general not possess an inverse that is also a stochastic map. The only case for which the inverse of P is also stochastic, is when P is a permutation matrix [17]. The semi-group structure for the transition functions immediately gives rise to a differential equation for the evolution of $P(t)$.

Lemma 8 (Kolmogorov equation). *The transition functions $P_{ij}(t)$ of a finite, homogeneous time continuous Markov process satisfy the following set of differential equations. These equations are called backward (BW) and forward (FW) Kolmogorov equation respectively.*

$$\begin{aligned}\partial_t P_{ij}(t) &= \sum_{k \neq i} P_{ik}(t) q_{kj} - P_{ij}(t) \sum_{k \neq i} q_{kj} \quad (BW) \\ \partial_t P_{ij}(t) &= \sum_{k \neq i} q_{ik} P_{kj}(t) - \sum_{k \neq i} q_{ik} P_{ij}(t) \quad (FW),\end{aligned}\tag{1.22}$$

where the different q_{ik} denote transition rates to go from state i to j in a time unit. Hence, the q_{ij} are not probabilities any longer. If we define the map

$$L = \sum_{ij; i \neq j} q_{ij} |i\rangle \langle j| - \sum_{ij; i \neq j} q_{ij} |i\rangle \langle i|,\tag{1.23}$$

we can write the backward as well as the forward equation as

$$\begin{aligned}\partial_t P(t) &= P(t)L \quad (BW) \\ \partial_t P(t) &= LP(t) \quad (FW).\end{aligned}\tag{1.24}$$

PROOF: The lemma rests on the existence of the following limits.

$$\lim_{h \rightarrow 0} \frac{1 - P_{ii}(h)}{h} = v_i \quad \text{and} \quad \lim_{h \rightarrow 0} \frac{P_{ij}(h)}{h} = q_{ij}\tag{1.25}$$

We have argued earlier that we will assume that these limits exist and are moreover finite for a finite state space Ω . It is in general not necessary to assume that this is true, since it can be proved based on the continuity and the Markov property of the process, cf. [25]. However, since the proof is rather lengthy, we will omit it here. Furthermore, note that we have that $\sum_i P_{ij}(t) = 1$ for all j . Hence, we can easily see that since $\sum_{i \neq j} P_{ij}(t) = 1 - P_{jj}(t)$ we have that due to the above limits the v 's and q 's are related via, $\sum_{k \neq i} q_{ki} = v_i$. It is now very easy to derive the form of the Kolmogorov equations. Let us consider

$$\frac{1}{h} (P_{ij}(t+h) - P_{ij}(t)) = \frac{1}{h} \sum_k P_{ik}(t) P_{kj}(h) - P_{ij}(t) = \sum_k P_{ik}(t) \frac{1}{h} (P_{kj}(h) - \delta_{ik}).$$

If we now take the limit $h \rightarrow 0$, and make use of the limits in (1.25), we obtain the Kolmogorov backward equation. The forward equation can be derived similarly by taking the right differential quotient. □

Let us now take a closer look at the Kolmogorov equations. We know from our previous discussion that the differential equation has to satisfy the initial condition $P(0) = \mathbb{1}$. With this at hand it is straight forward to find the formal solution, which is a solution for both the forward as well as the backward equation. The solution is simply given by the exponential

$$P(t) = \exp(Lt). \quad (1.26)$$

Let us now consider the time evolution of an arbitrary initial probability distribution $|\mu(0)\rangle$. Given the time-dependent transition function $P(t)$, we know that the time evolution is

$$|\mu(t)\rangle = P(t) |\mu(0)\rangle. \quad (1.27)$$

It is possible to rewrite the Kolmogorov equation as a dynamical equation, known as the *master equation* for the probability distribution directly by multiplying (1.24) from the right with the initial probability distribution $|\mu(0)\rangle$. The master equation then is,

$$\partial_t |\mu(t)\rangle = L |\mu(t)\rangle. \quad (1.28)$$

The evolution of a continuous time Markov process is completely determined by the generator of the semi group L . To describe the dynamical evolution of a physical system it suffices to provide the generator. In chapter 4 we will encounter a stochastic non-equilibrium process known as the asymmetric exclusion process which is described by such a master equation. Let us pause and investigate the generator more closely. Given the Definition of L in (1.23), we immediately see that the generator has the left eigenvector $\langle I | L = 0$. Furthermore, a steady state of $\exp(Lt)$ has to obey that $|\sigma\rangle = P(t) |\sigma\rangle$, which translates to the condition that the steady state has to be the right eigenvector of L corresponding to the eigenvalue 0, i.e. $L |\sigma\rangle = 0$. It is possible to apply the results of the spectral behavior, discussed in the previous section 1.2.1, directly to the matrix $P(t)$. Note, that we already know, that we have at least one eigenvalue $\lambda = 0$. Furthermore, since the rates q_{ij} are real, we know that the eigenvalues of L have to come in complex conjugate pairs. Moreover, we can immediately infer from the formal solution (1.26) a bound on the spectrum of L . Since $P(t)$ is a stochastic matrix, its spectrum has to be contained in the unit disc of the complex plane. Therefore the real part of the spectrum of L has to be negative or zero.

1.4 Quantum Markov chains

So far we have only been talking about classical stochastic systems. The central topic of this thesis is, however, the investigation of quantum mechanical Markov processes. We therefore need a natural extension of such processes to the quantum domain. We have to find some form of a stochastic transition matrix for non-commuting probability spaces. For quantum mechanical systems so-called trace-preserving completely-positive maps, in short tcp-maps or

just quantum channels, will take the role of the previously discussed stochastic maps [27, 28]. As will be discussed in the following, the justification of the analogy stems from the fact, that the trace-preserving completely-positive maps exhibit the same spectral characteristics as stochastic matrices. The concept of completely-positive maps is actually more general. In fact, any permissible operation on a quantum mechanical system can be represented by a completely-positive map. Let us discuss these maps in more detail.

1.4.1 Completely positive maps

For these maps to be physically meaningful [15], we shall require the following points:

Definition 9 (Quantum channels). *Let $A, B \in \mathcal{M}_d$ and $\lambda \in \mathbb{C}$ and furthermore let id_n denote the identity map on the space $\mathcal{M}(\mathcal{H}_n)$. Here \mathcal{H}^n denotes an arbitrary Hilbert space of dimension $n < \infty$. Then, a map $\mathcal{T} : \mathcal{M}_d \rightarrow \mathcal{M}_d$ is called trace preserving and completely-positive if it obeys the following conditions:*

1. *Linearity:*

$$\mathcal{T}(A + \lambda B) = \mathcal{T}(A) + \lambda \mathcal{T}(B) \quad (1.29)$$

2. *Preservation of trace and Hermiticity:*

$$\text{tr}[\mathcal{T}(A)] = \text{tr}[A] \quad \text{and} \quad \mathcal{T}(A)^\dagger = \mathcal{T}(A^\dagger). \quad (1.30)$$

3. *Complete positivity:*

$$\mathcal{T} \otimes id_n(A^\dagger A) \geq 0 \quad \forall A \in \mathcal{M}(\mathcal{H} \otimes \mathcal{H}^n) \quad \text{and} \quad n \in \mathbb{N} \quad (1.31)$$

The requirement of linearity is inherent to quantum mechanics. It can be shown [15, 28], that non-linear transformations would allow to transmit signals instantaneously. The second and third requirement ensure that states, i.e. $\rho \in \mathcal{S}^+(\mathcal{H})$ are again mapped to states. The requirement of positivity of the map \mathcal{T} alone does not guarantee this. Consider for instance a state $\rho \in \mathcal{S}^+(\mathcal{H}_1)$ that is part of a larger bipartite state $\phi \in \mathcal{S}(\mathcal{H}_1 \otimes \mathcal{H}_2)$. If we only apply \mathcal{T} on the subsystem described by ρ and act trivially, i.e. with the identity id on the complementary subsystem, then positivity of \mathcal{T} alone does not ensure that the total map $\mathcal{T} \otimes id$ is positive. An example of such a map is easily constructed by fixing a basis in \mathcal{H}_1 and defining \mathcal{T} as the transpose on this Hilbertspace. Even though \mathcal{T} is then a positive map, $\mathcal{T} \otimes id$ will be no longer. If one drops the requirement that the map has to be trace preserving, one refers to the maps, which still obey the last condition, only as completely-positive maps. Such maps can be implemented on the system in the context of so-called instruments [28] and can for instance be used to model the post selection of some measurement.

We stated earlier, that \mathcal{M}_d together with $\text{tr}[A^\dagger B] = \langle A|B \rangle_{HS}$ is a Hilbert space. It is possible to define the dual of \mathcal{T} , that is $\mathcal{T}^* : \mathcal{M}_d \rightarrow \mathcal{M}_d$ as the conjugate map of \mathcal{T} with respect to the Hilbert-Schmidt scalar product, i.e. $\text{tr}[\mathcal{T}(\rho)A] = \text{tr}[\rho\mathcal{T}^*(A)]$. Therefore \mathcal{T}^* acts on the observables, rather than on the states and can be seen to implement a quantum evolution in the Heisenberg picture. The requirement that \mathcal{T} is trace preserving now simply reads $\mathcal{T}^*(\mathbb{1}) = \mathbb{1}$, i.e. the identity is the fixed point of the dual channel. We call a map \mathcal{T} that has the identity as a fixed point *unital*.

Let us now find a way to characterize such completely-positive maps on a finite dimensional state space. We need a simple criterion to verify whether a linear map \mathcal{T} is indeed a tcp-map. Consider therefore the following state that is obtained by applying \mathcal{T} only to one subsystem of a maximally entangled state.

Definition 10 (Choi - Jamiolkowski). *Let $|\Omega\rangle = \frac{1}{\sqrt{d}} \sum_i^d |ii\rangle$ denote the maximally entangled state on $\mathbb{C}^d \otimes \mathbb{C}^d$, we define a state $\tau \in \mathcal{S}^+(\mathbb{C}^d \otimes \mathbb{C}^d)$ as the Choi - Jamiolkowski state associated with \mathcal{T} as*

$$\tau = (\mathcal{T} \otimes id_d) (|\Omega\rangle \langle\Omega|). \quad (1.32)$$

Between the map \mathcal{T} and the state τ exists a simple one - to - one correspondence. This correspondence and its consequences are expressed in the following lemma.

Lemma 11 (Jamiolkowski isomorphism). *A linear map $\mathcal{T} : \mathcal{M}(\mathbb{C}^d) \rightarrow \mathcal{M}(\mathbb{C}^d)$ is related to an operator $\tau \in \mathcal{M}(\mathbb{C}^d \otimes \mathbb{C}^d)$ via the identity*

$$\text{tr}[AT(B)] = d \text{tr}[\tau A \otimes B^T] \quad \text{where} \quad \tau = (\mathcal{T} \otimes id_d) (|\Omega\rangle \langle\Omega|) \quad (1.33)$$

where $A, B \in \mathcal{M}(\mathbb{C}^d)$ and $|\Omega\rangle$ again the maximally entangled state. The above relations lead to the following correspondence between τ and \mathcal{T} .

1. *Complete positivity: \mathcal{T} is completely positive if and only if $\tau \geq 0$.*
2. *Preservation of Hermiticity: $\tau = \tau^\dagger$ if and only if $\mathcal{T}(A)^\dagger = \mathcal{T}(A^\dagger)$.*
3. *Preservation of trace: $\mathcal{T}^*(\mathbb{1}) = \mathbb{1}$ if and only if $\text{tr}_{\mathcal{H}_1}[\tau] = \frac{1}{d}\mathbb{1}$.*
4. *Unitality: $\mathcal{T}(\mathbb{1}) = \mathbb{1}$ if and only if $\text{tr}_{\mathcal{H}_2}[\tau] = \frac{1}{d}\mathbb{1}$.*

PROOF: The first thing we need to show is the correspondence given in (1.33). Note that

$$d \text{tr}[\tau A \otimes B^T] = d \langle\Omega| \mathcal{T}^*(A) \otimes B^T |\Omega\rangle = \text{tr}[AT(B)]. \quad (1.34)$$

Furthermore, we know that τ has to be positive by definition, since \mathcal{T} is completely positive. So we only need to show the converse, namely that for $\tau \geq 0$ the map \mathcal{T} is completely positive. To proof this let us consider an arbitrary density matrix $\rho \in \mathcal{M}(\mathbb{C}^d \otimes \mathbb{C}^n)$ for some $n \in \mathbb{N}$. We know that $(\mathcal{T} \otimes id_n)(\rho) \geq 0$ if this statement holds for all pure states in the decomposition

of $\rho = \sum_i p_i |\psi_i\rangle \langle \psi_i|$, that is if $(\mathcal{T} \otimes id_n)(|\psi_i\rangle \langle \psi_i|) \geq 0$ for all i . However, we have already seen that every pure state $|\psi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^n$ can be written as $|\psi\rangle = \mathbb{1} \otimes R |\Omega\rangle$ for some $R \in \mathcal{M}(\mathbb{C}^d, \mathbb{C}^n)$. Thus we can rewrite $(\mathcal{T} \otimes id_n)(|\psi_i\rangle \langle \psi_i|) = (\mathbb{1} \otimes R)\tau(\mathbb{1} \otimes R^\dagger)$. Since $\tau \geq 0$, we see directly that some matrix of the form $A\tau A^\dagger$ is also positive and by that the positivity of $(\mathcal{T} \otimes id_n)$ for all n . The remaining relations are easily proved by direct observation. \square

The Choi-Jamiolkowski correspondence allows one to translate between the properties of states on a bipartite Hilbert space and completely-positive maps. It is also straight forward to recover the action of the map \mathcal{T} on some $A \in \mathcal{M}_d$. By working in a suitable basis of \mathcal{M}_d , choose for instance the product basis $\{|i\rangle \langle j|\}_{i,j=1\dots d}$, we can write

$$\mathcal{T}(A) = d \sum_{i,j,k,l} \langle ij | \tau | kl \rangle |i\rangle \langle j | A | k \rangle \langle l|. \quad (1.35)$$

It is also possible to derive a generic form for completely positive maps based on the discussed correspondence. We will shortly see, that every completely-positive map can be written in the so called Kraus representation.

Theorem 12 (Kraus representation). *A linear map $\mathcal{T} : \mathcal{M}_d \rightarrow \mathcal{M}_d$ is completely positive, if and only if it can be written in the Kraus form*

$$\mathcal{T}(\rho) = \sum_{j=1}^r A_j \rho A_j^\dagger, \quad (1.36)$$

where the operators A_j are the so-called Kraus operators and satisfy:

1. *Normalization:*

\mathcal{T} is trace preserving if and only if $\sum_j A_j^\dagger A_j = \mathbb{1}$ and unital if and only if $\sum_j A_j A_j^\dagger = \mathbb{1}$.

2. *Kraus rank:*

The minimal number of Kraus operators is $r = \text{rank}(\tau) \leq \dim(\mathcal{H})^2$.

3. *Orthogonality:*

Furthermore, there always exists a decomposition with $r = \text{rank}(\tau)$ Kraus operators that are orthogonal with respect to the Hilbert - Schmidt scalar product, i.e. $\text{tr}[A_j^\dagger A_i] \propto \delta_{ji}$.

4. *Freedom of the representation:*

The Kraus decomposition is not unique. However, it is possible to always find a unitary relationship between different sets of Kraus operators for the same map \mathcal{T} . We can always relate the two sets $\{A_i\}$ and $\{\tilde{A}_j\}$ by a unitary with entries U_{ij} via $A_i = \sum_j U_{ij} \tilde{A}_j$.

PROOF: We start by assuming that \mathcal{T} is completely positive and show that the Kraus representation follows from $\tau \geq 0$. Since τ is positive it always permits a spectral decomposition into r unnormalized rank 1 projectors

$$\tau = \sum_{i=1}^r |\psi_i\rangle\langle\psi_i| = \sum_{i=1}^r (A_i \otimes \mathbb{1}) |\Omega\rangle\langle\Omega| (A_i^\dagger \otimes \mathbb{1}). \quad (1.37)$$

Recall that τ was defined as $(\mathcal{T} \otimes \text{id})(|\Omega\rangle\langle\Omega|)$. From this we read off, that \mathcal{T} acts as (1.36) on one half of the maximally entangled state. Furthermore, the Kraus rank r and the orthogonality of the A_i immediately follow from the spectral decomposition of the state τ . Conversely, if the map \mathcal{T} is chosen as in (1.36), positivity of τ follows directly. The normalization of the partial traces of $\text{tr}_{A/B}[\tau]$ translate to the constraints $\sum_i A_i^\dagger A_i = \mathbb{1}$ and $\sum_i A_i A_i^\dagger = \mathbb{1}$ for the Kraus operators. Recall that the decomposition of τ into rank-1 projectors is not unique. Other decomposition with $k > r$ projectors are also possible,

$$\tau = \sum_{i=1}^r |\psi_i\rangle\langle\psi_i| = \sum_{j=1}^k |\tilde{\psi}_j\rangle\langle\tilde{\psi}_j|. \quad (1.38)$$

if the set of vectors $\{|\psi_i\rangle\}_i$ and $\{|\tilde{\psi}_j\rangle\}_j$ are related by a unitary U , so that $|\psi_i\rangle = \sum_j U_{ij} |\tilde{\psi}_j\rangle$. This follows directly from considering all possible purifications $|\Psi\rangle_{AB}$ of $\tau = \text{tr}_B[|\Psi\rangle\langle\Psi|]$, because two different purifications $|\Psi\rangle_{AB} = \sum_{i=1}^r |\psi_i\rangle |i\rangle$ and $|\tilde{\Psi}\rangle_{AB} = \sum_{j=1}^k |\tilde{\psi}_j\rangle |j\rangle$ only differ by a unitary of the form $(\mathbb{1}_A \otimes U_B)$. We can therefore conclude, that also all possible Kraus operators are related by a unitary transformation. \square

The general perception is that completely-positive maps arise due to *open system dynamics*. The dynamics of an open system is modeled by only considering the dynamics of a chosen subsystem of a larger system that evolves unitarily. With the Kraus representation at hand it is indeed possible to see, that every tcp-map can be understood in this sense.

Theorem 13 (Open-system representation). *Let $\mathcal{T} : \mathcal{M}_d \rightarrow \mathcal{M}_d$ be a completely-positive trace-preserving map, then there exists a unitary U acting on $\mathcal{H} \otimes \mathcal{H}_E$ and a normalized vector $|\phi\rangle \in \mathcal{H}_E$, so that we can write*

$$\mathcal{T}(\rho) = \text{tr} \left[U(\rho \otimes |\phi\rangle\langle\phi|) U^\dagger \right], \quad (1.39)$$

where tr_E is taken as the partial trace over the ancilla space \mathcal{H}_E .

PROOF: Let us first prove that the application of a unitary followed by a partial trace gives rise to a tcp-map. We choose a basis $\{|i\rangle\}_i$ of the ancilla space \mathcal{H}_E and write

$$\text{tr}_E \left[U(\rho \otimes |\phi\rangle\langle\phi|) U^\dagger \right] = \sum_i \langle i | U | \phi \rangle \rho \langle \phi | U^\dagger | i \rangle \equiv \sum_i A_i \rho A_i^\dagger, \quad (1.40)$$

where we defined the operators $A_i \equiv \langle i | U | \phi \rangle$, which act only the space \mathcal{M}_d and satisfy the normalization $\sum_i A_i^\dagger A_i = \langle \phi | U^\dagger U | \phi \rangle = \langle \phi | \mathbb{1} \otimes \mathbb{1}_E | \phi \rangle = \mathbb{1}$. Thus, these operators are the

Kraus operators of a tcp-map. Conversely, consider the Kraus decomposition of the tcp-map $\mathcal{T}(\rho) = \sum_i A_i \rho A_i^\dagger$. From this we construct the isometry $V = \sum_i A_i \otimes |i\rangle$ on $\mathcal{H} \otimes \mathcal{H}_E$. We have that $\mathcal{T}(\rho) = \text{tr}_E[V\rho V^\dagger]$. It is possible to find a unitary U that acts with respect to a reference state $|\phi\rangle$ like the isometry $V = U(\mathbb{1} \otimes |\phi\rangle)$. \square

Matrix representation A convenient way of expressing the action of a linear map on the space \mathcal{M}_d is via its matrix representation. Recall that \mathcal{M}_d itself is a linear space with an endowed natural scalar product, the Hilbert-Schmidt scalar product as discussed earlier. Furthermore, recall that by this \mathcal{M}_d is naturally isomorphic to $\mathcal{M}_d \simeq \mathcal{H}^d \otimes \mathcal{H}^d$. The action of a linear map on this space can be represented by a matrix, if a suitable basis is chosen. We will denote the matrix representation of \mathcal{T} on $\mathcal{H}^d \otimes \mathcal{H}^d$, which we will call the transfer matrix, by a hat, i.e. $\hat{\mathcal{T}}$. A concrete matrix representation of \mathcal{T} is obtained by choosing an orthonormal operator basis $\{F_\alpha \in \mathcal{M}_d\}_{\alpha=1\dots d^2}$ of \mathcal{M}_d . The matrix elements can then be computed as

$$\hat{\mathcal{T}}_{\alpha,\beta} = \langle F_\alpha | \mathcal{T} | F_\beta \rangle_{HS} = \text{tr} \left[F_\alpha^\dagger \mathcal{T}(F_\beta) \right]. \quad (1.41)$$

Therefore, we have by construction that the composition of two channels $\mathcal{T}^3 = \mathcal{T}^2 \circ \mathcal{T}^1$ corresponds to the common matrix product of $\hat{\mathcal{T}}^3 = \hat{\mathcal{T}}^2 \hat{\mathcal{T}}^1$. Furthermore, the matrix representation of the conjugate channel \mathcal{T}^* is given by the adjoint matrix $\hat{\mathcal{T}}^\dagger$. The simplest orthonormal operator basis of \mathcal{M}_d is given by the standard product basis $\{|i\rangle\langle j|\}_{i,j=1\dots d^2}$. With respect to this basis, a tcp-map with Kraus representation $\mathcal{T}(\rho) = \sum_i A_i \rho A_i^\dagger$ is simply given by $\hat{\mathcal{T}} = \sum_i A_i \otimes \bar{A}_i$. In this basis, it is straight forward to relate the Choi-Jamiolkowski operator τ to the matrix representation by writing

$$\hat{\mathcal{T}} = d \tau^{T_B}, \quad (1.42)$$

where T_B denotes the partial transposition with respect to the second Hilbert space, i.e. $\langle m n | \tau^{T_B} | i j \rangle \equiv \langle m i | \tau | n j \rangle$. It, however, may turn out that some other matrix basis can be more useful, such as choosing the F_α hermitian or unitary. This generally depends on the problem. A more exhaustive list of operator Bases can be found in [29, 28].

1.4.2 Perron Frobenius and irreducibility

The fact that the trace-preserving completely-positive maps can be seen as the natural quantum generalization of classical stochastic matrices is due to their spectral properties. A quantum channel does indeed exhibit features very similar to those of a classical transition matrix. Apart from the obvious fact, that both maps are linear and map probability distributions to probability distribution, that is, states to states in the quantum setting, we will see that the quantum channels possess the same spectral behavior as their classical counterparts. Furthermore, the concept of irreducibility and ergodicity can also be defined in an analogous fashion. In chapter 2 we carry this analogy further. Let us first investigate the spectral and fixed point properties

of tcp-maps. We can define the spectrum $\text{spec}(\mathcal{T})$ of some tcp-map $\mathcal{T} : \mathcal{M}_d \rightarrow \mathcal{M}_d$ by the set of $\{\lambda \in \mathbb{C}\}$ so that for each of the λ 's there is a $X \in \mathcal{M}_d$ for which

$$\mathcal{T}(X) = \lambda X. \quad (1.43)$$

The operator X is often referred to as the eigenvector of \mathcal{T} . When we consider \mathcal{M}_d as a d^2 -dimensional vector space, the eigenvalues of \mathcal{T} can be seen as just the eigenvalues of the $d^2 \times d^2$ -matrix representation $\hat{\mathcal{T}}$. Note, that since we are considering maps that are Hermiticity preserving, i.e. $\mathcal{T}(A)^\dagger = \mathcal{T}(A^\dagger)$, we are ensured that the eigenvalues are real or come in complex pairs. It is easy to see, that the spectrum of a completely-positive map is similarly confined as that of stochastic matrices.

Lemma 14 (Spectral radius). *If \mathcal{T} is a positive map on \mathcal{M}_d , then its spectral radius $\rho(\mathcal{T}) = \sup \{|\lambda| \mid \lambda \in \text{spec}(\mathcal{T})\}$ satisfies*

$$\rho(\mathcal{T}) \leq \|\mathcal{T}(\mathbb{1})\|_\infty, \quad (1.44)$$

where $\|\cdot\|_\infty$ denotes the infinity norm on the matrix space \mathcal{M}_d , cf. [17]. If in addition \mathcal{T} is either unital or trace preserving, there exists an eigenvalue $\lambda_0 = 1$ and we have that $\rho(\mathcal{T}) = 1$. So all eigenvalues lie in the unit disc of the complex plain.

PROOF: The proof is almost identical to that of stochastic matrices. We only need to make use of the Russo - Dye Theorem [17] $\|\mathcal{T}(X)\|_\infty \leq \|\mathcal{T}(\mathbb{1})\|_\infty \|X\|_\infty$ so we can write $|\lambda| \|X\|_\infty = \|\mathcal{T}(X)\|_\infty \leq \|\mathcal{T}(\mathbb{1})\|_\infty \|X\|_\infty$, which implies (1.44). Furthermore we have that if \mathcal{T} is unital, then $\|\mathcal{T}(\mathbb{1})\|_\infty = 1$. Since if \mathcal{T} is trace preserving, we have that \mathcal{T}^* is unital and the spectra of both maps coincide. \square

We have seen that the spectral properties of stochastic maps can be understood on the account of the Perron-Frobenius theorem. In order to derive a statement for completely-positive maps we first have to think about a suitable classification scheme of what irreducibility and primitivity means for completely-positive maps. Such results were first proved in [30] for completely-positive maps. The following Theorem defines *irreducible* positive maps and shows that this can be done in various different ways which turn out to be equivalent.

Theorem 15 (Irreducibility of positive maps). *Let $\mathcal{T} : \mathcal{M}_d \rightarrow \mathcal{M}_d$ be a positive linear map. The following properties are equivalent:*

1. *If $P \in \mathcal{M}_d$ is a Hermitian projector such that $\mathcal{T}(P \mathcal{M}_d P) \subseteq P \mathcal{M}_d P$, then $P \in \{0, \mathbb{1}\}$.*
2. *For every non-zero $A \geq 0$ we have $(id + \mathcal{T})^{d-1}(A) > 0$.*
3. *For every non-zero $A \geq 0$ and every $t \in (0, \infty)$ we have $\exp(t\mathcal{T})(A) > 0$.*
4. *For every orthogonal pair of non-zero, positive semi-definite matrices $A, B \in \mathcal{M}_d$, there is an integer $n \in \{1, \dots, d-1\}$ such that $\text{tr}[BT^n(A)] > 0$.*

PROOF: 1. \rightarrow 2.: From $\mathcal{T}(A) \geq 0$ we get an inclusion for the kernels $\ker(\text{id} + \mathcal{T})(A) \subseteq \ker(A)$. Suppose equality holds in this inclusion, then $\text{supp}(\mathcal{T}(A)) \subseteq \text{supp}(A)$. Therefore $\mathcal{T}(P \mathcal{M}_d P) \subseteq P \mathcal{M}_d P$, if we take P as the hermitian operator that projects on the support space of A . Since \mathcal{T} is irreducible this can only be if we have that $A > 0$. The application of $(\text{id} + \mathcal{T})$ to A thus has to increase the rank until there is no kernel left, which happens at least after $d - 1$ steps. 2. \rightarrow 3.: Comparing the Taylor expansion of $\exp(t\mathcal{T}(A)) > 0$ with that of $(\text{id} + \mathcal{T})^{d-1}(A) > 0$, we see that the expansion of $(\text{id} + \mathcal{T})^{d-1}(A) > 0$ is part of the expansion of the exponential. Moreover, all terms in the expansion are positive, thus we have $\exp(t\mathcal{T}(A)) \geq (\text{id} + \mathcal{T})^{d-1}(A)$. 3. \rightarrow 1.: Suppose \mathcal{T} is irreducible, then there is some $P \notin \{0, \mathbb{1}\}$ so that $\mathcal{T}(P) \leq cP$ for some constant c . Then, however, $\exp(t\mathcal{T})(P) \leq \exp(ct)(P)$ in contradiction with 3. 4. \rightarrow 1.: If \mathcal{T} is reducible we have a projection operator P so that $\text{tr}[(1 - P)\mathcal{T}^n(P)]$ for all n . 2. \rightarrow 4.: Choose B, A with $\text{tr}[BA] = 0$ and expand $\text{tr}[B(\text{id} + \mathcal{T})^{d-1}(A)] > 0$. Since all terms are positive, at least one term has to be $\text{tr}[BT^n(A)] > 0$ for n in the specified subset. \square

In order to relate *irreducibility* to the spectral properties of a positive map \mathcal{T} it is useful to consider the following functionals defined on the cone of positive semi-definite operators:

$$r(X) \equiv \sup\{\lambda \in \mathbb{R} \mid (\mathcal{T} - \lambda \text{id})(X) \geq 0\}, \quad (1.45)$$

$$\tilde{r}(X) \equiv \inf\{\lambda \in \mathbb{R} \mid (\mathcal{T} - \lambda \text{id})(X) \leq 0\}. \quad (1.46)$$

We are especially interested in the maxima $r \equiv \sup_{X \geq 0} r(X)$ and $\tilde{r} \equiv \sup_{X \geq 0} \tilde{r}(X)$ which obviously satisfy $r \geq \tilde{r}$ and have to coincide for irreducible maps.

Theorem 16 (Perron-Frobenius for positive maps). *Let $\mathcal{T} : \mathcal{M}_d \rightarrow \mathcal{M}_d$ be an irreducible positive map. Then:*

1. *We have that $r = \tilde{r}$ for the quantities defined in (1.45) and (1.46).*
2. *r is a non-degenerate eigenvalue of \mathcal{T} and the corresponding eigenvector is strictly positive, i.e. $\mathcal{T}(X) = rX > 0$.*
3. *If there is any $\lambda > 0$ which is an eigenvalue of \mathcal{T} with positive eigenvector, i.e. $\mathcal{T}(Y) = \lambda Y \geq 0$, then $\lambda = r$.*
4. *r is the spectral radius of \mathcal{T} .*

PROOF: We observe that we can be assured that the inf and sup in the equations (1.45) and (1.46) will be obtained, since we can work on a compact set by requiring that $\text{tr}[X] = 1$. We begin by showing that r is attained for a $X > 0$ so that $\mathcal{T}(X) = rX$. Consider any non-zero $X \geq 0$ for which $\lambda = r(X) > 0$. Since

$$(\mathcal{T} + \text{id})^{d-1}(\mathcal{T} - \lambda \text{id})(X) = (\mathcal{T} - \lambda \text{id})(\mathcal{T} + \text{id})^{d-1}(X) \quad (1.47)$$

we can infer two things. First, the supremum of (1.45) is attained for a strictly positive $X > 0$ since we have that irreducibility is equivalent to $(\mathcal{T} + \text{id})^{d-1}(X) > 0$. Second, the X achieving the supremum must satisfy $(\mathcal{T} - r\text{id})(X) = 0$. Otherwise the expression in (1.47) would be positive definite and a large multiple of the identity could be subtracted, which is in contradiction to the supposed maximality. Also, we have that $r(X) = \tilde{r}(X)$ for any eigenvector $X \geq 0$, the mentioned observation together with $r \geq \tilde{r}(X)$ proves statement 1 of the theorem. We need to show non-degeneracy to prove the second point of the theorem. To this end, assume there is an X' which is linearly independent from X and also an eigenvector to the eigenvalue r . We can always choose $X'^{\dagger} = X'$. It is always possible to choose a $c \in \mathbb{R}$ so that $X + cX' \geq 0$ has a non-vanishing kernel. This is, however, in contradiction to $0 < (\mathcal{T} + \text{id})^{d-1}(X + cX') = (r + 1)^{d-1}(X + cX')$. Thus r has to be non-degenerate. Let us proof 3: Assume $Y \geq 0$ is an eigenvector of \mathcal{T} with eigenvalue $\lambda > 0$. Since r is also the eigenvalue of the dual \mathcal{T}^* for some eigenvector $X' > 0$, we have that $r\text{tr}[X'Y] = \text{tr}[\mathcal{T}^*(X')Y] = \text{tr}[X'\mathcal{T}(Y)] = \lambda\text{tr}[X'Y]$. Since $\text{tr}[X'Y] > 0$ we have that $r = \lambda$. Furthermore, since $X > 0$ as the eigenvector corresponding to r , we can always define a unital map $\mathcal{T}'(\cdot) \equiv 1/r X^{-1/2} \mathcal{T}(X^{1/2} \cdot X^{1/2}) X^{-1/2}$ which is similar to \mathcal{T} . We have that the spectral radius $\rho(\mathcal{T}') = 1$ and thus $\rho(\mathcal{T}) = r$. \square

This non-commutative version of the Perron-Frobenius Theorem immediately gives rise to some statements about the fixed point structure of positive maps. This is expressed in the following proposition.

Proposition 17 (Irreducibility from spectral properties). *Let $\mathcal{T} : \mathcal{M}_d \rightarrow \mathcal{M}_d$ be a positive map with spectral radius $\rho(\mathcal{T})$. Then the following statements are equivalent:*

1. \mathcal{T} is irreducible
2. The spectral radius $\rho(\mathcal{T})$ is a non-degenerate eigenvalue and the corresponding left and right eigenvectors are positive definite. That is $\mathcal{T}^*(Y) = \rho(\mathcal{T})Y$ and $\mathcal{T}(X) = \rho(\mathcal{T})X$.

PROOF: The direction $1 \rightarrow 2$ is a direct consequence of the previous Theorem 16 when applied to \mathcal{T} and \mathcal{T}^* respectively. For the converse, observe that if $Q \in \mathcal{M}_d$ is some invertible matrix we have that for an irreducible \mathcal{T} the map $\mathcal{T}'(\cdot) = cQ^{-1}\mathcal{T}(Q \cdot Q^{\dagger})[Q^{\dagger}]^{-1}$ is also irreducible. This is true, since if \mathcal{T}' were reducible, with some projection P' , some P acting as a projection on the support of $QP'Q^{\dagger}$ would also reduce \mathcal{T} . If we now choose $c = 1/\rho(\mathcal{T})$ and $Q = Y^{-1/2}$ the map \mathcal{T}' becomes trace preserving and we have that $\mathcal{T}'(X') = X' > 0$ where $X' = Y^{1/2}XY^{1/2}$. Moreover, we have now that the eigenvalue 1 of \mathcal{T}' is non-degenerate. Now assume that \mathcal{T}' and thus \mathcal{T} would be reducible, then there is some hermitian projection $P \notin \{0, \mathbb{1}\}$ so that $\mathcal{T}'(P\mathcal{M}_dP) = P\mathcal{M}_dP$ and there has to be a right eigenvector $\sigma \geq 0$ that does not have full rank. Hence, σ and $X > 0$ would be linear independent and the eigenvalue would be degenerate, which leads to a contradiction \square

We now focus on a special class of trace-preserving completely-positive maps that are called *primitive*. A primitive map is an irreducible tcp-map that has only one eigenvalue of magnitude 1. The Perron-Frobenius Theorem for tcp - maps only ensures that the eigenvalue $\lambda_0 = 1$ is non-degenerate. It, however, does not exclude some form of ‘oscillatory’ behavior with respect to some other eigenvalue, e.g. $\lambda_- = -1$ or some other phase factor. The presence of other eigenvalues of magnitude $|\lambda| = 1$ hinder the convergence to a unique fixed point. We will see that primitive maps are not plagued by this oscillatory behavior. The following Theorem was first proved in [31].

Theorem 18 (Primitive maps). *Let $\mathcal{T} : \mathcal{M}_d \rightarrow \mathcal{M}_d$ denote a tcp-map. Then the following statements are equivalent:*

1. *There exists an $n \in \mathbb{N}$ so that for any state $\rho \in \mathcal{S}_d^+$ we have that $\mathcal{T}^n(\rho) > 0$, i.e. after n applications the resulting state has full rank.*
2. *The composed channel \mathcal{T}^k is irreducible for any $k \in \mathbb{N}$.*
3. *There exists a $\sigma \in \mathcal{S}_d^+$ with $\sigma > 0$ so that for all $\rho \in \mathcal{S}_d^+$ we have that $\lim_{n \rightarrow \infty} \mathcal{T}^n(\rho) = \sigma$.*

PROOF: We prove 2. \rightarrow 3. Since \mathcal{T} is irreducible there is a unique fixed point $\mathcal{T}(\sigma) = \sigma > 0$ and due to the trace preservation the spectral radius of \mathcal{T} is one. Yet, since \mathcal{T}^k is irreducible for any k we have that all other eigenvalues have to be strictly smaller than one. Therefore the limit in 3 exists. Now 3. \rightarrow 2.: Suppose that \mathcal{T}^k is reducible for some k , then there is some projection $P \notin \{0, \mathbb{1}\}$ for which $\mathcal{T}^{kl}(P) \not\propto 0$ for all l which is in contradiction to 3. 1. \rightarrow 2. follows by the same argument. Finally 3. \rightarrow 1.: Assume that $\mathcal{T}^n(\rho)$ has a kernel with eigenvector $|\psi\rangle$. Then $\lambda_{\min}(\sigma) \leq |\langle \psi | \mathcal{T}^n(\rho) - \sigma | \psi \rangle| \|\mathcal{T}^n(\rho) - \sigma\| = \|(\mathcal{T}^n - \mathcal{T}_\infty)(\rho - \sigma)\| \leq |\lambda_1|^n c \|\rho - \sigma\|$, where λ_{\min} denotes the smallest eigenvalue and λ_2 is the second largest eigenvalue in magnitude of \mathcal{T} . Here $\|\cdot\|$ denotes the operator norm. The constant c depends on \mathcal{T} , but is independent on n and $\mathcal{T}_\infty(\cdot) = \sigma \text{tr}[\cdot]$. Taking $n \rightarrow \infty$ leads to a contradiction. \square

1.5 The Lindblad equation

In section 1.3 we have been discussing the description of classical continuous-time Markov processes. It is of course desirable to have a similar description for continuous time quantum Markov processes. Indeed, such a formalism has been derived in [32, 33]. The formalism gives the correct description of the irreversible evolution of an open quantum system that is in contact with a heat bath provided that the relaxation time of the correlations with the bath, is typically much shorter than the decay times of the system. The formalism is encountered in a variety of of physical problem such as in quantum optics [34] and in terms of the description of decoherence in open quantum systems [35]. We will be making use of the Lindblad equation in

chapter 5 to investigate transport properties of driven quantum systems. The starting point for the formal derivation is the dynamical semi-group of completely-positive maps. This follows the reasoning we have already encountered in the derivation of the classical master equation, cf. Lemma 8. We have seen that the Markov property is essentially equivalent to the structure of a semi-group as can be seen in the Chapman - Kolmogorov Lemma 7.

Definition 19 (Dynamical semi-group). *For a set of states $\mathcal{S}^+(\mathcal{H})$ a family of maps $\mathcal{T}_t : \mathcal{S}^+(\mathcal{H}) \rightarrow \mathcal{S}^+(\mathcal{H})$, parametrized by $t \in [0, \infty)$ is called a **dynamical semi-group**, if for all $t, s \in [0, \infty)$ we have that*

$$\mathcal{T}_{t+s} = \mathcal{T}_t \circ \mathcal{T}_s \quad \text{and} \quad \mathcal{T}_0 = \text{id}. \quad (1.48)$$

Note, that this semi-group property implies that the underlying process is both Markovian and time homogeneous, so the evolution depends neither on the history nor on the current point in time. Furthermore, we have not yet required that the family of maps \mathcal{T}_t has to be *completely positive*! In this section we are concerned with deriving the necessary and sufficient condition for a dynamical law known as the quantum master equation that generates the evolution so that the corresponding semi-group is indeed completely positive. We will show that the semi-group can always be written as $\mathcal{T}_t = \exp(\mathcal{L}t)$, if the semi-group is continuous in t . Here, continuity is typically assumed with respect to the trace distance $\|\cdot\|_{tr}$. The *complete positivity* follows from the special form of the generator.

Proposition 20 (Form of the semi-group). *Let \mathcal{T}_t be a dynamical semi-group, which is continuous in $t \in [0, \infty)$, then \mathcal{T}_t is differentiable and has to be of the form $\mathcal{T}_t = \exp(\mathcal{L}t)$ for some generator $\mathcal{L} : \mathcal{M}_d \rightarrow \mathcal{M}_d$. Thus, \mathcal{T}_t satisfies the differential equation $\partial_t \mathcal{T}_t = \mathcal{L} \mathcal{T}_t$.*

PROOF: Since \mathcal{T}_t is by assumption continuous in t and furthermore $\mathcal{T}_0 = \text{id}$,

$$M_\epsilon = \int_0^\epsilon \mathcal{T}_s ds \quad (1.49)$$

is also invertible for sufficiently small $\epsilon > 0$. We now express \mathcal{T}_t as an integral, which shows that it is differentiable. To this end, note that

$$\begin{aligned} \mathcal{T}_t &= M_\epsilon^{-1} M_\epsilon \mathcal{T}_t = M_\epsilon^{-1} \int_0^\epsilon \mathcal{T}_{s+t} ds \\ &= M_\epsilon^{-1} \int_t^{t+\epsilon} \mathcal{T}_s ds = M_\epsilon^{-1} (M_{t+\epsilon} - M_t). \end{aligned} \quad (1.50)$$

Hence \mathcal{T}_t is differentiable and we define the generator as $\lim_{\epsilon \rightarrow 0} \|\mathcal{L}(\rho) - \epsilon^{-1} (\mathcal{T}_\epsilon(\rho) - \rho)\|_{tr}$ for all $\rho \in \mathcal{S}^+(\mathbb{C})$. We therefore have, that \mathcal{T}_t is of the desired form and hence satisfies the differential equation $\partial_t \mathcal{T}_t = \mathcal{L} \mathcal{T}_t$. \square

Let us now consider the requirements the generator \mathcal{L} needs to satisfy to ensure that the generated semi-group is actually completely positive. We need to make use of the two following propositions which give one general criteria to check for complete positivity.

Proposition 21 (Conditional positivity). *Let $\mathcal{L} : \mathcal{M}_d \rightarrow \mathcal{M}_d$ denote a linear map. Then the following two statements are equivalent:*

1. *There is a completely-positive map $\phi : \mathcal{M}_d \rightarrow \mathcal{M}_d$ and some matrix $\kappa \in \mathcal{M}_d$, so that*

$$\mathcal{L}(\rho) = \phi(\rho) - \kappa\rho - \rho\kappa^\dagger. \quad (1.51)$$

2. *We denote by $|\Omega\rangle$ a maximally entangled state in $\mathbb{C}^d \otimes \mathbb{C}^d$. Then L is Hermiticity preserving and we have*

$$P(\mathcal{L} \otimes \text{id})(|\Omega\rangle\langle\Omega|)P \geq 0, \quad (1.52)$$

where $P = \mathbb{1} - |\Omega\rangle\langle\Omega|$ is the projection on the orthogonal complement of the maximally entangled state.

PROOF: We prove that 1. \rightarrow 2. Upon inserting (1.51) into (1.52) we get that the condition reduces to $P(\mathcal{L} \otimes \text{id})(|\Omega\rangle\langle\Omega|)P = P(\phi \otimes \text{id})(|\Omega\rangle\langle\Omega|)P$, since $P|\Omega\rangle = 0$. This is of course positive since ϕ is completely positive. Moreover, \mathcal{L} is Hermiticity preserving. Conversely 2 follows from 1; since \mathcal{L} is Hermiticity preserving, we have that $\omega \equiv (\mathcal{L} \otimes \text{id})(|\Omega\rangle\langle\Omega|)$ is Hermitian. But since $P\tau P \geq 0$ we can write $\omega = Q - |\psi\rangle\langle\Omega| - |\Omega\rangle\langle\psi|$ where $Q \geq 0$ has, when written in a basis containing $|\Omega\rangle$, non-zero entries only in columns and rows orthogonal to $|\Omega\rangle$. $|\psi\rangle$ then contains all elements of ω in the column and row corresponding to $|\Omega\rangle$. By Lemma 11, we can now identify $Q = (\phi \otimes \text{id})(|\Omega\rangle\langle\Omega|)$ and $(\mathbb{1} \otimes \kappa)(|\Omega\rangle) = |\psi\rangle$. \square

The second proposition links the *conditional positivity* of the generator \mathcal{L} to the complete positivity of \mathcal{T}_t .

Proposition 22 (Completely positive dynamical semi-groups). *Consider a family of linear maps $\mathcal{T}_t : \mathcal{M}_d \rightarrow \mathcal{M}_d$ for $t \in [0, \infty)$, then the following two statements are equivalent:*

1. *\mathcal{T}_t is a dynamical semi-group of completely-positive maps which is continuous in t .*
2. *We can write $\mathcal{T}_t = \exp(\mathcal{L}t)$ for some conditionally completely-positive map (cf. prop. 21) $\mathcal{L} : \mathcal{M}_d \rightarrow \mathcal{M}_d$.*

PROOF: The first statement implies the second statement: By Proposition 20 we know that the semi-group can be written as $\exp(\mathcal{L}t)$. \mathcal{T}_t is by assumption completely positive and we consider the expansion

$$0 \leq (\exp(\mathcal{L}t))(|\Omega\rangle\langle\Omega|) = |\Omega\rangle\langle\Omega| + t(\mathcal{L} \otimes \text{id})(|\Omega\rangle\langle\Omega|) + \mathcal{O}(t^2) \quad (1.53)$$

Applying on both sides the projector $P = \mathbb{1} - |\Omega\rangle\langle\Omega|$, dividing by t and taking the limit $t \rightarrow 0$, we see that $P(\mathcal{L} \otimes \text{id})(|\Omega\rangle\langle\Omega|)P$ is Hermiticity preserving and positive. Conversely we can see that the second statement implies the first by observing that $\exp(\mathcal{L}t)$ is a dynamical

semi-group. The complete positivity can be seen when we decompose the generator into two parts $L = \phi + \phi_\kappa$, where $\phi_\kappa(\rho) = -\kappa\rho - \rho\kappa$. From the Lie-Trotter formula we get that

$$\exp(\mathcal{L}t) = \lim_{n \rightarrow \infty} (\exp(t/n\phi) \exp(t/n\phi_\kappa))^n. \quad (1.54)$$

Since concatenations of completely-positive maps are again completely positive, it is sufficient to show that both $\exp(t/n\phi)$ and $\exp(t/n\phi_\kappa)$ are completely positive. The complete positivity of $\exp(t/n\phi)$ follows from that of ϕ by Taylor expansion. For $\exp(t/n\phi_\kappa)$ we invoke the matrix representation of ϕ_κ in the natural basis $|i\rangle |j\rangle$. This yields $\hat{\phi}_\kappa = -\kappa \otimes \mathbb{1} - \mathbb{1} \otimes \kappa$ and thus we get $\exp(t/n\hat{\phi}_\kappa) = A \otimes \bar{A}$ where $A = \exp(-t/n\kappa)$. Hence $\exp(t/n\phi_\kappa)$ has a Kraus representation with the Kraus operator K and is thereby completely positive. \square

As we already noted, \mathcal{L} is the generator of the semi-group. Therefore, if we consider a time dependent density matrix $\rho(t) = \mathcal{T}_t\rho(0)$, we can interpret \mathcal{T}_t as the integrated form of a Markovian quantum master equation, which is referred to as the Lindblad equation

$$\partial_t \rho = \mathcal{L}(\rho). \quad (1.55)$$

We now proceed to determine the structure of the generator of a completely-positive dynamical semi-group. This Theorem was proved by Lindblad in [32] for infinite dimensional Hilbert spaces. However, we will only prove the form of the generator for finite dimensional spaces [33].

Theorem 23 (Lindblad generator). *A linear operator $\mathcal{L} : \mathcal{M}_d \rightarrow \mathcal{M}_d$ is the generator of a completely-positive dynamical semi-group on \mathcal{M}_d , if and only if it can be expressed as*

$$\mathcal{L}(\rho) = -i[H, \rho] + \sum_k \gamma_k \left(L_k \rho L_k^\dagger - \frac{1}{2} \left\{ L_k^\dagger L_k, \rho \right\}_+ \right) \quad (1.56)$$

where $H = H^\dagger$ is a Hermitian operator (the Hamiltonian) and $L_k \in \mathcal{M}(\mathbb{C})$ are arbitrary matrices called Lindblad operators. Moreover, the rates γ_k have to be positive $\gamma_k \geq 0$.

PROOF: We start by first proving the "if" part. Given a completely-positive semi-group, we can write in Kraus representation $\mathcal{T}_t(\rho) = \sum_\alpha A_\alpha(t)\rho A_\alpha^\dagger(t)$. We denote by $\{F_\alpha\}_{\alpha=0 \dots d^2-1}$ an orthonormal operator basis of \mathcal{M}_d that contains $F_0 = \sqrt{1/N}\mathbb{1}$. Expanding the Kraus-operators $K_\alpha = \sum_i \langle F_i | K_\alpha \rangle_{HS} F_i$ in this basis we obtain $\mathcal{T}_t(\rho) = \sum_{ij} c_{ij}(t) F_i \rho F_j^\dagger$. Note, that the $c_{ij}(t) = \sum_\alpha \langle F_i | A_\alpha \rangle_{HS} \overline{\langle F_j | A_\alpha \rangle_{HS}}$ are the entries of a positive matrix c , since for all vectors $|v\rangle = \sum_i v_i |i\rangle$ in \mathbb{C}^{d^2} we have that $\langle v | c | v \rangle = \sum_{ij} \sum_\alpha \bar{v}_i v_j \sum_\alpha \langle F_i | A_\alpha \rangle_{HS} \overline{\langle F_j | A_\alpha \rangle_{HS}} = \sum_\alpha \left\| \sum_i v_i F_i A_\alpha \right\|_2^2 \geq 0$. Now, since we have a continuous semi-group we know it is differentiable due to Proposition 20. So we obtain

$$\mathcal{L}(\rho) = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} (\mathcal{T}_\epsilon(\rho) - \rho) = \frac{1}{N} a_{00} \rho + \frac{1}{\sqrt{N}} \sum_{i=1}^{d^2-1} a_{i0} F_i \rho + a_{0i} \rho F_i^\dagger + \sum_{i,j=1}^{d^2-1} a_{ij} F_i \rho F_j^\dagger, \quad (1.57)$$

where we have defined the following limits

$$a_{00} = \lim_{\epsilon \rightarrow 0} \frac{c_{00}(\epsilon) - N}{\epsilon} \quad \text{and} \quad a_{ij} = \lim_{\epsilon \rightarrow 0} \frac{c_{ij}(\epsilon)}{\epsilon} = \bar{a}_{ji} \quad \text{for} \quad (i, j) \neq (0, 0). \quad (1.58)$$

If we now define the operators

$$F = \frac{1}{\sqrt{N}} \sum_{i=1}^{d^2-1} a_{i0} F_i, \quad G = \frac{1}{1N} a_{00} \mathbb{1} + \frac{1}{2} (F + F^\dagger) \quad \text{and} \quad H = \frac{1}{2i} (F - F^\dagger), \quad (1.59)$$

we obtain

$$\mathcal{L}(\rho) = -i[H, \rho] + \{G, \rho\} + \sum_{i,j=1}^{d^2-1} a_{ij} F_i \rho F_j^\dagger, \quad (1.60)$$

for which by trace preservation $\text{tr}[\mathcal{L}(\rho)] = 0$ we have that $G = -1/2 \sum_{i,j=1}^{d^2-1} a_{ij} F_i^\dagger F_j$. Now, since the a_{ij} were derived from the c_{ij} they too comprise the entries of a positive matrix a which can be diagonalized with some unitary $UaU^\dagger = \text{diag}(\gamma_1, \dots, \gamma_{d^2}) \geq 0$. If we denote by u_{ij} the entries of the unitary matrix and define the Lindblad operators via $F_i = \sum_k u_{ik} L_k$, we obtain (1.56).

Conversely we show the "only if" part of the Theorem by observing that we only need to show conditional complete positivity (cf. Proposition 22) of the generator (1.56). We have that

$$\begin{aligned} \langle \psi | P(\mathcal{L} \otimes \mathbb{1})(|\Omega\rangle\langle\Omega|) P | \psi \rangle &\geq \sum_k \gamma_k \langle \psi | P(L_k \otimes \mathbb{1}) |\Omega\rangle\langle\Omega| (L_k^\dagger \otimes \mathbb{1}) P | \psi \rangle \\ &= \sum_k \gamma_k |\langle \psi | P(L_k \otimes \mathbb{1}) |\Omega\rangle|^2 \geq 0, \end{aligned} \quad (1.61)$$

where we made use of the fact that $P|\Omega\rangle = 0$ and all $\gamma_k \geq 0$. \square

One may call $-i[H, \cdot]$ the "Hamiltonian" part of the generator, which generates the coherent evolution of the state, and the remaining summands $\gamma_k(L_k \cdot L_k - 1/2\{L_k^\dagger L_k, \cdot\})$ its "dissipative parts". That can, depending on the model, describe the dissipation or the damping of the dynamics. We have chosen a rather formal introduction of the generator in this section. The generator can, however, be derived from a physical picture in terms of perturbation theory as well. The general starting point is the open-system description, cf. Lemma 13, of the dynamics [36, 34]. One can always obtain the dynamics of an open system by considering the closed dynamics of a larger system described by some Hamiltonian $H_{tot} = H_{sys} \otimes \mathbb{1} + \mathbb{1} \otimes H_{bath} + H_{int}$. This Hamiltonian has the system evolution H_{sys} separated from the bath dynamics H_{bath} and the interaction between both parties is only via the Hamiltonian H_{int} . The system dynamics of the system state ρ_{sys} can under very restrictive circumstances be derived by taking the partial trace over the bath degrees of freedom of the von Neuman equation $\partial_t \rho_{sys} = \text{tr}_{bath} [[H_{tot}, \rho_{sys} \otimes \rho_{bath}]]$. One obtains a closed equation for the system evolution by the Nakajima Zwanzig [37, 38] projection operator technique. However, in general this approach will not give rise to a Lindblad type master equation, since for many realistic physical systems the dynamics of the physical subsystem are non-Markovian. Nevertheless, a

Markovian approximation to the dynamics can often be made for certain systems. For these systems, a quantum master equation can be obtained perturbatively [36, 34].

Recall, that the fact that we are dealing with a family of trace-preserving completely-positive maps $\mathcal{T}_t = \exp(\mathcal{L}t)$ leads to several properties of the generator. Since the map \mathcal{T}_t is trace preserving we have that the trace of $\text{tr}[\mathcal{L}(\rho)] = 0$ for all ρ . We can write alternatively that the dual has $\mathbb{1}$ as a fixed point $\mathcal{L}^*(\mathbb{1}) = 0$. This can be checked directly when considering the form (23). Furthermore, since \mathcal{T}_t is a tcp - map, we know that the the real part of the spectrum of \mathcal{L} cannot be positive $\text{Re}(\lambda_i(\mathcal{L})) \leq 0$ and that at least one eigenvalues $\lambda_0 = 0$ exists.

1.6 Matrix product states

In this section we will briefly introduce the matrix product state representation (MPS) for pure multipartite states [5, 6]. We will consider a representation of states on some Hilbertspace $\mathcal{H} = \mathbb{C}^d \otimes \dots \otimes \mathbb{C}^d$ that have a bounded amount of entanglement. We will see in chapter 5, how these states can be generalized to so-called matrix product operators and by that to mixed states. Matrix product states prove utterly useful in the simulation of quantum many-body systems and are intimately connected to the density matrix renormalization group (DMRG) algorithm [39, 40]. The notorious complexity of quantum many-body systems stems to a great extend from the exponential growth of the underlying Hilbert space. If we consider a system comprised of N systems of dimension d we have $\dim(\mathcal{H}) = d^N$ and we need exponentially many parameters just to write down the state. Matrix product states, however, have only a polynomial number of parameters and explore only a small subset of the larger Hilbertspace. Expectation values with respect to these states can be computed efficiently. Furthermore, MPS also give great insight to the general entanglement structure present in strongly correlated quantum system and are in this regard interesting in their own right. It can be argued that for several physical systems MPS approximate the ground states of the Hamiltonian faithfully [41], namely those with only local interactions. In this section we will only focus on a brief introduction, as a faithful account of the work done in this field would go beyond the scope of this chapter. A good introduction to these states and their higher dimensional generalizations can be found in [3, 4]. We will consider pure states $|\psi_N\rangle \in \mathbb{C}^{d^N}$ characterizing a ‘one-dimensional’ chain of N local d - level systems. The most general construction can be given in the valence bond picture. Consider a ring of N sites labeled by s and assign two virtual spins of dimension D_s to each of them. Assume that every pair of neighboring D_s -level systems share an (unnormalized) maximally entangled state $|\tilde{\Omega}\rangle = \sum_{k=1}^{D_s} |kk\rangle$. The last site N and the first site 1 are also connected by a maximally entangled state. Then apply the projection

$$\mathbf{A}^{[s]} = \sum_{i=1}^d \sum_{\alpha,\beta=1}^{D_s} \left[A_{i_s}^{[s]} \right]_{\alpha\beta} |i\rangle \langle\alpha\beta|. \quad (1.62)$$

to each of the N -sites. This operation projects the ‘virtual’ spins at each side on a single physical spin. Here, the Greek indices refer to the virtual D_s -level system and the d matrices $A_{i_s}^{[s]}$ are $D_s \times D_{s+1}$ dimensional. This construction is depicted in Fig. 1.1.

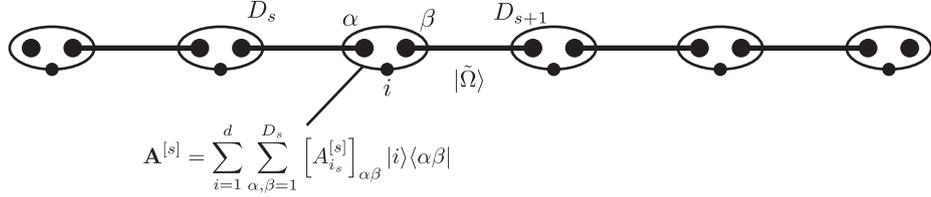


Figure 1.1: Construction of a matrix product states. The virtual spins at each site are depicted by the two larger dots and a maximally entangled state is represented by a line between different sites.

The resulting multipartite state will then have the form

$$|\psi_N\rangle = \sum_{i_1, \dots, i_N=1}^d \text{tr} [A_{i_1}^{[1]} \dots A_{i_N}^{[N]}] |i_1, \dots, i_N\rangle. \quad (1.63)$$

and are called matrix product states. In fact, it is shown in [42] that every state in the space \mathbb{C}^{d^N} can be cast into this form by making use of subsequent singular value decomposition between all bi-partitions along the chain. Albeit with matrices $A_{i_s}^{[s]}$ with a dimension $D = \max_s D_s$ exponential in the system size. It is immediately clear from the above valence bond construction, in which one interprets the bond-state $|\Omega\rangle$ as a resource state that mediates the quantum correlations between the sites, that the amount of entanglement these states carry is determined by the bond - dimension D . Hence, the advantage of the form of these states is that the entanglement can be bounded by choosing a lower bond-dimension D which makes the states become tractable for a classical computer. It is easy to see that all expectation values of product operators can be computed by multiplying N matrices of dimension at most D , since

$$\begin{aligned} \langle \psi_N | \bigotimes_{s=1}^N O_s | \psi_N \rangle &= \text{tr} [E_{O_1}^{[1]} E_{O_2}^{[2]}, \dots, E_{O_N}^{[N]}], \quad \text{with} \\ E_{O_s}^{[s]} &= \sum_{i, j=1}^d \langle i | O_s | j \rangle \bar{A}_i^s \otimes A_j^s, \end{aligned} \quad (1.64)$$

where the O_s are local operators acting only on a space \mathbb{C}^d at a single site s . It is because the expectation values of simple operators can be computed efficiently that one can use MPS as a variational ansatz for ground states of local Hamiltonians $H = \sum_k h_k$. The ground state is approximated by the minimum of the Rayleigh - Ritz coefficient

$$E \lesssim \min_{\{A\}} \frac{\langle \psi_N | H | \psi_N \rangle}{\langle \psi_N | \psi_N \rangle}, \quad (1.65)$$

which can be found by sweeping through all sites and optimizing the matrices $A_{i_k}^k$ locally. This algorithm is known as the variational formulation of DMRG, i.e. density matrix renormalization algorithm [39, 40].

In the construction (1.63) of the MPS we have placed an entangled bond between the first and the last site, so this state can be seen as a state with periodic boundary conditions. It is, however, also possible to leave this bond open, which corresponds to choosing bond - dimensions $D_1 = D_N = 1$ in the equation (1.63). In this case we speak of a MPS with open boundary conditions.

Gauge transformations We note, that there is no unique correspondence between the matrices $A_{i_s}^{[s]}$ and the state $|\psi_N\rangle$ they comprise. This can be seen easily by considering the following transformation. Consider two adjacent sites with the corresponding sets of matrices $A_{i_s}^{[s]}$ and $A_{i_{s+1}}^{[s+1]}$. We call a transformation of the A s on the virtual level, between sites s and $s + 1$, that leaves the overall state invariant, a *gauge transformation*. We can insert a partition of unity $\mathbb{1} = X_s^{-1} X_s$ so that

$$A_{i_s}^{[s]} A_{i_{s+1}}^{[s+1]} = A_{i_s}^{[s]} X_s^{-1} X_s A_{i_{s+1}}^{[s+1]} \equiv \tilde{A}_{i_s}^{[s]} \tilde{A}_{i_{s+1}}^{[s+1]}. \quad (1.66)$$

So in general we can see that the transformation

$$\tilde{A}_{i_s}^{[s]} = X_{s-1} A_{i_s}^{[s]} X_s^{-1} \quad (1.67)$$

leaves the overall state $|\psi_N\rangle$ invariant and the state is described now by the matrices $\tilde{A}_{i_s}^{[s]}$. This property is of great use when considering MPS with open boundary conditions. For open boundary conditions the state can be cast into its canonical form [42, 3, 4]. To see this, let us write an arbitrary MPS that is comprised of some matrices B with open boundary conditions as,

$$|\psi_N\rangle = \sum_{i_1, \dots, i_N=1}^d B_{i_1}^{[1]} \dots B_{i_N}^{[N]} |i_1, \dots, i_N\rangle, \quad (1.68)$$

where now $B_{i_1}^{[1]}$ and $B_{i_N}^{[N]}$ are $1 \times D_2$ and $D_N \times 1$ dimensional respectively. To construct the normal form consider the following procedure. We start, for instance, at site N and proceed to the left of the chain. For each set of matrices $B_{i_s}^{[s]}$ we group the indices β and i_s together so that we now write $[\mathcal{B}^s]_{\alpha, (i_s \beta)} \equiv [B_{i_s}^{[s]}]_{\alpha, \beta}$ and perform a singular value decomposition [17] on the larger matrix \mathcal{B}^s so that

$$[\mathcal{B}^s]_{\alpha, (i_s \beta)} = \sum_{\alpha', \alpha''} U_{\alpha, \alpha'} \Sigma_{\alpha', \alpha''}^s \bar{V}_{\alpha'', (i_s \beta)}. \quad (1.69)$$

Here Σ denotes a diagonal matrix of the singular values and U and V are isometries that arise from the standard singular value decomposition. This decomposition is depicted in Fig . 1.2(a). We proceed now by defining the set of matrices $[A_{i_s}^{[s]}]_{\alpha, \beta} = \bar{V}_{\alpha'', (i_s \beta)}$ and absorb the isometry U acting only on the virtual indices into the new $\tilde{B}_{i_{s-1}}^{[s-1]} = B_{i_{s-1}}^{[s-1]} U$. We now repeat this procedure for the site $s - 1$ until we reach the end of the chain.

This leads to a decomposition of the state that is of the form

$$|\psi_N\rangle = \sum_{i_1, \dots, i_N=1}^d A_{i_1}^{[1]} \Sigma^1 \dots \Sigma^N A_{i_N}^{[N]} |i_1, \dots, i_N\rangle \quad (1.70)$$

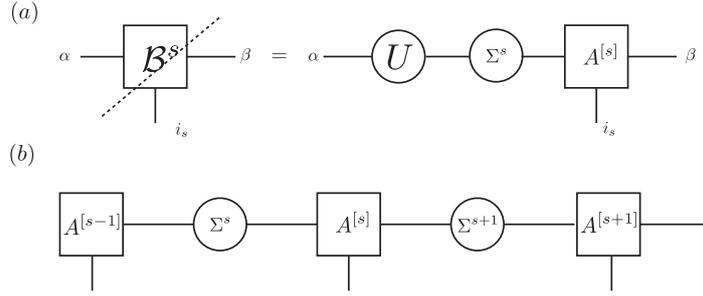


Figure 1.2: Fig (a): Depiction of the grouping of indices in the tensor $[B_i]_{\alpha\beta}$ to which the singular value decomposition is applied. Fig (b): Graphical representation of the normal form.

This form can be conveniently depicted as in Fig . 1.2(b). Note that since V is an isometry, the matrices A_i satisfy the constraint

$$\sum_{i_s} A_{i_s}^{[s]} A_{i_s}^{[s]\dagger} = \mathbb{1}_{D_s}. \quad (1.71)$$

This is the normal form of the MPS as it was given in [42]. As already mentioned, it is possible to arrive at the same expression by starting with a general multi-partite quantum states and performing singular value decompositions along every bi-partition. We see that the $\Sigma_{\alpha,\beta} = \sigma_\alpha \delta_{\alpha\beta}$ - matrices correspond to the Schmidt coefficients along each bi-partition.

Matrix product states and completely positive maps We would like to point out a close and interesting relationship between translationally invariant matrix product states and completely positive maps. A MPS is said to be translationally invariant when all matrices $A_{i_1}^{[1]} = \dots = A_{i_N}^{[N]} = A_i$ coincide. When we take a closer look at Eq. (1.64) and make the choice $O = \mathbb{1}$, we see that the matrix $E_{\mathbb{1}}$ reduces to $E_{\mathbb{1}} = \sum \bar{A}_i \otimes A_i$. We note, that due this structure, $E_{\mathbb{1}}$ can be seen as the matrix representation of a completely positive map (1.36) with Kraus operators A_i . We therefore associate a cp-map \mathcal{E} to any translationally invariant MPS so that

$$\mathcal{E}^*(X) = \sum_i A_i^\dagger X A_i \quad (1.72)$$

is acting as the dual cp-map on some operator X , that has support on the virtual system \mathbb{C}^D . We can understand the action of \mathcal{E} on the virtual level by the mapping $\langle \tilde{\Omega} | \mathcal{E}^*(X) = \langle X | E_{\mathbb{1}}$, where $\langle X | = \langle \tilde{\Omega} | \mathbb{1} \otimes X$. We have already discussed that such maps, if they are irreducible, always have an eigenvalue $\lambda_0 > 0$ and a positive fixed point $\sigma > 0$ that is associated with this eigenvalue. Note, that it is always possible to rescale the matrices A_i so that the largest eigenvalue can be taken as $\lambda_0 = 1$. One can show [6, 3] that the correlation length ξ of the MPS is related to the second largest eigenvalue λ_1 of the map \mathcal{E} . Here, for the sake of argument we make some simplifying assumptions. We assume that the eigenvalues $1 > \lambda_1 > \lambda_2 \dots$ are all

real and positive. Let us compute the correlation function in the thermodynamic limit $N \rightarrow \infty$:

$$\begin{aligned} & \lim_{N \rightarrow \infty} \langle \psi_N | O_1 O_n | \psi_N \rangle - \langle \psi_N | O_1 | \psi_N \rangle \langle \psi_N | O_n | \psi_N \rangle \\ &= \langle \phi_L^0 | E_{O_1} E_{\mathbb{1}}^n E_{O_n} | \phi_R^0 \rangle - \langle \phi_L^0 | E_{O_1} | \phi_R^0 \rangle \langle \phi_L^0 | E_{O_n} | \phi_R^0 \rangle, \end{aligned} \quad (1.73)$$

where $\langle \phi_L^k |$ and $|\phi_R^k \rangle$ denote the left and right eigenvectors corresponding to λ_k of $E_{\mathbb{1}}$ respectively. Now, when we consider the correlation function for $n \gg 1$ we can write

$$\begin{aligned} & \lim_{N \rightarrow \infty} \langle \psi_N | O_1 O_n | \psi_N \rangle - \langle \psi_N | O_1 | \psi_N \rangle \langle \psi_N | O_n | \psi_N \rangle \\ &= \langle \phi_L^0 | E_{O_1} \left[\sum_{k=0}^{D-1} \lambda_k^n |\phi_R^k \rangle \langle \phi_L^k | - |\phi_R^0 \rangle \langle \phi_L^0 | \right]^n E_{O_n} | \phi_R^0 \rangle \\ &\approx_{n \gg 1} \lambda_1^n \langle \phi_L^0 | E_{O_1} | \phi_R^1 \rangle \langle \phi_L^1 | E_{O_n} | \phi_R^0 \rangle = c e^{-\frac{n}{\xi}}. \end{aligned} \quad (1.74)$$

If we now define $c = \langle \phi_L^0 | E_{O_1} | \phi_R^1 \rangle \langle \phi_L^1 | E_{O_n} | \phi_R^0 \rangle$, we see that the correlation length is related to the second largest eigenvalue via $\xi = -1/\ln(\lambda_1)$. The only purpose of this exercise was to point out that there exists an interesting relationship between the correlations of one dimensional multi-partite state, with a bounded amount of correlations, and the convergence properties of quantum Markov chains. There also exists a formulation of a Kadanoff blocking-type renormalization group [43] on the MPS, which can be understood as a quantum Markov chain on this level, as is shown in [44]. The fixed point MPS of the renormalization group is determined completely by the fixed point of the cp - map \mathcal{E} .

This formal relationship is no accident. In fact, an alternative construction of MPS, that was first proposed in [6], is based on this connection. Let us therefore discuss a different approach to the construction of MPS that is along the lines of [6] and actually leads to a way of generating multi-partite entangled states in an experiment [45].

Consider therefore the following scenario: We denote by $\mathcal{H}_A = \mathbb{C}^D$ a D -dimensional Hilbert space of some ancilla system, one could think for instance of an atom in a cavity, and another system with a d -dimensional Hilbert space $\mathcal{H}_B = \mathbb{C}^d$. For $d = 2$ we could think of a photon. Now imagine that we have a set of N “photons” that each pass through the cavity one by one and interact with the atom in the cavity one at a time, cf Fig. 1.3. This interaction generates a unitary evolution for each $U : \mathcal{H}_A \otimes \mathcal{H}_B \mapsto \mathcal{H}_A \otimes \mathcal{H}_B$. We assume that all the “photons”, i.e. in general the d - level systems, are initially prepared in the same state $|0\rangle$, before they enter the cavity and that the cavity is initially in the state $|\varphi_I\rangle$. Since all d - level systems are initially in the same state we disregard them at the input and only consider the isometry $V = U(\mathbb{1} \otimes |0\rangle)$. We now write for the isometry

$$V = \sum_{i=0}^{d-1} \sum_{\alpha, \beta=1}^D [V_i]_{\alpha, \beta} |\alpha, i\rangle \langle \beta|, \quad (1.75)$$

where the Greek indices label a basis in \mathcal{H}_A and the Latin indices label one in \mathcal{H}_B . The fact that we require V to be an isometry translates to the fact that $\sum_{i=0}^{d-1} V_i^\dagger V_i = \mathbb{1}_D$, where we

treat each of the V_i as a $D \times D$ matrix with entries $[V_i]_{\alpha,\beta}$. After the N “photons” have passed through the system we observe that the total system on the Hilbert space $\mathcal{H}_A \otimes \mathbb{C}^{d^N}$ is in the state $|\Psi\rangle = V^N \dots V^1 |\varphi_I\rangle$, where we labeled each of the different isometries by V^k . Suppose, we could tune the last interaction V^N in such a way that the complete state between the cavity, i.e. H_A , and the N particle state factorizes as $|\Psi\rangle = |\varphi_F\rangle \otimes |\psi_N\rangle$. Then we find the remaining N - particle system in the state

$$|\psi\rangle = \sum_{i_N, \dots, i_1} \langle \varphi_F | V_{i_N}^N \dots V_{i_1}^1 | \varphi_I \rangle |i_N \dots i_1\rangle. \quad (1.76)$$

We observe, that the state generated this way is nothing but a standard matrix product state as introduced previously.

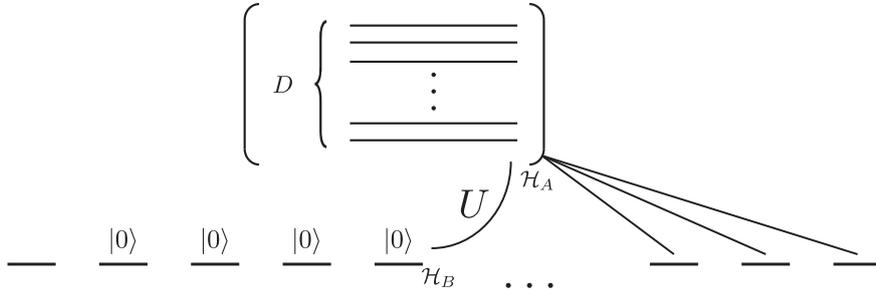


Figure 1.3: A larger quantum system with Hilbert space $\mathcal{H}_A = \mathbb{C}^D$ interacts with a sequence of N particles, each with state space $\mathcal{H}_B = \mathbb{C}^d$, one at a time via some unitary U . All the particles are initially in state $|0\rangle$.

We see that at each step we effectively implement a transformation on \mathcal{H}_A , which is of the form $\mathcal{E}(\rho_A^{k+1}) = \text{tr} [U \rho_A^k \otimes |0\rangle \langle 0| U^\dagger]$. Due to Theorem 13 it is clear, that we implement a tcp - map on the ancilla system each time a single particle passes through the system. The correlations between the individual particles can therefore be understood by the internal dynamics of the ancilla system. In chapter 4, we will introduce a class of classical multi-partite probability distributions that can be constructed in a similar fashion. Here, however, a classical stochastic Markov process will take the place of the quantum processes.

Chapter 2

Mixing time analysis of quantum Markov chains

Synopsis:

In this chapter, we introduce quantum versions of the χ^2 -divergence, provide a detailed analysis of their properties, and apply them to the investigation of mixing times of quantum Markov processes. An approach similar to the one presented in [46, 47, 48] for classical Markov chains is taken to bound the trace-distance from the steady state of a quantum processes. A strict spectral bound to the convergence rate can be given for time-discrete as well as for time-continuous quantum Markov processes. Furthermore, the contractive behavior of the χ^2 -divergence under the action of a completely positive map is investigated and contrasted to the contraction of the trace norm. In this context we analyze different versions of quantum detailed balance and, finally, give a geometric conductance bound to the convergence rate for unital quantum Markov processes.

Based on:

K. Temme, M. J. Kastoryano, M. B. Ruskai, M. M. Wolf and F. Verstraete,
J. Math. Phys. (accepted), *e-print* arXiv: 1005.2358, (2010).

2.1 Mixing time in Markov chains

The mixing time of a classical Markov chain is the time it takes for the chain to be close to its steady state distribution, starting from an arbitrary initial state. The ability to bound the mixing time is important, for example in the field of computer science, where the bound can be used to give an estimate for the running time of some probabilistic algorithm such as the Monte Carlo algorithm. The mixing time for a classical Markov process P_{ij} , with $\sum_i P_{ij} = 1$ on the space of probability measures $L^1(\Omega)$ is commonly defined in terms of the one norm, $\|p\|_1 = \sum_i |p_i|$. Let π denote the fixed point of the classical Markov process, i.e. $P\pi = \pi$, then the mixing time is defined as:

$$t_{mix}(\epsilon) = \min \{n \mid \forall q \in \mathcal{S}, \|P^n q - \pi\|_1 < \epsilon\}. \quad (2.1)$$

A large set of tools has emerged over the years that allows to investigate the convergence rate of classical Markov chains [22]. One of the most prominent approaches [46, 47, 48] to bounding the mixing time of a Markov chain is based on the χ^2 -divergence [49]. This divergence is defined for two probability distributions $p, q \in L^1(\Omega)$ as:

$$\chi^2(p, q) = \sum_i \frac{(p_i - q_i)^2}{q_i}. \quad (2.2)$$

The usefulness of the χ^2 -divergence for finding bounds to the mixing time of classical Markov chains arises from the fact that it serves as an upper bound to the one norm difference between two probability distributions

$$\|p - q\|_1^2 \leq \chi^2(p, q) \quad (2.3)$$

and allows for an easier access to the spectral properties of the Markov chain. The χ^2 -divergence is intimately related to the Kullback-Leibler divergence, or relative entropy, $H(p, q) = \sum_i p_i (\log p_i - \log q_i)$. In fact, it can be obtained directly from the relative entropy as the approximating quadratic form, i.e. as the Hessian, of the latter:

$$\chi^2(p, q) = -\frac{\partial^2}{\partial \alpha \partial \beta} H(q + \alpha(p - q), q + \beta(p - q)) \Big|_{\alpha=\beta=0}. \quad (2.4)$$

The χ^2 divergence was first introduced by Karl Pearson in the context of statistical inference tests, the most widely used of which is the "Pearson's χ^2 test". Its computational simplicity and its clear relation to other distance measures have made it one of the most studied divergence measures in the literature.

In this chapter, we find convergence bounds for arbitrary quantum Markov chains, also called quantum channels, with a technique that can be seen as a generalization of the work of [46, 47, 48] to non-commutative probability spaces. A prototypical example of mixing time in physics is the decoherence time of the underlying quantum process, i.e. the time in which quantum states decohere to an (often classical) mixture given a specific underlying noise model. The ability to bound the mixing time for quantum processes also turns out to be relevant when

one seeks to give bounds on the runtime of quantum algorithms that are based on quantum Markov chains [14, 50]. Other applications of such bounds can be found in the framework of matrix product states [6, 5], where the correlation length of the quantum state is connected to the convergence of the corresponding transfer operator that can be interpreted as a quantum channel. In this chapter, we introduce the mathematical framework necessary to extend the classical mixing time results to the quantum setting. In particular, we introduce a new divergence measure - the quantum χ^2 -divergence - for quantum states and use it to obtain some basic convergence bounds that mirror existing classical ones. Furthermore, we extend the classical concept of detailed balance to the quantum setting and discuss its relevance in general terms.

2.2 The quantum χ^2 -divergence

We want to define a generalization of the classical χ^2 -divergence to the case when we are working on spaces with non-commuting density matrices. We shall require that any generalization to the setting of density matrices satisfies the condition that when the inputs are diagonal, the classical χ^2 -divergence is recovered. The first observation we make, reading straight off from (2.2), is that the classical χ^2 -divergence can be seen as an inner product on the probability space weighted with the inversion of the distribution q_i . Due to the non-commutative nature of density matrices there is no unique generalization of this inversion. Consider for instance a generalization for two density matrices $\rho, \sigma \in S_d$, where for now we assume σ to be full rank, that is given by

$$\chi_\alpha^2(\rho, \sigma) = \text{tr} [(\rho - \sigma)\sigma^{-\alpha}(\rho - \sigma)\sigma^{\alpha-1}] = \text{tr} [\rho\sigma^{-\alpha}\rho\sigma^{\alpha-1}] - 1. \quad (2.5)$$

This gives rise to an entire family of χ^2 -divergences with (as we see below) special properties, for every $\alpha \in [0, 1]$. The natural question of whether there exists a classification of all possible inversions of σ , was investigated in a series of papers by Morozova and Chentsov [51] Petz [52, 53, 54], in the context of information geometry. They considered the characterization of monotone Riemannian metrics on matrix spaces. Their general definition is based on the modular operator formalism of Araki [55, 56], which we will also consider here. In order to classify the valid inversions, we first need to define the following set of functions, each of which gives rise to a possible inversion:

$$\mathcal{K} = \{k \mid -k \text{ is operator monotone, } k(w^{-1}) = wk(w), \text{ and } k(1) = 1\}. \quad (2.6)$$

Now, we define left and right multiplication operators as $\mathcal{L}_Y(X) = YX$ and $\mathcal{R}_Y(X) = XY$ respectively. The *modular operator* is defined as

$$\Delta_{\rho, \sigma} = \mathcal{L}_\rho \mathcal{R}_\sigma^{-1}, \quad (2.7)$$

for all $\rho, \sigma \in \mathcal{S}_d$, $\sigma > 0$. Note, that \mathcal{R}_σ and \mathcal{L}_ρ commute and inherit hermicity and positivity from ρ, σ . The above should be read as follows: acting on some $A \in \mathcal{M}_d$, $\Delta_{\rho, \sigma}(A) = \rho A \sigma^{-1}$. When manipulating the modular operators it is often convenient to write them in matrix form, in which case, they read: $\hat{\Delta}_{\rho, \sigma} |A\rangle = \rho \otimes \bar{\sigma}^{-1} |A\rangle$, where $|A\rangle = A \otimes \mathbb{1} |\tilde{\Omega}\rangle$, and $|\tilde{\Omega}\rangle = \sum_{i=1}^d |ii\rangle$ corresponds to \sqrt{d} times the maximally mixed state. This formalism gives rise to a more general quantum χ^2 -divergence.

Definition 24. For $\rho, \sigma \in \mathcal{S}_d$, and $k \in \mathcal{K}$ we define the the quantum χ^2 -divergence

$$\chi_k^2(\rho, \sigma) = \left\langle \rho - \sigma, \Omega_\sigma^k(\rho - \sigma) \right\rangle, \quad (2.8)$$

when $\text{supp}(\rho) \subseteq \text{supp}(\sigma)$, and infinity otherwise. Here, \langle, \rangle denotes the standard Hilbert-Schmidt scalar product. The inversion inversion of σ is defined only when $\text{supp}(\rho) \subseteq \text{supp}(\sigma)$, and given by

$$\Omega_\sigma^k = \mathcal{R}_\sigma^{-1} k(\Delta_{\sigma, \sigma}). \quad (2.9)$$

Families of divergences The functions $k_\alpha(w) = \frac{1}{2}(w^{-\alpha} + w^{\alpha-1})$ yield the family of χ_α^2 -divergences given in (2.5) which we call the mean α -divergences to distinguish them from the well-known family of Wigner-Yanase-Dyson (WYD) α -divergences, which we will discuss shortly along with several other families. Although we focus on the family (2.5), most of our results hold for any divergence given by (2.8) with $k \in \mathcal{K}$ with the exceptions of Theorem 37.

The most widely used family of divergences, often called α -divergence [57, Chapter 7], is associated with the functions

$$k_\alpha^{\text{WYD}}(w) = \frac{(1-w^\alpha)(1-w^{1-\alpha})}{\alpha(1-\alpha)(1-w)^2} \quad \text{for} \quad \alpha \in [-1, 2] \quad (2.10)$$

This family is sometimes called the WYD divergences, because it arises from an extension of the Wigner-Yanase-Dyson entropy [58, 59] associated with the (unsymmetrized) function $g(w) = \frac{1}{\alpha(1-\alpha)}(w - w^\alpha)$. In the limit $\alpha \rightarrow 1$ this yields [60] the familiar (asymmetric) relative entropy $H(\rho, \sigma) = \text{tr} \rho(\log \rho - \log \sigma)$ and Ω_P^{\log} given by (2.56). Like the family of divergences introduced here, the minimal WYD divergence occurs for $\alpha = 1/2$, it is convex in α , symmetric around $\alpha = 1/2$ and yields the maximal $\frac{1+w}{2w}$ when $\alpha = -1$ or 2. However, $\alpha = 1/2$ gives $k_{1/2}^{\text{WYD}}(w) = 4(1 + \sqrt{w})^{-2}$ which is quite different from $k_{1/2}^{\text{mean}}(w) = w^{-1/2}$. The WYD family is often studied only for $\alpha \in (0, 1)$; it was first observed by Hasegawa in [61] that it yields a monotone metric if and only if $\alpha \in [-1, 2]$.

The metrics associated with $k_\alpha^{\text{mean}}(w)$ and $k_\alpha^{\text{WYD}}(w)$ both give increasing families for $\alpha \geq \frac{1}{2}$ and both yield the maximal metric $k(w) = (1+w)(2w)$ for α the maximal values of 1 and 2 respectively. However, neither reduces to the minimal metric $k(w) = 2/(1+w)$. The measure $\delta(s-a)$ in (2.16) leads to the family $k_a(w) = \frac{(1+a)^2}{2} \frac{(1+w)}{(1+wa)(w+a)}$ for $a \in [0, 1]$ which reduces to the the maximal and minimal functions for $a = 0, 1$. However, this family is

neither increasing nor decreasing. Hansen [62] has found families of functions which increase monotonically from the smallest to the largest of which we mention only

$$k_a(w) = w^{-a} \left(\frac{1+w}{2} \right)^{2a-1} \quad \text{for } a \in [0, 1]. \quad (2.11)$$

2.2.1 Monotone Riemannian metrics and generalized relative entropies.

This definition of the χ^2 -divergence stems from the analysis of monotone Riemannian metrics. By Riemannian metric, we mean a positive definite bilinear form $M_\sigma(A, B)$ on the hermitian tangent hyperplane $\mathcal{TP} = \{A \in \mathcal{M}_d : A = A^\dagger, \text{tr}[A] = 0\}$. The metric is monotone if for all quantum channels $\mathcal{T} : \mathcal{M}_d \mapsto \mathcal{M}_d$, states $\sigma \in \mathcal{S}_d^+$ and $A \in \mathcal{TP}$, $M_{\mathcal{T}(\sigma)}(\mathcal{T}(A), \mathcal{T}(A)) \leq M_\sigma(A, A)$. Petz showed that there is a one-to-one correspondence between the above metrics and a special class of convex operator functions, which correspond to $1/k$ in our notation. He furthermore was able to relate several generalized relative entropies (which he defined much earlier [63] and referred to as quasi-entropies) to monotone Riemannian metrics [53, 54, 64]. The reverse implication, that every monotone Riemannian metric stems from a generalized relative entropy was first proved by Lesniewski and Ruskai [59]. Taking advantage of the well-known integral representations of operator monotone and convex functions [16] one can express the χ^2 -divergences as well as the relative entropies explicitly. We shall briefly repeat the key points of the analysis that are necessary for our understanding of the mixing-time and contraction analysis for tcp-maps.

We need to consider the class of functions \mathcal{G} by which we denote the set of continuous operator convex functions from \mathbb{R}^+ to \mathbb{R} that satisfy $g(1) = 0$. Note that these functions can all be classified in terms of the integral representation:

$$g(w) = a(w-1) + b(w-1)^2 + c \frac{(w-1)^2}{w} + \int_0^\infty \frac{(w-1)^2}{w+s} d\nu(s), \quad (2.12)$$

where $a, b, c > 0$ and the integral of the positive measure $d\nu(s)$ on $(0, \infty)$ is bounded. The generalized relative entropy for states $\rho, \sigma \in \mathcal{S}_d^+$ was first defined in [65, 66].

Definition 25. *Let $g \in \mathcal{G}$. The generalized quantum relative entropy is given by*

$$H_g(\rho, \sigma) = \text{tr}[\rho^{1/2} g(\Delta_{\sigma, \rho})(\rho^{1/2})] \quad (2.13)$$

when $\text{supp}(\rho) \subseteq \text{supp}(\sigma)$, and infinity otherwise, and where $\Delta_{\rho, \sigma}$ is again the modular operator.

We now recall without proof a Theorem [53, 54, 59] relating the relative entropy and the monotone Riemannian metric, mirroring the classical result (2.4):

Theorem 26. For every $k \in \mathcal{K}$, there is a $g \in \mathcal{G}$ such that for a given $\sigma \in \mathcal{S}_d$, and A, B hermitian traceless, we get:

$$\begin{aligned} M_\sigma^k(A, B) &= -\frac{\partial^2}{\partial\alpha\partial\beta} H_g(\sigma + \alpha A, \sigma + \beta B) \Big|_{\alpha=\beta=0} \\ &= \langle A, \Omega_\sigma^k(B) \rangle. \end{aligned} \quad (2.14)$$

and, k is related to g by

$$k(w) = \frac{g(w) + wg(w^{-1})}{(w-1)^2} \quad (2.15)$$

From this Theorem follows a convenient integral representation of the inversion Ω_σ^k , which is equivalent to (2.9) [59].

$$\Omega_\sigma^k = \int_0^\infty \left(\frac{1}{s\mathcal{R}_\sigma + \mathcal{L}_\sigma} + \frac{1}{\mathcal{R}_\sigma + s\mathcal{L}_\sigma} \right) N_g(s) ds, \quad (2.16)$$

where N_g denotes the singular measure $N_g(s)ds = (b_g + c_g)\delta(s)ds + d\nu_g(s)$. Note, that the relationship between k and g is not one-to-one. Indeed, by setting $\hat{g}(w) = wg(w^{-1})$, we get back the above relation. However, there is a one-to-one correspondence between each k and a symmetric $g_s(w) = g(w) + wg(w^{-1})$, and hence between each metric and a symmetric relative entropy.

Note that the α -subfamily of (2.5) has the associated symmetric relative entropy: $g^{sym}(x) = \frac{(1-w)^2}{2} (w^{\alpha-1} + w^{-\alpha})$, so that

$$H_\alpha^{sym}(\rho, \sigma) = \frac{1}{2}(H_\alpha(\rho, \sigma) + H_\alpha(\sigma, \rho)) \quad (2.17)$$

where,

$$H_\alpha(\rho, \sigma) = \text{tr}[\rho^{2-\alpha}\sigma^{\alpha-1} + \rho^{1+\alpha}\sigma^{-\alpha} - 2\rho^\alpha\sigma^{1-\alpha}].$$

The integral representation (2.16) of the inversion Ω_σ^k allows for a partial ordering of different monotone Riemannian metrics that follows from the set of inequalities:

$$\frac{2}{x+1} \leq \frac{1+s}{2} \left(\frac{1}{s+x} + \frac{1}{sx+1} \right) \leq \frac{x+1}{2x}. \quad (2.18)$$

for $s \in [0, 1]$, and $x \in \mathbb{R}^+$. We therefore see that there exists a partial ordering for the inversions, with a lowest and highest element in the hierarchy. The lowest element gives rise to the so called Bures metric. Thus,

$$\Omega_\sigma^{Bures} = 2(\mathcal{R}_\sigma + \mathcal{L}_\sigma)^{-1} \leq \Omega_\sigma^k \leq (\mathcal{L}_\sigma^{-1} + \mathcal{R}_\sigma^{-1})/2 = \Omega_\sigma^{\alpha=0} \quad (2.19)$$

The χ^2 -divergence is recovered from the metric upon setting $\chi_k^2(\rho, \sigma) \equiv M_\sigma^k(\rho - \sigma, \rho - \sigma)$. We are therefore left with a partial order for all possible χ^2 -divergences with a smallest and largest element according to,

$$\chi_{Bures}^2(\rho, \sigma) \leq \chi_k^2(\rho, \sigma) \leq \chi_{\alpha=0}^2(\rho, \sigma). \quad (2.20)$$

The defining attribute of the above set of metrics is their monotonicity under the action of quantum channels. This was first shown by Petz [54], and later a proof based on the integral representation of Ω_σ^k (2.16), and on Schwarz-type inequalities, was provided by Ruskai and Lesniewski in [59]. Due to its importance for the mixing time analysis we shall repeat it here.

Theorem 27. *For all $\sigma \in \mathcal{S}_d$, M_σ^k is monotone under the action of a quantum channel $\mathcal{T} : \mathcal{M}_d \rightarrow \mathcal{M}_d$ for all $k \in \mathcal{K}$ and $A \in \mathcal{M}_d$, i.e.*

$$M_\sigma^k(A, A) \geq M_{\mathcal{T}(\sigma)}^k(\mathcal{T}(A), \mathcal{T}(A)) \quad (2.21)$$

PROOF: The monotonicity follows immediately from the integral representation of the inversion Ω_σ^k in (2.16), and an argument proved in the following Theorem 28. \square

The proof of the contractivity of a general Riemannian metric is based on the following Theorem first proved in [59].

Theorem 28. *For a channel $\mathcal{T} : \mathcal{M}_d \rightarrow \mathcal{M}_d$, we have that,*

$$\begin{aligned} \operatorname{tr} \left[A^\dagger \frac{1}{\mathcal{R}_\sigma + s\mathcal{L}_\rho} A \right] &= \operatorname{tr} \left[\mathcal{T} \left(A^\dagger \frac{1}{\mathcal{R}_\sigma + s\mathcal{L}_\rho} A \right) \right] \\ &\geq \operatorname{tr} \left[\mathcal{T}(A)^\dagger \frac{1}{\mathcal{R}_{\mathcal{T}(\sigma)} + s\mathcal{L}_{\mathcal{T}(\rho)}} \mathcal{T}(A) \right]. \end{aligned} \quad (2.22)$$

PROOF: Let $\sigma > 0$, then $\operatorname{tr}[A^\dagger \sigma A] \geq 0$, and $\operatorname{tr}[A^\dagger A \sigma] \geq 0$ so that \mathcal{L}_σ as well as \mathcal{R}_σ are both positive semi definite super operators on the matrix space. Therefore we infer, that for a positive $\rho > 0$ the operator $\mathcal{R}_\sigma + s\mathcal{L}_\rho$ is also positive. We define a matrix $X = [\mathcal{R}_\sigma + s\mathcal{L}_\rho]^{-1/2}(A) + [\mathcal{R}_\sigma + s\mathcal{L}_\rho]^{1/2}T^*(A)$ and furthermore $B = [\mathcal{R}_{\mathcal{T}(\sigma)} + s\mathcal{L}_{\mathcal{T}(\rho)}]^{-1}\mathcal{T}(A)$. Since $\operatorname{tr}[X^\dagger X] \geq 0$, we have that

$$\operatorname{tr} \left[A^\dagger \frac{1}{\mathcal{R}_\sigma + s\mathcal{L}_\rho} A \right] - \operatorname{tr} \left[T^*(B^\dagger)A \right] - \operatorname{tr} \left[A^\dagger T^*(B) \right] + \operatorname{tr} \left[T^*(B^\dagger)[\mathcal{R}_\sigma + s\mathcal{L}_\rho]T^*(B) \right] \geq 0. \quad (2.23)$$

Furthermore note, that

$$-\operatorname{tr} \left[A^\dagger T^*(B) \right] - \operatorname{tr} \left[T^*(B^\dagger)A \right] = -2\operatorname{tr} \left[\mathcal{T}(A^\dagger) \frac{1}{\mathcal{R}_{\mathcal{T}(\sigma)} + s\mathcal{L}_{\mathcal{T}(\rho)}} \mathcal{T}(A) \right]. \quad (2.24)$$

It therefore suffices to show that we are able to bound the last term in (2.23) by the right side of the inequality (2.22). Note, that

$$\begin{aligned} \operatorname{tr} \left[T^*(B^\dagger)[\mathcal{R}_\sigma + s\mathcal{L}_\rho]T^*(B) \right] &= \operatorname{tr} \left[T^*(B^\dagger)T^*(B)\sigma + sT^*(B^\dagger)\rho T^*(B) \right] \\ &\leq \operatorname{tr} \left[T^*(B^\dagger B)\sigma + sT^*(BB^\dagger)\rho \right], \end{aligned} \quad (2.25)$$

since $\rho, \sigma > 0$ and due to the operator inequality $T^*(B^\dagger)T^*(B) \leq T^*(B^\dagger B)$. This inequality holds for any B since \mathcal{T} is a channel and by that trace preserving, hence $\mathcal{T}^*(\mathbb{1}) = \mathbb{1}$. With

$\text{tr} [\mathcal{T}^*(B^\dagger B)\sigma] = \text{tr} [B^\dagger B\mathcal{T}(\sigma)]$ we can write

$$\begin{aligned} & \text{tr} \left[\mathcal{T}^*(B^\dagger) [\mathcal{R}_\sigma + s\mathcal{L}_\rho] \mathcal{T}^*(B) \right] \leq \text{tr} \left[B^\dagger B\mathcal{T}(\sigma) + sB^\dagger B\mathcal{T}(\rho) \right] \\ & = \text{tr} \left[B^\dagger [\mathcal{R}_{\mathcal{T}(\sigma)} + s\mathcal{L}_{\mathcal{T}(\rho)}] B \right] = \text{tr} \left[B^\dagger \mathcal{T}(A) \right] = \text{tr} \left[\mathcal{T}(A^\dagger) \frac{1}{\mathcal{R}_{\mathcal{T}(\sigma)} + s\mathcal{L}_{\mathcal{T}(\rho)}} \mathcal{T}(A) \right]. \end{aligned} \quad (2.26)$$

□

2.2.2 Properties of the quantum χ^2 -divergence

The fact that the quantum χ_k^2 -divergence can be used to bound the mixing time lies in the following Lemma, that upper bounds the trace distance which is the relevant distance measure in the mixing time definition.

Lemma 29. *For every pair of density operators $\rho, \sigma \in \mathcal{S}_d$, we have that*

$$\|\rho - \sigma\|_{\text{tr}}^2 \leq \chi_k^2(\rho, \sigma) \quad (2.27)$$

PROOF: If the support of ρ is not contained in the support of σ , then the right hand side is ∞ . We can therefore assume w.l.o.g. that $\sigma > 0$ by restricting the analysis to the support space of σ . The trace norm $\|A\|_{\text{tr}}$ of some matrix $A \in \mathcal{M}_d$ can be expressed as [17] $\|A\|_{\text{tr}} = \max_{U \in U(d)} |\text{tr}[UA]|$, where the maximum is taken over all unitaries acting on the d -dimensional Hilbert space. Thus, for any inversion Ω_σ^k :

$$\begin{aligned} \|A\|_{\text{tr}}^2 &= \max_{U \in U(d)} |\text{tr}[UA]|^2 = \max_{U \in U(d)} \left| \text{tr} \left[U [\Omega_\sigma^k]^{-1/2} \circ [\Omega_\sigma^k]^{1/2}(A) \right] \right|^2 \\ &= \max_{U \in U(d)} \left| \text{tr} \left[[\Omega_\sigma^k]^{-1/2}(U) [\Omega_\sigma^k]^{1/2}(A) \right] \right|^2 \\ &\leq \text{tr} \left[A^\dagger \Omega_\sigma^k(A) \right] \max_{U \in U(d)} \text{tr} \left[U^\dagger [\Omega_\sigma^k]^{-1}(U) \right] \end{aligned} \quad (2.28)$$

Let us consider the Bures inversion given by $\Omega_\sigma^{\text{Bures}} = 2[\mathcal{L}_\sigma + \mathcal{R}_\sigma]^{-1}$. Clearly, its inverse is $[\Omega_\sigma^{\text{Bures}}]^{-1} = \frac{1}{2}[\mathcal{L}_\sigma + \mathcal{R}_\sigma]$. Therefore, for any unitary U ,

$$\text{tr} \left[U^\dagger [\Omega_\sigma^{\text{Bures}}]^{-1}(U) \right] = \frac{1}{2} \left(\text{tr}[U^\dagger \sigma U] + \text{tr}[U^\dagger U \sigma] \right) = 1. \quad (2.29)$$

Setting $A = \rho - \sigma$ and observing that $\chi_{\text{Bures}}^2 \leq \chi_k^2$ for all $k \in \mathcal{K}$ completes the proof. □

We are also able to bound the relative entropy in terms of the α -subfamily of χ^2 -divergences.

Lemma 30. *For every pair of density operators ρ and σ and every $\alpha \in (0, 1]$ we have that*

$$\chi_\alpha^2(\rho, \sigma) \geq S(\rho, \sigma), \quad (2.30)$$

where $S(\rho, \sigma) = \text{tr} \rho(\log \rho - \log \sigma)$ is the usual relative entropy.

PROOF: It was shown in [67] that for $\gamma \in (0, 1]$, the following holds:

$$S(\rho, \sigma) \leq \frac{1}{\gamma} (\text{tr} \rho^{1+\gamma} \sigma^{-\gamma} - 1) \quad (2.31)$$

Then consider,

$$\begin{aligned} \chi_\alpha^2(\rho, \sigma) - S(\rho, \sigma) &\geq \text{tr} \rho \sigma^{-1/2} \rho \sigma^{-1/2} - 2 \text{tr} \rho^{3/2} \sigma^{-1/2} + 1 \\ &= \text{tr} (\rho^{1/2} \sigma^{-1/2} \rho^{1/2} - \rho^{1/2})^2 \geq 0 \end{aligned} \quad (2.32)$$

where the first inequality comes from taking $\gamma = 1/2$ in (2.2.2), and $\alpha = 1/2$ for χ_α^2 , and the last line is obtained from rearranging terms. \square

Furthermore, we note that this subfamily also has a natural ordering.

Proposition 31. *For every $\rho, \sigma \in \mathcal{S}_d$, χ_α^2 is convex in α , and reaches a minimum for $\alpha = 1/2$.*

PROOF: First note that $\chi_{\alpha=0}^2(\rho, \sigma) = \chi_{\alpha=1}^2(\rho, \sigma)$. That the minimum is reached for $\alpha = 1/2$ follows directly from the Cauchy-Schwarz inequality. Applied to our problem we get

$$\begin{aligned} \text{tr} \left[\rho \sigma^{-1/2} \rho \sigma^{-1/2} \right]^2 &= \text{tr} \left[\rho \sigma^{(\alpha-1)/2} \sigma^{-\alpha/2} \rho \sigma^{(\alpha-1)/2} \sigma^{-\alpha/2} \right]^2 \\ &\leq \text{tr} \left[\rho \sigma^{-\alpha} \rho \sigma^{\alpha-1} \right]^2 \end{aligned} \quad (2.33)$$

To see convexity, consider the second partial derivative of χ_α^2 with respect to α :

$$\begin{aligned} \frac{\partial^2}{\partial \alpha^2} \chi_\alpha^2(\rho, \sigma) &= \text{tr} \sigma^{\alpha-1} \rho \sigma^{-\alpha} (\rho \log^2 \sigma + \log^2 \sigma \rho - 2 \log \sigma \rho \log \sigma) \\ &= \sum_{kl} \mu_k^{\alpha-1} \mu_l^{-\alpha} (\log \mu_k - \log \mu_l)^2 |\langle k | \rho | l \rangle|^2 \geq 0 \end{aligned} \quad (2.34)$$

where we used $\sigma = \sum_k \mu_k |k\rangle \langle k|$. \square

2.3 Mixing time bounds and contraction of the χ^2 -divergence under tcp-maps

2.3.1 Mixing time bounds

The χ^2 -divergence is an essential tool in the study of Markov chain mixing times, because on the one hand it bounds the trace distance, and on the other it allows easy access to the spectral properties of the map. The subsequent analysis can be seen as a generalization of the work presented in [46, 47] to the non-commutative setting.

Theorem 32 (Mixing time bound). *Let $\mathcal{T} : \mathcal{M}_d \mapsto \mathcal{M}_d$ be an ergodic quantum channel with fixed point $\sigma \in \mathcal{S}_d$, for any $\rho \in \mathcal{S}_d$ and any $k \in \mathcal{K}$, we can bound*

$$\|T^n(\rho) - \sigma\|_{\text{tr}} \leq (s_1^k)^n \sqrt{\chi_k^2(\rho, \sigma)}. \quad (2.35)$$

Here s_1^k denotes the second largest singular value (the largest being 1) of the map

$$\mathcal{Q}_k = [\Omega_\sigma^k]^{1/2} \circ \mathcal{T} \circ [\Omega_\sigma^k]^{-1/2} \quad (2.36)$$

Before we prove Theorem (32), we would like to point out an important fact that regards the singular values of \mathcal{Q}_k . The monotonicity of the χ^2 -divergence ensures, that the singular values s_i^k of \mathcal{Q}_k are always contained in $[0, 1]$ irrespectively of the choice of $k \in \mathcal{K}$. Let us therefore prove the following:

Lemma 33 (spectral interval). *The spectrum of the map $\mathcal{S}_k \equiv \mathcal{Q}_k^* \circ \mathcal{Q}_k = [\Omega_\sigma^k]^{-1/2} \circ T^* \circ \Omega_\sigma^k \circ \mathcal{T} \circ [\Omega_\sigma^k]^{-1/2}$ is contained in $[0, 1]$.*

PROOF: Let us first note, that the map \mathcal{S}_k is Hermitian and positive by construction. Furthermore, the monotonicity of the χ^2 -divergence, as stated in Theorem (27) ensures that the Rayleigh-Ritz quotient is bounded by 1. This holds, since $\forall B$

$$\begin{aligned} \langle B, \mathcal{S}_k(B) \rangle &= \langle A, T^* \circ \Omega_\sigma^k \circ \mathcal{T}(A) \rangle = M_{\mathcal{T}(\sigma)}^k(\mathcal{T}(A), \mathcal{T}(A)) \\ &\leq M_\sigma^k(A, A) = \langle A, \Omega_\sigma^k(A) \rangle = \langle B, B \rangle, \end{aligned} \quad (2.37)$$

where we defined the intermediate state $A = [\Omega_\sigma^k]^{-1/2}(B)$. Note that we made use of the fact that $\sigma = \mathcal{T}(\sigma)$ is the fixed point of the map. Therefore

$$\lambda_{max} = \max_{B \in \mathcal{M}_d} \frac{\langle B, \mathcal{S}_k(B) \rangle}{\langle B, B \rangle} \leq 1 \quad (2.38)$$

and the maximum is attained for $\lambda_{max} = 1$ and $B_{max} = [\Omega_\sigma^k]^{1/2}(\sigma)$. \square

With the bound on the spectrum at hand, it is now straight forward to prove Theorem (32)

PROOF: Define $e(n) \in \mathcal{M}_d$, as $e(n) = T^n(\rho - \sigma)$. By Lemma 29, we get $\|e(n)\|_{\text{tr}}^2 \leq \chi_k^2(T^n(\rho), T^n(\sigma)) \equiv \chi_k^2(n)$. In the matrix representation, $|e(n)\rangle = e(n) \otimes \mathbb{1} | \tilde{\Omega} \rangle$, we can rewrite $\chi_k^2(n) = \langle e(n) | \hat{\Omega}_\sigma^k | e(n) \rangle$. Note that also, $|e(n+1)\rangle = \hat{\mathcal{T}} | e(n) \rangle$ and so,

$$\chi_k^2(n) - \chi_k^2(n+1) = \langle e(n) | \hat{\Omega}_\sigma^k | e(n) \rangle - \langle e(n) | \hat{\mathcal{T}}^\dagger \hat{\Omega}_\sigma^k \hat{\mathcal{T}} | e(n) \rangle \quad (2.39)$$

$$= \langle e(n) | [\hat{\Omega}_\sigma^k]^{1/2} \left(\mathbb{1} - \hat{\mathcal{Q}}_k^\dagger \hat{\mathcal{Q}}_k \right) [\hat{\Omega}_\sigma^k]^{1/2} | e(n) \rangle. \quad (2.40)$$

Due to Lemma (33) we know that the spectrum of $\hat{S}_k = \hat{\mathcal{Q}}_k^\dagger \hat{\mathcal{Q}}_k$, which is equal to the square of the singular values of $\hat{\mathcal{Q}}_k$, is contained in the interval $[0, 1]$. Hence,

$$\langle e(n) | [\hat{\Omega}_\sigma^k]^{1/2} \left(\mathbb{1} - \hat{S}_k \right) [\hat{\Omega}_\sigma^k]^{1/2} | e(n) \rangle \quad (2.41)$$

$$\geq (1 - s_1^2) \langle e(n) | [\hat{\Omega}_\sigma^k]^{1/2} \sum_{\alpha \neq 0} P_\alpha [\hat{\Omega}_\sigma^k]^{1/2} | e(n) \rangle. \quad (2.42)$$

The sum is taken over spectral projectors P_α^k of $\hat{S}_k = \sum_\alpha (s_\alpha^k)^2 P_\alpha$, apart from P_0^k which projects onto $[\hat{\Omega}_\sigma^k]^{-1/2} | \sigma \rangle$. In particular, $P_0^k = [\hat{\Omega}_\sigma^k]^{-1/2} | \sigma \rangle \langle \mathbb{1} | [\hat{\Omega}_\sigma^k]^{-1/2}$, so that

$\langle e(n) | [\hat{\Omega}_\sigma^k]^{1/2} P_0^k [\hat{\Omega}_\sigma^k]^{1/2} | e(n) \rangle = \langle e(n) | \sigma \rangle \text{tr}[T^n(\rho - \sigma)] = 0$, by trace preservation of \mathcal{T} .

We can write,

$$\chi_k^2(n) - \chi_k^2(n+1) \geq (1 - (s_1^k)^2) \chi_k^2(n). \quad (2.43)$$

Rearranging terms completes the theorem. \square

Remark: The fact, that the singular values of Q_k are always smaller or equal to one justifies the use of the generalized χ^2 -divergence as the appropriate distance measure to bound the convergence of an arbitrary channel. It is tempting to use the Hilbert-Schmidt inner product to give an upper bound to the trace norm. This can always be done at the cost of a dimension dependent prefactor, since on finite dimensional spaces all norms are equivalent. However, when doing so a problem arises if one tries to bound the convergence in terms of the spectral properties of the map $S_{HS} = \mathcal{T}^* \circ \mathcal{T}$. It is in general not ensured that the spectrum will be bounded by one. In fact, for every non-unital channel \mathcal{T} , S_{HS} will have an eigenvalue larger than one [68]. The similarity transformation of the channel \mathcal{T} with $[\Omega_\sigma^k]^{1/2}$ alters the singular values, but of course leaves the spectrum invariant. Furthermore, it is a well known fact [17] that the singular values of a square matrix log-majorize the absolute value of the eigenvalues. As the spectrum of Q_k is bounded by one (and equal that of $\hat{\mathcal{T}}$ by similarity), we conclude that its second largest eigenvalue is always smaller or equal to its second largest singular value.

For some instances of the inversion Ω_σ^k it becomes immediately evident that the symmetrization S_k has the desired spectral properties without making use of the monotonicity of the χ_k^2 -divergence. It can occur, that S_k is again similar to a quantum channel that is of the form $\mathcal{T}_s^k = [\Omega_\sigma^k]^{-1/2} \circ S_k \circ [\Omega_\sigma^k]^{1/2}$. A possible example of such an inversions is $\Omega_\sigma^{\alpha=1/2} = \mathcal{L}_\sigma^{-1/2} \mathcal{R}_\sigma^{-1/2}$. This is however not the generic case, most inversions will lead to maps that are not completely positive any longer. It would be very desirable to find other such examples, as they mirror the classical situation where the symmetrized maps are always probability transition matrices, and because these specific inversions allow for clean contraction bounds as seen in section III.B.

It is clear from the discussion above that the singular values of Q_k play a crucial role in the mixing time analysis presented here. This seems to contradict the general understanding that the convergence is determined by the spectral properties of the channel \mathcal{T} in the asymptotic limit. This can however be understood as follows: the matrix \hat{Q}_k is similar to $\hat{\mathcal{T}}$, i.e. $\hat{Q}_k = [\hat{\Omega}_\sigma^k]^{1/2} \cdot \hat{\mathcal{T}} \cdot [\hat{\Omega}_\sigma^k]^{-1/2}$, so the spectra of Q_k and \mathcal{T} coincide. The following lemma establishes a relation between the singular values and the eigenvalues in the asymptotic limit. For a proof, see e.g. [69] pg.180.

Lemma 34 (Singular values). *Let $\hat{Q}_k \in \mathcal{M}_{d^2}$ be given, and let $s_0(\hat{Q}_k) \geq \dots \geq s_{d^2-1}(\hat{Q}_k)$ and $\{\lambda_i(\hat{Q}_k)\}_{i=0\dots d^2-1}$ denote its singular values and eigenvalues, respectively with $|\lambda_0(\hat{Q}_k)| \geq \dots \geq |\lambda_{d^2-1}(\hat{Q}_k)|$. Then*

$$\lim_{n \rightarrow \infty} [s_i(\hat{Q}_k^n)]^{1/n} = |\lambda_i(\hat{Q}_k)| \quad \forall i = 0 \dots d^2 - 1 \quad (2.44)$$

In the limit of $n \rightarrow \infty$ applications of the quantum channel, we can start blocking the channel in m subsequent applications $\mathcal{T}^{(m)} \equiv \mathcal{T}^m$ and bound the convergence rate as a function of the singular values of the corresponding $\hat{\mathcal{Q}}_k^{(m)}$, which indeed converge to the eigenvalues of the original cp-map. Convergence following the eigenvalue is therefore only guaranteed in the limit of $n \rightarrow \infty$, and this would indeed be the case, when e.g. the eigenstructure of the original cp-maps contains a Jordan block associated to the second largest eigenvalue. Note, that convergence in the above lemma goes typically as $1/n$, which is very slow. Hence for finite n , convergence is governed by the singular values of $\hat{\mathcal{Q}}_k$ as opposed to the eigenvalues. The bound derived in (32) is an absolute bound for finite n and clearly leads to a strictly monotonic decay. Note that in the case that the second largest singular value is also equal to 1, this can then always be cured by blocking the cp-maps together. Finally, it is worth mentioning that the convergence can in fact be much more rapid if one starts in a state "closer" to the fixed point. In particular, if the initial state is such that $\rho - \sigma \propto Y_k$, $k \geq 2$, where Y_k is the eigenvector corresponding to λ_k , then the convergence will be governed by the magnitude of λ_k . Furthermore, if instead of a single fixed point, we have a fixed subspace, or a collection of fixed subspaces (with or without rotating points), then the convergence to this fixed subspace will be governed by the largest eigenvalue whose magnitude is strictly smaller than one.

Thus far we have only considered the time-discrete case, it is however straightforward to give a similar bound for time-continuous Markov processes, that are described by a one parameter semi-group. The following lemma bounds the trace-distance as a function of $t \in \mathbb{R}_0^+$: The proof of the following lemma is very similar to the proof of the time discrete case, we will therefore omit it here.

Lemma 35 (Time-continuous bound). *Let \mathcal{L} denote the generator of a time continuous Markov process, described by the master equation $\partial_t \rho = \mathcal{L}(\rho)$, with solution $\rho(t) \in \mathcal{S}_d \forall t \in [0, +\infty)$. Furthermore let $\sigma \in \mathcal{S}_d^+$ denote the fixed-point $\mathcal{L}(\sigma) = 0$, then*

$$\|\rho(t) - \sigma\|_{\text{tr}}^2 \leq e^{l_1^k t} \chi_k^2(\rho(0), \sigma). \quad (2.45)$$

Here, $l_1^k \leq 0$ refers to the second largest eigenvalue of

$$\Lambda_k = [\Omega_\sigma^k]^{1/2} \circ \mathcal{L}^* \circ [\Omega_\sigma^k]^{-1/2} + [\Omega_\sigma^k]^{-1/2} \circ \mathcal{L} \circ [\Omega_\sigma^k]^{1/2}. \quad (2.46)$$

The symmetrization for the generator of the time continuous Markov process is additive as would be expected. Furthermore, we note that the monotonicity of the χ^2 -divergence ensures that the spectrum of Λ_k is never positive, based on a similar reasoning as given in Lemma (33).

2.3.2 Contraction coefficients

In the following we study the contraction of the χ^2 -divergences under quantum channels, and its relation to the trace norm contraction. We consider general contraction rather than contraction to the fixed point because analytic results are more readily available, and because these

bounds are in a sense the *most* stringent one can require. We focus primarily on the mean α -subfamily of χ^2 -divergences.

Let us define the following contraction coefficients which we call the χ^2 - and trace norm-contraction respectively:

$$\eta_{\chi}^{\alpha}(\mathcal{T}) = \sup_{\rho, \sigma \in \mathcal{S}_d} \frac{\chi_{\alpha}^2(\mathcal{T}(\rho), \mathcal{T}(\sigma))}{\chi_{\alpha}^2(\rho, \sigma)} \quad (2.47)$$

and

$$\eta_{\text{tr}}(\mathcal{T}) = \sup_{\rho, \sigma \in \mathcal{S}_d} \frac{\|\mathcal{T}(\rho - \sigma)\|_{\text{tr}}}{\|\rho - \sigma\|_{\text{tr}}} = \sup_{\phi, \psi \in \mathcal{S}_d^1, \langle \phi | \psi \rangle = 0} \frac{1}{2} \|\mathcal{T}(\psi) - \mathcal{T}(\phi)\|_{\text{tr}}, \quad (2.48)$$

where $\mathcal{T} : \mathcal{M}_d \rightarrow \mathcal{M}_d$ is a quantum channel, and the last equality is seen simply by convexity of the trace norm.

We first upper bound the trace-norm contraction in terms of the χ^2 contraction, which is a generalization of a result in [70]:

Lemma 36. *For all $\alpha \in (0, 1]$, and a quantum channel $\mathcal{T} : \mathcal{M}_d \rightarrow \mathcal{M}_d$,*

$$\eta_{\text{tr}}(\mathcal{T}) \leq \sqrt{\eta_{\chi}^{\alpha}(\mathcal{T})}. \quad (2.49)$$

PROOF: From Lemma 29, we have that $\|\mathcal{T}(\rho - \sigma)\|_{\text{tr}}^2 \leq \chi_{\alpha}^2(\mathcal{T}(\rho), \mathcal{T}(\sigma))$, for all $\rho, \sigma \in \mathcal{S}_d$. Let N be traceless and hermitian, and note that it can be written as $N = N_+ - N_-$, where N_+, N_- are positive definite and orthogonal in their support. Now let $P = |N|/\|N\|_{\text{tr}}$ and recall that $|N| = N_+ + N_-$, then we get $\text{tr}[NP^{-\alpha}NP^{\alpha-1}] = \|N\|_{\text{tr}}^2$, for every $\alpha \in (0, 1]$. Also,

$$\frac{\|\mathcal{T}(N)\|_{\text{tr}}^2}{\|N\|_{\text{tr}}^2} \leq \frac{\text{tr}[\mathcal{T}(N)\mathcal{T}(P)^{-\alpha}\mathcal{T}(N)\mathcal{T}(P)^{\alpha-1}]}{\text{tr}[NP^{-\alpha}NP^{\alpha-1}]} \quad (2.50)$$

where the inequality is in the numerator, and the denominators are equal, by the previous observation. Taking the supremum over all traceless hermitian N on the left hand side and identifying $\rho - \sigma = N/\|N\|_{\text{tr}}$, $P = \sigma$ then gives desired result. \square

We now provide a lower bound to the trace norm contraction for primitive channels:

Lemma 37. *Given a quantum channel $\mathcal{T} : \mathcal{M}_d \rightarrow \mathcal{M}_d$,*

$$\eta_{\chi}^{\alpha=1/2}(\mathcal{T}) \leq \eta_{\text{tr}}(\mathcal{T}) \quad (2.51)$$

First we introduce an eigenvalue type min-max characterization of the χ^2 -contraction, and then show that this eigenvalue must be smaller than the trace norm-contraction.

Let $P > 0$, and consider the following eigenvalue equation:

$$\hat{\Gamma} | A \rangle \equiv \hat{\Omega}_P^{-1} \hat{\mathcal{T}}^\dagger \hat{\Omega}_{\mathcal{T}(P)} \hat{\mathcal{T}} | A \rangle = \lambda | A \rangle, \quad (2.52)$$

where $\Omega_X \equiv \Omega_X^{\alpha=1/2}$. If \mathcal{T} has a non-trivial kernel, then $\Omega_{\mathcal{T}(P)}$ should be understood in terms of the pseudo-inverse. First note that Γ is a quantum channel, so its spectrum is bounded

by one, and that it reaches one for $A = P$. Also note that Γ is similar to a hermitian operator, so it has all real eigenvalues, so we can take the eigenvectors to be hermitian. Then rewriting (2.52) as $\hat{T}^\dagger \hat{\Omega}_{\mathcal{T}(P)} \hat{T} | A \rangle = \lambda \hat{\Omega}_P | A \rangle$, we can express the second largest eigenvalue as:

$$\begin{aligned} \lambda_1(\mathcal{T}, P) &= \sup_{\langle N | \Omega_P(P) \rangle = 0, N = N^\dagger} \frac{\langle N | \hat{T}^\dagger \hat{\Omega}_{\mathcal{T}(P)} \hat{T} | N \rangle}{\langle N | \hat{\Omega}_P | N \rangle} \\ &= \sup_{\text{tr} N = 0, N = N^\dagger} \frac{\text{tr}[\mathcal{T}(N) \mathcal{T}(P)^{-1/2} \mathcal{T}(N) \mathcal{T}(P)^{-1/2}]}{\text{tr}[NP^{-1/2} NP^{-1/2}]} \end{aligned} \quad (2.53)$$

Clearly, by maximizing over all P , one recovers $\eta_\chi^{1/2}(\mathcal{T})$. We now prove the above theorem:

PROOF: Let N_1 be the eigenvector for which λ_1 satisfies the eigenvalue equation (2.52), and recall that N_1 is Hermitian and traceless. Then,

$$\lambda_1 \|N_1\|_{\text{tr}} = \|\Gamma(\mathcal{T}(N_1))\|_{\text{tr}} \leq \|\mathcal{T}(N_1)\|_{\text{tr}} \quad (2.54)$$

because Γ is a channel, and

$$\lambda_1 \leq \frac{\|\mathcal{T}(N_1)\|_{\text{tr}}}{\|N_1\|_{\text{tr}}} \leq \sup_{\text{tr} N = 0, N^\dagger = N} \frac{\|\mathcal{T}(N)\|_{\text{tr}}}{\|N\|_{\text{tr}}} = \eta_{\text{tr}}, \quad (2.55)$$

taking the supremum over positive P completes the proof. \square

Remark: Theorem 37 gives a computable lower bound to the trace norm contraction. A key subtlety in the argument is that $[\Omega_P(A)]^{-1} = \sqrt{P} A \sqrt{P}$ is a completely positive, but not trace preserving, (CP) map (with a single Kraus operator \sqrt{P}) which implies that Γ is a quantum channel. In general, Ω_P is not even positivity preserving. Another exception is the monotone metric associated with the usual logarithmic relative entropy for which $k(w) = \frac{\log w}{w-1}$. It is well-known [57, 63, 59] that $\Omega_P^{\log}(A)$ can be written as

$$\Omega_P^{\log}(A) = \int_0^\infty \frac{1}{P + xI} A \frac{1}{P + xI} dx \quad (2.56)$$

which is clearly CP. An analogous lower bound was shown in [59] for this map using a similar argument. Clearly, this can be extended to any monotone metric for which Ω_P is CP; however, we do not know of any other examples.

Very little is known about the ordering of the general η_k contraction coefficients. In particular, We do not know whether whether η_χ^{\log} is smaller or larger than $\eta_\chi^{\alpha=1/2}$. However, it is known [59] that η_k are not all identical for different $k \in \mathcal{K}$; because examples can be constructed using non-unital qubit channels. Theorem 36 can readily be extended to any metric associated with $k \in \mathcal{K}$. However, it seems unlikely that Theorem 37 holds in general,. Thus, we can conclude

$$\max\{\eta_\chi^{\alpha=1/2}(\mathcal{T}), \eta_\chi^{\log}(\mathcal{T})\} \leq \eta_{\text{tr}}(\mathcal{T}) \leq \inf_{k \in \mathcal{K}} \sqrt{\eta_\chi^k(\mathcal{T})}. \quad (2.57)$$

Note that if instead of maximizing over all P we only consider contraction of the map to the steady state, and denote it $\bar{\eta}(\mathcal{T}) = \eta(\mathcal{T})_{P=\sigma}$, then from the above arguments one immediately gets:

$$\bar{\eta}_X^\alpha(\mathcal{T}) \leq \bar{\eta}_{\text{tr}}(\mathcal{T}) \leq \eta_{\text{tr}}(\mathcal{T}) \leq 1 \quad (2.58)$$

Combing this with the previous bounds above, we have

$$\lambda_1 \leq s_1^{\alpha=1/2} = \bar{\eta}_X^{\alpha=1/2} \leq \eta_X^{\alpha=1/2} \leq \eta_{\text{tr}} \leq \sqrt{\eta_X^{\alpha=1/2}}. \quad (2.59)$$

Moreover, $k(w) = w^{-1/2}$ on the right can be replaced by any $k \in \mathcal{K}$, and that on the left by $k(w) = (w-1)^{-1} \log w$. It is very tempting to conjecture that $\bar{\eta}_{\text{tr}}^2 \leq \bar{\eta}_X^\alpha$, and/or that $\eta_{\text{tr}} \leq \sqrt{\bar{\eta}_X^{\alpha=1/2}}$, but simple numerical counterexamples show these to be false.

2.4 Quantum detailed balance

The detailed balance condition is often crucial in the analysis of classical Markov chain mixing times, as it ensures several convenient properties of the Markov chain. In particular, it implies that the classical probability distribution with respect to which the stochastic map is detailed balanced is a fixed point of the chain. Furthermore, detailed balanced stochastic maps have a real spectrum. In this section we generalize the notion of classical detailed balance to quantum Markov chains. Alternative definitions of quantum detailed balance have been given in the literature: [71, 72, 73, 74] and references therein. Central to our approach is the operator \mathcal{Q}_k as previously introduced in Lemma 32. In the literature for classical Markov chains an analogous matrix exists and is often referred to as the discriminant.

Definition 38. For a channel $\mathcal{T} : \mathcal{M}_d \rightarrow \mathcal{M}_d$ and a state $\sigma \in \mathcal{S}_d^+$ with corresponding inversion Ω_σ^k as defined in (2.9), we define the quantum discriminant of \mathcal{T} as,

$$\mathcal{Q}_k = [\Omega_\sigma^k]^{1/2} \circ \mathcal{T} \circ [\Omega_\sigma^k]^{-1/2}. \quad (2.60)$$

We recall that the convergence of an arbitrary quantum Markov process can be bounded by the singular values of $\hat{\mathcal{Q}}_k$. Classical detailed balanced Markov chains have the property that the corresponding discriminant becomes symmetric. We shall therefore define the quantum generalization by requiring that for a quantum detailed balanced process

$$\mathcal{Q}_k^* = \mathcal{Q}_k. \quad (2.61)$$

This immediately allows to make a statement about the spectrum of quantum detailed balanced maps. Due to the hermicity of the matrix representation of the map (2.60) we can immediately deduce, just as for classically case, that the quantum channel \mathcal{T} has a real spectrum. For

detailed balanced maps, the second largest eigenvalue in magnitude coincides with the second largest singular value. Furthermore, we would like to point out that this is actually not just a single condition for quantum detailed balance but a whole family. Hence every different inversion Ω_σ^k gives rise to a different condition for detailed balance. We therefore define as the quantum generalization of detailed balance:

Definition 39. For a channel $\mathcal{T} : \mathcal{M}_d \rightarrow \mathcal{M}_d$ and a state $\sigma \in \mathcal{S}_d^+$, we say that \mathcal{T} obeys k -detailed balance with respect to σ with $k \in \mathcal{K}$, when

$$[\Omega_\sigma^k]^{-1} \circ \mathcal{T}^* = \mathcal{T} \circ [\Omega_\sigma^k]^{-1}. \quad (2.62)$$

A consequence of this definition is that σ is a fixed point of \mathcal{T} .

Lemma 40. Let $\sigma \in \mathcal{S}_d$ be a state and \mathcal{T} a channel that satisfies the detailed balance Definition 39 with respect to Ω_σ^k , then σ is a steady state of \mathcal{T} .

PROOF: Recall that the inverse is given by $[\Omega_\sigma^k]^{-1} = \mathcal{R}_\sigma f(\Delta_{\sigma,\sigma})$, where $f(w) = 1/k(w)$. Hence, since $k(1) = f(1) = 1$, we have

$$[\Omega_\sigma^k]^{-1}(\mathbb{1}) = \mathcal{R}_\sigma f(\Delta_{\sigma,\sigma})\mathbb{1} = \mathcal{R}_\sigma \mathbb{1} = \sigma. \quad (2.63)$$

Now, since furthermore $\mathcal{T}^*(\mathbb{1}) = \mathbb{1}$, we have that

$$\mathcal{T}(\sigma) = \mathcal{T} \circ [\Omega_\sigma^k]^{-1}(\mathbb{1}) = [\Omega_\sigma^k]^{-1} \circ \mathcal{T}^*(\mathbb{1}) = [\Omega_\sigma^k]^{-1}(\mathbb{1}) = \sigma. \quad (2.64)$$

□

Given a probability distribution on some set of states, it is desirable to have a simple criterium to check whether a completely positive map obeys detailed balance with respect to the state generated from the distribution. This criterium may then serve to set up a Markov chain that actually converges to the desired steady state.

Proposition 41. Let $\{|i\rangle\}_i$ be a complete orthonormal basis of \mathcal{H} and let $\{\mu_i\}_i$ be a probability distribution on this basis. Furthermore, assume that a quantum channel \mathcal{T} obeys

$$\frac{\mu_n}{k(\mu_m/\mu_n)} \langle i | \mathcal{T}(|n\rangle\langle m|) | j \rangle = \frac{\mu_i}{k(\mu_j/\mu_i)} \langle m | \mathcal{T}(|j\rangle\langle i|) | n \rangle, \quad (2.65)$$

then $\sigma = \sum_i \mu_i |i\rangle\langle i|$ and \mathcal{T} obey the detailed balance condition with respect to Ω_σ^k .

PROOF: Note that $\{|i\rangle\langle j|\}_{ij}$ forms a complete and orthonormal basis in the space \mathcal{M}_d with respect to the Hilbert-Schmidt scalar product. We can therefore express equation (2.62) in this basis. The individual entries are equal due to

$$\begin{aligned} \text{tr} \left[(|m\rangle\langle n|)^\dagger [\Omega_\sigma^k]^{-1} \circ \mathcal{T}^*(|j\rangle\langle i|) \right] &= \mu_n k^{-1} (\mu_m/\mu_n) \text{tr} \left[\mathcal{T}(|m\rangle\langle n|)^\dagger (|j\rangle\langle i|) \right] = \\ &= \mu_n k^{-1} (\mu_m/\mu_n) \langle i | \mathcal{T}(|n\rangle\langle m|) | j \rangle = \mu_i k^{-1} (\mu_j/\mu_i) \langle m | \mathcal{T}(|j\rangle\langle i|) | n \rangle = \\ &= \mu_i k^{-1} (\mu_j/\mu_i) \text{tr} \left[(|m\rangle\langle n|)^\dagger \mathcal{T}(|j\rangle\langle i|) \right] = \text{tr} \left[(|m\rangle\langle n|)^\dagger \mathcal{T} \circ [\Omega_\sigma^k]^{-1}(|j\rangle\langle i|) \right]. \end{aligned} \quad (2.66)$$

□

Remark: We note that the different quantum detailed balance conditions coincide for classical channels, i.e. for stochastic processes that are included in the framework of quantum channels. Define the following "classical" Kraus operators:

$$A_{ij}^{cl} = \sqrt{P_{ij}} |i\rangle \langle j| \quad \text{and a state, } \sigma = \sum_i \mu_i |i\rangle \langle i|. \quad (2.67)$$

In this case, the condition of Proposition 41 reduces to the classical condition. This can be seen when considering the channel $\mathcal{T}^{cl}(\rho) = \sum_{ij} A_{ij}^{cl} \rho A_{ij}^{cl\dagger}$ and checking for detailed balance with respect to sigma, since

$$\frac{\mu_m}{k(\mu_n/\mu_m)} \langle i | \mathcal{T}^{cl}(|n\rangle \langle m|) |j\rangle = \frac{\mu_m}{k(\mu_n/\mu_m)} \delta_{nm} \delta_{ij} P_{in}$$

and

$$\frac{\mu_i}{k(\mu_j/\mu_i)} \langle i | \mathcal{T}^{cl}(|n\rangle \langle m|) |j\rangle = \frac{\mu_i}{k(\mu_j/\mu_i)} \delta_{nm} \delta_{ij} P_{ni}. \quad (2.68)$$

However since $k(1) = 1$ we are just left with the classical detailed balance condition $\mu_i P_{ni} = \mu_n P_{in}$ for all pairs i, n .

A natural question to ask is therefore, whether the different detailed balance conditions are all identical. To see that this is not the case, consider the example given by the Kraus operators of a single qubit, i.e. $\mathcal{H} = \mathbb{C}^2$,

$$A_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad A_2 = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}. \quad (2.69)$$

This channel has the unique fixed point

$$\sigma = \frac{1}{6} \begin{pmatrix} 5 & 1 \\ 1 & 1 \end{pmatrix}. \quad (2.70)$$

From this channel it is now possible to construct a channel that obeys detailed balance with respect to the inversion given by choosing $k(w) = w^{-1/2}$, that is the inversion reads $\Omega_\sigma^{\alpha=1/2} = \mathcal{L}_\sigma^{-1/2} \mathcal{R}_\sigma^{-1/2}$. We consider therefore the symmetrized map,

$$\mathcal{T}_s = \left[\Omega_\sigma^{\alpha=1/2} \right]^{-1} \circ \mathcal{T}^* \circ \Omega_\sigma^{\alpha=1/2} \circ \mathcal{T}. \quad (2.71)$$

For the specific instance where $\Omega_\sigma^{\alpha=1/2}$ is given as above, we are assured that the map \mathcal{T}_s is again a quantum channel, because one immediately finds the Kraus representation for $\mathcal{T}_s(\rho) = \sum_{ij} B_{ij} \rho B_{ij}^\dagger$ as $B_{ij} = \sqrt{\sigma} A_i^\dagger [\sqrt{\sigma}]^{-1} A_j$. The individual Kraus operators read,

$$B_{11} = \frac{3}{5} \begin{pmatrix} 1 & 1 \\ 1/2 & 1/2 \end{pmatrix} \quad \text{and} \quad B_{12} = \frac{\sqrt{2}}{5} \begin{pmatrix} 1 & -1 \\ 1/2 & -1/2 \end{pmatrix}, \quad (2.72)$$

$$B_{21} = \frac{\sqrt{2}}{20} \begin{pmatrix} 3 & 3 \\ -1 & -1 \end{pmatrix} \quad \text{and} \quad B_{22} = \frac{1}{5} \begin{pmatrix} 3 & -3 \\ -1 & 1 \end{pmatrix}.$$

The channel \mathcal{T}_s satisfies detailed balance with respect to $\Omega_\sigma^{\alpha=1/2}$ by construction. This channel however does not satisfy detailed balance with respect to the inversion $\Omega_\sigma^{Bures} = 2[\mathcal{L}_\sigma + \mathcal{R}_\sigma]^{-1}$ as can be seen directly by evaluating the detailed balance condition in terms of the matrix representations,

$$\left[\hat{\Omega}_\sigma^{Bures}\right]^{-1} \cdot \hat{\mathcal{T}}_s^\dagger - \hat{\mathcal{T}}_s \cdot \left[\hat{\Omega}_\sigma^{Bures}\right]^{-1} = \frac{7}{600} [\mathbb{1} \otimes Y + Y \otimes \mathbb{1}], \quad (2.73)$$

where

$$Y = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \quad (2.74)$$

The family of quantum detailed balance conditions is therefore much richer than the classical counterpart.

2.5 Quantum Cheeger's inequality

In the context of classical stochastic processes a very powerful formalism has been developed, often referred to as the conductance bound or Cheeger's inequality, to bound convergence rates of stochastic processes. We will generalize this to the quantum setting in this section. Similar results have appeared in [75]. The gap of the map \mathcal{S}_k is defined as the difference between the largest and second largest eigenvalue, $\Delta = 1 - \lambda_1$. The gap can be characterized in a variational fashion [17].

Proposition 42. *The gap of the map $\mathcal{S}_k = [\Omega_\sigma^k]^{-1/2} \circ \mathcal{T}^* \circ \Omega_\sigma^k \circ \mathcal{T} \circ [\Omega_\sigma^k]^{-1/2}$ is given by*

$$\Delta = \min_{X \in \mathcal{M}_d} \frac{\langle X, (\text{id} - \mathcal{S}_k)X \rangle}{\frac{1}{2} \|(X \otimes \sqrt{\sigma} - \sqrt{\sigma} \otimes X)\|_{HS}^2}, \quad (2.75)$$

where $\|A\|_{HS}^2 = \text{tr}[A^\dagger A]$ denotes the standard Hilbert-Schmidt norm and $\langle \cdot, \cdot \rangle$ the corresponding Hilbert-Schmidt scalar product.

PROOF: The eigenvector that corresponds to the eigenvalue $\lambda_0 = 1$ of \mathcal{S}_k is given by $\sqrt{\sigma}$. The gap can therefore be written as[17]:

$$\begin{aligned} \Delta &= \min_{X \in \mathcal{M}_d; \text{tr}[X\sqrt{\sigma}] = 0} 1 - \frac{\text{tr}[X^\dagger \mathcal{S}_k(X)]}{\text{tr}[X^\dagger X]} \\ &= \min_{X \in \mathcal{M}_d; \text{tr}[X\sqrt{\sigma}] = 0} \frac{\text{tr}[X^\dagger (X - \mathcal{S}_k(X))]}{\text{tr}[X^\dagger X] - \text{tr}[X\sqrt{\sigma}]^2} \\ &= \min_{X \in \mathcal{M}_d} \frac{\text{tr}[X^\dagger (X - \mathcal{S}_k(X))]}{\frac{1}{2} \|(X \otimes \sqrt{\sigma} - \sqrt{\sigma} \otimes X)\|_{HS}^2}, \end{aligned} \quad (2.76)$$

Note that the constrained $\text{tr}[X\sqrt{\sigma}] = 0$ can be dropped in the last line. Suppose that $\text{tr}[X\sqrt{\sigma}] = c$, we can then define $X' = X - c\sqrt{\sigma}$ and vary X' since the equation is invariant under such shifts. \square

Throughout the remainder of this section we consider unital quantum channels, i.e. maps which obey $\mathcal{T}(\mathbb{1}) = \mathbb{1}$. For this case it is ensured that already the simple map $\mathcal{S} = \mathcal{T}^* \circ \mathcal{T}$ has a spectrum that is contained in $[0, 1]$, since all Ω_σ^k coincide and correspond to the identity map. The χ^2 -divergence just reduces to the standard Hilbert-Schmidt inner product times a prefactor given by $d = \dim(\mathcal{H})$. In the case of a detailed balanced stochastic map it even suffices to just consider the map itself. In either case we will denote the corresponding map as S from now on. The variational characterization of the gap Δ now allows us to give an upper as well as a lower bound to the second largest eigenvalue of S .

Lemma 43. *Let $\mathcal{T} : \mathcal{M}_d \rightarrow \mathcal{M}_d$ be a unital quantum channel. Then the second largest eigenvalue λ_1 of its symmetrization $\mathcal{S} = \mathcal{T}^* \circ \mathcal{T}$, is bounded by,*

$$1 - 2h \leq \lambda_1 \leq 1 - \frac{1}{2}h^2, \quad (2.77)$$

where h is Cheeger's constant defined as,

$$h = \min_{\Pi_A, \text{tr}[\Pi_A] \leq d/2} \frac{\text{tr}[(\mathbb{1} - \Pi_A)\mathcal{S}(\Pi_A)]}{\text{tr}[\Pi_A]}. \quad (2.78)$$

The minimum is to be taken over all projectors Π_A on the space $A \subset \mathcal{H}$, so that $\text{tr}[\Pi_A] \leq d/2$.

PROOF: An upper bound to the gap is immediately found by choosing $X = \Pi_A$. Due Proposition (42) we can write:

$$\begin{aligned} \Delta &\leq \frac{\text{tr}[\Pi_A(\text{id} - \mathcal{S})(\Pi_A)]}{\text{tr}[\Pi_A^2] - \frac{1}{d}\text{tr}[\Pi_A]^2} \\ &= \frac{\text{tr}[(\mathbb{1} - \Pi_A)\mathcal{S}(\Pi_A)]}{\frac{1}{d}\text{tr}[(\mathbb{1} - \Pi_A)]\text{tr}[\Pi_A]} \leq 2h, \end{aligned} \quad (2.79)$$

where in the last line we have used that $\text{tr}[\mathbb{1} - \Pi_A] \geq d/2$.

For the lower bound, we can restrict the minimization in (2.78) to diagonal projections. Furthermore, when considering only unital quantum channels, it is possible to reduce the problem of bounding the gap Δ to that of bounding the gap of a classical stochastic process. To see this, let us work in the basis where the eigenvector $X_1 \in \mathcal{M}_d$ corresponding to λ_1 is diagonal. We shall assume wlog that $X_1^\dagger = X_1$. In this basis, we can write $X = \sum x_i |i\rangle \langle i|$. The numerator then becomes

$$\begin{aligned} \text{tr} \left[X^\dagger (X - \mathcal{S}(X)) \right] &= \sum_{ij} x_i x_j (\text{tr}[|i\rangle \langle i| |j\rangle \langle j|] - \text{tr}[|i\rangle \langle i| \mathcal{S}(|j\rangle \langle j|)]) \\ &= \sum_i x_i^2 - \sum_{ij} x_i x_j P_{ij} = \frac{1}{2} \sum_{ij} P_{ij} (x_i - x_j)^2. \end{aligned} \quad (2.80)$$

We introduced the matrix $P_{ij} = \langle i | \mathcal{S}(|j\rangle \langle j|) | i \rangle$, which is a symmetric non-negative matrix which obeys $P_{ij} \geq 0$, $\sum_i P_{ij} = 1$ and $P^T = P$. Hence P is doubly stochastic. Performing the same reduction in the denominator we obtain

$$\frac{1}{2d} \|(X \otimes \mathbb{1} - \mathbb{1} \otimes X)\|_{HS}^2 = \frac{1}{2d} \sum_{ij} (x_i - x_j)^2 \quad (2.81)$$

Hence, we arrive at the classical version of Mihail's Identity [48],

$$\Delta = \min_{\{x_i\}} \frac{\sum_{ij} P_{ij} (x_i - x_j)^2}{1/d \sum_{ij} (x_i - x_j)^2}. \quad (2.82)$$

Given the classical version of Mihail's identity, the proof of the lower bound is the same as in the classical case. For completeness we repeat it here. First, we define, $z_i \equiv |x_i| x_i$ and write,

$$\begin{aligned} \sum_{ij} P_{ij} |z_i - z_j| &= \sum_{ij} P_{ij} ||x_i| x_i - |x_j| x_j| \leq \sum_{ij} \sqrt{P_{ij}} \sqrt{P_{ij}} (|x_i| + |x_j|) (x_i - x_j) \\ &\leq \sqrt{\sum_{ij} P_{ij} (x_i - x_j)^2} \sqrt{\sum_{ij} P_{ij} (|x_i| + |x_j|)^2}, \end{aligned} \quad (2.83)$$

where we used Cauchy-Schwartz in the last step. Consider now,

$$\sum_{ij} P_{ij} (|x_i| + |x_j|)^2 = 2 \left(\sum_i x_i^2 + \sum_{ij} |x_i| P_{ij} |x_j| \right) \leq 4 \sum_i |x_i|^2. \quad (2.84)$$

Furthermore, note that we can bound,

$$1/d \sum_{ij} (x_i - x_j)^2 \leq 2/d \sum_{ij} x_i^2 = 2 \sum_i |z_i|. \quad (2.85)$$

We are therefore left with a lower bound to Mihail's identity, which holds for all choices of $\{x_i\}$

$$\frac{1}{2} \left(\frac{\sum_{ij} P_{ij} |z_i - z_j|}{2 \sum_i |z_i|} \right)^2 \leq \frac{\sum_{ij} P_{ij} (x_i - x_j)^2}{1/d \sum_{ij} (x_i - x_j)^2}. \quad (2.86)$$

We shall now assume, that $x_i \geq 0$ everywhere and we can hence drop the absolute values in the definition for the z_i . This is assumption is valid since we are free in adding an arbitrary constant $x_i \rightarrow x_i + c$ to make all x_i positive. Note that we therefore are left with a lower bound to the gap of the form,

$$\Delta \geq \frac{1}{2} \left(\frac{\sum_{ij} P_{ij} |x_i^2 - x_j^2|}{2 \sum_i x_i^2} \right)^2 \quad (2.87)$$

Let's focus on the right side of the inequality. Since,

$$2 \sum_{i,j : x_i \geq x_j} P_{ij} (x_i^2 - x_j^2) = 4 \sum_{i,j : x_i \geq x_j} P_{ij} \int_{x_j}^{x_i} t dt = 4 \int_0^\infty t \sum_{i,j : x_i > t \geq x_j} P_{ij} dt, \quad (2.88)$$

and furthermore,

$$\sum_{i,j : x_i > t \geq x_j} P_{ij} = \sum_{i \in A(t)} \sum_{j \in A^c(t)} P_{ij} \quad \text{where, } A(t) \equiv \{i | x_i \geq t\}, \quad (2.89)$$

we can bound,

$$4 \int_0^\infty t \sum_{i,j : x_i > t \geq x_j} P_{ij} dt \geq h 4 \int_0^\infty t \sum_{i \in A(t)} \Theta(t - x_i) dt = 2 h \left(\sum_i x_i^2 \right), \quad (2.90)$$

where we defined h as in the same fashion as above. We have therefore found the desired lower bound for the spectral gap of the map \mathcal{S} . \square

2.5.1 Example: Conductance bound for unital qubit channels

A convenient basis for the matrix space \mathcal{M}_2 associated with the Hilbert space $\mathcal{H} = \mathbb{C}^2$ is given in terms of the Pauli basis $\{\mathbb{1}, \sigma_x, \sigma_y, \sigma_z\}$. In this basis a density matrix $\rho \in \mathcal{S}_2$ can be parametrized in terms of its Bloch vector $\mathbf{r} \in \mathbb{R}^3$. In the Bloch representation the density matrix reads $\rho = \frac{1}{2}(\mathbb{1} + \mathbf{r} \cdot \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma} = (\sigma_x, \sigma_y, \sigma_z)$. It is also straight forward to determine the matrix representation of a quantum channel $\mathcal{T} : \mathcal{M}_2 \rightarrow \mathcal{M}_2$ with respect to the Pauli basis. A general channel can be written as a matrix $\hat{\mathcal{T}} \in \mathcal{M}_4$.

$$\hat{\mathcal{T}} = \begin{pmatrix} 1 & 0 \\ \mathbf{t} & \mathbf{L} \end{pmatrix}. \quad (2.91)$$

The channel acts on a density matrix via $\mathcal{T}(\rho) = \mathcal{T}(\frac{1}{2}(\mathbb{1} + \mathbf{r} \cdot \boldsymbol{\Sigma})) = \frac{1}{2}(\mathbb{1} + (\mathbf{t} + \mathbf{L}\mathbf{r}) \cdot \boldsymbol{\Sigma})$. It can be shown, that the map \mathcal{T} is unital if and only if $\mathbf{t} = 0$. Let us now consider the optimization for Cheeger's constant h as given in Lemma (43). Given the constraint, we have to vary all one dimensional projectors $\Pi_A = |\psi\rangle\langle\psi|$ with $\| |\psi\rangle \|_2 = 1$, so that

$$h = \min_{|\psi\rangle \in \mathbb{C}^2} \text{tr} [(\mathbb{1} - |\psi\rangle\langle\psi|) \mathcal{S}(|\psi\rangle\langle\psi|)]. \quad (2.92)$$

The symmetrized map S of the unital channel \mathcal{T} , with $\mathbf{t} = 0$, now assumes the matrix representation,

$$\hat{S} = \begin{pmatrix} 1 & 0 \\ 0 & \mathbf{L}^\dagger \mathbf{L} \end{pmatrix}. \quad (2.93)$$

Furthermore note, that any projector $|\psi\rangle\langle\psi| \in \mathcal{S}_2$ can be parametrized via a Bloch vector $\mathbf{a} \in \mathbb{R}^3$ that obeys $\|\mathbf{a}\|_2 = 1$. The minimization for Cheeger's constant reduces therefore to

$$h = \min_{\|\mathbf{a}\|_2=1} 1 - \langle \mathbf{a} | \mathbf{L}^\dagger \mathbf{L} | \mathbf{a} \rangle, \quad (2.94)$$

where $\langle \mathbf{a} | \mathbf{b} \rangle$ denotes the canonical scalar product in \mathbb{R}^3 . The minimum is attained when \mathbf{a} is the eigenvector associated with the largest eigenvalue s_1^2 of the matrix $\mathbf{L}^\dagger \mathbf{L}$. Hence for an arbitrary single qubit unital channel, Cheeger's constant is given by $h = 1 - s_1^2$, where s_1 is the largest singular value of the matrix \mathbf{L} and hence the second largest singular value of the channel \mathcal{T} . We see that the conductance bound as stated in Lemma 43 is indeed satisfied, since

$$2s_1^2 - 1 \leq s_1^2 \leq \frac{1}{2}(1 + s_1^2). \quad (2.95)$$

Chapter 3

Quantum Metropolis sampling

Synopsis:

In this chapter we propose a direct quantum generalization of the classical Metropolis algorithm and show how a single iteration of the algorithm can be implemented in polynomial time on a quantum computer. The original motivation to build a quantum computer came from Feynman [12], who envisaged a machine capable of simulating generic quantum mechanical systems, a task that is believed to be intractable for classical computers. Such a machine would have a wide range of applications in the simulation of many-body quantum physics, including condensed matter physics, chemistry, and high energy physics. Part of Feynman's challenge was met by Lloyd [13], who showed how to approximately decompose the time-evolution operator of interacting quantum particles into a short sequence of elementary gates, suitable for operation on a quantum computer. However, this left open the problem of how to simulate the equilibrium and static properties of quantum systems. This requires the preparation of ground and Gibbs states on a quantum computer. For classical systems, this problem is solved by the ubiquitous Metropolis algorithm [7], a method that basically acquired a monopoly for the simulation of interacting particles. Here, we demonstrate how to implement a quantum version of the Metropolis algorithm on a quantum computer. This algorithm permits to sample directly from the eigenstates of the Hamiltonian and thus evades the sign problem present in classical simulations and can be used to prepare ground and thermal states of generic quantum many-body systems, both bosonic and fermionic. A small scale implementation of this algorithm can already be achieved with today's technology.

Based on:

K. Temme, T.J. Osborne, K.G. Vollbrecht, D. Poulin and F. Verstraete,
Nature (accepted), *e-print arXiv: 0911.3635*, (2009)

3.1 Summary of the algorithm

In this section, we present a sketch of how the quantum Metropolis algorithm works. Details and generalizations will be worked out in the following sections.

Ground states could in principle be prepared using the quantum phase estimation algorithm [76, 77], but this method is in general not scalable, because it requires a variational state with a large overlap with the ground state. Methods are known for systems with frustration-free interactions [78] or systems that are adiabatically connected to trivial Hamiltonians [79], but such conditions are not generically satisfied. Terhal and DiVincenzo [80] suggested two approaches of how a quantum computer could sample from the thermal state of a system. The first suggestion is also related to the Metropolis rule, yet left open the problem of how one could get around the no-cloning result and could construct local updates, which can be rejected. This shortcoming immediately leads to an exponential running time of the algorithm, as already discussed in the said paper. The second approach shows, how thermal states can be prepared by simulating the system's interaction with a heat bath. However, this procedure seems to produce rather large errors when run on a quantum computer with finite resources, and a precise framework to describe these errors seems to be out of reach. Moreover, certain systems like polymers [81], binary mixtures [82] and critical spin chains [83, 84] experience extremely slow relaxation, when put into interaction with a heat bath. The Metropolis dynamics solve this problem by allowing transformations that are not physically achievable, speeding up relaxation by many orders of magnitude and bridging the microscopic and relaxation time scales; this freedom is to a large extent responsible for the tremendous empirical success of the Metropolis method. It is therefore desirable to have generalization of the Metropolis algorithm for quantum Hamiltonians.

To set the stage for the quantum Metropolis algorithm, let us briefly recall the classical Metropolis algorithm we introduced in chapter 1. We can assume for definiteness that the system is composed of n two-level particles, i.e., Ising spins. A lattice of 100 spins has 2^{100} different configurations, so it is inconceivable to average them all. The key insight of Metropolis *et. al.* was to set up a rapidly mixing *Markov chain* obeying detailed balance that samples from the configurations with the most significant probabilities. This can be achieved by randomly transforming an initial configuration to a new one (e.g. by flipping a randomly selected spin): if the energy of the new configuration is lower than the original, we retain the move, but if the energy is larger, we only retain the move with probability $\exp(\beta(E_{old} - E_{new}))$, where E is the energy of the configurations and β the inverse temperature.

The challenge we address is to set up a similar process in the quantum case, i.e., to initiate an ergodic random walk on the eigenstates of a given quantum Hamiltonian with the appropriate Boltzmann weights. In analogy to a spin flip, the random walk can be realized by a random local unitary, and the *move* should be accepted or rejected following the Metropolis rule. There are, however, three obvious complications:

1. We do not know what the eigenvectors of the Hamiltonian are (this is precisely one of the problems we want to solve).
2. Certain operations, such as energy measurements, are fundamentally irreversible in quantum mechanics, but the Metropolis method requires rejecting, hence undoing, certain transformations.
3. One has to devise a criterion that proves that the fixed point of the quantum random walk is the Gibbs state.

To address the first obstacle, we assume for simplicity that the Hamiltonian has non-degenerate eigenvalues E_i , and denote the corresponding eigenvectors $|\psi_i\rangle$. In the following sections, it is shown that those conditions are unnecessary. We can make use of the phase estimation algorithm [85, 86, 76, 87] to prepare a random energy eigenstate and measure the energy of a given eigenstate. Then, each quantum Metropolis step (depicted in Fig. 3.1) takes as input an energy eigenstate $|\psi_i\rangle$ with known energy E_i , and applies a random local unitary transformation C , creating the superposition $C|\psi_i\rangle = \sum_k x_k^i |\psi_k\rangle$. C could be a bit-flip at a random location like in the classical setting, or some other simple transformation. The phase estimation algorithm is now used in a coherent way, producing $\sum_k x_k^i |\psi_k\rangle |E_k\rangle$. At this point, we could measure the second register to read out the energy E_k and accept or reject the move following the Metropolis prescription. However, such an energy measurement would involve an irreversible collapse of the wave function, which will make it impossible to return to the original configuration in the case of a reject step.

Classically, we get around this second obstacle by keeping a *copy* of the original configuration in the computer's memory, so a rejected move can be easily undone. Unfortunately, this solution is ruled out in the quantum setting by the no-cloning Theorem [88]. The key to the solution is to engineer a measurement that reveals as little information as possible about the new state, and therefore only slightly disturbs it. This can be achieved by a measurement that only reveals one bit of information—accept or reject the move—rather than a full energy measurement. The circuit that generates this binary measurement is shown at Fig. 3.1. It transforms the initial state $|\psi_i\rangle$ into

$$\underbrace{\sum_k x_k^i \sqrt{f_k^i} |\psi_k\rangle |E_i\rangle |E_k\rangle |1\rangle}_{|\psi_i^+\rangle} + \underbrace{\sum_k x_k^i \sqrt{1 - f_k^i} |\psi_k\rangle |E_i\rangle |E_k\rangle |0\rangle}_{|\psi_i^-\rangle} \quad (3.1)$$

where $f_k^i = \min(1, \exp(-\beta(E_k - E_i)))$. The state can be seen as a coherent superposition of accepting the update or rejecting it. The amplitudes $x_k^i \sqrt{f_k^i}$ correspond exactly to the transition probabilities $|x_k^i|^2 f_k^i$ of the classical Metropolis rule. The measurement is completed by measuring the last qubit in the computational basis. The outcome $|1\rangle$ will project the other registers in the state $|\psi_i^+\rangle$. Upon obtaining this outcome, we can measure the second register to

learn the new energy E_k and use the resulting energy eigenstate as input to the next Metropolis step.

A measurement outcome $|0\rangle$ signals that the move must be rejected, so we must return to the input state $|\psi_i\rangle$. As $|\psi_i^+\rangle$ is orthogonal to $|\psi_i^-\rangle$ we actually work in a simple 2-dimensional subspace, i.e. a qubit. In such a case, it is possible to go back to the initial state by an iterative scheme similar to the one employed by Marriott and Watrous in the context of quantum Merlin Arthur amplification [89]. The circuit implementing this process is shown in Fig. 3.2. In essence, it repeatedly implements two binary measurements. The first is the one described in the previous paragraph. The second one, after a basis change, determines if the computer is in the eigenstate $|\psi_i\rangle$ or not. A positive outcome to the latter measurement implies that we have returned to the input state, completing the rejection; in the case of a negative outcome, we repeat both measurements. Every sequence of these two measurements has a constant probability of achieving the rejection, so repeating recursively yields a success probability exponentially close to 1.

The quantum Metropolis algorithm can be used to generate a sequence of m states $|\phi_j\rangle$, $j = 1, \dots, m$ that reproduce the statistical averages of the thermal state $\rho_G = e^{-\beta H} / \mathcal{Z}$ for any observable X :

$$\frac{1}{m} \sum_{j=1}^m \langle \phi_j | X | \phi_j \rangle = \text{Tr} X \rho + \mathcal{O}(1/\sqrt{m}). \quad (3.2)$$

To show that the fixed point of the quantum random walk is the Gibbs state, we developed the theory of *quantum detailed balance* in section 2.4. We choose a specific inversion, with $k(w) = 1/\sqrt{w}$, cf. (2.9), that, according to Proposition 41, gives rise to the following condition. Let $\{|\psi_i\rangle\}$ be a complete basis of the physical Hilbert space and let $\{p_i\}$ be a probability distribution on this basis. Assume that a completely positive map \mathcal{E} obeys the condition

$$\sqrt{p_n p_m} \langle \psi_i | \mathcal{E}(|\psi_n\rangle\langle\psi_m|) | \psi_j \rangle = \sqrt{p_i p_j} \langle \psi_m | \mathcal{E}(|\psi_j\rangle\langle\psi_i|) | \psi_n \rangle. \quad (3.3)$$

Then $\sigma = \sum_i p_i |\psi_i\rangle\langle\psi_i|$ is a fixed point of \mathcal{E} .

The *quantum detailed balance* condition only ensures that the thermal state ρ_G is a possible fixed point of the quantum Metropolis algorithm. The uniqueness of this fixed point as well as the convergence rate to it depend on the choice of the set of random unitaries $\{C\}$. If the set of moves are chosen such that the map \mathcal{E} is primitive, cf. Theorem 18 and [31], the uniqueness of the fixed point is ensured. This condition can be satisfied by choosing $\{C\}$ to be a universal gate set [15]. The Metropolis step obeys the *quantum detailed balance* condition, if the probability of applying a specific C is equal to the probability of applying its conjugate C^\dagger . This can be seen as the quantum analogue of the classical symmetry condition for the update probability. In some cases it even suffices to just apply the same local unitary C at every step of the algorithm (see Fig. 3.4). In this case, the single unitary C has to be Hermitian and has to ensure ergodicity. The local unitary can be seen to induce ‘non-local’ transitions between the

eigenstates because it is followed by a phase estimation procedure. Like the classical Metropolis algorithm, the quantum Metropolis algorithm is not expected to reach the ground state of an arbitrary Hamiltonian in polynomial time. The ability to prepare the ground state of a general Hamiltonian in polynomial time would allow to solve QMA-complete problems. However, as a rule of thumb it always seems possible to define an update strategy for which the Metropolis algorithm thermalizes efficiently if the physical system thermalizes in polynomial time. There are no obvious reasons why the same should not be true for the quantum Metropolis algorithm. It also inherits all the flexibility and versatility of the classical method, leading, for instance, to a quantum generalization of simulated annealing [10].

3.2 Implementation

In this section we describe how to efficiently implement the quantum gates required by our algorithm on a quantum computer. As is now standard in the literature, we assume that we can implement single-qubit operations, measurements of the observables σ^α , and elementary two-qubit gates, such as the CNOT gate with unit cost. The core element of the algorithm is the quantum phase estimation procedure [85, 86, 76, 87]. This procedure requires a means to simulate the unitary dynamics e^{-itH} generated by a k -particle Hamiltonian H . We assume that $H = \sum_{j=1}^m H_j$ can be written as the sum of m terms, each of which is easy to simulate on a quantum computer. The best way to do this follows the method described by Berry *et. al.* [87] and by Childs [90]: This procedure provides a simulation of the dynamics e^{-itH} for $0 \leq t \leq t_0$ using a quantum circuit of length T_H , where

$$T_H \leq 2c m^2 \tau e^{2\sqrt{\ln(5) \ln(m\tau/\epsilon_H)}}, \quad (3.4)$$

and c is a constant, m denotes the number of summands in H , $\tau = \|H\|t_0$, and ϵ_H is the desired error. Now, for a typical Hamiltonian encountered in condensed matter physics or quantum chemistry, the number of terms m scales as a polynomial with N , the number of particles. Thus the length T_H of the circuit scales better than any power of $1/\epsilon_H$ and is almost linear with t_0 and scales slightly worse than a polynomial in N . Thus we can simulate e^{-itH} for a length of time $t \sim p(N)$ and to precision $\epsilon_H \sim 1/q(N)$ with an effort scaling polynomially with N , where p and q are polynomials.

Our algorithm requires a method to measure the observable H . This can be done by making use of the quantum phase estimation, which is a discretization of von Neumann's prescription to measure a Hermitian observable.

First adjoin an ancilla – the *pointer* – which is a continuous quantum variable initialized in the state $|0\rangle$, so that the system+pointer is initialized in the state $|\psi\rangle|0\rangle$, where $|\psi\rangle$ is the initial state of the system. Then evolve according to the new Hamiltonian $K = H \otimes \hat{p}$ for a time t , so the evolution is given by

$$e^{-itH \otimes \hat{p}} = \sum_{j=1}^{2^N} |\psi_j\rangle\langle\psi_j| \otimes e^{-itE_j\hat{p}}. \quad (3.5)$$

Supposing that $|\psi\rangle$ is an eigenstate $|\psi_j\rangle$ of H we find that the system evolves to

$$e^{-itH \otimes \hat{p}}|\psi_j\rangle|0\rangle = |\psi_j\rangle|x = tE_j\rangle. \quad (3.6)$$

A measurement of the position of the pointer with sufficiently high accuracy will provide an approximation to E_j .

To carry out the above operation efficiently on a quantum computer we discretize the pointer using r qubits, replacing the continuous quantum variable with a 2^r -dimensional space,

where the computational basis states $|z\rangle$ of the pointer represent the basis of *momentum* eigenstates of the original continuous quantum variable. The label z is the binary representation of the integers 0 through $2^r - 1$. In this representation the discretization of the momentum operator becomes

$$\hat{p} = \sum_{j=1}^r 2^{-j} \frac{\mathbb{1} - \sigma_j^z}{2}. \quad (3.7)$$

With this normalization $\hat{p}|z\rangle = \frac{z}{2^r}|z\rangle$. Now the discretized Hamiltonian $K = H \otimes \hat{p}$ is a sum of terms involving at most $k + 1$ particles, if H is a k -particle system. Thus we can simulate the dynamics of K using the method described above.

In terms of the momentum eigenbasis the initial (discretized) state of the pointer is written

$$|x = 0\rangle = \frac{1}{2^{r/2}} \sum_{z=0}^{2^r-1} |z\rangle. \quad (3.8)$$

This state can be prepared efficiently on quantum computer by first initializing the qubits of the pointer in the state $|0\rangle \cdots |0\rangle$ and applying an (inverse) quantum Fourier transform. The discretized evolution of the system+pointer now can be written

$$e^{-itH \otimes \hat{p}} |\psi_j\rangle |x = 0\rangle = \frac{1}{2^{r/2}} \sum_{z=0}^{2^r-1} e^{-iE_j z t / 2^r} |\psi_j\rangle |z\rangle. \quad (3.9)$$

Performing an inverse quantum Fourier transform on the pointer leaves the system in the state $|\psi_j\rangle \otimes |\phi\rangle$, where

$$|\phi\rangle = \sum_{x=0}^{2^r-1} \left(\frac{1}{2^r} \sum_{z=0}^{2^r-1} e^{\frac{2\pi i}{2^r} \left(x - \frac{E_j t}{2\pi}\right) z} \right) |x\rangle. \quad (3.10)$$

Thus we find that

$$|\phi\rangle = \sum_{x=0}^{2^r-1} f(E_j, x) |x\rangle, \quad (3.11)$$

where

$$|f(E_j, x)|^2 = \frac{1}{4^r} \frac{\sin^2 \left(\pi \left(x - \frac{E_j t}{2\pi} \right) \right)}{\sin^2 \left(\frac{\pi}{2^r} \left(x - \frac{E_j t}{2\pi} \right) \right)}, \quad (3.12)$$

which is strongly peaked near $x = \lfloor \frac{E_j t}{2\pi} \rfloor$. To ensure that there are no overflow errors we need to choose $t < \frac{2\pi}{\|H\|}$. (We assume here, for simplicity, that $H \geq 0$.)

It is easy to see that actually performing the simulation of K for $t = 1$ using the method of [87] requires a product of r simulations of the evolution according to $\frac{1}{2^r} H \otimes \frac{\mathbb{1} - \sigma_k^z}{2}$ for $1, 2, 2^2, \dots, 2^{r-1}$ units of time, respectively.

Thus far we have only discussed the action of the quantum phase estimation procedure on a predetermined input state of the form $|\psi_i\rangle |0\rangle$. For the algorithm, however, it is important that we apply the quantum phase estimation procedure subsequently on the same register. Since the

phase estimation algorithm does not act deterministically, we find that the pointer register can return to a different value $|y \neq 0\rangle$ after the procedure. It is therefore important to understand how the full unitary Φ of the complete phase estimation procedure acts on the system. The full unitary consists of three individual steps: First we perform a quantum Fourier transform on the pointer register. This is followed by the simulation of the Hamiltonian $K = H \otimes \hat{p}$. In the last step we then apply an inverse quantum Fourier transform. The unitary acts on the N -qubit register that stores the state of the simulated system and a single r -qubit ancilla register that is used to read out the phase information. We write

$$\Phi = \sum_{y=0}^{2^r-1} \sum_{x=0}^{2^r-1} M_x^y \otimes |x\rangle\langle y|, \quad \text{where} \quad M_x^y = \sum_{j=1}^{2^N} f(E_j, x - y) |\psi_j\rangle\langle\psi_j|. \quad (3.13)$$

Note that the function

$$f(E_j, x - y) = \frac{1}{2^r} \frac{e^{i\pi(x - \frac{E_j t}{2\pi} - y)}}{e^{i\frac{\pi}{2^r}(x - \frac{E_j t}{2\pi} - y)}} \left(\frac{\sin\left(\pi\left(x - \frac{E_j t}{2\pi} - y\right)\right)}{\sin\left(\frac{\pi}{2^r}\left(x - \frac{E_j t}{2\pi} - y\right)\right)} \right) \quad (3.14)$$

is complex valued. The operators $M_x^{y=0}$ constitute the measurement generated on the system state by the phase estimation procedure. The label x denotes the r -bit approximation to the energy generated by the phase estimation procedure, whereas y corresponds to the initial value of the ancilla register. Note, that since we only resolve the energy to r bits of precision, it is not correct to suppose, that quantum phase estimation always outputs the closest r -bit approximation to the energy of the eigenstate. Rather, it outputs a random energy distributed according to Eq. (3.12), sharply peaked around the exact energy. Thus, what we described earlier as projectors onto energy bins are not truly von Neumann projective measurements, but rather correspond to generalized (positive operator valued measure, POVM) measurements on the system. It can easily be verified, that the operators M_x^y constitute a general POVM and are only projectors, when either the pointer register size is infinite $r \rightarrow \infty$, or the energies of the Hamiltonian are spaced at integer levels. For the construction of the measurement we assumed earlier, that the pointer register is always initialized into the state $|0\rangle = |0, \dots, 0\rangle$. However, in our algorithm it can happen that the pointer register will differ from that state prior to the application of a quantum phase estimation procedure due to imperfections. Nevertheless, due to (3.14) it becomes clear that the estimate x of the eigenvalue E_i only gets shifted by an amount of y , if the ancilla register is initially in the state $|y\rangle$.

The median method The distribution $|f(E_j, x)|^2$ can be sharpened by employing a method developed in [91]: the idea is to adjoin $\eta + 1$ separate pointers, each comprising r qubits, and to perform quantum phase estimation η times on the system using each of the first η pointer systems in turn for the readout. Then the *median* of the results in the η pointers is computed

in a coherent way and written into the $(\eta + 1)$ th pointer. The probability that the median value deviates from the true energy by more than 2^{-r} is less than $2^{-\eta}$ [91]. Given an eigenstate of H , this leaves two possible phase estimation outcomes, corresponding to the r -bit energy values directly below and directly above the true energy. Hence, the high confidence phase estimation algorithm acts as

$$|\psi_i\rangle|0\rangle \rightarrow |\psi_i\rangle (\alpha_{E_i}(\lfloor E_i \rfloor) |\lfloor E_i \rfloor\rangle + \alpha_{E_i}(\lceil E_i \rceil) |\lceil E_i \rceil\rangle) + \mathcal{O}(2^{-\eta}), \quad (3.15)$$

where $|\alpha_{E_i}(\lfloor E_i \rfloor)|^2 + |\alpha_{E_i}(\lceil E_i \rceil)|^2 \approx 1$ and $\lfloor E_i \rfloor$ and $\lceil E_i \rceil$ are the two closest r -bit approximations to E_i . Despite this improvement, it is not possible to make the outcome of the quantum phase estimation procedure deterministic. In the worst case, where the exact energy for a given eigenstate falls exactly between two r -bit values, the two measurement outcomes will be equally likely. So the enhanced quantum phase estimation algorithm still implements a non-projective measurement on the system just like the standard procedure. In either case, we will denote the unitary corresponding to the quantum phase estimation algorithm by Φ and write $M_x^y = \sum_{j=1}^{2^N} \alpha_{E_j}(x-y) |\psi_j\rangle \langle \psi_j|$, when we refer to the high confidence phase estimation procedure.

3.3 Description of the quantum Metropolis algorithm

In this section, we provide a more elaborate description of the quantum Metropolis algorithm. The fundamental building block is the quantum phase estimation algorithm (see section 3.2); throughout this section we assume that the phase estimation algorithm works perfectly, i.e. given an eigenstate $|\psi_i\rangle$ of the Hamiltonian H with energy E_i , we assume that the quantum phase estimation circuit Φ implements the transformation

$$|\psi_i\rangle|0\rangle \rightarrow |\psi_i\rangle|E_i\rangle,$$

where E_i is encoded with r bits of precision. The fact that errors inevitably occur during quantum phase estimation will be dealt with in section 3.7. The algorithm runs through a number of steps 0..4 and, just as in the classical case, the total number of iterations of this procedure is related to the autocorrelation times of the underlying stochastic map. As analyzed in the section 3.6, this procedure obeys the quantum detailed balance condition and hence allows to sample from the Gibbs state. The different steps are also depicted in Fig. 3.3.

0 Initialization

Initialize the quantum computer in a convenient state, e.g. $|00\dots 0\rangle$. We need 4 quantum registers in total. The first one will encode the quantum states of the simulated system, while the other 3 registers are ancillas that will be traced out after every individual Metropolis step. The second register consists of r qubits and encodes the energy of

the incoming quantum state with r bits of precision (bottom register in Fig. 3.1a). The third register is the one used to implement the quantum phase estimation algorithm, also with r qubits (top register 3.1a). The fourth register is a single qubit that will provide the randomness for accepting or rejecting the Metropolis step.

1 State preparation:

Re-initialize the three ancilla registers and implement the quantum phase estimation based circuit depicted in Fig. 3.1a followed by a measurement of the second register. This prepares an eigenstate $|\psi_i\rangle$ with energy E_i and associated energy register $|E_i\rangle$. The upper ancillas are left in the state $|0\rangle^r$ as we assumed perfect phase estimation. The global state is now

$$|\psi_i\rangle|E_i\rangle|0\rangle|0\rangle$$

2 Propose update:

The next step is depicted in Fig. 3.1b. Assume that we have defined a set of unitaries $\mathcal{C} = \{C\}$ that can be implemented efficiently; those will correspond to the proposed moves or updates of the algorithm, just like one does for instance spin flips in the case of classical Monte Carlo. Just as in the classical case, the exact choice of this set of unitaries does not really matter as long as it is rich enough to generate all possible transitions; the convergence time will, however, depend on the particular choice of moves. The unitary C is drawn randomly from the set \mathcal{C} according to some probability measure $d\mu(C)$. It is only necessary that the probability of choosing a C is equal to the probability of choosing C^\dagger , i.e. $d\mu(C) = d\mu(C^\dagger)$, as this is dictated by the requirement that the process obeys detailed balance, cf. section 3.6.

The new state can be written as a superposition of the eigenstates:

$$C|\psi_i\rangle = \sum_k x_k^i |\psi_k\rangle$$

Implement the coherent quantum phase estimation step specified in Fig. 3.1b, which results in the state

$$\sum_k x_k^i |\psi_k\rangle \rightarrow \sum_k x_k^i |\psi_k\rangle |E_i\rangle |E_k\rangle |0\rangle.$$

Note that E_k is only encoded with a precision of r bits, so that in practice there will be a lot of degeneracies.

Finally, implement the unitary $W(E_k, E_i)$ (Fig. Fig. 3.1b) which is a one-qubit operation conditioned on the value of the 2 energy registers:

$$W(E_k, E_i) = \begin{pmatrix} \sqrt{1-f_{ik}} & \sqrt{f_{ik}} \\ \sqrt{f_{ik}} & -\sqrt{1-f_{ik}} \end{pmatrix} \quad (3.16)$$

$$f_{ik} = \min(1, \exp(-\beta(E_k - E_i))). \quad (3.17)$$

The system is now in the state

$$\sum_k x_k^i \sqrt{f_k^i} |\psi_k\rangle |E_i\rangle |E_k\rangle |1\rangle + \sum_k x_k^i \sqrt{1-f_k^i} |\psi_k\rangle |E_i\rangle |E_k\rangle |0\rangle.$$

For later reference, the product of the three unitaries C , the phase estimation step, and W is called U (see Fig. 3.1b).

3 Accept instance:

Measure the single ancilla qubit in the computational basis. A measurement outcome 1 corresponds to an acceptance of the move and collapses the state into

$$\sum_k x_k^i \sqrt{f_k^i} |\psi_k\rangle |E_i\rangle |E_k\rangle |1\rangle.$$

In the case of this accept move, we can next measure the third register which prepares a new eigenstate $|\psi_k\rangle$, and follow that by an inverse quantum phase estimation step. This leads to the state

$$|\psi_k\rangle |E_i\rangle |0\rangle |1\rangle$$

with probability proportional to $\left|x_k^i \sqrt{f_k^i}\right|^2$. This state will be the input for the next step in the iteration of the Metropolis algorithm: go back to step 1 for this next iteration. Note that the sequence $E \rightarrow Q_1 \rightarrow L$ depicted in Fig. 3.3 exactly corresponds to this sequence of gates.

A measurement $|0\rangle$ in the single ancilla qubit signals a reject of the update. In this case, first apply the gate U^\dagger , and then go to step 4.

4 Reject instance:

Let us first define the Hermitian projectors Q_0 and Q_1 , made up of the gates defined in step 2 – 3 including the measurement on the ancilla:

$$Q_0 = U^\dagger (\mathbb{1} \otimes \mathbb{1} \otimes \mathbb{1} \otimes |0\rangle\langle 0|) U$$

$$Q_1 = U^\dagger (\mathbb{1} \otimes \mathbb{1} \otimes \mathbb{1} \otimes |1\rangle\langle 1|) U$$

Let us also define the Hermitian projectors P_0 and P_1 as

$$P_0 = \sum_i \sum_{E_\alpha \neq E_i} |\psi_\alpha\rangle\langle\psi_\alpha| \otimes |E_i\rangle\langle E_i| \otimes \mathbb{1} \otimes \mathbb{1}$$

$$P_1 = \sum_i \sum_{E_\alpha = E_i} |\psi_\alpha\rangle\langle\psi_\alpha| \otimes |E_i\rangle\langle E_i| \otimes \mathbb{1} \otimes \mathbb{1}$$

Here equality (or inequality) means that the first r bits of the energies do (not) coincide. This measurement P_α can easily be implemented by a phase estimation step depicted in Fig. 3.1c.

The fourth step of the algorithm now consists of a sequence of measurements (see Fig. 3.2). First we implement the von Neumann measurement defined by P_α . If the outcome is P_1 , then we managed to prepare a new eigenstate $|\psi_\alpha\rangle$ with the same energy as the initial one $|\psi_i\rangle$, and therefore succeeded in undoing the measurement. Go to step 1. If the outcome is P_0 , we do the von Neumann measurement Q_α . Independent of the outcome, we again measure P_α , and if the outcome is P_1 , we achieved our goal, otherwise we continue the recursion (see Fig. 3.3). It happens that the probability of failure decreases exponentially with the number of iterations (see section 3.4.1), and therefore we have a very good probability of achieving our goal. In the rare occasion when we do not converge after a pre-specified number of steps, we abort the whole Monte Carlo simulation and start all over.

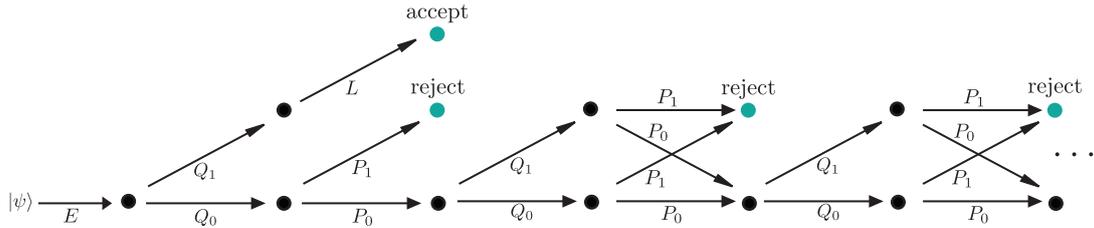


Figure 3.3: Given an input state $|\psi\rangle$, we first perform a quantum phase estimation to collapse to an eigenstate with known energy E . This graph represents the plan of action conditioned on the different measurement outcomes of the binary P and Q measurements. Each node in the graph corresponds to an intermediate state in the algorithm. One iteration of the map is completed when we reach one of the final leafs labelled by either accept or reject. The sequence $E \rightarrow Q_1 \rightarrow L$ corresponds to accepting the update, all other leafs to a rejection. The individual operations are defined in section 3.5

This finishes the description of the steps in the algorithm. A single iteration of the quantum Metropolis algorithm corresponds to a single application of the Metropolis tcp-map \mathcal{E} . This

map will be defined in section 3.5. Finally, let us briefly discuss how to implement the unitary gate $W(E_k, E_i)$. This is a single qubit unitary conditioned on two energy registers. That this conditional unitary can be performed efficiently follows by observing that one can efficiently compute the angle $\theta = \arcsin(e^{\frac{\beta}{2}(E_k - E_i)})$ into a scratchpad register, conditionally rotate the answer qubit by this angle, and uncompute θ .

3.4 Running time of the quantum Metropolis algorithm

Let us discuss the runtime scaling of the full Metropolis algorithm. In general, there are three types of error one has to deal with when we consider the the runtime scaling of the algorithm.

First, we are dealing with a Markov chain and hence there is an associated mixing error ϵ^{mix} . We have introduced the mixing error and the associated mixing time in chapter 2. The mixing error of the Markov chain is defined with respect to trace norm distance, as $\|\mathcal{E}^{m_{mix}}(\rho) - \sigma^*\|_{\text{tr}} \leq \epsilon^{mix}$. Where m_{mix} denotes the mixing time, i.e. the number of times the completely positive map has to be applied starting from an initial state ρ to be ϵ^{mix} close to the steady state σ^* of the Markov chain. The mixing time is determined by the the gap $\Delta = 1 - \lambda_1$ between the two largest eigenvalues in magnitude of the corresponding completely positive map, if the map obeys quantum detailed balance. We have shown in section 2.3 and in [92] that the trace norm is bounded by

$$\|\mathcal{E}^m(\rho) - \sigma^*\|_{\text{tr}} \leq C_{\text{exp}} (1 - \Delta)^m, \quad (3.18)$$

For some constant $C_{\text{exp}} = \max_{\rho} \chi_k^2(\rho, \sigma^*)$, which is typically in the order of the total Hilbert space dimensions, i.e. in our case $C_{\text{exp}} = \mathcal{O}(2^N)$. The runtime, or the mixing time, scales therefore as

$$m_{mix} \geq \mathcal{O}\left(\frac{\ln(1/\epsilon^{mix}) + N}{\Delta}\right). \quad (3.19)$$

Just as for classical stochastic maps one needs to prove that the gap is bounded by a polynomial in the system size for each problem instance individually to ensure that the chain is rapidly mixing. It is generally believed, that to prove rapid mixing for a realistic Hamiltonian is hard. However, the convergence rate of the classical Metropolis algorithm is often good for many realistic physical systems and it is conceivable that the same will be true for the quantum Metropolis algorithm as well. In the following, cf. section 3.4.2, we will provide a simple example system for which the scaling of the gap can be estimated numerically.

The second type of imperfection relates to the fact, that the reject part of a local move cannot be implemented deterministically. However, we will show, cf. section 3.4.1, that this probability can be made arbitrary small by increasing the number of iterations in the reject move. For all realistic applications one would choose a fixed n^* so that one only attempts to perform $n \leq n^*$ reject moves before discarding the sample. We want to achieve an overall

success probability of preparing a valid sample that is bounded by some constant c . What do we mean by that? As already stated the Metropolis algorithm allows one to sample from the eigenstates $|\psi_i\rangle$ with a given probability $p_i \propto \exp(-\beta E_i)$. Since our reject procedure can only be implemented probabilistically we have to choose a fixed number of times n^* we try to reject a proposed update. The probability of failure $p_{\text{fail}}(n)$ of rejecting a proposed update after n steps is bounded by $p_{\text{fail}}(n) \leq \frac{1}{2e(n+1)}$, see (3.30). For the algorithm to work, we want the algorithm to produce a sample after m_{mix} applications of the map \mathcal{E} with a probability that is larger than a constant c . Hence the probability of failure after m_{mix} steps should obey $(1 - p_{\text{fail}}(n^*))^{m_{\text{mix}}} \geq c$. This condition is met if we choose

$$n^* > \frac{m_{\text{mix}}}{2e(1-c)} \quad (3.20)$$

This means, that we have to implement for each Metropolis step at most n^* measurements P_i and Q_i , before we discard the sample and start over again. Note, that this is a very loose upper bound for the actual number of reject attempts, since the probability of failure actually decays exponentially in n , with some unknown constant ensured to be smaller than unity.

The third error relates to the fact that we are implementing the algorithm on a quantum computer with finite resources, e.g. a finite register to store the energy eigenvalues in the phase estimation procedure. This leads to a modification of the completely positive map \mathcal{E} , whose fixed point σ^* now deviates from the Gibbs state ρ_G by $\|\sigma^* - \rho_G\|_{\text{tr}} \leq \epsilon^*$. This error will be discussed in section 3.7.

3.4.1 The rejection procedure

Let us discuss the convergence of the reject step more closely. As already explained, the algorithm should prepare a new state with the same energy as the original one E_i in the case of a reject move. As shown in Fig. 3.3, we will do this by repeating a sequence of two different binary measurements P_i and Q_i . The recursion stops, whenever the measurement outcome P_1 is obtained, where P_1 is the projector on the subspace of energy E_i . Note that it is crucial for the algorithm that the initially prepared state $E|\psi_i\rangle|0^{2r+1}\rangle$ is an eigenstate of the projection P_1 . This is indeed the case, even if we take into account the fluctuations in the quantum phase estimation step discussed in the section 3.7: the error that is generated by the fluctuations of the pointer variable can be accounted for if we verify the equality of the energy in P only up to $\tilde{r} < r$ bits of precision. This allows to enlarge the eigenspace of P with approximate energy E_i , encompassing the fluctuations of the pointer variable.

Here we will calculate the expected running time. The probability of failure to reject the move, given that we start in some state $|\psi_i\rangle$ in the energy E_i subspace, after $n \geq 2$ steps, is given by the probability of measuring P_0 after n subsequent binary measurements, see Fig. 3.3. Note, that the commutator $[P_0 Q_s P_0, P_0 Q_{s'} P_0] = 0$ for all s, s' , therefore the probability of

failure can be cast into the form

$$p_i^{fail}(n) = \sum_{m=0}^n \binom{n}{m} \text{tr} [(P_0 Q_0 P_0)^{n-m} (P_0 Q_1 P_0)^m P_0 Q_0 E] \langle \psi_i | \langle 0^{2r+1} | \otimes | 0^{2r+1} \rangle E Q_0 P_0 (P_0 Q_1 P_0)^m (P_0 Q_0 P_0)^{n-m} | \psi_i \rangle | 0^{2r+1} \rangle. \quad (3.21)$$

The full expression can conveniently be summed to a single term:

$$p_i^{fail}(n) = \langle \psi_i | \langle 0^{2r+1} | E Q_0 P_0 \left[P_0 \left(\sum_{s=0}^1 Q_s P_0 Q_s \right) P_0 \right]^n P_0 Q_0 E | \psi_i \rangle | 0^{2r+1} \rangle. \quad (3.22)$$

We now make use of the following Lemma 44. The key technical reason why it is possible to implement the reject move in the quantum Metropolis algorithm is related to a very special normal form in which two (non-commuting) Hermitian projectors can be brought. The Lemma states, that there is a basis in which both projectors P_i and Q_i are block diagonal.

Lemma 44 (Jordan 1875). *Let P_1 and Q_1 be two projectors of rank(Q_1) = q and rank(P_1) = p on a Hilbert space $\mathcal{H} = \mathbb{C}^k$ with $p + q \leq k$. We assume w.l.o.g, that $q \geq p$. Then there exists a basis of \mathcal{H} in which P_1 and Q_1 can be written in the form*

$$P_1 = \begin{pmatrix} \mathbb{1}_p & 0_{k-p,p} \\ 0_{p,k-p} & 0_{k-p,k-p} \end{pmatrix} \quad (3.23)$$

$$Q_1 = \begin{pmatrix} D_p & \sqrt{D_p(\mathbb{1}_p - D_p)} & 0 & 0 \\ \sqrt{D_p(\mathbb{1}_p - D_p)} & \mathbb{1}_p - D_p & 0 & 0 \\ 0 & 0 & \mathbb{1}_{q-p} & 0 \\ 0 & 0 & 0 & 0_{k-(q+p),k-(q+p)} \end{pmatrix}.$$

Here, D is a $p \times p$ diagonal matrix with real entries $0 \leq d_1 \leq \dots \leq d_p \leq 1$.

PROOF: We can always choose a basis of \mathcal{H} in which the projector P_1 can be written as

$$P_1 = \begin{pmatrix} \mathbb{1}_p & 0_{k-p,p} \\ 0_{p,k-p} & 0_{k-p,k-p} \end{pmatrix}. \quad (3.24)$$

In any basis, a general rank q projector Q_1 can be written in the form

$$Q_1 = \begin{pmatrix} A_{pq} \\ B_{k-p,q} \end{pmatrix} \begin{pmatrix} A_{pq}^\dagger & B_{k-p,q}^\dagger \end{pmatrix} \quad (3.25)$$

Here A_{pq} and $B_{n-p,q}$ are rectangular matrices over \mathbb{C} . We require that Q_1 is a projector: $Q_1^2 = Q_1$ leads to the constraint

$$A_{pq}^\dagger A_{pq} + B_{k-p,q}^\dagger B_{k-p,q} = \mathbb{1}_q. \quad (3.26)$$

We can now choose to perform a singular value decomposition of $A_{pq} = U_A \Sigma_A V_A^\dagger$ and $B_{n-p,q} = U_B \Sigma_B V_B^\dagger$. The projector can thus be written as

$$Q_1 = \begin{pmatrix} U_A & 0 \\ 0 & U_B \end{pmatrix} \begin{pmatrix} \Sigma_A \Sigma_A^\dagger & \Sigma_A V_A^\dagger V_B \Sigma_B \\ \Sigma_B V_B^\dagger V_A \Sigma_A & \Sigma_B \Sigma_B^\dagger \end{pmatrix} \begin{pmatrix} U_A^\dagger & 0 \\ 0 & U_B^\dagger \end{pmatrix} \quad (3.27)$$

Note, that U_A and U_B are p - and $(k-p)$ -dimensional unitary matrices respectively. Therefore the total block diagonal Unitary $U_A \oplus U_B$ leaves the projector P_1 invariant. If we turn to equation (3.26), we see that upon inserting the singular value decomposition, the matrix $V = V_A^\dagger V_B$ must satisfy

$$\Sigma_A^\dagger \Sigma_A = V(\mathbb{1}_q - \Sigma_B^\dagger \Sigma_B)V^\dagger \quad (3.28)$$

Note, that both $\Sigma_A^\dagger \Sigma_A$ and $\mathbb{1}_q - \Sigma_B^\dagger \Sigma_B$ are diagonal matrices, which are according to (3.28) similar. If we assume w.l.o.g, that the singular values are non-degenerate, we conclude that V can only be a permutation matrix. The degenerate case can be covered by a continuity argument. If we define $D = \Sigma_A \Sigma_A^\dagger$ and apply the appropriate permutations to the remaining sub matrices, we are left with the desired expression for Q_1 . \square

Note, that we reuse the same two pointer registers at each phase estimation step in the algorithm. This means that even though a realistic phase estimation procedure does not necessarily act as a projective measurement on the physical subsystem, the binary measurements P_i and Q_i are still projectors on the full circuit. Therefore, Lemma (44) can still be employed, even for a realistic phase estimation procedure. Without loss of generality, we assume that the rank of $rank(P_1) = p$ is smaller than the rank of Q_1 , which is equal to half the dimension of the complete Hilbert space (note that P_1 projects on a single energy subspace). Assume that the unitary U_J brings P and Q to this desired form. This allows us to rewrite (3.22) as $p_i^{\text{fail}}(n) = \langle \psi_i | \langle 0^{2r+1} | E U_J^\dagger D_{\text{fail}}(n) U_J E | \psi_i \rangle | 0^{2r+1} \rangle$ with

$$D_{\text{fail}}(n) = \begin{pmatrix} D(\mathbb{1} - D)(D^2 + (\mathbb{1} - D)^2)^n & -\sqrt{D(\mathbb{1} - D)}(D^2 + (\mathbb{1} - D)^2)^n & 0 & 0 \\ -\sqrt{D(\mathbb{1} - D)}(D^2 + (\mathbb{1} - D)^2)^n & D^2(D^2 + (\mathbb{1} - D)^2)^n & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Here, D denotes a p -dimensional diagonal matrix with only positive entries. Note that the state $U_J E | \psi_i \rangle | 0^{2r+1} \rangle$ has complete support on the projection operator P_1 . That is, as we stated earlier, the state is an eigenstate of P_1 . this means that it only acts on the first upper left block. If we denote by $0 \leq d^* \leq 1$ the diagonal entry of D that gives rise to the largest entry in the upper left block of the matrix $D_{\text{fail}}(n)$, we can bound

$$p_i^{\text{fail}}(n) \leq d^*(1 - d^*)(d^{*2} + (1 - d^*)^2)^n. \quad (3.29)$$

We observe, that the probability of failure decays exponentially in n , for a n -independent d^* . Let us maximize this expression over all possible values of d^* , in order to obtain an absolute upper bound to the failure probability. Defining $x = d^{*2} + (1 - d^*)^2 = 1 - 2d^*(1 - d^*)$, we see that this probability may be bounded by $\frac{1-x}{2}x^n$. This expression is maximized by choosing $x = \frac{n}{n+1}$, for which we have

$$p_{\text{fail}}(n) \leq \frac{1}{2(n+1)} \left(\frac{1}{1 + \frac{1}{n}} \right)^n \approx \frac{1}{2e(n+1)}. \quad (3.30)$$

Hence, choosing $n = O(1/\epsilon)$ recursion steps is sufficient to reduce the probability of failure to below ϵ . We have to choose this ϵ in such a manner, that the probability of failure during a complete cycle of the Metropolis algorithm is bounded by a small constant number.

3.4.2 Mixing time scaling for an example system

We are not able to make a general statement about the mixing times of the algorithm for arbitrary Hamiltonians. In fact, as we argued earlier it is expected that the gap and by that the runtime will scale exponentially in the system size for Hamiltonians that can encode QMA-hard instances. But such a runtime scaling is present also in the classical Metropolis algorithm if one investigates NP-hard problems, such as classical spin glasses. However, the classical Metropolis algorithm is a powerful tool for simulating physical problems, because one observes that the algorithm does indeed converge very fast for most of the physical Hamiltonians. In this section we investigate the scaling of the gap Δ of the quantum Markov chain for a simple example system numerically. Of course, we need to investigate a model that is fit for the simulation on a classical computer, that is we need to be able to compute the eigenbasis of the system analytically in order to describe the transitions between the states by classical means. To this end we investigate the XX -Hamiltonian with an external magnetic field of strength g . The spin Hamiltonian of the one-dimensional chain is given by

$$H = \sum_{n=1}^{N-1} \sigma_n^x \otimes \sigma_{n+1}^x + \sigma_n^y \otimes \sigma_{n+1}^y + g \sum_{n=1}^N \sigma_n^z. \quad (3.31)$$

This Hamiltonian can conveniently be diagonalized by making use of the Jordan-Wigner transformation $a_n^\dagger = -(\otimes_{r=1}^{n-1} \sigma_r^z) \sigma_n^+$, which transforms the Hamiltonian to a free fermionic Hamiltonian. If we furthermore apply a transformation on the single mode level transforming the fermionic creation and annihilation operators as $a_n^\dagger = \sum_{k=1}^N v_n^k c^\dagger(k)$, we can diagonalize the full Hamiltonian so that

$$\begin{aligned} H &= -2 \sum_{n=1}^{N-1} \left(a_{n+1}^\dagger a_n + a_n^\dagger a_{n+1} \right) + 2g \sum_{n=1}^N a_n^\dagger a_n - Ng \\ &= \sum_{k=1}^N 2 \left(g - 2 \cos \left(\frac{k\pi}{N+1} \right) \right) c^\dagger(k) c(k) - Ng, \end{aligned} \quad (3.32)$$

where we have defined $v_n^k = \sqrt{\frac{2}{N+1}} \sin\left(\frac{\pi}{N+1} k n\right)$ as the single particle modes. Since we are dealing with a free fermion system, the eigenstates of the Hamiltonian can be generated from the Fock vacuum $|0\rangle$ as $|k_1, \dots, k_n\rangle = c^\dagger(k_1) \dots c^\dagger(k_n) |0\rangle$ and are nothing but Slater determinants. Here, each of the individual momenta assumes the values $k_i = 1 \dots N$.

We investigate the runtime scaling of the Metropolis for $T = 0$, i.e. preparing the ground state of the above Hamiltonian. To this end, we need to choose an update that can be simulated classically. We choose $C = \sigma_1^x$ that is acting as a single bit flip on the first qubit. The update acts on the eigenbasis via

$$C = \sigma_1^x = \left(a_1^\dagger + a_1\right) = \sum_{k=1}^N \sqrt{\frac{2}{N+1}} \sin\left(\frac{\pi}{N+1} k\right) \left(c^\dagger(k) + c(k)\right). \quad (3.33)$$

We observe, that a single eigenstate $|k_1, \dots, k_n\rangle$ gets mapped to at most $2N$ other eigenstates. Hence, the unitary C is very sparse in the eigenbasis of the Hamiltonian. We can therefore set up a classical Markov process that only keeps track of the $2N$ possible transitions between the eigenstates which we identify by their momentum labels. We have plotted the energy as a function of the number n of applications of the tcp-map in Fig. 3.4(a). From the plots and an exponential fit to the plot we can infer the gap, which scales as $\Delta \propto 1/N$ Fig. 3.4(b). The runtime of the Metropolis algorithm therefore scales linear in the system size for this particular system. The observed linear scaling indicates that, at least in the case of 1D spin chains with nearest - neighbor Hamiltonians, the quantum Metropolis algorithm appears to converge in polynomial time. Proving this remains an interesting open problem.

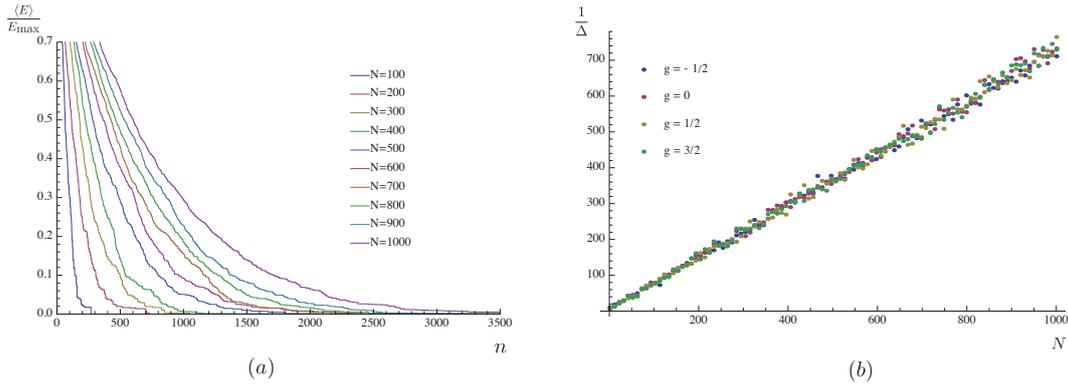


Figure 3.4: Plot (a): This plot depicts the average energy of the state as a function of the number of applications of the Metropolis map at $T = 0$ for spins in a chain with XX -Hamiltonian and $g = 1$ for different system sizes $N = 100 \dots 1000$. The update rule is a single-spin flip σ_1^x . Plot (b): The inverse gap of the quantum Metropolis map as a function of the number of sites N for different values of the magnetic field strength g . The plot clearly indicates a linear scaling of the inverse of the gap.

3.5 The completely positive map

We now investigate the actual completely positive map \mathcal{E} generated by all unitaries and measurements in more detail. The full map can be understood as an initialization step denoted by E followed by successive P and Q measurements, as discussed in section 3.3 and illustrated in Fig. 3.3. Note, that the projectors Q_i depend on the random unitary C . For each application of the map we draw a random unitary C from the set $\mathcal{C} = \{C\}$ according to the probability measure $d\mu(C)$. We therefore have to average over the set \mathcal{C} . The tcp-map on the system is obtained by tracing out all ancilla registers. This construction corresponds to the open system representation as introduced in Theorem 13, and we therefore are ensured that the map we are considering is indeed trace preserving and completely positive. As shown in the previous section 3.4.1, the error obtained by cutting the number of iterations in the reject case to n^* can be made arbitrarily small; we can therefore approximate the full map as an infinite sum

$$\begin{aligned} \mathcal{E}(\rho) &= \int_{\mathcal{C}} \text{tr}_A \left[LQ_1 E (\rho \otimes |0^{2r+1}\rangle\langle 0^{2r+1}|) EQ_1 L^\dagger \right] \\ &+ \text{tr}_A \left[P_1 Q_0 E (\rho \otimes |0^{2r+1}\rangle\langle 0^{2r+1}|) EQ_0 P_1 \right] \\ &+ \sum_{n=1}^{\infty} \sum_{s_1 \dots s_n=0}^1 \text{tr}_A \left[P_1 Q_{s_n} P_0 \dots P_0 Q_{s_1} P_0 Q_0 E \right. \\ &\quad \left. (\rho \otimes |0^{2r+1}\rangle\langle 0^{2r+1}|) EQ_0 P_0 Q_{s_1} P_0 \dots P_0 Q_{s_n} P_1 \right] d\mu(C). \end{aligned} \quad (3.34)$$

The projective measurements P_s and Q_s are comprised of several individual operations. We adopt a new notation: an unmarked sum over the indices written as small Latin letters, e.g. k_1, p_1, \dots is taken to run over all 2^r integer values of the phase estimation ancilla register. The projectors can be written as

$$\begin{aligned} Q_s &= \sum_{k_1, k_2} \sum_{p_1, p_2} C^\dagger M_{k_2}^{p_1 \dagger} M_{k_2}^{p_2} C \otimes |k_1\rangle\langle k_1| \otimes |p_1\rangle\langle p_2| \otimes R^s(k_1, k_2), \quad (3.35) \\ P_0 &= \sum_{k_1 \neq k_2} \sum_{p_1, p_2} M_{k_2}^{p_1 \dagger} M_{k_2}^{p_2} \otimes |k_1\rangle\langle k_1| \otimes |p_1\rangle\langle p_2| \otimes \mathbb{1}, \\ P_1 &= \sum_{k_1 = k_2} \sum_{p_1, p_2} M_{k_2}^{p_1 \dagger} M_{k_2}^{p_2} \otimes |k_1\rangle\langle k_1| \otimes |p_1\rangle\langle p_2| \otimes \mathbb{1}. \end{aligned}$$

As before, we used the convention that the first register contains the physical state of the system. The second register of r -qubits corresponds to the register that stores the eigenvalue estimates of the first phase estimation, the third register is again used for phase estimation and the last register sets the single condition bit. The last matrix is defined as

$$R^s(k_1, k_2) = W(k_1, k_2)^\dagger |s\rangle\langle s| W(k_1, k_2), \quad (3.36)$$

with W defined in (3.17). Furthermore, the first operation in the circuit, that prepares an eigenstate and copies its energy eigenvalue to the lowest register, is denoted by

$$E = \sum_{k_1, k_2} \sum_{p_1, p_2} M_{k_2}^{p_1 \dagger} M_{k_2}^{p_2} \otimes |k_1 \oplus_r k_2\rangle \langle k_1| \otimes |p_1\rangle \langle p_2| \otimes \mathbb{1}, \quad (3.37)$$

where \oplus_r denotes an addition modulo 2^r . For notational purposes we introduced another operation

$$L = \sum_{k_1, k_2} \sum_{p_1, p_2} M_{k_2}^{p_1 \dagger} M_{k_2}^{p_2} C \otimes |k_1\rangle \langle k_1| \otimes |p_1\rangle \langle p_2| \otimes W(k_1, k_2). \quad (3.38)$$

A successful measurement of Q_1 at the beginning of the circuit, Fig. 3.2, followed by the operation L corresponds to an acceptance of the Metropolis update and a further clean-up operation that becomes necessary, when considering a realistic phase estimation procedure.

If we define new super-operators $\mathcal{A}(\rho)$ and $\mathcal{B}_n^{\{s_n\}}(\rho)$, the tcp-map on the physical system can be written as

$$\mathcal{E}(\rho) = \mathcal{A}(\rho) + \mathcal{B}_0(\rho) + \sum_{n=1}^{\infty} \sum_{s_1 \dots s_n=0}^1 \mathcal{B}_n^{\{s_n\}}(\rho). \quad (3.39)$$

Here, \mathcal{A} denotes the contribution to the tcp-map that corresponds to the instance, where the suggested Metropolis move is accepted. Each of the \mathcal{B}_n correspond to a rejection of the update after $n + 1$ subsequent Q and P measurements. These superoperators can be expressed as follows:

$$\begin{aligned} \mathcal{A}(\rho) = & \sum_{k_1, k_2} \sum_{d, p_1, q_1} \int_{\mathcal{C}} d\mu(C) \min\left(1, e^{-\beta \frac{2\pi}{t} (k_2 - k_1)}\right) \\ & M_{k_2}^d \dagger M_{k_2}^{p_1} C M_{k_1}^{p_1 \dagger} M_{k_1}^0 \rho M_{k_1}^0 \dagger M_{k_1}^{q_1} C \dagger M_{k_2}^{q_1 \dagger} M_{k_2}^d. \end{aligned} \quad (3.40)$$

Furthermore,

$$\begin{aligned} \mathcal{B}_0(\rho) = & \sum_{k_1} \sum_{l_1, r_1} \sum_{d; p_1, p_2; q_1, q_2} \int_{\mathcal{C}} d\mu(C) \langle 0 | R^0(k_1, r_1) R^0(k_1, l_1) | 0 \rangle \\ & M_{k_1}^d \dagger M_{k_1}^{p_2} C \dagger M_{l_1}^{p_2 \dagger} M_{l_1}^{p_1} C M_{k_1}^{p_1 \dagger} M_{k_1}^0 \rho M_{k_1}^0 \dagger M_{k_1}^{q_1} C \dagger M_{r_1}^{q_1 \dagger} M_{r_1}^{q_2} C M_{k_1}^{q_2 \dagger} M_{k_1}^d, \end{aligned} \quad (3.41)$$

and

$$\begin{aligned} \mathcal{B}_n^{\{s_n\}}(\rho) = & \sum_{k_1} \sum_{d, \{l_{n+1}\}; \{r_{n+1}\}} \int_{\mathcal{C}} d\mu(C) g_{k_1}(\{s_n\}, \{l_{n+1}\}, \{r_{n+1}\}) \\ & D_{k_1}^d(\{l_{n+1}\}) \rho D_{k_1}^d \dagger(\{r_{n+1}\}). \end{aligned} \quad (3.42)$$

The operators D and the scalar function g in the definition of $\mathcal{B}_n^{\{s_n\}}$ are given by

$$\begin{aligned} g_{k_1}(\{s_n\}, \{l_{n+1}\}, \{r_{n+1}\}) = & \langle 0 | R^0(k_1, r_1) R^{s_1}(k_1, r_2) \dots R^{s_n}(k_1, r_{n+1}) \\ & R^{s_n}(k_1, l_{n+1}) \dots R^{s_1}(k_1, l_2) R^0(k_1, l_1) | 0 \rangle \end{aligned} \quad (3.43)$$

and

$$D_{k_1}^d(\{l_{n+1}\}) = \sum_{\{a_{n+1}\} \neq k_1} \sum_{\{p_{2n}\}} M_{k_1}^{d \dagger} M_{k_1}^{p_{2n}} C^\dagger M_{l_{n+1}}^{p_{2n} \dagger} M_{l_{n+1}}^{p_{2n-1}} C M_{a_{n+1}}^{p_{2n-1} \dagger} M_{a_{n+1}}^{p_{2n-2}} C^\dagger \quad (3.44)$$

$$\dots M_{a_1}^{p_3 \dagger} M_{a_1}^{p_2} C^\dagger M_{l_1}^{p_2 \dagger} M_{l_1}^{p_1} C M_{k_1}^{p_1 \dagger} M_{k_1}^0.$$

This concludes the description of the completely positive map corresponding to one iteration of the Metropolis algorithm.

3.6 Fixed point of the ideal chain

In the previous descriptions of the algorithm we only considered the idealized case when we are able to identify each eigenstate by its energy label. When this is the case, the algorithm can be interpreted as a classical Metropolis random walk where the configurations of the system are replaced by the eigenstates of a quantum Hamiltonian. However, this picture falls short if we consider the more realistic scenario of a Hamiltonian with degenerate energy subspaces. The rejection procedure ensures in this case only that we end up in the same energy subspace we started from. We therefore need to investigate the fixed point of the actual completely positive map that is generated by the circuit. We will see that the quantum Metropolis algorithm yields the exact Gibbs state as its fixed point, if the quantum phase estimation algorithm resolves the energies of all eigenstates exactly. To be able to make statements about the fixed point of this quantum Markov chain, we have introduced (see section 2.4) a quantum generalization of the detailed balance concept. As for classical Markov chains, this criterion only ensures that the state with respect to which the chain is detailed balanced is a fixed point. However, it does not ensure that this fixed point is unique. The uniqueness follows from the ergodicity of the Markov chain [30, 28], as discussed in the preliminaries section 1, and thus depends in our case on the choice of updates $\{C\}$, which can be chosen depending on the problem Hamiltonian. A sufficient (but not necessary) condition for ergodicity can easily be obtained by enforcing $\{C\}$ to form a universal gate set, as will be shown below.

In section 2.4 we show in Proposition 41, that a quantum Markov chain obeys quantum detailed balance for a specific inversion $k(w) = 1/\sqrt{w}$, cf. (2.9), if there exists a probability distribution $\{p_i\}$ and a complete set of orthonormal vectors $\{|\psi_i\rangle\}$ such that the following condition holds

$$\sqrt{p_n p_m} \langle \psi_i | \mathcal{E}[|\psi_n\rangle\langle\psi_m|] | \psi_j \rangle = \sqrt{p_i p_j} \langle \psi_m | \mathcal{E}[|\psi_j\rangle\langle\psi_i|] | \psi_n \rangle. \quad (3.45)$$

This condition together with the ergodicity of the updates $\{C\}$ ensures that the unique fixed point of the quantum Markov chain is

$$\sigma = \sum_{i=1}^{2^N} p_i |\psi_i\rangle\langle\psi_i|. \quad (3.46)$$

We therefore would like to verify whether condition (3.45) is satisfied when we choose the p_i equal to the Boltzmann weights of H and the vectors equal to the eigenvectors $|\psi_i\rangle$.

The condition (3.45) is linear in the superoperators. We can therefore conclude that, when each of the summands \mathcal{A} and all the \mathcal{B} 's in (3.39) individually satisfy this condition, the total tcp-map \mathcal{E} is detailed balanced.

The idealized case would be met if we could simulate a Hamiltonian H with eigenvalues E_i that are r -bit integer multiples of $\frac{2\pi}{t}$, or if we had an infinitely large ancilla register for the phase estimation. In this case, the operators M_E^p would reduce to simple projectors Π_{E+p} on the energy subspace labeled by $E + p$. Hence

$$M_E^p{}^\dagger M_E^q = \delta_{p,q} \Pi_{E+p}.$$

Note that the $\delta_{p,q}$ ensures that after each P and Q measurement the second ancilla register used for phase estimation is again completely disentangled and returns to its original value.

Furthermore, in the special case when the eigenvalues of the Hamiltonian are non-degenerate the projectors reduce to $\Pi_{E_i} = |\psi_i\rangle\langle\psi_i|$. In this case it can be seen that the dynamics of the algorithm reduce to the standard classical Metropolis algorithm that is described by a classical stochastic matrix that can be computed as

$$S_{ij} = \langle\psi_j|\mathcal{E}[|\psi_i\rangle\langle\psi_i||\psi_j\rangle].$$

For this special case it is obvious that the detailed balance condition is met.

Let us now turn to the more generic case, when the energy eigenvalues are degenerate. We investigate each of the contributions to the completely positive map (3.39).

The accept instance: We first investigate the accept instance described by the operator $\mathcal{A}(\rho)$.

$$\mathcal{A}(\rho) = \sum_{E_1, E_2} \int_{\mathcal{C}} d\mu(C) \min\left(1, e^{-\beta(E_2 - E_1)}\right) \Pi_{E_2} C \Pi_{E_1} \rho \Pi_{E_1} C^\dagger \Pi_{E_2}. \quad (3.47)$$

The detailed balance criterion (3.45) for $p_i = \frac{1}{Z} e^{-\beta E_i}$ and $|\psi_i\rangle$ reads

$$\frac{1}{Z} e^{-\beta(E_i + E_j)/2} \langle\psi_l|\mathcal{A}(|\psi_i\rangle\langle\psi_j||\psi_m\rangle) = \frac{1}{Z} e^{-\beta(E_l + E_m)/2} \langle\psi_j|\mathcal{A}(|\psi_m\rangle\langle\psi_l||\psi_i\rangle). \quad (3.48)$$

Note that the chain of operators begins with a projector Π_{E_1} and ends with a projector Π_{E_2} .

The detailed balance condition reads therefore

$$\begin{aligned} & \frac{1}{Z} e^{-\beta(E_i+E_j)/2} \int_{\mathcal{C}} d\mu(C) \min\left(1, e^{-\beta(E_l-E_i)}\right) \delta_{E_l, E_m} \delta_{E_i, E_j} \langle \psi_l | C | \psi_i \rangle \langle \psi_j | C^\dagger | \psi_m \rangle \\ &= \frac{1}{Z} e^{-\beta(E_l+E_m)/2} \int_{\mathcal{C}} d\mu(C) \min\left(1, e^{-\beta(E_j-E_m)}\right) \delta_{E_l, E_m} \delta_{E_i, E_j} \langle \psi_j | C | \psi_m \rangle \langle \psi_l | C^\dagger | \psi_i \rangle. \end{aligned} \quad (3.49)$$

Due to the fact that $\frac{1}{Z} e^{-\beta E_l} \min\left(1, e^{-\beta(E_i-E_l)}\right) = \frac{1}{Z} e^{-\beta E_i} \min\left(1, e^{-\beta(E_l-E_i)}\right)$, this reduces to

$$\int_{\mathcal{C}} d\mu(C) \langle \psi_l | C | \psi_i \rangle \langle \psi_j | C^\dagger | \psi_m \rangle = \int_{\mathcal{C}} d\mu(C) \langle \psi_j | C | \psi_m \rangle \langle \psi_l | C^\dagger | \psi_i \rangle, \quad (3.50)$$

where the energies of the eigenstates have to satisfy $E_l = E_m$ and $E_i = E_j$.

One sees that (3.47) is satisfied when the probability measure obeys

$$d\mu(C) = d\mu(C^\dagger). \quad (3.51)$$

If we consider an implementation that only makes use of a single unitary C for every update, we have to ensure that this unitary is Hermitian, i.e. $C = C^\dagger$. This symmetry constraint on the measure can be seen as the quantum analogue of the fact, that we need to choose a symmetric update rule for the classical Metropolis scheme.

The reject instance: We now turn to the reject case described by the operators $\mathcal{B}_n^{\{s_n\}}(\rho)$. The rejecting operators also simplify greatly when we consider the case of perfect phase estimation. After each phase estimation step the second register disentangles due to the $\delta_{p_i, p_{i+1}}$, we get

$$\mathcal{B}_n^{\{s_n\}}(\rho) = \sum_E \sum_{\{l_{n+1}\}; \{r_{n+1}\}} g_E(\{s_n\}, \{l_{n+1}\}, \{r_{n+1}\}) \int_{\mathcal{C}} d\mu(C) D_E^0(\{l_{n+1}\}) \rho D_E^{0\dagger}(\{r_{n+1}\}). \quad (3.52)$$

The chain of unitaries and measurement operators in the operator D (3.44) reduces to

$$D_E^0(\{l_{n+1}\}) = \Pi_E C^\dagger \Pi_{l_{n+1}} C \Pi_E^\perp C^\dagger \dots \Pi_E^\perp C^\dagger \Pi_{l_1} C \Pi_E, \quad (3.53)$$

where Π_E^\perp is the projector on to the orthogonal complement of energy subspace E . Note that the first and the last projector in each chain of operators is Π_E . Hence, all elements

$$\langle \psi_l | \mathcal{B}_n^{\{s_n\}}(|\psi_i\rangle \langle \psi_j|) | \psi_m \rangle$$

vanish, if all energies are not equal $E_l = E_i = E_j = E_m$. We can therefore disregard the probabilities p_i on either side of the detailed balance equation (3.45). The detailed balance condition thus reads

$$\langle \psi_l | \mathcal{B}_n^{\{s_n\}}(|\psi_i\rangle \langle \psi_j|) | \psi_m \rangle = \langle \psi_j | \mathcal{B}_n^{\{s_n\}}(|\psi_m\rangle \langle \psi_l|) | \psi_i \rangle. \quad (3.54)$$

It is important that the function $g_E(\{s_n\}, \{l_{n+1}\}, \{r_{n+1}\})$ (3.43) is real. Due to this fact and furthermore, since all the individual operators $R^s(E, k)$ are Hermitian, we may exchange the ordering of the indices $\{l_{n+1}\}, \{r_{n+1}\}$. That is, we may write

$$\begin{aligned} g_E(\{s_n\}, \{l_{n+1}\}, \{r_{n+1}\}) &= g_E(\{s_n\}, \{l_{n+1}\}, \{r_{n+1}\})^* \\ &= \langle 0 | R^0(k_1, l_1)^\dagger R^{s_1}(k_1, l_2)^\dagger \dots R^{s_n}(k_1, l_{n+1})^\dagger R^{s_n}(k_1, r_{n+1})^\dagger \dots R^{s_1}(k_1, r_2)^\dagger R^0(k_1, r_1)^\dagger | 0 \rangle \\ &= g_E(\{s_n\}, \{r_{n+1}\}, \{l_{n+1}\}) \end{aligned} \quad (3.55)$$

Furthermore, since the individual projectors Π_{l_i} and Π_E^\perp are of course Hermitian, we may write

$$\begin{aligned} &\langle \psi_l | \mathcal{B}_n^{\{s_n\}}(|\psi_i\rangle\langle\psi_j|) | \psi_m \rangle \\ &= \sum_{\{l_{n+1}\}; \{r_{n+1}\}} g_E(\{s_n\}, \{l_{n+1}\}, \{r_{n+1}\}) \int_C d\mu(C) \langle \psi_l | D_{E_l}^0(\{l_{n+1}\}) |\psi_i\rangle \langle \psi_j | D_{E_l}^0(\{r_{n+1}\}) | \psi_m \rangle \\ &= \sum_{\{l_{n+1}\}; \{r_{n+1}\}} g_E(\{s_n\}, \{r_{n+1}\}, \{l_{n+1}\}) \int_C d\mu(C) \langle \psi_j | D_{E_l}^0(\{r_{n+1}\}) | \psi_m \rangle \langle \psi_l | D_{E_l}^0(\{l_{n+1}\}) |\psi_i\rangle \\ &= \langle \psi_j | \mathcal{B}_n^{\{s_n\}}(|\psi_m\rangle\langle\psi_l|) | \psi_i \rangle. \end{aligned} \quad (3.56)$$

The last equality in (3.56) is precisely due to the fact that we can reorder the indices as previously discussed and that we are dealing with projectors on the energy subspaces.

As already said, a possible set of updates that will ensure ergodicity in general is given by choosing $\{C\}$ equal to a universal gate set. So for instance the set of all possible single qubit unitaries augmented with the CNOT gate would suffice to ensure ergodicity for an arbitrary Hamiltonian. Recall the Theorem 18 about primitive maps in section 1.4.2. With this Lemma at hand, it is straight forward to prove the uniqueness of the fixed point. All we need to show is that the tcp-map \mathcal{E} is primitive.

Lemma 45 (Uniqueness of the Fixed point). *If we choose the set of all possible updates $\{C\}$ equal to a set of universal gates, then the Metropolis Markov chain has a unique full rank fixed point for all finite $\beta < \infty$.*

PROOF: If \mathcal{E} denotes the map defined in (3.39), according to Theorem 18, cf section 1.4.2, all we need to show is that there is an m such that for every $|\psi\rangle$ and every ρ $\langle \psi | \mathcal{E}^m[\rho] | \psi \rangle > 0$. Since ρ can always be written as a convex combination of rank 1 projectors it suffices to choose $\rho = |\varphi\rangle\langle\varphi|$. Furthermore, we observe that all \mathcal{B}_n defined in (3.39) are positive, i.e.

$$\langle \psi | \mathcal{B}_n^{\{s_n\}}(\tilde{\rho}) | \psi \rangle \geq 0, \quad (3.57)$$

since this expression can always be written as the trace over the product of positive semi-definite operators for any $\tilde{\rho}$ and $|\psi\rangle$, see (3.34). We can therefore disregard the contributions from the \mathcal{B}_n and focus only on the accept instance \mathcal{A} of the map \mathcal{E} , since by virtue of (3.57) we have

$$\langle \psi | \mathcal{E}^m(|\varphi\rangle\langle\varphi|) | \psi \rangle \geq \langle \psi | \mathcal{A}^m(|\varphi\rangle\langle\varphi|) | \psi \rangle. \quad (3.58)$$

We can thus write

$$\begin{aligned}
\langle \psi | \mathcal{A}^m(|\varphi\rangle\langle\varphi|) | \psi \rangle &= \\
\int d\mu(C_1) \dots d\mu(C_m) \sum_{E_1 \dots E_{m+1}} \prod_{i=1}^m \min(1, e^{-\beta(E_{i+1}-E_i)}) & |\langle \psi | \Pi_{E_{m+1}} C_m \dots C_1 \Pi_{E_1} |\varphi \rangle|^2 \\
\geq e^{-\beta(E_{max}-E_{min})} \int d\mu(C_1) \dots d\mu(C_m) F_{\psi,\phi}(C_1, \dots, C_m). &
\end{aligned} \tag{3.59}$$

Here E_{max} and E_{min} denote the largest and the smallest eigenvalues of the problem Hamiltonian H respectively, and we defined the integrand F as

$$F_{\psi,\phi}(C_1, \dots, C_m) = \sum_{E_1 \dots E_{m+1}} |\langle \psi | \Pi_{E_{m+1}} C_m \dots C_1 \Pi_{E_1} |\varphi \rangle|^2. \tag{3.60}$$

Note that the prefactor $e^{-\beta(E_{max}-E_{min})}$ does not vanish for all finite β . Since the integrand F is non-negative, we only need to prove that F does not vanish. Since we are drawing the $C_1 \dots C_m$ from a set of universal gates we can always find a finite m , by virtue of the Solovay – Kitaev Theorem [93], so that there exists a sequence of gates C_i that ensures that there is a sufficiently large overlap between $|\psi\rangle$ and $C_m \dots C_1 |\psi\rangle$. That is, for a given ϵ_m , there exists a sequence of m gates, so that

$$|\langle \psi | C_m \dots C_1 |\varphi \rangle|^2 = \left| \sum_{E_1 \dots E_{m+1}} \langle \psi | \Pi_{E_{m+1}} C_m \dots C_1 \Pi_{E_1} |\varphi \rangle \right|^2 \geq 1 - \epsilon_m, \tag{3.61}$$

where we inserted resolutions of the identity $\sum_{E_i} \Pi_{E_i}$. Hence, at least one of summands in (3.61) has to be non-zero and thus $F_{\psi,\phi}$ is strictly positive and does not vanish. Therefore, there exists an integer m so that the integral in the last line of (3.59) is strictly positive. Since (3.59) acts as a lower bound to $\langle \psi | \mathcal{E}^m(|\varphi\rangle\langle\varphi|) | \psi \rangle$ we can conclude that \mathcal{E} is primitive. \square

3.7 The influence of imperfections to the fixed point

We have seen that the idealized quantum Metropolis algorithm yields the exact Gibbs state as its fixed point, if the quantum phase estimation algorithm resolves the energies of all eigenstates exactly. This is obviously impossible for non integer eigenvalues as one would need infinitely many bits just to write down the energies in binary arithmetic. However, we will show that this is not a real problem. A polynomial resolution will yield samples that approximate the Gibbs state very well, if the Markov chain converges sufficiently fast. For the error analysis we will assume that the ergodicity condition is met, and that the problem Hamiltonian we are trying to simulate is such that the Markov chain is rapidly mixing. To be precise, for the error analysis we assume that the Markov chain is trace-norm contracting, see section 3.7.1. We previously discussed the errors that arise due to the finite runtime of the algorithm in section 3.4 and the error due to the indeterministic rejection scheme, cf. section 3.4.1. In this section we consider

the error that is related to the implementation of the algorithm. Due to the implementation on a quantum computer three types of error arise.

1. **Simulation errors.** The quantum phase estimation algorithm requires implementing the dynamics $U = e^{-iHt}$ generated by the system's Hamiltonian for various times t . This can only be done within a finite accuracy.
2. **Round-off errors.** The quantum phase estimation algorithm represents the system's energy in binary arithmetic with r bits. This unavoidably implies that the energy is rounded off to r bits of accuracy.
3. **Phase estimation fluctuations.** As seen in Eq. (3.12), given an energy eigenstate of the system, the quantum phase estimation procedure outputs a random r -bit estimate of the corresponding energy. The output distribution is highly peaked around the true energy, but fluctuations are important and cannot be ignored.

The first error is related to the fact that $\exp(itH)$ has to be approximated by a Trotter-Suzuki unitary. This error can be ignored as long as the necessary effort in the simulation time T_H to make this small, scales better than any power of $1/\epsilon_H$ with ϵ_H being this simulation error [87]. This first source of error can be suppressed at polynomial cost. Another way to tackle this error is to adopt the analysis done in [94].

The second type of error is not a problem on its own. Suppose that each eigenvalue of H is replaced by its closest r -bit approximation. The corresponding thermal state would differ from the exact one by factors of $\exp(\beta 2^{-r})$. By choosing $r \gg \log \beta$, this error can be made arbitrarily small. Note that the simulation cost grows exponentially with r , which implies that our Metropolis algorithm has complexity increasing linearly with β .

The third type of error is more delicate and is intimately related to the second type. Indeed, it is not correct to suppose, as we did in the previous paragraph, that quantum phase estimation outputs the closest r -bit approximation to the energy of the eigenstate. Rather, it outputs a random energy distributed according to Eq. (3.12), sharply peaked around the exact energy. Hence, we are implementing a POVM on the system state and not a projective measurement as we already discussed in section 3.2. We have furthermore introduced a method to sharpen the distribution of the energy pointers by employing a method developed in [91]. In the following we will therefore compute the error bounds of the algorithm when we employ this enhanced quantum phase estimation procedure.

3.7.1 Error bounds and realistic phase estimation

Let us next return to a more general Hamiltonian that has a realistic spectrum. As was discussed earlier, a realistic phase estimation procedure introduces errors not only due to the rounding of the energy values, but more importantly due to the fluctuations of the pointer variable. For a completely positive map with realistic phase estimation the detailed balance condition (3.45)

will not be met exactly, but we can show that the condition is satisfied approximately. This will be sufficient for our purposes.

In order to bound this error we adopt a standard procedure also used for classical Markov chains [95]. Throughout this analysis we assume that the completely positive map is well behaved and is contracting. Whether this assumption is satisfied depends on the mixing properties of the problem we consider and on the choice of updates. Therefore, these properties have to be verified for every problem instance individually. A quantum Markov chain is trace - norm contracting if it satisfies

$$\|\mathcal{E}(\rho - \sigma)\|_{\text{tr}} \leq \eta_{\text{tr}} \|\rho - \sigma\|_{\text{tr}}, \quad (3.62)$$

where the constant $\eta_{\text{tr}} < 1$ is the smallest constant, so that this inequality holds [95]. We introduced the contraction coefficient η_{tr} already in chapter 2 and related it to the contraction coefficient of the χ_k^2 - divergence. Note, that the map is considered contracting only when the constant is strictly smaller than unity. It can occur, for some pathologically behaved maps, that this constant is not strictly smaller than unity even though the map is rapidly mixing. However, this can be cured by blocking several applications of the channel together, leading to a new constant smaller than unity [96].

Lemma 46 (Error bound). *The error ϵ^* between the exact fixed point σ^* of the map \mathcal{E} and the Gibbs state $\rho_G = \frac{1}{Z} \exp(-\beta H)$ can be bounded by*

$$\|\sigma^* - \rho_G\|_{\text{tr}} \leq \frac{\epsilon^{sg}}{1 - \eta_{\text{tr}}}. \quad (3.63)$$

Here $\eta_{\text{tr}} < 1$ is the ergodicity coefficient of \mathcal{E} and ϵ^{sg} the error that arises due to a single application of the map on ρ_G , i.e. $\|\mathcal{E}(\rho_G) - \rho_G\|_{\text{tr}} \leq \epsilon^{sg}$.

PROOF: The error ϵ^* can be written as

$$\begin{aligned} \|\sigma^* - \rho_G\|_{\text{tr}} &= \lim_{m \rightarrow \infty} \|\mathcal{E}^m(\rho_G) - \rho_G\|_{\text{tr}} \leq \lim_{m \rightarrow \infty} \sum_{k=1}^m \|\mathcal{E}^k(\rho_G) - \mathcal{E}^{k-1}(\rho_G)\|_{\text{tr}} \\ &\leq \lim_{m \rightarrow \infty} \sum_{k=1}^m \eta_{\text{tr}}^{k-1} \|\mathcal{E}(\rho_G) - \rho_G\|_{\text{tr}} = \frac{\|\mathcal{E}(\rho_G) - \rho_G\|_{\text{tr}}}{1 - \eta_{\text{tr}}}. \end{aligned} \quad (3.64)$$

□

Thus we only need to bound the error that occurs when we apply the map \mathcal{E} to the Gibbs state ρ_G once. In order to bound this error, we will make use of the fact that the completely positive map satisfies the detailed balance condition (3.45) at least approximately. Let us discuss what it means to satisfy detailed balance approximately.

Lemma 47 (Approximate detailed balance). *Suppose we are given a completely positive map \mathcal{E} and an orthonormal basis $\{|\psi_i\rangle\}$. To each state we assign a Boltzmann weight of the form*

$\{p_i = \frac{1}{Z} e^{-\beta E_i}\}$. If this cp-map does not precisely satisfy detailed balance, but only an approximate form such as

$$\sqrt{p_n p_m} \langle \psi_i | \mathcal{E}(|\psi_n\rangle \langle \psi_m|) | \psi_j \rangle = \sqrt{p_i p_j} \langle \psi_m | \mathcal{E}(|\psi_j\rangle \langle \psi_i|) | \psi_n \rangle (1 + \mathcal{O}(\epsilon^{sg})), \quad (3.65)$$

we can give the following bound on the error, measured in the trace - norm, that occurs upon a single application of the completely positive map.

$$\|\mathcal{E}(\rho_G) - \rho_G\|_{\text{tr}} \leq \mathcal{O}(\epsilon^{sg}) \quad (3.66)$$

PROOF: Let us define $\rho = \sum_i p_i |\psi_i\rangle \langle \psi_i|$. Then due to (3.65) we have

$$\begin{aligned} \langle \psi_l | \mathcal{E}(\rho_G) | \psi_m \rangle &= \sum_i p_i \langle \psi_l | \mathcal{E}(|\psi_i\rangle \langle \psi_i|) | \psi_m \rangle = \\ &= \sqrt{p_l p_m} (1 + \mathcal{O}(\epsilon^{sg})) \text{tr} [\mathcal{E}(|\psi_m\rangle \langle \psi_l|)] = p_m (1 + \mathcal{O}(\epsilon^{sg})) \delta_{ml}. \end{aligned} \quad (3.67)$$

So the application of \mathcal{E} yields $\mathcal{E}(\rho_G) = \tilde{\rho}_G$. Note that the state $\tilde{\rho}_G$ is still diagonal in the same basis as ρ_G and both of the probabilities \tilde{p}_i of $\tilde{\rho}_G$ relate to the original probabilities via $\tilde{p}_i = p_i (1 + \mathcal{O}(\epsilon^{sg}))$. Since ρ_G and $\tilde{\rho}_G$ are both diagonal in the same basis, it is straightforward to compute that $\|\tilde{\rho}_G - \rho_G\|_{\text{tr}} \leq \mathcal{O}(\epsilon^{sg})$. \square

Let us now verify the approximate detailed balance condition (3.65) of the completely positive map (3.39) for a realistic spectrum of the Hamiltonian H . First let us consider the standard phase estimation procedure. Since the actual eigenvalues may have arbitrary real values, we may not assume that the individual M_x^y act as projectors on the system. Note that even the combination of $M_k^{p\dagger} M_k^q$ is not Hermitian anymore when $p \neq q$. This is precisely due to the fact that the function $f(E_j, k - p)$ (3.14) is complex valued. An additional phase is imprinted on the system state. At first sight this seems to hinder any form of detailed balance in the eigenbasis of the Hamiltonian. It turns out, however, that the total expression on either side of the detailed balance equation is still real. Note that $M_k^{p\dagger} M_k^q$ is diagonal in the eigenbasis of H and assumes the form

$$M_k^{p\dagger} M_k^q = \sum_{j=1}^{2^N} f(E_j, k - p)^* f(E_j, k - q) |\psi_j\rangle \langle \psi_j|. \quad (3.68)$$

Hence, the phases in $f(E_j, k - p)^* f(E_j, k - q)$ cancel up to a total phase factor $\frac{e^{i\pi(p-q)}}{e^{i\frac{\pi}{2^r}(p-q)}}$, which is independent of both k and E_j . This allows us to write

$$M_k^{p\dagger} M_k^q \equiv \frac{e^{i\pi(p-q)}}{e^{i\frac{\pi}{2^r}(p-q)}} S_k^{pq}, \quad (3.69)$$

where now $S_k^{pq\dagger} = S_k^{pq}$. Let us have look at a segment of the chain of operators as they typically appear in the superoperators A or B (3.39). The typical sequences look like

$$\dots M_{k_2}^{p_3\dagger} M_{k_2}^{p_2} C M_{k_1}^{p_2\dagger} M_{k_1}^{p_1} \dots \rightarrow \dots \frac{e^{i\pi(p_3-p_1)}}{e^{i\frac{\pi}{2^r}(p_3-p_1)}} S_{k_2}^{p_3 p_2} C S_{k_1}^{p_2 p_1} \dots \quad (3.70)$$

This leads us to the conclusion that in each of the operator sequences the phases that arise due to imperfect phase procedure cancel. The first phase associated to p_0 is 0 due to the initialization, whereas the last phase associated with d is canceled due to the measurement. This gives an additional explanation of why it is necessary to reuse the same pointer register for the phase estimation procedure each time. However, this comes at a cost as the realistic phase estimation procedure doesn't naturally disentangle the pointer register used for the next phase estimation anymore. Hence, the initial state of the ancilla register for the next phase estimation step may be altered. So after subsequent measurements using the same register the distribution function of the pointer variable spreads.

We now consider what happens in the case where we use the high confidence phase estimation based on the *median* - method [91]. As already stated, this method allows us to perform phase estimation where the pointer variable fluctuates at most in the order of 2^{-r} . All other fluctuations are suppressed by a factor of $2^{-\eta}$ and will therefore be neglected in the following. According to (3.15) we can replace the function $f(E_j, k - p)$ by its enhanced counterpart $\alpha_{E_j}(k - p)$, which acts as a binary amplitude for the two closest r -bit integers to the actual energy E_j . As discussed earlier, the phases that arise due to the imperfect phase estimation algorithm cancel, if for each of the η phase estimations the corresponding registers are reused. We are therefore left again with operators S_k^{pq} acting on the physical system that are diagonal and have only real entries. We will thus regard the amplitudes $\alpha_{E_i}(k - p)$ as real from now on. We will therefore write

$$S_k^{pq} = \sum_{j=1}^{2^N} \alpha_{E_j}(k - p) \alpha_{E_j}(k - q) |\psi_j\rangle \langle \psi_j|. \quad (3.71)$$

Let us pause for a minute and have a closer look at the operators S_k^{pq} . As stated previously the S_k^{pq} are diagonal in the Hamiltonians eigenbasis and have only real entries. Hence, these operators are Hermitian. Furthermore, since $\alpha_{E_j}^2$ acts as a binary probability distribution on the two $\delta = 2^{-r}$ closest integers to $\frac{E_j t}{2\pi}$, we see that for a fixed E_j and a fixed q , the only possible two values for k are

$$k^\uparrow = \left\lfloor \frac{E_j t}{2\pi} \right\rfloor_{2^{-r}} + q \quad \text{and} \quad k^\downarrow = \left\lceil \frac{E_j t}{2\pi} \right\rceil_{2^{-r}} + q.$$

Conversely, the operator S_k^{pq} has only support on the subspace spanned by the eigenvectors $|\psi_j\rangle$ whose energies lie in the interval

$$E_j \in [(k + q) - 2^{-r}; (k + q) + 2^{-r}] \cap [(k + p) - 2^{-r}; (k + p) + 2^{-r}].$$

This allows a further conclusion. For a fixed k and q the operator does not vanish only if

$$p \in [q - 2^{-r+1}; q + 2^{-r+1}].$$

The interpretation is as follows: the operator S_k^{pq} implements the action of a phase estimation and its conjugate on the system. If the ancilla register was initially in the state $|q\rangle$ the full phase estimation process does not disentangle the ancilla register afterwards, if we have performed in an intermediate operation. We have seen previously in the analysis for the idealized phase estimation procedure, see section 3.6, that the inverse phase estimation procedure returns the ancilla register to its original value $|q\rangle$. Since the pointer variable fluctuates now, this is not the case anymore and the pointer register remains entangled with the simulated system. However, since we perform an enhanced phase estimation procedure, the allowed values for the ancilla register are bounded by $p^\pm = q \pm 2^{-r+1}$. Thus even though S_k^{pq} is not a projector anymore, the previously discussed conditions suffice to ensure approximate detailed balance.

Let us now verify the approximate detailed balance condition for each of the summands in (3.39).

The accept instance: We analyze what happens in the accept case indicated by the operator $\mathcal{A}(\rho)$. Due to the cancellation of the spurious phases (3.70) this operator has the form

$$\mathcal{A}(\rho) = \sum_{k_1, k_2} \sum_{d, p_1, q_1} \int_{\mathcal{C}} d\mu(C) \min\left(1, e^{-\beta \frac{2\pi}{t}(k_2 - k_1)}\right) S_{k_2}^{dp_1} C S_{k_1}^{p_1 0} \rho S_{k_1}^{0q_1} C^\dagger S_{k_2}^{q_1 d}. \quad (3.72)$$

We now want to verify whether the approximate detailed balance condition is met, when we choose again $p_i = \frac{1}{Z} e^{-\beta E_i}$ and $|\psi_i\rangle$ as the eigenstate of H . We choose a symmetric measure, i.e. $d\mu(C^\dagger) = d\mu(C)$, and verify the approximate detailed balance condition (3.65). The left side of the equation reads

$$\begin{aligned} & \frac{1}{Z} e^{-\beta(E_i + E_j)/2} \langle \psi_l | \mathcal{A}(|\psi_i\rangle \langle \psi_j|) | \psi_m \rangle \quad (3.73) \\ &= \sum_{k_1, k_2} \sum_{d, p_1, q_1} \frac{1}{Z} e^{-\beta(E_i + E_j)/2} \int_{\mathcal{C}} d\mu(C) \min\left(1, e^{-\beta \frac{2\pi}{t}(k_2 - k_1)}\right) \langle \psi_l | S_{k_2}^{dp_1} C S_{k_1}^{p_1 0} | \psi_i \rangle \\ & \quad \langle \psi_j | S_{k_1}^{0q_1} C S_{k_2}^{q_1 d} | \psi_m \rangle \\ &= \sum_{k_1, k_2} \sum_{d, p_1, q_1} \frac{1}{Z} e^{-\beta(E_i + E_j)/2} \int_{\mathcal{C}} d\mu(C) \min\left(1, e^{-\beta \frac{2\pi}{t}(k_2 - k_1)}\right) \langle \psi_l | C | \psi_i \rangle \langle \psi_j | C | \psi_m \rangle \\ & \quad \alpha_{E_i}(k_2 - d) \alpha_{E_l}(k_2 - p_1) \alpha_{E_i}(k_1 - p_1) \alpha_{E_i}(k_1) \alpha_{E_m}(k_2 - d) \alpha_{E_m}(k_2 - q_1) \\ & \quad \alpha_{E_j}(k_1 - q_1) \alpha_{E_j}(k_1). \end{aligned}$$

We are free to relabel all the summation indices k_1, k_2, d, \dots to match it with the other side of the equation. The sequence

$$k_2 = k'_1 + d \rightarrow \left\{ \begin{array}{l} p_1 = q'_1 + d \\ q_1 = p'_1 + d \end{array} \right\} \rightarrow k_1 = k'_2 + d \rightarrow d = 2^r - d' \quad (3.74)$$

does exactly this. Note that since $\alpha_{E_j}(k + 2^r) = \alpha_{E_j}(k)$ the constant 2^r in the last step can be dropped. If we now consider the worst case scenario of the fluctuations of $\alpha_{E_i}(k_1)$, we see that k_1 deviates at most as much as $k_1 \approx \frac{E_i t}{2\pi} \pm 2^{-r+1}$. The same is also true for k_2 and k'_2, k'_1 respectively. Hence we can conclude

$$\frac{1}{Z} e^{-\beta E_i} \min \left(1, e^{(-\beta \frac{2\pi}{t} (k_2 - k_1))} \right) = \frac{1}{Z} e^{-\beta E_i} \min \left(1, e^{(-\beta \frac{2\pi}{t} (k'_1 - k'_2))} \right) \left(1 + \mathcal{O}(\beta \frac{4\pi}{t} 2^{-r}) \right). \quad (3.75)$$

We can therefore establish, that

$$\frac{1}{Z} e^{-\beta(E_i + E_j)/2} \langle \psi_l | \mathcal{A}(|\psi_i\rangle \langle \psi_j|) | \psi_m \rangle = \frac{1}{Z} e^{-\beta(E_l + E_m)/2} \langle \psi_j | \mathcal{A}(|\psi_m\rangle \langle \psi_l|) | \psi_i \rangle (1 + \mathcal{O}(\epsilon)) \quad (3.76)$$

with $\epsilon = \beta \frac{4\pi}{t} 2^{-r}$ which can be fully controlled by adjusting the relevant free parameters.

The reject instance We now turn to the reject case. The operators change accordingly. We consider the detailed balance condition for each of the full $\mathcal{B}_n^{\{s_n\}}(\rho)$. Note that due to the previously discussed phase cancellations the operators $D_{k_1}^d(\{l_{n+1}\})$ as defined in (3.44) assume the form

$$D_{k_1}^d(\{l_{n+1}\}) = \sum_{\{a_{n+1}\} \neq k_1} \sum_{\{p_{2n}\}} S_{k_1}^{dp_{2n}} C^\dagger S_{l_{n+1}}^{p_{2n} p_{2n-1}} C S_{a_{n+1}}^{p_{2n-1} p_{2n-2}} C^\dagger \dots S_{a_1}^{p_3 p_2} C^\dagger S_{l_1}^{p_2 p_1} C S_{k_1}^{p_1 0}. \quad (3.77)$$

The analysis of the reject case is very similar to the exact case. We make use of the fact that all the functions $g_{k_1}(\{s_n\}, \{l_{n+1}\}, \{r_{n+1}\})$ and $\alpha_{E_i}(k - p)$ are real, and that we can relabel the indices like we did in the exact analysis. We have to establish that

$$\begin{aligned} & \frac{1}{Z} e^{-\beta(E_i + E_j)/2} \langle \psi_l | \mathcal{B}_n^{\{s_n\}}(|\psi_i\rangle \langle \psi_j|) | \psi_m \rangle \\ &= \frac{1}{Z} e^{-\beta(E_l + E_m)/2} \langle \psi_j | \mathcal{B}_n^{\{s_n\}}(|\psi_m\rangle \langle \psi_l|) | \psi_i \rangle (1 + \mathcal{O}(\epsilon)), \end{aligned} \quad (3.78)$$

up to some ϵ , that will turn out to be $\epsilon = n \frac{4\pi}{t} \beta 2^{-r}$. We again start by considering the left side of (3.78) and show that it will be equal to the right side up the specified ϵ .

$$\begin{aligned} & \frac{1}{Z} e^{-\beta(E_i + E_j)/2} \langle \psi_l | \mathcal{B}_n^{\{s_n\}}(|\psi_i\rangle \langle \psi_j|) | \psi_m \rangle \\ &= \sum_{k_1} \sum_{d; \{l_{n+1}\}; \{r_{n+1}\}} g_{k_1}(\{s_n\}, \{l_{n+1}\}, \{r_{n+1}\}) \int_C d\mu(C) \frac{1}{Z} e^{-\beta(E_i + E_j)/2} \\ & \quad \langle \psi_j | D_{k_1}^{d \dagger}(\{r_{n+1}\}) | \psi_m \rangle \langle \psi_l | D_{k_1}^d(\{l_{n+1}\}) | \psi_i \rangle. \end{aligned} \quad (3.79)$$

We will first exchange the index sets $\{r_{n+1}\}$ and $\{l_{n+1}\}$. This is possible since the function g_{k_1} is real and we follow the same analysis we already performed in the case of the idealized phase estimation. Now we turn to the sequence of the relabeling of the index set $d, k_1, l_1, r_1, a_1, b_1, \dots$. Note that a_i and b_i are part of the definition of $D_{k_1}^d(\{l_{n+1}\})$ and

$D_{k_1}^{d \dagger}(\{r_{n+1}\})$ respectively (3.77). The relabeling sequence that does what we want reads

$$\begin{aligned}
k_1 = k'_1 + d &\rightarrow \tag{3.80} \\
\left\{ \begin{array}{l} p_{2n} = q'_{2n} + d \\ q_{2n} = p'_{2n} + d \end{array} \right\} &\rightarrow \left\{ \begin{array}{l} l_{n+1} = l'_{n+1} + d \\ r_{n+1} = r'_{n+1} + d \end{array} \right\} \rightarrow \left\{ \begin{array}{l} p_{2n-1} = q'_{2n-1} + d \\ q_{2n-1} = p'_{2n-1} + d \end{array} \right\} \rightarrow \\
\left\{ \begin{array}{l} a_{n+1} = b'_{n+1} + d \\ b_{n+1} = a'_{n+1} + d \end{array} \right\} &\rightarrow \dots \left\{ \begin{array}{l} l_1 = l'_1 + d \\ r_1 = r'_1 + d \end{array} \right\} \rightarrow \left\{ \begin{array}{l} p_1 = q'_1 + d \\ q_1 = p'_1 + d \end{array} \right\} \rightarrow d = 2^r - d'.
\end{aligned}$$

For these replacements to work, it is important to note that the operators $R^s(k_1, l_i)$ depend only on the differences, i.e. $R^s(k_1 - l_i)$. The sequence of replacements therefore leaves the function $g_{k_1}(\{s_n\}, \{l_{n+1}\}, \{r_{n+1}\})$ unchanged. However, since we do perform $2n$ phase estimation processes for each of the superoperators $\mathcal{B}_n^{\{s_n\}}$, the variable k_1 in the last process may fluctuate in the order of $n2^{-r+1}$, as was discussed earlier, and we may no longer assume that the statistical weights on either side of the equation are equal. Hence we know that for the worst instance k_1 is $\delta = \pm n2^{-r+1}$ close to either energy E_i, E_j, E_l, E_m . We can therefore see, upon evaluating (3.78), that the detailed balance condition for each individual \mathcal{B}_n is met up to an $\epsilon = n \frac{4\pi}{t} \beta 2^{-r}$.

We observe that the ϵ increases linearly in the number n of subsequent P and Q measurements we make to reject the proposed update. For all realistic applications, as discussed in section 3.4, one would choose a fixed n^* so that one only would attempt to perform $n \leq n^*$ reject moves before discarding the sample. Since we want to achieve an overall success probability of preparing a valid sample that is lower bounded by a constant c , we have to choose $n^* > \frac{m}{2e(1-c)}$. Here m denotes the number of times we have to apply the map \mathcal{E} to be sufficiently close to the desired steady-state. This is related to the gap Δ of the map \mathcal{E} , cf. section 3.4. Hence in the end we can give an error estimate for a single application of the map, which is of the order

$$\epsilon^{sg} = \mathcal{O}\left(\frac{m}{2e(1-c)} \frac{4\pi}{t} \beta 2^{-r}\right). \tag{3.81}$$

3.8 An experimental implementation

It is possible to implement the quantum Metropolis algorithm with today's technology for a simple 2 qubit example system. Here, we will show how the different building blocks of the quantum Metropolis algorithm can be represented with simple quantum circuits. For this we need to consider a quantum computer of 5 qubits. Let's assume that we want to simulate the Gibbs state of the Heisenberg ferromagnet on 2 spin $1/2$'s, i.e.

$$H_2 = -\frac{1}{2}(\sigma_1^x \otimes \sigma_2^x + \sigma_1^y \otimes \sigma_2^y + \sigma_1^z \otimes \sigma_2^z), \tag{3.82}$$

which is certainly one of the most interesting Hamiltonians for 2 qubits. With the appropriate energy offset, this Hamiltonian has the spectrum $\{0, 2\}$, where the eigenvalue 0 is threefold

degenerate. This is very good news, as it means that an exact phase estimation algorithm can be set up with just a single (qu)bit of accuracy. Such a phase estimation requires simulating the Hamiltonian for a time $t = \pi/2$. One sees that this unitary corresponds exactly to the SWAP gate, that is,

$$U\left(\frac{\pi}{2}\right) = e^{-i\frac{\pi}{2}H_2} = \text{SWAP}. \quad (3.83)$$

In the quantum Metropolis algorithm, we need to implement the controlled version of this SWAP, which is the Fredkin gate. In [97], it has been shown how this Fredkin gate can be implemented efficiently using optics. A related gate, the so-called Toffoli gate, was recently realized in the group of R. Blatt with an ion trap computer [98]. The second gate to be implemented is the controlled Metropolis unitary W . The Metropolis unitary can be implemented with two controlled R_y rotations:

$$W(\theta_\beta) = R_y(-\theta_\beta)_C X R_y(\theta_\beta)_C, \quad (3.84)$$

where we have made use of the standard single qubit unitary, $R_y(\theta) = \exp(-i\frac{\theta}{2}\sigma^y)$ and wrote $X = \sigma^x$. The temperature can be controlled by the angle θ_β . Comparison with the original Metropolis unitary (3.16) shows that we have to set $\cos(\theta_\beta) = e^{-\beta}$. The full circuit is depicted in Fig. 3.5. Note that this circuit can be simplified if we regard the lowest qubit as a classical bit, which is determined by the first phase estimation. It is then possible to condition the remainder of the circuit on the first phase estimation result, then the controlled Metropolis unitary W can be implemented by a single CNOT operation.

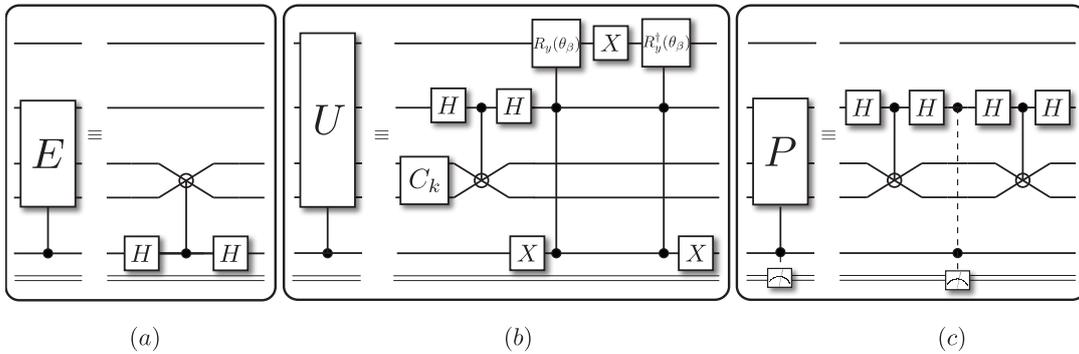


Figure 3.5: Fig. (a) describes the first phase estimation step of the circuit. Since the phase estimation of the two-qubit Heisenberg Hamiltonian can be implemented exactly by the Fredkin gate, a single phase estimation operation is sufficient. In Fig. (b) the elementary unitary of the circuit is depicted. The angle of the controlled-controlled $R_y(\theta_\beta)$ rotation needs to be chosen such that $\cos(\theta_\beta) = e^{-\beta}$. The final measurement P is depicted in Fig. (c). The first phase estimation has to be followed by a measurement which verifies that the two phase estimation bits are equal.

Chapter 4

Stochastic matrix product states

Synopsis:

This chapter explores parallels between the many-body description of non-equilibrium steady states of classical stochastic processes and ground states of strongly correlated quantum many-body systems. Classical non-equilibrium steady states are typically much richer than their equilibrium counterparts and can exhibit interesting behavior such as the presence of a current, non-equilibrium phase transitions and entire phases with a diverging correlation length [99, 100], features also found in the context of ground states of quantum many-body Hamiltonians. It has indeed long been observed that there are strong parallels between the many-body description of non-equilibrium classical stochastic spin systems such as the asymmetric exclusion process and the physics of equilibrium quantum Hamiltonians such as the ferromagnetic Heisenberg spin chain [101], and it turns out that matrix product states (MPS) play a very important role in both fields [102, 4]. We introduce the concept of stochastic matrix product states and derive a natural form for the states. This allows us to define the analogue of Schmidt coefficients for steady states of non-equilibrium stochastic processes. We discuss a new measure for correlations which is analogous to the entanglement entropy, the entropy cost S_C , and show that this measure quantifies the bond dimension needed to represent a steady state as a matrix product state. We illustrate these concepts by means of the asymmetric exclusion process.

Based on:

K. Temme and F. Verstraete,

Phys. Rev. Lett. **104**(21), 210502 (2010)

4.1 Quantum states and correlation measures

Before we turn to the discussion of the correlation properties of non-equilibrium steady states, we would briefly like to state some of the results obtained in the context of the application of quantum information theory to strongly correlated quantum many-body systems. The theory of entanglement has recently proven to yield valuable new insights into the nature of the wavefunctions arising in such systems [103]. Quantum states are ultimately characterized by the observable correlations they exhibit. Consider for instance the ground state of a quantum spin Hamiltonian. An observable such as the correlation function between two spins at different lattice sites typically decays exponentially as a function of the distance separating them. However, when the system undergoes a phase transition, the state becomes scale invariant and thus the correlation function decays algebraically. We note, that by investigating correlation functions, the emphasis has shifted from the study of Hamiltonians to states. It is therefore reasonable to investigate the entanglement properties of quantum states per se. This is indeed a common scheme found in the context of quantum information theory. One of the contributions to the field of strongly correlated systems has been made by providing a universal figure of merit for the quantum correlations that is model independent. The entanglement measures provided by quantum information yield a universal currency of quantum correlation which is independent of the model specific observables. The most prominent measure for pure quantum states is entanglement entropy. For a pure state $|\psi\rangle$ on some bipartite Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$ the entanglement entropy is defined by the von Neumann entropy of the reduced density matrix $\rho_A = \text{tr}_B [|\psi\rangle\langle\psi|]$, i.e.

$$S(A) = -\text{tr} [\rho_A \log_2(\rho_A)]. \quad (4.1)$$

Since any bipartite quantum state can be written in the Schmidt basis as $|\psi\rangle = \sum_i \sigma_i |\alpha_i\rangle |\beta_i\rangle$, we have that $S(A) = S(B) = \sum_i \sigma_i^2 \ln(\sigma_i^2)$. In the case of such pure states, this entropy has multiple operational meanings, ranging from the amount of Bell states that can be distilled from it using local operations to the maximum amount of secret information that can be sent from one side to the other in a cryptographic setting [104]. The usefulness of the entanglement entropy for many-body systems stems to some extent from its behavior at criticality and provides a good characterization for the universal aspect of quantum states. It allows for a cleaner access to fundamental properties of critical neighborhoods such as the central charge of the corresponding conformal field theory [105]. In the previous chapter we have introduced so-called matrix product states. A matrix product state can be cast into the normal form $|\psi\rangle = \sum_{\{i\}} A_{i_1}^1 \Sigma^1 \dots \Sigma^{N-1} A_{i_N}^N |\{i\}\rangle$ [42], where the Σ^i denote diagonal matrices of Schmidt coefficients. Upon choosing a bipartition of the chain at some site k the entanglement entropy S_A of a MPS can easily be calculated and we immediately get the bound $S_k \leq \log_2(D_k)$. We observe that the entanglement entropy is bounded by the logarithm of the bond dimension. Hence, when we increase the partition A the entanglement entropy scales at most as $S(A) \sim \log_2(D_{max})$, i.e. it is constant. This behavior is known as an area law

[106, 107, 108]. An area law just states that the entropy scales as $S(A) \sim |\partial A|$, i.e. only as the boundary of the subsystem, as opposed to, for instance, its volume. In one dimension $|\partial A|$ is trivially just a constant. Higher dimensional generalizations of MPS exist and are known as tensor network states or PEPS [109] and these states also obey an area law by construction. More interesting, however, is the question of whether a MPS can be used to approximate the ground states of local Hamiltonians. It turns out that this is indeed the case. This is a consequence of the fact that, in the case of ground states of gapped local Hamiltonians, the associated Schmidt coefficients decay very fast [75]. A MPS approximation with finite bond dimension D can be obtained by setting the smallest Schmidt coefficients σ_I equal to zero, and it has been proven that this approximation is justified whenever an area law is satisfied [41]. This is the precisely the reason why numerical techniques such as DMRG [39, 40] have given such great results. For critical quantum systems the entanglement entropy behaves quite differently. For second-order phase transitions the correlation length diverges and the system becomes scale invariant. This scaling symmetry gets enlarged to the conformal group [110] and the critical system is described by a $1 + 1$ dimensional conformal field theory. The development of conformal field theory is a remarkable achievement that goes beyond the scope of this thesis and we refer the reader to other sources [111]. The scaling of the entropy for a block of length $A = L$ in a conformal field theory of $1 + 1$ dimensions was proven to diverge logarithmically [105] as

$$S(L) \sim \frac{c + \bar{c}}{6} \log_2(L), \quad (4.2)$$

where c and \bar{c} are the so-called central charges of the holomorphic and anti-holomorphic sectors of the conformal field theory. These charges specify the universal description of the model at criticality completely and are only dependent on the symmetries of the underlying model. The scaling of the entanglement entropy can therefore be used as an order parameter for pure quantum states that is expected to obey an area law for non-critical systems which is modified by logarithmic corrections at criticality.

The situation is much more complicated, however, in the case of mixed quantum states or in the presence of classical correlations, and a lot of research in quantum information theory has concentrated on resolving the relationship between entanglement and classical correlations. A better understanding of those measures will be necessary for describing the classical and quantum correlations in thermal or non-equilibrium states of quantum spin systems. A measure for mixed state correlations that is of particular importance to our approach is the entanglement of purification [112]. This measure quantifies both quantum and classical correlations and has a clear operational meaning in terms of the number of maximally entangled states needed to asymptotically generate a quantum state.

4.2 Entropy cost and stochastic matrix products states

The main topic of this chapter is to explore the classical non-equilibrium analogues of the quantum notions of entanglement entropy, area laws and the density matrix renormalization group, which justify the use of MPS in the quantum setting. The main technical difficulty in achieving this is the fact that classical probability distributions are normalized in the L_1 norm ($\sum_i |p_i| = 1$), while quantum wavefunctions are normalized in the L_2 norm ($\sum_i |\psi_i|^2 = 1$). This difficulty can partly be overcome by working with a subclass of MPS where all matrices only contain non-negative entries; we will define such MPS as stochastic matrix product states (sMPS). The concept of mutual information, defined for classical bipartite distributions p_{AB} , where A and B will represent the variables or spins on both halves of a chain, plays a role analogous to the entanglement entropy:

$$I(A : B) = \sum_{AB} p_{AB} \log_2 \left(\frac{p_{AB}}{p_A p_B} \right). \quad (4.3)$$

It immediately gives an upper bound to the error made when approximating p_{AB} by a product of its marginals, since $\|p_{AB} - p_A p_B\|_1^2 \leq 2 \ln(2) I(A : B)$ [113]. Just as in the quantum case, one would expect that the global non-equilibrium steady state probability distribution of the stochastic process can be represented as a stochastic matrix product state (sMPS) with small bond dimension, if this mutual information is small. However, more subtle measures are needed in the case of stochastic processes and we will introduce the notion of entropy cost to quantify the bond dimension needed for the corresponding sMPS.

Definition 48 (Entropy cost). *The entropy cost S_C for a bipartite probability distribution $P(x, y)$ is given by:*

$$S_C = \min_{p_\lambda, P_A, P_B} S(\{p_\lambda\}) \quad (4.4)$$

where $S(\{p_\lambda\}) = -\sum_\lambda p_\lambda \log_2(p_\lambda)$ is the Shannon information of $\{p_\lambda\}$, and where the optimization is over all probability distributions p_λ and over all conditional probabilities P_A and P_B for which $P(x, y) = \sum_\lambda P_A(x|\lambda) P_B(y|\lambda) p_\lambda$.

The entropy cost bears a lot of resemblance to the notion of common information introduced by Wyner in the context of cryptography and classical information theory [114], and this entropy cost serves as an upper bound to the common information. The entropy cost can be thought of as the classical analog of the entanglement of purification, and the probability distribution p_λ plays a role analogous to the Schmidt coefficients in the quantum case.

Let us next define a D -dimensional sMPS describing a classical probability distribution of N -spins each of dimension d ; obviously, those sMPS were already extensively used in the literature, and we will just formalize the definition here.

Definition 49 (Stochastic matrix product state). *A stochastic matrix product state (sMPS) is given by:*

$$|p_D\rangle = \sum_{i_1, \dots, i_N=1}^d \langle l | B_{i_1}^1 \dots B_{i_N}^N | r \rangle | i_1 \dots i_N \rangle, \quad (4.5)$$

where we only consider real matrices that are $D_k \times D_{k+1}$ dimensional, with $D_k \leq D$, and additionally fulfill the requirement $[B_{i_k}^k]_{\delta}^{\gamma} \geq 0$ for every element individually.

This ensures that all the weights of the distribution are positive after contraction. The left and right vector $\langle l |$ and $| r \rangle$ are also element wise positive and can be absorbed into the matrices $B_{i_1}^1$ and $B_{i_N}^N$, which corresponds to choosing $D_1 = D_N = 1$. Furthermore, we require $|p_D\rangle$ to be normalized in the L_1 norm, $\| |p_D\rangle \|_1 = 1$.

Every multipartite probability distribution of a chain of discrete variables can obviously be written in the form (4.5) if we allow for a sufficiently large matrix dimension D^{\max} , i.e. exponential in the number of sites: let $|r\rangle$, a vector, correspond to the original distribution $|p\rangle$ with the set of spin indices $\{i_k\}$ relabeled as one index $\alpha = \sum_k i_k d^k$. Now all the matrices read $B_{i_k}^k = \left(\otimes_{n=1}^{k-1} \sum_{n=0}^{d-1} |n\rangle \langle n| \right) \otimes |i_k\rangle$. This way the full distribution can be reconstructed in its original form.

4.3 Normal form and bounds on the entropy cost

It would be desirable to have a way of computing the entropy cost S_C for a given stochastic matrix product state. It will turn out, however, that this is a task that seems to be very hard for more general distributions, as the minimization (4.4) over all possible decomposition proves very challenging. In most cases we will therefore have to work with an upper bound to the entropy cost. Such an upper bound can be computed easily. We can write any sMPS that is of the form (4.5) as

$$|p\rangle = \sum_{\lambda} \sum_{\{i_n\}} \langle l | B_{i_1}^1, \dots, B_{i_k}^k | \lambda \rangle \langle \lambda | B_{i_{k+1}}^{k+1}, \dots, B_{i_N}^N | r \rangle | \{i_n\} \rangle. \quad (4.6)$$

upon inserting a partition of unity $\mathbb{1} = \sum_{\lambda=1}^{D_k} | \lambda \rangle \langle \lambda |$ in (4.5). Observe that

$$p_{\lambda} = \langle l | \prod_{n=1}^k C^{[n]} | \lambda \rangle \langle \lambda | \prod_{n=k+1}^N C^{[n]} | r \rangle \quad (4.7)$$

defines a new probability distribution if we define $C^{[n]} = \sum_{i_n} B_{i_n}^{[n]}$ (i.e. the transfer matrix). The probability distribution $\{p_{\lambda}\}$ sums up to one due to the normalization we require for (4.5). This allows us to rewrite the MPS as

$$|p\rangle = \sum_{\lambda=1}^{D_k} \sum_{i_1 \dots i_N} P_A(\{i_n\}_{n \in A} | \lambda) p_{\lambda} P_B(\{i_n\}_{n \in B} | \lambda) | \{i_n\}_{n \in A} \rangle | \{i_n\}_{n \in B} \rangle, \quad (4.8)$$

where,

$$\begin{aligned} P_A(\{i_n\}_{n \in A} | \lambda) &= \frac{\langle l | B_{i_1}^1 \dots B_{i_k}^k | \lambda \rangle}{\langle l | \prod_{l=1}^k C^{[l]} | \lambda \rangle} \geq 0, \\ P_B(\{i_n\}_{n \in B} | \lambda) &= \frac{\langle \lambda | B_{i_{k+1}}^{k+1} \dots B_{i_N}^N | r \rangle}{\langle \lambda | \prod_{l=k+1}^N C^{[l]} | r \rangle} \geq 0. \end{aligned} \quad (4.9)$$

We observe that $P_A(\{i_n\} | \lambda)$ and $P_B(\{i_n\} | \lambda)$ can be interpreted as information channels due to their normalization and the positivity of the sMPS matrices. That is, both channels obtain a symbol λ from the source described by (4.7) and transform it into a spin configuration on either side of the bipartition. Note that there are several partitions of unity that will give rise to valid P_A, P_B and p_λ , since any partition of unity is allowed just as long as it preserves the positivity of the individual elements in P_A and P_B . Therefore the decomposition (4.8) is not unique. It is probably a NP-hard problem to find the optimal decomposition that minimizes the entropy cost, and in practice we will therefore rely on the construction that was just given by choosing a simple partition for finding upper bounds to it. An instructive pictorial representation of this decomposition is given in Fig. 4.1(a).

Example: To give an example where the entropy cost S_C can be computed *exactly*, consider the classical Ising model defined by the Hamiltonian

$$-\beta H = K \sum_{i=1}^{N-1} s_i s_{i+1}, \quad (4.10)$$

where $s_i = \pm 1$. The equilibrium distribution $p(\{s_i\}) = 1/Z \exp(-\beta H)$ can be written in terms of a MPS with $D = 2$ [109]. In this model S_C can be calculated for $N = 2$. In this case, the distribution is written in terms of the matrices

$$\begin{aligned} B_{s_1=-1}^1 &= \frac{1}{\sqrt{2q}} \begin{pmatrix} a^+ & a^- \end{pmatrix} \\ B_{s_1=+1}^1 &= \frac{1}{\sqrt{2q}} \begin{pmatrix} a^- & a^+ \end{pmatrix}, \end{aligned} \quad (4.11)$$

as $p(s_1, s_2) = B_{s_1}^1 B_{s_2}^2$. The second set of matrices is given simply by transposition, i.e. $B_{s_2}^2 = B_{s_2}^1{}^T$. Here we have written $a^\pm = 1/\sqrt{2} \left(\sqrt{\cosh(K)} \pm \sqrt{\sinh(K)} \right)$ as well as $q = a^+ + a^-$. The 2d resolution of the identity is chosen as $\mathbb{1} = |\hat{0}\rangle \langle \hat{0}| + |\hat{1}\rangle \langle \hat{1}|$ with $|\hat{0}\rangle = (\cos(\varphi) \quad \sin(\varphi))$ and $|\hat{1}\rangle = (-\sin(\varphi) \quad \cos(\varphi))$. The resulting two source probabilities are

$$\begin{aligned} p_0(\varphi) &= \frac{1}{2} (\cos \varphi + \sin \varphi)^2 \\ p_1(\varphi) &= \frac{1}{2} (\cos \varphi - \sin \varphi)^2. \end{aligned} \quad (4.12)$$

Due to the periodicity, it suffices to consider only the interval $\varphi \in [0; \frac{\pi}{4}]$. Since $S(\{p_\lambda(\varphi)\})$ has no local minima on this interval, the entropy assumes its extremal values at the boundary determined by the constraints (4.9). The allowed values for φ are constrained by the fact that all elements of $P_{A/B}(\{s_i\}|\lambda) \geq 0$ have to be well defined and positive. This leads to the relevant inequality $\frac{a^-}{a^+} \geq \tan \varphi$, since by fulfilling this, we automatically fulfill all other inequalities. The minimum value for $S(\{p_\lambda(\varphi)\})$ is assumed when this inequality is saturated. The entropy cost reads:

$$S_C = - (e^{-K} \cosh(K) \log_2 (e^{-K} \cosh(K)) + e^{-K} \sinh(K) \log_2 (e^{-K} \sinh(K))) \quad (4.13)$$

As expected, this function monotonously increases from 0 to 1, i.e. from the paramagnetic $K \ll 1$ without correlations to the ferromagnetic region $K \gg 1$ with strong correlations. The entropy cost is of course bounded by $\log_2(D = 2) = 1$ for the Ising model. A different measure that was recently investigated in the context of a non-equilibrium model is the so-called shared information [115]. It has been shown that it obeys an area law for several non-critical stochastic models and that critical behavior can be identified by logarithmic corrections.

It is also easy to find lower bounds to the entropy cost; S_C is directly related to the mutual information due to the data processing inequality [113].

Lemma 50 (Mutual information bound). *For a given distribution $|p\rangle$ the mutual information $I(A : B)$ is bounded by the entropy cost S_C , i.e. $I(A : B) \leq S_C$.*

PROOF: By virtue of (4.8) we can focus on calculating the mutual information $I(\lambda : \mu)$ of the distribution $P(\lambda, \mu) = p_\lambda \delta_{\lambda, \mu}$, since the full distribution can be read as

$$P_{AB} = \sum_{\lambda, \mu} P_A(\{i_A\}|\lambda) P_B(\{i_B\}|\mu) P(\lambda, \mu). \quad (4.14)$$

This corresponds to a source that generates two outputs, which are then transformed by the channels P_A and P_B . The data-processing inequality [113] guarantees that the mutual information of the processed source $I(A : B) = I(p_A(P) : p_B(P)) \leq I(\lambda : \mu)$ is smaller than the mutual information of the source itself which is equal to its entropy $I(\lambda : \mu) = S(\{p_\lambda\})$. Since the decomposition $[P_A(\{i_A\}|\lambda), P_B(\{i_B\}|\lambda), p(\lambda)]$ is not unique the bound is improved by taking the minimum over all decompositions. \square

The decomposition of the sMPS as given in (4.8) suggests the existence of the following (non-unique) normal form:

$$|p\rangle = \sum_{i_1 \dots i_N} A_{i_1}^{[1]} P^{[1]} A_{i_2}^{[2]} \dots P^{[N-1]} A_{i_N}^{[N]} |i_1 \dots i_N\rangle. \quad (4.15)$$

Here the matrices $P^{[k]}$ represent diagonal matrices with probabilities $\{p_{\lambda_k}\}$ sorted in decreasing order. In this form, which has a nice pictorial representation in Fig. 4.1(b) the analogy to the

quantum matrix product states becomes evident, cf. Eq. (1.70). The matrices $C^{[k]} = \sum_{i_k} A^{[k]}$, where $[A_{i_k}^{[k]}]_{\delta}^{\gamma} \geq 0$, combined with the $P^{[k]}$'s behave as stochastic matrices. This is due to the fact, that along each bipartition the two channels sum up to $\sum_{i_{A/B}} P_{A/B}(\{i_{A/B}\}|\lambda) = 1$. Now, since we move from one partition to the next by applying $P^{k-1}C^{[k]} = S^k$ to the left or $C^k P^k = S_k^T$ to the right, we are ensured that the S^k and S_k^T have the left eigenvector $\langle I | = \sum_i \langle i |$. Thus we can conclude that the S^k and S_k denote different stochastic matrices.

To see that every MPS distribution can be written this way, consider the following scenario: We start by introducing the first bipartitioning between the first two sites. After the necessary normalization we proceed to the next site and perform the same procedure renormalizing the resulting matrices by the total contraction of the two halves of the chain. Proceeding along the chain results in the desired form.

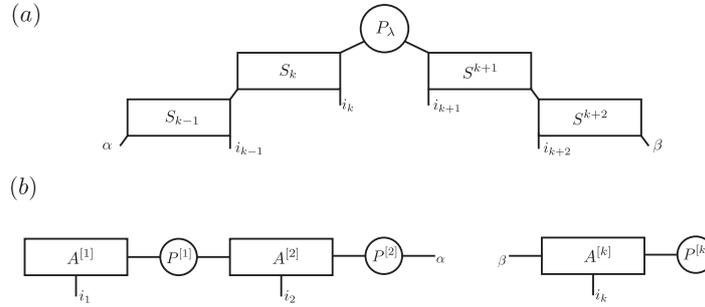


Figure 4.1: Pictorial representation of the natural sMPS decomposition: Image (a) can be seen as the graphical representation of eqn. (4.8). From a given source P_λ , the correlations are distributed via the two channels on the left and right. The normalizing factor is included in the A s. In (b) the analogy to the quantum MPS becomes evident for the decomposition as given in eqn. (4.15). The probabilities in the matrices $P^{[k]}$ are the analogues of the singular values which arise upon a Schmidt decomposition of the quantum state [42].

This representation (4.15) enables us to give a good estimate on the error measured in the L_1 norm, which is made upon truncating the dimension of the source space, i.e. neglecting probabilities smaller than a given value along each bipartition:

Lemma 51 (Error bound). *For every multipartite distribution $|p\rangle$ there exists a sMPS $|p_D\rangle$ of the form (4.15) with dimension D , such that*

$$\| |p\rangle - |p_D\rangle \|_1 \leq 2 \sum_{k=1}^{N-1} \epsilon_k(D),$$

where $\epsilon_k(D) = \sum_{\lambda=D+1}^{D_k^{\max}} p_\lambda^{[k]}$.

PROOF: We can always write $|p\rangle$ as a distribution of the form (4.15) with a $D_k^{\max} = d^N$. We now introduce another MPS $|p_D\rangle$ in natural form with a bond dimension of D . Let $|p_D\rangle = |p_D^*\rangle / \| |p_D^*\rangle \|_1$, where $|p_D^*\rangle$ is the pseudo, i.e. unnormalized, probability distribution which arises from neglecting along each cut all the probabilities $\{p_{\lambda_k}\}_{D+1}^{D_k^{\max}}$. We

write $|p_D^*\rangle = \sum_{\{i_k\}} A_{i_1}^{[1]} P^{*[1]} \dots A_{i_N}^{[N]} |\{i_k\}_{k=1}^N\rangle$. Note, that if $\| |p\rangle - |p_D^*\rangle \|_1 \leq \epsilon$, then $\| |p\rangle - |p_D\rangle \|_1 \leq 2\epsilon$, since

$$\begin{aligned} \| |p\rangle - |p_D\rangle \|_1 &\leq \| |p\rangle - |p_D^*\rangle \|_1 + \left\| |p_D^*\rangle - \frac{|p_D^*\rangle}{\| |p_D^*\rangle \|_1} \right\|_1 \\ &\leq \epsilon + \| |p_D^*\rangle \|_1 - 1 \leq 2\epsilon \end{aligned} \quad (4.16)$$

Since $|p_D^*\rangle$ arose by only neglecting positive numbers, we may write:

$$\begin{aligned} &\| |p\rangle - |p_D^*\rangle \|_1 \\ &= \sum_{\{i_k\}} \left| A_{i_1}^{[1]} P^1 \dots A_{i_N}^{[N]} - A_{i_1}^{[1]} P^{*1} \dots A_{i_N}^{[N]} \right| \\ &= C^{[1]} P^1 \dots C^{[N]} - C^{[1]} P^{*1} \dots C^{[N]} \\ &= \langle l^{N-1} | - \langle l^{N-1} |^* \rangle C^{[N]} = \| \langle l^{N-1} | - \langle l^{N-1} |^* \rangle \|_1. \end{aligned} \quad (4.17)$$

Going from the second to the third line we have used the fact that all summands are positive. Here we defined $\langle l^k | = C^{[1]} P^{[1]} \dots P^{[k]}$ as well as $\langle l^k |^* = C^{[1]} P^{*[1]} \dots P^{*[k]}$. The difference $\| \langle l^1 | - \langle l^1 |^* \|_1 = \| \langle I | (P^{[1]} - P^{*[1]}) \|_1 = \sum_{\alpha_1=D+1}^{D_1^{\max}} p_{\alpha_1}$ is simply given by $\epsilon_1(D)$. Note that due to (4.15) $\langle l^{k-1} | C^{[k]} = \langle I |$. Proceeding to calculate the difference for other k we find:

$$\begin{aligned} &\| \langle l^{[k]} | - \langle l^{*[k]} | \|_1 \\ &\leq \| \left(\langle l^{[k-1]} | - \langle l^{*[k-1]} | \right) C^{[k]} P^{[k]} \|_1 + \| \langle l^{*[k-1]} | C^{[k]} (P^{[k]} - P^{*[k]}) \|_1 \\ &\leq \sum_{n=1}^{k-1} \epsilon_n(D) + \| \langle I | (P^{[k]} - P^{*[k]}) \|_1. \end{aligned} \quad (4.18)$$

The last summand corresponds exactly to $\sum_{\alpha_k=D+1}^{D_k^{\max}} p_{\alpha_k} = \epsilon_k(D)$, which completes the proof. \square

We have therefore proven that an efficient parametrization of the steady state exists in terms of a sMPS with low bond dimension, if there exists a parametrization of this steady state for which the entropy cost with respect to all bipartite cuts is small. If this is the case, then $\sum_{k=1}^{N-1} \epsilon_k(D)$ can be made small by following the arguments outlined in [41]. This is analogous to the quantum case for which the existence of an area law implies the existence of an efficient representation in terms of MPS. Note, however, that the classical statement is a bit weaker, as the same normal form has to be used with respect to all bipartite cuts, and there is no guarantee that the same parametrization is optimal for all of bipartitions.

4.4 Application to the asymmetric exclusion process

To make the investigations concrete, we consider the non-equilibrium steady state of the asymmetric exclusion process (ASEP) [116]. This classical non-equilibrium process is modelled

by a chain of sites labeled $k = 1 \dots N$ occupiable by hardcore particles, i.e. classical spins $i_k \in \{0, 1\}$. The particles are only allowed to hop to the right, and this only if the next site is empty. To drive the system, particles at the left are injected with a given rate α and removed on the right with a rate β .



Figure 4.2: Chain of hardcore particles; each black dot can be either occupied or empty. Particles enter on the left with rate α and hop to the right with a rate of unity, before they are extracted on the right with rate β .

The dynamics of the ASEP are described by a classical master equation $\partial_t |p\rangle = L |p\rangle$, where the Liouvillian generator is

$$L = L_1 + \sum_{k=1}^{N-1} L_{k,k+1} + L_N, \quad (4.19)$$

with

$$L_1 = \begin{pmatrix} \alpha & 0 \\ -\alpha & 0 \end{pmatrix}, \quad L_{k,k+1} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad L_N = \begin{pmatrix} 0 & -\beta \\ 0 & \beta \end{pmatrix}. \quad (4.20)$$

Here the operators L_1 , L_N , and $L_{k,k+1}$ act on the sites $1, N$ and on the two sites indicated by $k, k+1$ respectively. On all other sides they act as the identity.

The ASEP is one of the most studied non-equilibrium processes in classical statistical mechanics and is soluble [116]. This process exhibits three phases determined by the inflow α and outflow β as shown in Fig. 4.3. The different phases are characterized by the particle density and the particle current as functions of the driving parameters. In the region where $\alpha + \beta \leq 1$ and when $\alpha > \beta$ we find the so-called high density phase (HD). This phase is indicated by the fact that one has a high particle density throughout the system that drops to small values only at the right boundary of the chain. In the thermodynamic limit $N \rightarrow \infty$ the particle current j_{st} is given by $j_{st} = \alpha(1 - \alpha)$. The low density phase (LD) $\alpha + \beta \leq 1$ and $\alpha < \beta$ is related to the (HD) phase by a reflection of the chain and a particle hole transformation. The particle density behaves correspondingly and the current can likewise be computed as $j_{st} = \beta(1 - \beta)$. The third phase $\alpha > 1/2, \beta > 1/2$ is known as the maximum current phase (MC). The particle density assumes its maximal value in the bulk of the chain, decaying at the edges. The current is $j_{st} = 1/4$ throughout the diagram. A peculiarity of this phase is its diverging correlation length. The entire non-equilibrium phase exhibits correlations that decay according to a power

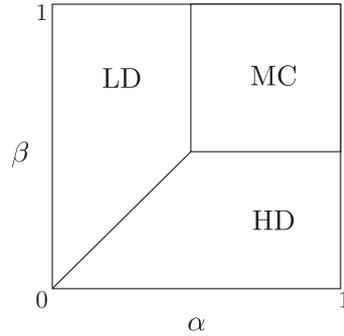


Figure 4.3: phase diagram of the asymmetric exclusion process. The ASEP processes three phases, LD (low-density), HD (high-density) and the maximum current phase MC.

law as opposed to a generally expected exponential decay [117]. This can be taken as an indication that this phase should have a steady state that is strongly correlated and we will see in the following that this is indeed the case.

As was shown in [116] the steady state $|p\rangle$ of the corresponding master equation can be found exactly in terms of a MPS, albeit one for which the matrices are infinite dimensional. The steady state solution can be found by the following ansatz [116]: We assume a distribution $|p\rangle$ that is of MPS form (4.5) with site independent matrices $B_0^1 = \dots = B_0^N = E$ and $B_1^1 = \dots = B_1^N = G$ and boundary vectors $\langle l|, |r\rangle$. We will write for convenience

$$|p\rangle = \langle l| \left(\begin{array}{c} E \\ G \end{array} \right)^{\otimes N} |r\rangle, \quad (4.21)$$

where the vectors $\langle l|, |r\rangle$ act only on the matrices E, G and leave the local basis $|0\rangle, |1\rangle$ untouched. This state has to be normalized by the constant $Z = \langle l| C^{\otimes N} |r\rangle$ with $C = E + G$. Let us see what happens when we impose the following algebraic relations on the matrices E, G and some ancilla matrices \hat{E}, \hat{G}

$$\begin{aligned} \langle l| L_1 \left[\begin{array}{c} E \\ G \end{array} \right] &= \langle l| \begin{array}{c} \hat{E} \\ \hat{G} \end{array}, & L_N \left[\begin{array}{c} E \\ G \end{array} \right] |r\rangle &= - \begin{array}{c} \hat{E} \\ \hat{G} \end{array} |r\rangle, \\ L_{k,k+1} \left[\begin{array}{c} E \\ G \end{array} \right] \otimes \begin{array}{c} E \\ G \end{array} &= \begin{array}{c} E \\ G \end{array} \otimes \begin{array}{c} \hat{E} \\ \hat{G} \end{array} - \begin{array}{c} \hat{E} \\ \hat{G} \end{array} \otimes \begin{array}{c} E \\ G \end{array}. \end{aligned} \quad (4.22)$$

The Liouvillian (4.19) is given as a sum of local terms. When this Liouvillian acts on a distribution that is of the form (4.21), where the matrices satisfy (4.22), the individual terms telescope to zero, and hence (4.21) is a steady state of the system. Thus we find a steady state solution of the master equation if we find a representation $\langle l|, |r\rangle$, E , and G for the algebra (4.22). It is easy to see, that the algebraic constraints reduce to

$$GE = E + G, \quad \langle l| E = \frac{1}{\alpha} \langle l| \quad \text{and} \quad G |r\rangle = \frac{1}{\beta} |r\rangle, \quad (4.23)$$

when we choose $\hat{E} = -\hat{G} = \mathbb{1}$.

Except for the special case when $\alpha + \beta = 1$, the representations obeying those constraints are infinite dimensional [116]. If $\alpha + \beta = 1$ the algebra can be chosen so that the matrices commute. In this case we are left with $G = 1/\beta$ and $E = 1/\alpha$. The resulting state becomes a mean - field state and all correlations vanish. We will refer to this special line as the *mean field line*. For all other values of α and β the representation of the algebra is necessarily infinite dimensional. However, as the total occupation number of particles is limited by N , it is possible to construct truncated representations of a given dimension $D = N + 1$ that still reproduce the exact solution for a chain of length N , even though the algebra is not satisfied exactly, since we multiply at most N - times the same matrix E or G . Hence, the entropy cost can immediately be upper bounded by the logarithm of the system size, just as in the case of critical quantum spin chain, because $S_C \leq \ln(D) = \ln(N + 1)$. By that we see, that the mutual information can also diverge at most logarithmically. We construct a family of truncated representations that are of the form

$$\begin{aligned} E &= \sum_{i,j=0}^1 [B]_{i,j} |i\rangle \langle j| + \sum_{n=2}^N |n\rangle \langle n| + |n\rangle \langle n-1| \\ G &= \sum_{i,j=0}^1 [A]_{i,j} |i\rangle \langle j| + \sum_{n=2}^N |n\rangle \langle n| + |n-1\rangle \langle n|. \end{aligned} \quad (4.24)$$

With the left and right vector

$$\langle l| = \sum_{n=0}^1 w_n \langle n| \quad |r\rangle = \sum_{n=0}^1 v_n |n\rangle. \quad (4.25)$$

Here A and B are 2-dimensional matrices fulfilling $AB + \sigma^- \sigma^+ = A + B$ (here σ^+ and σ^- are the Pauli raising and lowering operators), and the 2-component vectors $|v\rangle$ and $\langle w|$ must be chosen to be eigenvectors: $A|v\rangle = 1/\beta|v\rangle$, $\langle w|B = 1/\alpha\langle w|$. Since this representation of the steady state only depends on a small number of parameters, it is possible to carry out an optimization over all states belonging to this subclass in order to find a good bound on the actual entropy cost.

Correlations and the ASEP phase diagram We want to see whether the correlations of the ASEP steady state give an insight to the phase diagram of the process. To this end we want to investigate the previously introduced correlation measures in dependence of the in- and outflow. We begin with the mutual information. The mutual information of the steady state of a chain of length $N = 20$, where the bipartitioning cut is chosen in the middle, can be calculated numerically, see Fig. 4.4(a). Even though we don't have direct access to the entropy cost for even small chains, reasonable upper bounds can be found by optimizing the solution (4.24) subject to the discussed algebraic constraints. The minimum of the entropy $S(\{p_\lambda\})$, see Fig. 4.4(b), for p_λ as constructed in (4.7) is obtained for 3 different solutions depending on the parameter range of α and β .

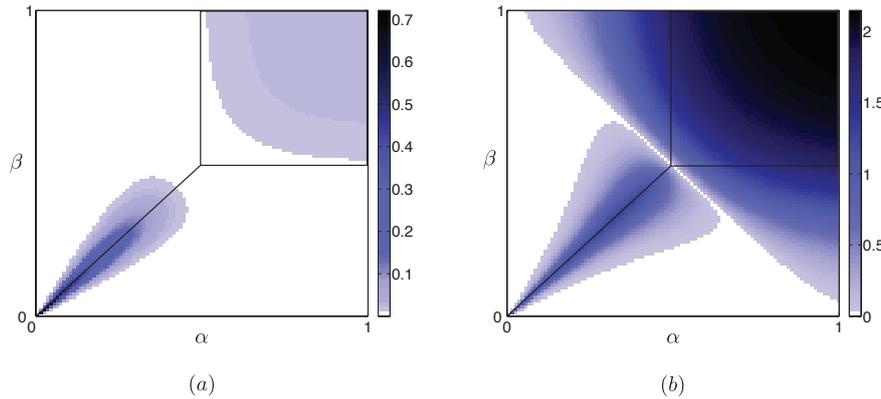


Figure 4.4: (color online) Figure (a): The mutual information of a chain length $N = 20$ for different inflow parameters α and β has been calculated numerically. Figure (b): entropy cost for a representation with $D^{\max} = 21$. Note the different scales of the two plots.

- If $\alpha + \beta \leq 1$ and $\beta \leq \alpha$, then

$$A = \begin{pmatrix} \frac{1}{\beta} & 0 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} \frac{1}{1-\beta} & 0 \\ b & \frac{1}{\alpha} \end{pmatrix}$$

$$\langle w | = \left(\alpha(1-\beta) \quad b \right) \quad \text{and} \quad |v\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (4.26)$$

with $b = \sqrt{1 - \alpha - \beta}$.

- However, if $\alpha + \beta \leq 1$ but $\beta \geq \alpha$, the optimal solution can be obtained from the previous one by the replacements,

$$A_{new} = B^T(\alpha \rightleftharpoons \beta), \quad B_{new} = A^T(\alpha \rightleftharpoons \beta), \quad \text{and}$$

$$\langle w |_{new} = \langle v |, \quad |v\rangle_{new} = |w(\alpha \rightleftharpoons \beta)\rangle. \quad (4.27)$$

- For $\alpha + \beta \geq 1$, the optimal solution is given by

$$A = \begin{pmatrix} \frac{1}{\beta} & a \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} \frac{1}{\alpha} & a \\ 0 & 1 \end{pmatrix},$$

$$\langle v | = \langle w | = \left(1 \quad 0 \right) \quad (4.28)$$

where we defined $a = \sqrt{1/\alpha + 1/\beta - 1/\alpha\beta}$.

The resulting plot Fig. 4.4 clearly reflects the underlying phase diagram of the ASEP, see Fig. 4.3 for comparison. The upper bound to the entropy cost Fig. 4.4(b) as well as the actual mutual information Fig. 4.4(a) drop to zero along the mean-field line $\alpha + \beta = 1$, as is expected for a mean field solution where no correlations are present. We observe, that the mutual information is considerably low throughout the diagram. This explains why the first approaches

with mean-field theories have already given such good results [99]. As expected, the mutual information as well as the upper bound to the entropy cost is largest at the phase transition between the high density phase (HD) and the low density phase (LD). Furthermore, we see an increased amount of correlations in the maximum current phase. This is in accordance with the fact, that the correlation length diverges in the entire phase corresponding to a power law for the correlation functions [117]. Note, that the upper bound for the entropy cost is significantly larger than the mutual information, the qualitative behavior, however, is quite similar. To a certain extent this is expected, since we are optimizing only over a very small subclass of all possible steady state parametrizations.

Mutual information scaling As we discussed previously, in quantum states critical behavior is heralded by a logarithmic correction to the area law of the entanglement entropy. We would like to see, whether a similar behavior occurs for the mutual information in non-equilibrium steady states. We therefore need to investigate the scaling behavior of the mutual information for larger system sizes. To this end we made use of a simple Monte Carlo simulation to compute the mutual information of the steady state. We picked three points in the phase diagram Fig. 4.3, corresponding to the low density phase $\alpha = 1/4, \beta = 1/2$ marked by the red line in the plots Fig. 4.5, the maximum current phase $\alpha = \beta = 1/2$ (blue line), and the phase transition between the low and high density phase $\alpha = 1/4, \beta = 1/4$ (black line). The mutual information is computed by sampling the function $\log_2(p/p_{APB})$ according to the steady state distribution $|p\rangle$, where p_A and p_B are obtained by computing the marginals of subsystem $A = 1 \dots L$ and $B = L + 1 \dots N$ respectively. The mutual information, Fig. 4.5(a), remains constant in the low density phase (red line) as well as in the maximum current phase (blue line). One sees that the correlations in the maximum current phase first increase and then saturate at a finite value. The mutual information at the coexistence line, however, seems to diverge logarithmically. Motivated by the logarithmic divergence of the entanglement entropy in critical quantum systems, we make the following ansatz which we fit to the numerical data

$$I(L) = c \log_2 \left(\frac{L}{s} - t \right). \quad (4.29)$$

The numerical fit indeed suggests that this is the behavior of the function and we obtain for the scaling coefficients

$$c \approx 0.2055 \quad , \quad s \approx 7.9532 \quad \text{and} \quad t \approx -0.9344. \quad (4.30)$$

We conclude that this non-equilibrium steady state also exhibits a logarithmic correction to the otherwise fulfilled area law at criticality, just like the ground states of critical quantum Hamiltonians [106, 108]. It would be interesting to see whether this critical point is also related to some conformal field theory, as is the case for critical quantum systems. Investigations in this direction have indeed been undertaken for an exclusion process with periodic boundary

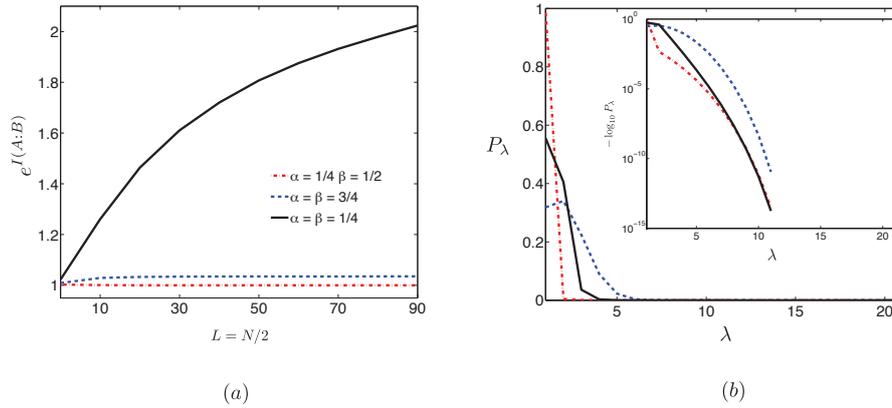


Figure 4.5: (color online) Plot (a): Monte Carlo simulations of the mutual information for different α and β values. The system size was varied from $N = 2 \dots 180$. The chain was cut in the middle, the length L of the individual blocks is given by half of the total system size. The plot shows the exponential of the mutual information $\exp I(A : B)$. The simulations suggest that the mutual information grows only logarithmically along the coexistence line $\alpha = \beta$, $\alpha + \beta \leq 1$; Plot (b): The logarithm of the probability distribution $\{p_\lambda\}$ plotted for different values of α and β . The dimension of the matrices \mathcal{E} and \mathcal{D} is $D^{\max} = 21$, corresponding to a chain of length $N = 20$. The distributions decay super-exponentially. For all $\lambda \geq 11$, $p_\lambda = 0$.

conditions in [115]. The plot Fig. 4.5(b) depicts the source probabilities in each of the different phases for the optimal solution (4.24). The color coding is the same as in Fig. 4.5(a). We see that all source probabilities decay super-exponentially. We therefore conclude that a much smaller bond dimension than $D = N + 1$ would suffice to represent the state faithfully. However, we would probably lose the “semi-translationally” invariant description where all matrices are equal.

We observe that classical non-equilibrium states exhibit a quite similar behavior to that of ground states of local spin Hamiltonians. The definition of the sMPS gives a reasonable starting point for a construction of a DMRG like algorithm for classical non-equilibrium states. Note, however, that since the Liouvillian is in general not a symmetric matrix, the Ritz variational principle, that has led to such a convenient formulation of DMRG for quantum states, falls short. In order to find a DMRG-like algorithm it is therefore advisable to look for an algorithm that is along the lines of the original DMRG or even NRG formulation.

Chapter 5

Stochastic exclusion processes versus coherent transport

Synopsis:

The asymmetric exclusion process we considered in the previous section is a prime example for a stochastic exclusion process. Stochastic exclusion processes play an integral role in the physics of non-equilibrium statistical mechanics. These models are Markovian processes, described by a classical master equation. In this chapter a quantum mechanical version of a stochastic hopping process in one dimension is formulated in terms of a quantum master equation. This allows the investigation of coherent and stochastic evolution in the same formal framework. The focus lies on the non-equilibrium steady state. Two stochastic model systems are considered, the totally asymmetric exclusion process and the fully symmetric exclusion process. We compare the transport properties of these two classical models to the transport properties of a system that has in addition to the classical stochastic hopping a means of coherent transport generated by the H_{XX} Hamiltonian. First, we introduce the hopping model and formulate the problem as a quantum master equation. Then, we investigate the symmetric process. For this process the two-point correlation functions can be calculated exactly in the steady state. The scaling of the current for larger lattice sizes is investigated. In the following, the quantum analog of the asymmetric process is treated numerically in the framework of matrix product density operators. The master equation for a chain of $N = 40$ sites is evolved in time, until the steady state is reached. The current, the particle density, as well as the particle density-density correlations are computed.

Based on:

K. Temme, M. M. Wolf and F. Verstraete,
e-print arXiv: 0912.0858, (2009)

5.1 Exclusion processes

Stochastic exclusion processes have been studied in statistical mechanics for a long time [25, 20]. These are simplified one-dimensional hopping models, that allow the study of non-equilibrium phenomena in many-particle systems. The steady state of the process exhibits interesting non-equilibrium behavior, such as the presence of a current, non-equilibrium phase transitions and entire phases with a diverging correlation length [99, 100]. The presence of currents, such as the current of particles, energy or momentum, is a common feature of non-equilibrium steady states and can have profound effects on the correlations present in the system [118, 119]. Non-equilibrium systems can develop long-range correlations in the presence of a high current. The asymmetric exclusion process (ASEP), we already introduced in the previous section, as well as the symmetric exclusion process (SEP) are prime examples for such model systems [99, 21]. Both processes describe the hopping of hard-core particles in a one-dimensional chain, only driven by the inflow and the outflow of particles at the boundaries of the chain. Here one considers open boundary conditions, where particles are injected at the first site and are removed at the last site N of the chain. The dynamics of the particles in the bulk are given by translationally invariant hopping rates, that either constrain the hopping of particles to take place in only one direction (ASEP) or allow for a hopping in both directions (SEP).

Transport properties of open quantum mechanical systems, on the other hand, are subject to recent research activities. A general interest is placed on how external noise, generated by the environment, affects the coherent transport in the system. It has been found, that the presence of noise in quantum mechanical systems can actually aid the transport process of excitations through heterogeneous environments [120, 121], such as bio-molecules. An optimal ratio between coherent transport and dephasing noise can be found. The dynamics of open quantum systems are generally formulated in terms of a Markovian Lindblad master equation that describes the time evolution of the density matrix [34]. In this chapter we want to investigate the interplay between stochastic transport processes and coherent transport present in the same system. Here we consider only the steady state properties of the system. To treat both processes on equal footing, we incorporate the classical hopping terms into the quantum master equation. The stochastic hopping is modeled by appropriately chosen quantum jump operators. Such a construction has also been used to find quantum master equations that describe a quantization of kinetic Ising models [122]. These models obey detailed balance and allow for an exact solution. Considering hopping models in this more general quantum framework allows now for additional quantum transport, so to speak, on top of the classical hopping evolution. We can choose an arbitrary particle-number conserving Hamiltonian to mediate the coherent transport and investigate the effect this quantum perturbation has on the classical hopping process.

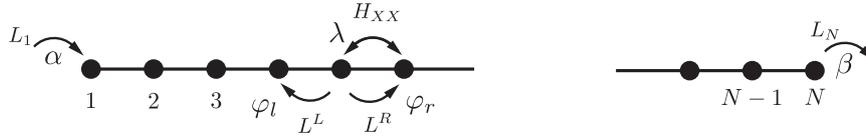


Figure 5.1: Spin-1/2 chain with jump-operators and coherent evolution. The inflow is given by rate α , the outflow by rate β , stochastic hopping to the left occurs with rate φ_L and to the right with φ_R . The strength of the coherent evolution is given by λ .

5.2 Formulation of the quantum master equation

We study a system of hard-core particles in a one dimensional chain of length N , where each site ($1 \leq k \leq N$) can be either occupied or empty. This can be cast into the formulation of a spin - 1/2 chain. In the spin chain picture this corresponds to either spin up $|1\rangle$ (occupied), or spin down $|0\rangle$ (empty). At the boundary $k = 1$, we allow for an inflow with a rate α and at $k = N$ for an outflow of particles, given by a rate β . The particles at each site are allowed to hop stochastically to the left with a rate φ_L and to the right with a rate φ_R , see Fig. 5.1 for comparison. In this chapter we consider only two cases: First, the fully symmetric case (SEP) where both hopping rates are equal $\varphi_L = \varphi_R = \varphi$. The stochastic hopping rates in the bulk are in this case completely symmetric. The only asymmetry that can generate driving in this model is due to the biased in- and out flow at the boundary. In turn, the fully asymmetric case (ASEP), that was already discussed in the previous section, is obtained by setting $\varphi_R = \varphi$ and $\varphi_L = 0$.

We seek to formulate the classical stochastic processes in terms of a quantum Lindblad master equation of the form (23). The stochastic particle jumps that correspond to the classical stochastic exclusion process can be formulated in terms of Lindblad operators L_μ . As already discussed, these operators govern the incoherent evolution of the quantum master equation and are typically responsible for the damping or decoherence of the quantum system. In our model, however, these terms generate the classical non-equilibrium dynamics. It is possible to formulate these jumps in terms of spin-flip operations:

$$\begin{aligned} L_1 &= \sqrt{\alpha} \sigma_1^+ & \text{and} & & L_{k,k+1}^R &= \sqrt{\varphi_R} \sigma_k^- \otimes \sigma_{k+1}^+, \\ L_N &= \sqrt{\beta} \sigma_N^- & \text{and} & & L_{k-1,k}^L &= \sqrt{\varphi_L} \sigma_{k-1}^+ \otimes \sigma_k^-. \end{aligned} \quad (5.1)$$

Here, the σ^\pm correspond to the Pauli raising and lowering operators. The master equation written with only these operators reproduces exactly the classical stochastic behavior, when one restricts oneself to density matrices diagonal in the computational basis.

In this generalized framework, we can now also allow for an additional coherent evolution of the system by choosing an appropriate Hamiltonian. The XX -Hamiltonian

$$H_{XX} = \frac{\lambda}{2} \sum_{k=1}^{N-1} \sigma_k^x \otimes \sigma_{k+1}^x + \sigma_k^y \otimes \sigma_{k+1}^y. \quad (5.2)$$

gives rise to the free coherent evolution of the hard-core particles. Furthermore, we will see that this coherent evolution has the property to conserve the total number of hard-core particles, since it satisfies a continuity equation. The full quantum master equation $\partial_t \rho = \mathcal{L}(\rho)$, including both coherent and stochastic evolution, can be written as:

$$\partial_t \rho = -i [\rho, H_{XX}] + \sum_{\mu} L_{\mu} \rho L_{\mu}^{\dagger} - \frac{1}{2} \left\{ L_{\mu}^{\dagger} L_{\mu}; \rho \right\}_+, \quad (5.3)$$

where the individual L_{μ} were defined in (5.1).

The central observables are the particle density $n_k = \sigma_k^+ \sigma_k^-$ and the current j_k of particles. To find the right expression for the particle current, we consider the continuity equation for the density n_k . The continuity equation is obtained in the Heisenberg picture, when the adjointed of \mathcal{L} is acting on n_k

$$\begin{aligned} \partial_t n_k &= \mathcal{L}^* [n_k] \\ &= -i [H, n_k] + \sum_{\mu} L_{\mu}^{\dagger} n_k L_{\mu} - \frac{1}{2} \left\{ L_{\mu}^{\dagger} L_{\mu}; n_k \right\}_+. \end{aligned} \quad (5.4)$$

If we compute the time evolution of the density operator n_k with respect to the full master equation (5.3) for all sites $k = 2, \dots, N-1$, we can cast the equation in the following form:

$$\partial_t n_k + (j_{k-1,k}^{co} + j_{k-1,k}^{st}) - (j_{k,k+1}^{co} + j_{k,k+1}^{st}) = 0. \quad (5.5)$$

This equation is the standard form of a discrete continuity equation. The additional coherent evolution with respect to the H_{XX} Hamiltonian can thus be seen to be particle number preserving. Furthermore, we can now interpret the sum of the terms $j_{k,k+1} = j_{k,k+1}^{co} + j_{k,k+1}^{st}$ given by

$$\begin{aligned} j_{k,k+1}^{co} &= \frac{\lambda}{i} (\sigma_k^- \sigma_{k+1}^+ - \sigma_k^+ \sigma_{k+1}^-) \\ j_{k,k+1}^{st} &= \varphi_R (n_k (1 - n_{k+1})) - \varphi_L ((1 - n_k) n_{k+1}), \end{aligned} \quad (5.6)$$

as the total current density of the system. Note, that there are two different contributions to the current, the coherent part j^{co} due to the dynamics generated by the Hamiltonian and the stochastic contribution j^{st} originating from the hopping induced by the Lindblad operators. We observe, that the stochastic contribution to the current corresponds exactly to the current present in the classical model [99, 116]. The continuity equation leads to a further conclusion. Since for the steady state of the master equation we have that $\partial_t \langle n_k \rangle = 0$ we can infer that the total current-density has to be constant throughout the system. Hence, $\langle j_{k-1,k} \rangle = \langle j_{k,k+1} \rangle$ and

therefore no spatial variations of the current are allowed. Here, we have defined the average with respect to the non-equilibrium steady state of the system. We will later see, however, that the individual contributions to the total current-density themselves actually do exhibit a special dependence in the steady state.

5.3 The symmetric exclusion process

Let us now consider a specific choice for the classical hopping rates that makes the model soluble, better said, a choice for which we can compute the ground state density and current exactly. If we allow for stochastic hopping in both directions with an equal rate $\varphi_R = \varphi_L = \varphi$ and turn off the coherent evolution, the model describes the classical symmetric exclusion process. The symmetric exclusion process is known to possess only a single classical phase [21] that is determined by a vanishing total current j in the thermodynamic limit $N \rightarrow \infty$. We would like to see, whether a quantum perturbation to the system would change this behavior. Note, that now, since the classical hopping rates are equal, the quantum-jump operators are related via $L_{k,k+1}^R = L_{k,k+1}^{L\dagger} \equiv L_{k,k+1}$. This allows us to rewrite the full master equation as,

$$\begin{aligned} \partial_t \rho &= -i[\rho, H] + \sum_{k=1}^{N-1} [[L_{k,k+1}; \rho], L_{k,k+1}^\dagger] + [[L_{k,k+1}^\dagger; \rho], L_{k,k+1}] \\ &+ L_1 \rho L_1^\dagger - \frac{1}{2} \{L_1^\dagger, L_1; \rho\}_+ + L_N \rho L_N^\dagger - \frac{1}{2} \{L_N^\dagger, L_N; \rho\}_+. \end{aligned} \quad (5.7)$$

Note, that now the dissipative terms in the bulk are given by the sum of two double commutators. It is possible to calculate the nearest neighbor two-point correlation functions exactly. To see why this is possible, we first transform the Pauli raising and lowering operators, σ^+ and σ^- to fermionic modes by means of the Jordan-Wigner transformation [123]. The fermionic modes read then,

$$a_k^\dagger = - \left(\bigotimes_{i=1}^{k-1} \sigma^z \right) \sigma_k^+ \quad \text{and} \quad a_k = - \left(\bigotimes_{i=1}^{k-1} \sigma^z \right) \sigma_k^-. \quad (5.8)$$

One can verify, that these modes now obey the fermionic anti-commutation relations, $\{a_k, a_l^\dagger\}_+ = \delta_{k,l}$ and $\{a_k, a_l\}_+ = \{a_k^\dagger, a_l^\dagger\}_+ = 0$. It is possible to calculate the evolution of the fermionic two-point function $\langle a_k^\dagger a_m \rangle$ from the master equation (5.7) via, $\partial_t \langle a_k^\dagger a_m \rangle = \text{tr} \left[\mathcal{L}^*(a_k^\dagger a_m) \rho \right]$. Here, the operators comprising the Lindblad operators also get transformed and now read

$$\begin{aligned} H &= -\lambda \sum_{k=1}^{N-1} a_k^\dagger a_{k+1} + a_{k+1}^\dagger a_k \quad \text{and} \quad L_{k,k+1} = \sqrt{\varphi} a_k a_{k+1}^\dagger, \\ L_1 &= -\sqrt{\alpha} a_1^\dagger \quad \text{and} \quad L_N = -\sqrt{\beta} \left(\prod_{k=1}^{N-1} (2a_k^\dagger a_k - 1) \right) a_N \end{aligned} \quad (5.9)$$

Since the commutator of two pairs of fermionic modes is again an operator made up from two fermionic modes, we see, that the time-evolution of the fermionic two-point functions again

only depends on two-point functions. So the two-point correlation functions of the steady state can be computed exactly. The equations for the correlation functions read,

$$\begin{aligned} \partial_t \langle a_m^\dagger a_l \rangle &= \frac{\alpha}{2} (\delta_{m,1} + \delta_{l,1}) \langle a_l a_m^\dagger \rangle - \frac{\beta}{2} (\delta_{m,N} + \delta_{l,N}) \langle a_m^\dagger a_l \rangle \\ &- \frac{\varphi}{2} \left([\langle a_{m+1} a_{m+1}^\dagger \rangle - \langle a_{m+1}^\dagger a_{m+1} \rangle + \langle a_{m-1} a_{m-1}^\dagger \rangle - \langle a_{m-1}^\dagger a_{m-1} \rangle] \delta_{l,m} \right. \\ &\left. + 2[\langle a_m^\dagger a_l \rangle - \langle a_l a_m^\dagger \rangle] - i\lambda \left([\langle a_m^\dagger a_{l+1} \rangle + \langle a_m^\dagger a_{l-1} \rangle] - [\langle a_{m+1}^\dagger a_l \rangle + \langle a_{m-1}^\dagger a_l \rangle] \right) \right), \end{aligned} \quad (5.10)$$

and similarly for the correlation function $\langle a_m^\dagger a_l^\dagger \rangle$, we obtain

$$\begin{aligned} \partial_t \langle a_m^\dagger a_l^\dagger \rangle &= -\frac{\alpha}{2} (\delta_{1,m} + \delta_{1,l}) \langle a_m^\dagger a_l^\dagger \rangle - \frac{\beta}{2} (\delta_{l,N} + \delta_{m,N}) \langle a_m^\dagger a_l^\dagger \rangle \\ &- \varphi \left([\delta_{l+1,m} \langle a_l^\dagger a_{l+1}^\dagger \rangle - \delta_{m+1,l} \langle a_m^\dagger a_{m+1}^\dagger \rangle] + 2 \langle a_m^\dagger a_l^\dagger \rangle \right) \\ &+ i\lambda \left(\langle a_m^\dagger a_{l-1}^\dagger \rangle + \langle a_m^\dagger a_{l+1}^\dagger \rangle + \langle a_{m-1}^\dagger a_l^\dagger \rangle + \langle a_{m+1}^\dagger a_l^\dagger \rangle \right). \end{aligned} \quad (5.11)$$

The other correlation functions are related to the correlation functions considered above by the identities imposed due to the anti-commutation relations of the fermionic modes, thus $\langle a_m^\dagger a_l^\dagger \rangle = \langle a_m a_l \rangle^*$ and $\langle a_m^\dagger a_l \rangle = \delta_{l,m} - \langle a_l a_m^\dagger \rangle$. The steady state correlations can be computed from these equations by requiring that $\partial_t \langle a_m^\dagger a_l \rangle = \partial_t \langle a_m^\dagger a_l^\dagger \rangle = 0$. This leads to a set of difference equations. The current density as well as the particle number density can be expressed in terms of these correlators. One finds for the particle number density $\langle n_k \rangle = \langle a_k^\dagger a_k \rangle$ and the two contributions to the current read,

$$\begin{aligned} j_{k,k+1}^{st} &= \varphi (\langle n_k \rangle - \langle n_{k+1} \rangle) \\ j_{k,k+1}^{co} &= \frac{\lambda}{i} \left(\langle a_{k+1}^\dagger a_k \rangle - \langle a_k^\dagger a_{k+1} \rangle \right). \end{aligned} \quad (5.12)$$

Note, that the stochastic current now only depends on the difference of the densities at adjacent sites and thus greatly simplifies with respect to (5.6). With these definitions at hand it is possible to compute the current density as well as the particle number density explicitly. We only need to restrict ourself to the equations (5.10) for the choices $l = m = k$ and $m = k, l = k + 1$ as well as $m = k + 1, l = k$ and we obtain the following difference equations.

$$\alpha \delta_{k,1} (1 - \langle n_k \rangle) + j_{k-1,k}^{st} + j_{k-1,k}^{co} = \beta \delta_{k,N} \langle n_k \rangle + j_{k,k+1}^{st} + j_{k,k+1}^{co} \quad (5.13)$$

$$\alpha \delta_{k,1} j_{k,k+1}^{co} + 4\varphi j_{k,k+1}^{co} = 4 \frac{\lambda^2}{\varphi} j_{k,k+1}^{st} - \beta \delta_{k+1,N} j_{k,k+1}^{co}. \quad (5.14)$$

Let us first consider the scenario, when $\varphi > 0$. Thus, we have to take into account the full set of equations. One sees, that in the bulk, i.e. $k \in \{2, \dots, N-1\}$, the density has to satisfy the difference equation $\langle n_{k+1} \rangle - 2\langle n_k \rangle + \langle n_{k-1} \rangle = 0$. We see that the assignment $\langle n_k \rangle = c_1 + kc_2$ satisfies this equation. We need to determine the two constants based on

the boundary conditions, i.e. $k = 1, N$. We obtain a linear system of equations that is easily soluble and we obtain

$$\begin{aligned} c_1 &= \frac{\alpha\beta\varphi N + \alpha\varphi\left(\varphi + \frac{4\lambda^2}{\beta+4\varphi}\right)}{\alpha\beta\varphi(N-1) + \alpha\varphi\left(\varphi + \frac{4\lambda^2}{\beta+4\varphi}\right) + \beta\varphi\left(\varphi + \frac{4\lambda^2}{\alpha+4\varphi}\right)} \\ c_2 &= -\frac{\alpha\beta\varphi}{\alpha\beta\varphi(N-1) + \alpha\varphi\left(\varphi + \frac{4\lambda^2}{\beta+4\varphi}\right) + \beta\varphi\left(\varphi + \frac{4\lambda^2}{\alpha+4\varphi}\right)}. \end{aligned} \quad (5.15)$$

Note, that $c_2 < 0$, so the density is a line that decreases from some fixed value $c_1 > 0$ on the left to $c_1 - N|c_2|$ on the right. From the density, we can immediately deduce the stochastic contribution to the current in the bulk, which reads $j_{k,k+1}^{st} = -\varphi c_2$. Due to the second equation we can also infer the coherent contribution, which is $j_{k,k+1}^{co} = -\frac{\lambda^2}{\varphi} c_2$. We now consider the thermodynamic limit $N \gg 1$. In this limit both the coherent contribution and the stochastic contribution behave as

$$j^{st} \approx \frac{\varphi}{N} \quad \text{and} \quad j^{co} \approx \frac{\lambda^2}{\varphi N}. \quad (5.16)$$

We recall that the SEP without any further driving, i.e. $\alpha = \beta = 0$, obeys the detailed balance condition and thus does not support a steady state current. When one allows for an external driving of the particles at the boundaries, as we do in our example, a current is induced in the SEP steady state. This current, however, vanishes as $\sim 1/N$ in the system size N . As we have shown, this behavior does not change when adding the coherent evolution on top. Both the coherent as well as the stochastic contribution to the current vanish in the same fashion. Furthermore, neither c_1 nor c_2 depend strongly on λ . The coherent evolution only seems to play a role for smaller system sizes, i.e. small N . We deduce from this, that for all finite φ the quantum perturbation to the SEP is an irrelevant perturbation and does not lead to a qualitatively different behavior of the system's transport properties. However, whether the quantum perturbation is completely irrelevant can not be deduced from just considering the steady state density and the current alone. One would also need to take higher order correlations into account, as for instance the current-current correlation function at unequal times.

The equations (5.13) do exhibit a phase transition, albeit a quite naive transition, upon choosing $\varphi = 0$. It is easy to see that for this value the system behavior changes abruptly. The model that is obtained by setting $\varphi = 0$ corresponds to limiting case of another model for quantum transport that was investigated recently [124, 125]. This model only has coherent transport in the bulk, and stochastic driving only occurs at the boundaries. The equations immediately yield that the current-density $j_{k,k+1}^{co} = j^B$, as well as the particle density $\langle n_k \rangle = n^B$, is constant in the bulk and only deviates at the boundaries from this constant value. With this at hand, the set of equations simplify greatly and turn into a set of algebraic equations. The resulting particle density and current in the bulk are given by

$$n^B = \frac{\alpha(\beta^2 + 4\lambda^2)}{\beta(\alpha^2 + 4\lambda^2) + \alpha(\beta^2 + 4\lambda^2)}, \quad j^B = \frac{4\alpha\beta\lambda^2}{\beta(\alpha^2 + 4\lambda^2) + \alpha(\beta^2 + 4\lambda^2)}. \quad (5.17)$$

The boundary densities n_1 and n_k , turn out to be different from the density in the bulk. For these densities, we obtain

$$n_1 = \frac{\beta\alpha^2 + \alpha(\beta^2 + 4\lambda^2)}{\beta(\alpha^2 + 4\lambda^2) + \alpha(\beta^2 + 4\lambda^2)}, \quad n_N = \frac{4\alpha\lambda^2}{\beta(\alpha^2 + 4\lambda^2) + \alpha(\beta^2 + 4\lambda^2)}. \quad (5.18)$$

We see that the current as well as the particle density is independent of the lattice size N . This model with only coherent transport has a non-vanishing current in the thermodynamic limit.

5.4 The asymmetric exclusion process

Before we now turn to the asymmetric exclusion process, introduced in section 4.4 of the previous chapter, we would like to briefly review the concept of matrix product operators (MPO), and in particular so-called matrix product density operators (MPDO) [126]. We have already introduced MPS and sMPS in the previous chapters. One can extend this formal construction to operators that act on a Hilbert space $\mathcal{H} = \mathbb{C}^{d^N}$ by defining an MPO as

$$O = \sum_{i_1, \dots, i_N=1}^{d^2} \text{tr} [M_{i_1}^1 \dots M_{i_N}^N] \sigma_{i_1}^1 \otimes \dots \otimes \sigma_{i_N}^N, \quad (5.19)$$

where the $\sigma_{i_k}^k$ constitute an operator basis of the local matrix space $\mathcal{M}(\mathbb{C}^d)$. For a qubit, for instance, they could resemble the Pauli operators augmented with the identity. Due to their construction MPOs can be treated like standard matrix product states on an enlarged Hilbert space, i.e. the tensor product of N local spaces which now are \mathbb{C}^{d^2} . If we now require in addition that an MPO is a positive operator with trace of unity, we are dealing with matrix product density operators. This way, mixed states of many-body systems can be approximated with lesser parameters. The requirement that some MPO has to be positive can be enforced by the following construction. Suppose we start with some arbitrary MPS on a Hilbert space that is doubled locally, we can see this state as the purification of some density matrix.

$$|\psi\rangle = \sum_{i_1, j_1, \dots, i_N, j_N} \text{tr} [A_{i_1, j_1}^1 \dots A_{i_N, j_N}^N] |i_1\rangle |j_1\rangle \dots |i_N\rangle |j_N\rangle \quad (5.20)$$

Then, if we trace out the doubled Hilbert space $\rho = \text{tr}_{j_1 \dots j_N} [|\psi\rangle \langle\psi|]$, we obtain a state ρ that is of the form of the MPO in (5.19), where now the matrices $M_{i_k}^k$ are given by $M_{i_k, j_k}^k = \sum_{l=1}^d A_{i_k, l}^k \otimes \bar{A}_{l, j_k}^k$ and we have the standard product basis $\sigma_{i_k j_k}^k = |i_k\rangle \langle j_k|$ as the local operator basis. This construction ensures that the resulting state is positive definite. However, not every matrix product density operator can be written this way.

5.4.1 The Derrida algebra

We now turn to the description of the steady state of the master equation (5.3) when we choose the stochastic hopping parameters to resemble those of the asymmetric exclusion process, that

is we choose $\varphi_L = 0$ and $\varphi_R = \varphi$. Recall, that we briefly discussed the Derrida algebra [116] already in section 4.4. This concept can of course also be generalized to MPOs. We define a translationally invariant density matrix with open boundary conditions, i.e. we choose the matrices site independent, and write:

$$\tilde{\rho} = \langle l | \left(\begin{array}{cc} G & B \\ A & E \end{array} \right)^{\otimes N} | r \rangle, \quad (5.21)$$

where we have to require that $A = \overline{B}$ and G, E real due to Hermiticity. Note, that we have adopted a change of notation with respect to the section 4.4 here, since we want the operator σ_k^+ to create a particle at site k and σ_k^- to annihilate one. We therefore have to perform a particle hole transformation, which leads to a left-right swap, in order to be consistent with the previous algebra. It is now possible to impose the same algebraic constraints on the MPO that we have imposed for the stochastic matrix product states. That is, if we split the master equation (5.3) into individual summands that constitute only two body interactions and write

$$\mathcal{L}[\rho] = \mathcal{L}_1[\rho] + \sum_{k=1}^{N-1} \mathcal{L}_{k,k+1}[\rho] + \mathcal{L}_N[\rho], \quad (5.22)$$

we can require that the matrices A, B, E, G together with some ancilla matrices that we mark by \hat{A}, \hat{B}, \dots , have to satisfy the constraints

$$\begin{aligned} \mathcal{L}_{k,k+1} \left[\left(\begin{array}{cc} G & B \\ A & E \end{array} \right) \otimes \left(\begin{array}{cc} G & B \\ A & E \end{array} \right) \right] = \\ \left(\begin{array}{cc} \hat{G} & \hat{B} \\ \hat{A} & \hat{E} \end{array} \right) \otimes \left(\begin{array}{cc} G & B \\ A & E \end{array} \right) - \left(\begin{array}{cc} G & B \\ A & E \end{array} \right) \otimes \left(\begin{array}{cc} \hat{G} & \hat{B} \\ \hat{A} & \hat{E} \end{array} \right). \end{aligned} \quad (5.23)$$

Furthermore, we require that the single-site operators at the boundaries have to satisfy

$$\begin{aligned} \langle l | \mathcal{L}_1 \left[\left(\begin{array}{cc} G & B \\ A & E \end{array} \right) \right] &= - \langle l | \left(\begin{array}{cc} \hat{G} & \hat{B} \\ \hat{A} & \hat{E} \end{array} \right) \\ \mathcal{L}_N \left[\left(\begin{array}{cc} G & B \\ A & E \end{array} \right) \right] | r \rangle &= \left(\begin{array}{cc} \hat{G} & \hat{B} \\ \hat{A} & \hat{E} \end{array} \right) | r \rangle. \end{aligned} \quad (5.24)$$

We see that the total sum (5.22) telescopes to zero and $\tilde{\rho}$ is the steady state solution of the equation. For a suitable decomposition into two body terms that correspond to $H_{k,k+1} = \lambda(\sigma_k^x \otimes \sigma_{k+1}^x + \sigma_k^y \otimes \sigma_{k+1}^y)$ and the Lindblad operators $L_{k,k+1}$ with the two boundary terms L_1, L_2 we can derive the following algebra for the steady state. The eight matrices need to satisfy sixteen equations in the bulk (5.23). First, all matrices have to commute with their ancilla counterpart, that is $[A, \hat{A}] = [B, \hat{B}] = [E, \hat{E}] = [G, \hat{G}] = 0$. The remaining equations are then as follows

$$-\frac{2\lambda}{i}BG - \frac{\varphi}{2}GB = \hat{G}B - G\hat{B} \quad \text{and} \quad -\frac{2\lambda}{i}GB = \hat{B}G - B\hat{G}, \quad (5.25)$$

$$-\frac{2\lambda}{i}EA - \frac{\varphi}{2}AE = \hat{A}E - A\hat{E} \quad \text{and} \quad -\frac{2\lambda}{i}AE = \hat{E}A - E\hat{A}, \quad (5.26)$$

as well as

$$\frac{2\lambda}{i}AG - \frac{\varphi}{2}GA = \hat{G}A - G\hat{A} \quad \text{and} \quad \frac{2\lambda}{i}GA = \hat{A}G - A\hat{G}, \quad (5.27)$$

$$\frac{2\lambda}{i}GE - \frac{\varphi}{2}BE = \hat{B}E - B\hat{E} \quad \text{and} \quad \frac{2\lambda}{i}BE = \hat{E}B - E\hat{B} \quad (5.28)$$

and finally we have also

$$\frac{2\lambda}{i}[A, G] - \varphi GE = \hat{G}E - G\hat{E} = E\hat{G} - \hat{E}G, \quad (5.29)$$

$$-\frac{2\lambda}{i}[G, E] - \frac{\varphi}{2}BA = \hat{B}A - B\hat{A} \quad \text{and} \quad \frac{2\lambda}{i}[G, E] - \frac{\varphi}{2}AB = \hat{A}B - A\hat{B}. \quad (5.30)$$

We furthermore need to satisfy the constraints set by the boundary terms that govern the inflow and outflow of the particles. The edge algebra derived from (5.24) then reads

$$\begin{aligned} \langle l | \begin{pmatrix} E & B \\ A & E \end{pmatrix} &= \langle l | \begin{pmatrix} -\frac{1}{\alpha}\hat{G} & \frac{2}{\alpha}\hat{B} \\ \frac{2}{\alpha}\hat{A} & \frac{1}{\alpha}\hat{E} \end{pmatrix} \\ \begin{pmatrix} G & B \\ A & G \end{pmatrix} | r \rangle &= \begin{pmatrix} -\frac{1}{\beta}\hat{G} & -\frac{2}{\beta}\hat{B} \\ -\frac{2}{\beta}\hat{A} & \frac{1}{\beta}\hat{E} \end{pmatrix} | r \rangle. \end{aligned} \quad (5.31)$$

As one sees, this algebra becomes significantly more complex in the quantum setting and it appears to be intractable to find an algebraic representation for the eight matrices. However, let us briefly consider the classical case setting $\lambda = 0$. Then the algebra simplifies greatly and we can set $A = \hat{A} = B = \hat{B} = 0$. Choosing furthermore $-\hat{G} = \hat{E} = \mathbb{1}$, we recover the classical algebra $GE = E + G$ and $\langle l | E = 1/\alpha \langle l |$ as well as $G | r \rangle = 1/\beta | r \rangle$. For this a representation can be found [116], such as for example the Fock representation

$$\begin{aligned} \langle l | &= \sum_{k=0}^N \frac{1}{\alpha^k} \langle k | \quad \text{and} \quad | r \rangle = | 0 \rangle \\ E &= \sum_{k=0}^{\infty} | k \rangle \langle k-1 | \quad \text{and} \quad G = \sum_{n=0}^{\infty} \frac{1}{\beta} | 0 \rangle \langle n | + \sum_{n=1}^{\infty} \sum_{m=1}^n | m \rangle \langle n |, \end{aligned} \quad (5.32)$$

which is a representation that is infinite dimensional. In fact, one can show easily that all representations have to be infinite dimensional unless $\alpha + \beta = 1$. From this representation, the normalization, or partition function, can be computed as $Z = \text{tr} [\tilde{\rho}]$ as done in [116].

However, we want to understand the system's response to a quantum mechanical perturbation. To this end suppose that $\lambda > 0$, so the algebra does not decouple any longer and the A and B terms mix with the matrix algebra for G and E . It is interesting to note, that for the choice $\alpha + \beta = 1$ the complete algebra is indeed soluble independently of the choice for λ . This domain is depicted in the phase diagram of the ASEP as the blue dashed line, see Fig. 5.2. The solution is then given by just the classical mean field solution. This is only possible, since E and G commute in equation (5.30). The full state is therefore given by

$$\rho = \bigotimes_{k=1}^N \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}. \quad (5.33)$$

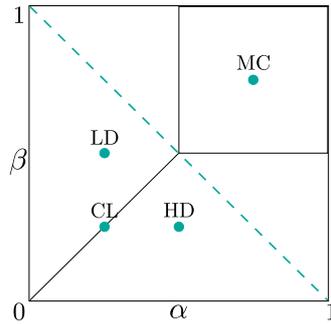


Figure 5.2: Classical phase diagram of the asymmetric exclusion process. The simulation was performed for four different points in the classical phase diagram. These points are depicted as blue dots which correspond to different values of (α, β) . For the coexistence line (CL) we have $(1/4, 1/4)$. For the maximum current phase (MC) we have $(3/4, 3/4)$ and for the high density (HD) and low density phase (LD) we have $(1/2, 1/4)$ and $(1/4, 1/2)$ respectively.

Note, that coherence can only build up, when the E and G do not commute any longer. Hence, we will only be able to find that the state has some form of coherence, when the system is also classically correlated.

5.4.2 Dynamical MPS approach to open quantum systems

In order to see that in the regime, where the stochastic steady state is correlated, the coherent evolution alters the steady state, we need to calculate the steady state of the system numerically by time-evolving the density matrix, until we reach the steady state. The numerical simulations of the real time evolution is performed by making use of an algorithm for the propagation of matrix product density operators [126, 127]. This algorithm works as follows: Starting from the initial density-matrix ρ_0 given as an MPDO, we apply the tcp-map $\mathcal{E}(\mathcal{L}, t) = \exp(t\mathcal{L})$ for a small time step Δt and approximate the resulting density operator, that has now an increased bond dimension, with an MPDO that has a bond-dimension D_k corresponding to that of the original MPDO. The approximation of the operator $\rho(t + \Delta t) = \mathcal{E}(\mathcal{L}, \Delta t) \rho(t)$ is chosen, such that the Hilbert-Schmidt norm $\|\rho(t + \Delta t) - \rho_{new}\|_{HS}^2 = \text{tr}[(\rho(t + \Delta t) - \rho_{new})^2]$ is minimized. This optimization can be performed efficiently by sweeping from left to right over the individual sites and optimizing the matrices M_{i_k, j_k}^k locally. For the application of the tcp-map to be computable, we perform a second-order Trotter expansion of the tcp-map as follows:

$$\mathcal{E}(\mathcal{L}, \Delta t) \simeq \mathcal{E}(\mathcal{L}_o, \Delta t/2) \mathcal{E}(\mathcal{L}_e, \Delta t) \mathcal{E}(\mathcal{L}_o, \Delta t/2), \quad (5.34)$$

where $\mathcal{L} = \mathcal{L}_e + \mathcal{L}_o$ corresponds to a splitting of the Liouvillian into commuting terms which act on the sites $(2k, 2k + 1)$ and $(2k - 1, 2k)$, respectively. The resulting MPDO ρ_{new} is then chosen as initial condition for the next step and the procedure is repeated. For a more detailed description of the algorithm, the reader is referred to [126, 127].

As initial state for the evolution we chose the classical steady state (4.24) from the previous chapter. The matrices of the steady state can thus be chosen with a bond dimension of

$D = N + 1$. We then changed the value of $\lambda = 0$ to $\lambda = 1/2$ and $\lambda = 1$, and evolved the MPDO, until the steady state was reached, i.e. until all considered observables did not change any more. Negative values of λ were also simulated and led to the same results. We conclude from this, that the systems response only depends on the absolute value of λ . The simulations were done for a lattice with $N = 40$ sites. The matrix bond dimension of the MPDO was chosen as $D = 60$ and we chose a Trotter step $\Delta t = 10^{-4}$. To get a better understanding of how the system responds to the quantum perturbation in each of the different phases, we computed the steady state at different values of α and β , which correspond to points lying in different phases. Four different points were selected, see Fig. 5.2. The point that corresponds to the coexistence line (CL), where $\alpha = \beta$ and $\beta + \alpha \leq 1$ was chosen as $\alpha = \beta = 1/4$. Recall that the classical steady state is critical along the coexistence line, which separates the high density phase from the low density phase. In the maximum current phase (MC), with $\alpha > 1/2$ and $\beta > 1/2$ we chose $\alpha = \beta = 3/4$. For the low-density (LD), with $\alpha > 1/2, \beta < 1/2$, and the high-density (HD) phase $\alpha < 1/2, \beta > 1/2$, we chose $\alpha = 1/4 \quad \beta = 1/2$ and $\alpha = 1/2 \quad \beta = 1/4$ respectively.

The observable we considered first was the density distribution $\langle n_k \rangle = \text{tr}[n_k \rho]$ as a function of the lattice site k , Fig 5.3. Furthermore, we calculated the values of the two-point correlation functions, of the densities $n_k = \sigma_k^+ \sigma_k^-$ for all pairs (i, j) of sites

$$\langle n_i n_j \rangle^c = \langle n_i n_j \rangle - \langle n_i \rangle \langle n_j \rangle. \quad (5.35)$$

The expectation values are taken with respect to the system's steady state. In the figures Fig. 5.4(a-d), the correlation functions are compared to the different contributions to the current j^{tot} defined in (5.6), for different values of $\lambda = 0, 1/2, 1$. The first observation to be made is that the individual contributions to the total current are no longer constant throughout the lattice anymore. They show a dependence on the lattice site. The total current, however, i.e. $j_{tot} = j_{co} + j_{st}$, is still constant at each site of the lattice, as is required since the system is in a steady state.

The low- and high-density phases Fig. 5.4 (a),(b) : These two phases are, just as in the classical set up, related by a particle-hole transformation by exchanging the ordering of the lattice sites from left to right. All the plots reflect this symmetry. One observes that the correlation functions in the classical regime, $\lambda = 0$, are already quite short ranged and decay rapidly. The quantum perturbation in both cases leads to a further decay of the correlations. The total current remains stable with respect to the quantum perturbation and does not change its value notably. The individual constituents to the current, however, change their behavior. At the boundaries the stochastic contributions are increased, whereas the coherent current gives rise to a flow in the opposite direction.

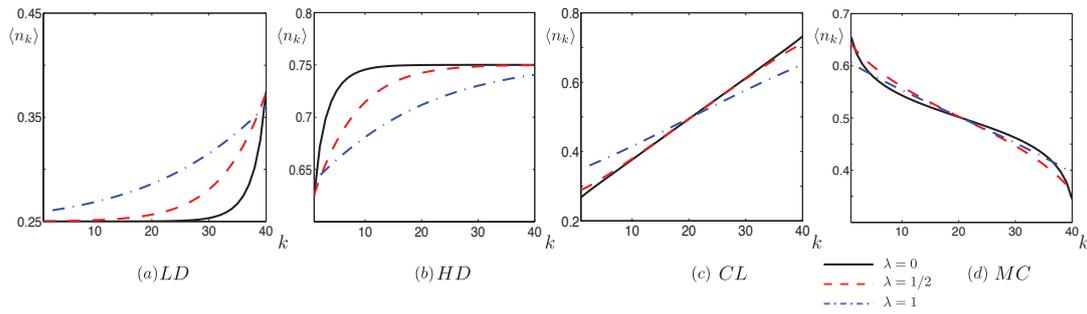


Figure 5.3: (color online) Density distribution $\langle n_k \rangle$ for the different points (α, β) , (a) HD $(1/2, 1/4)$, (b) LD $(1/4, 1/2)$ and (c) MC $(3/4, 3/4)$ as well as (d) CL $(1/4, 1/4)$. The black solid line corresponds to $\lambda = 0$, i.e. the classical solution. The red dashed line corresponds to a quantum perturbation with $\lambda = 1/2$, and the blue dashed-dotted line to a perturbation $\lambda = 1$.

The maximum current phase and coexistence line Fig. 5.4(c),(d): In the classical process, the MC phase corresponds to the maximum amount of current the system can carry. Allowing for a quantum perturbation, the system makes use of the additional transport capacity and increases its total current. For these boundary conditions the stochastic as well as the coherent contributions flow in the same direction. The classical correlation function initially assumes negative values close to the boundaries. The onset of the quantum perturbation also reduces the magnitude of the correlations in this phase, even though the total amount of current is increased. For the chosen boundary conditions that correspond to the coexistence line, the amount of correlations initially present in the steady state are decreased, when switching on the quantum perturbation. The final steady state, were $\lambda = 1$, however, still shows the presence of correlations to a higher degree than in the other phases. The total amount of current carried by the system, however, is decreased. The coherent contribution is negative throughout the system.

Conclusions and Outlook

Conclusions chapter 2: We have seen that by generalizing the χ^2 -divergence to the quantum setting, many of the classical results for the convergence of Markov processes can be recovered. The general perception, that the convergence should be governed by the spectral properties of the quantum channel could be verified in the asymptotic limit. The fact that we were working with non-commuting probabilities gave rise to a larger set of possibilities of defining an inversion of the fixed point density matrix, all of which give rise to a valid upper bound to the trace-distance. An interesting question is how the different singular values s_i^k of the corresponding quantum discriminant relate to each other. The generalization of the χ^2 -divergence also led to the definition of detailed balance for quantum channels. Again, not only a single condition for quantum detailed balance exists, but an entire family of conditions each determined by a different function $k \in \mathcal{K}$, all of which coincide in the case when we consider classical stochastic processes on a commuting subspace. The quantum concept of detailed balance therefore appears to be richer and allows for a wider set of channels to obey this definition. The conductance bound that was derived could only be shown for unital quantum channels. However, we would like to point out, that it is possible to give conductance bounds for classical maps when the Markov chain is not doubly stochastic. The fact that in general we may not assume that the fixed point of an arbitrary channel commutes with the eigenvector associated to the second largest eigenvalues seems to hinder a generalization for non-unital channels. Moreover, the classical conductance bound has a nice geometrical interpretation in terms of the cut-set analysis and the maximal flow on the graph associated to the stochastic matrix P_{ij} . When investigating general quantum channels such a nice geometric interpretation seems to be lacking. For unital quantum channels Cheeger's constant can also be viewed in terms of the minimal probability flow of one subspace to its complement.

Conclusions chapter 3: Even though an implementation of this algorithm for full scale quantum many-body problems may be out of reach for today's technological means, we have presented an algorithm that is indeed scalable to system sizes that are interesting for actual physical simulations. A small scale implementation of the algorithm that can be achieved with present day technology has been presented and we will include a discussion that sketches the basic steps necessary for a simulation of some notoriously hard quantum many-body problems in the following. As in the classical setting the convergence rate and hence the runtime of the algorithm is dictated by the spectral gap of the stochastic map. The scaling of the gap depends on the respective problem Hamiltonian and the choice of updates $\{C\}$. Just as for the classical Metropolis algorithm, efficient thermalization is of course not expected for an arbitrary Hamiltonian. This would allow one to solve QMA-complete problems in polynomial time

[128, 129, 78]. It is, however, expected that the algorithm will thermalize, if the physical system of interest thermalizes. We have presented a simple physical system, i.e. the XX -chain in a transverse magnetic field at $T = 0$, for which the inverse gap of the quantum Metropolis map, with a simple single spin flip update, scales like $\mathcal{O}(1/N)$ with N the number of spins, even at criticality. To prove a polynomial scaling of the gap for more complex Hamiltonians remains a challenging open problem. Also, it is well known that the choice of updates $\{C\}$ can have a dramatic impact on the convergence rate of the Markov chain in the classical setting. Finding good updates in the quantum setting is a very interesting open question, although the above example suggests that the problem might be simpler in the quantum than in the classical case. The algorithm can be seen as a classical random walk on the eigenstates of the Hamiltonian. All samples are thus computed with respect to the actual eigenstates. This is why our method is suitable for the simulation of fermionic systems by exploiting the Jordan - Wigner transformation [123] as discussed in [130]. The fermionic sign problem is therefore not an issue for the quantum Metropolis algorithm. It is worth noting that an additional quadratic speedup might be achievable using the methods of [131, 132, 94].

Simulation of quantum many-body systems: It would go far beyond the scope this thesis to give a faithful account on only the most eminent applications of the quantum Metropolis algorithm to the simulation of quantum many body systems. We will therefore give only a brief sketch on how we expect that the devised quantum algorithm will aid in the computation of static properties of some notoriously hard problems in quantum physics, that have eluded direct computation for large system sizes by classical means. Such problems are for instance the determination of the phase diagram of the Hubbard model, the computation of binding energies of complex molecules in quantum chemistry and the determination of the hadron masses in gauge theories. Common to these problems is that the particles are strongly interacting fermions and bosons. We expect that it is this class of problems where our algorithm will be able to give the strongest contributions.

In order to implement the quantum Metropolis algorithm for a specify many-body Hamiltonian H we need to be able to perform the phase estimation algorithm efficiently. The central subroutine that needs to be implemented is therefore the simulation of the time evolution for the Hamiltonian $H \otimes \hat{p}$, as was discussed previously in section 3.2. The simulation method described in [87] relies on the fact that we are able to decompose the Hamiltonian into a sum of local hamiltonians h_l with $H = \sum_l h_l$ that can by themselves be simulated on a quantum computer efficiently. A method to rephrase fermionic or bosonic degrees of freedom in terms of the quantum computational degrees of freedom, that is in terms of qubits, is therefore needed. Such a program was devised in [130, 133, 134] and we merely give a brief overview here and refer the reader to the corresponding references.

The Hubbard model: The Hubbard model [135] is based on a tight binding approximation that describes electrons in a periodic potential confined to move only in the lowest Bloch band. The Hubbard Hamiltonian consists of a hopping term and an interaction term written in form of fermionic creation $c_{i,\sigma}^\dagger$ and annihilation $c_{i,\sigma}$ operators that act on a lattice site i in a regular lattice of N sites.

$$H = -t \sum_{\langle i,j \rangle, \sigma} \left(c_{i,\sigma}^\dagger c_{j,\sigma} + c_{j,\sigma}^\dagger c_{i,\sigma} \right) + U \sum_i n_{i,\downarrow} n_{i,\uparrow} \quad (5.36)$$

This Hamiltonian has to be expressed in terms of spin degrees of freedom in order to be implemented in the standard quantum circuit formulation. The interaction term can be seen to be implementable directly since the particle density $n_{i,\sigma}$ operator acts only locally and is bosonic in nature. The implementation of the hopping term is a bit more challenging. Consider for simplicity the hopping term for a single electron spin only. This part can be expressed in terms of the Jordan-Wigner transformation, cf. Fig. 5.5, as

$$t \sum_{\langle i,j \rangle} \frac{1}{2} \left(\sigma_i^x (\otimes_{k=i+1}^{j-1} \sigma_k^z) \sigma_j^x + \sigma_i^y (\otimes_{k=i+1}^{j-1} \sigma_k^z) \sigma_j^y \right), \quad (5.37)$$

once a specific order of the N lattice sites has been chosen. As is shown in Fig. 5.5 the unitary evolution of each individual summand can be implemented with a cost that scales at most linearly with the total system size [133, 134]. More general fermionic Hamiltonians can be implemented in a similar fashion.

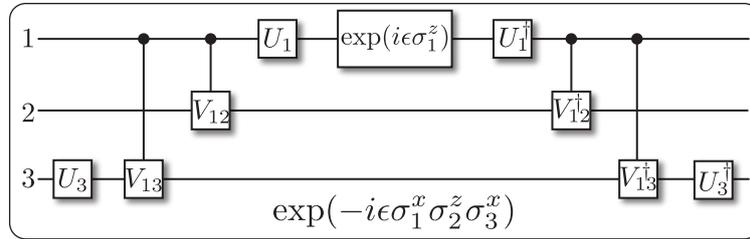


Figure 5.5: A fermionic many particle Hamiltonian can be simulated on a quantum computer by mapping the fermionic degrees of freedom to spin-1/2 particles [133, 134]. Such a mapping is given by the famous Jordan-Wigner transformation. Here the fermionic algebra can be expressed in terms of the $su(2)$ algebra via $c_k^\dagger = -(\otimes_{l=1}^{k-1} \sigma_l^z) \sigma_k^+$, where $\sigma_k^+ = \frac{1}{2} (\sigma_k^x + i\sigma_k^y)$. The dynamical part of the fermionic many-body Hamiltonian often contains terms of the form $h_{kj} = c_k^\dagger c_j + c_j^\dagger c_k$, which become non-local after the transformation. Operators that are not adjacent in terms of the labeling often contain a chain of Pauli σ^z operators in between them. A typical term of this kind that occurs after this transformation is $h_{kj}^X = \sigma_k^x (\otimes_{l=k+1}^{j-1} \sigma_l^z) \sigma_j^x$. To simulate the time evolution of such a non local term on a quantum computer we need to be able to decompose this unitary into two qubit gates. Given the two unitaries $V_{kl} = \exp(i\frac{\pi}{4} \sigma_k^z \sigma_l^z)$ and $U_l = \exp(i\frac{\pi}{4} \sigma_l^y)$ such a decomposition is indeed possible as depicted in the above circuit for the evolution of $\exp(-i\epsilon \sigma_1^x \sigma_2^z \sigma_3^x)$.

Quantum chemistry: A central problem in Quantum chemistry is the determination of molecule properties. The major challenge is the determination of the electron binding ener-

gies that need to be computed in dependence of the nuclei position. The general approach to this problem is to solve the approximate Hamiltonian of the electronic degrees of freedom that arises due to the Born-Oppenheimer approximation. In this approximation the nuclei positions are external parameters in the electronic Hamiltonian. The calculation of the molecule properties relies on the fact that the electronic energy can be determined efficiently in dependence of the nuclei position. In their paper [136], Kassal and Aspuru-Guzik show how a quantum computer could be used to determine molecule properties at a time that is a constant multiple of the time needed to compute the molecular energy. The quantum Metropolis algorithm would function here as a black box computing the energy. As is shown in [137], the phase estimation procedure can be implemented efficiently for a general second quantized chemical Hamiltonian.

Gauge theories: The current most common non-perturbative approach to QCD is Wilson's lattice gauge theory [138], which maps the problem to one of statistical mechanics where the Euclidean action now assumes the role of a classical Hamilton function. It is therefore reasonable to assume, that lattice gauge theories would also be the method of choice for the quantum Metropolis algorithm. However, the algorithm relies on a Hamiltonian formulation of the problem. Such a formulation is given by Kogut and Susskind's [139] Hamiltonian formulation of lattice gauge theories in $3 + 1$ dimensions. Here the 3-dimensional space is discretized and put on a cubic lattice, while time is left continuous. The fermions reside on the vertices of the lattice while the gauge degrees of freedom are put on the links. The physical subspace is required to be annihilated by the generators of the gauge transformation, i.e. all physical states need to satisfy Gauss's law.

It however turns out, that this approach seems to be very hard to implement on a quantum computer. This is due to the fact that each of the links carries a Hilbert space that is infinite dimensional, namely the space of all square integrable functions on the corresponding gauge group $SU(N)$. A finite approximation to this Hilbert space therefore leads immediately to a breakdown of the underlying symmetry. A different formulation of gauge theories, that does not suffer from this problem, is therefore needed. Such a formulation is given in terms of quantum link models introduced by Horn [140]. Brower et al. showed that QCD and in general any $SU(N)$ gauge theory can be expressed as a quantum link model [141]. In the quantum link formulation the classical statistical mechanics problem is replaced by a problem formulated in terms of quantum statistical mechanics in which the classical Euclidean action is replaced by a quantum Hamiltonian. The central feature is that the corresponding Hilbert space of the gauge degrees of freedom at each link is now finite. It suffices that each link of a $SU(N)$ link model carries a single, finite, representation of $SU(2N)$. This is achieved by formulating the problem in $4 + 1$ dimensions, where the four physical dimensions correspond to the actual physical Euclidean space time, while the additional dimension plays the role of an additional unphysical dimension. The 4-dimensional Euclidean space time is discretized and lives on a

cubic lattice. Furthermore it was shown by Brower et al. [141], that the continuum limit is obtained by sending the fifth unphysical Euclidean dimension to infinity, which corresponds to preparing the ground state of the lattice Hamiltonian. It can be seen, that the $4 + 1$ dimensional link models are related to standard gauge theories in 4 dimensions via dimensional reduction [142]. The full Hilbert space of the $SU(3)$ gauge theory can be written as the tensor product of a 20-dimensional Hilbert space for each link of the lattice and the finite dimensional fermionic Hilbert space that resemble the quarks. In contrast to the standard lattice gauge theories the configuration space of the quantum link model resembles the space of quantum spin models. The physical spectrum, and by that the Hadron masses, of the 4-dimensional theory can be obtained from computing the correlation functions in the Euclidean direction on the ground state of the 4-dimensional lattice Hamiltonian.

Conclusions chapter 4: We have revisited the notion of stochastic matrix product states, and showed that a low bond dimension suffices to efficiently parametrize steady states of non-equilibrium distribution, if the entropy cost in the system is low. This opens up the interesting question of how to characterize the conditions under which such steady states have a low entropy cost. It would be interesting to see to what extent this relates to the gap of the corresponding stochastic process. This also opens up novel ways for constructing numerical renormalization group methods for simulating non-equilibrium systems in the line of the MPS algorithms for quantum spin chains [39, 40, 4].

Conclusions chapter 5: We have investigated a quantum perturbation to the dynamics of the stochastic asymmetric exclusion process as well as to the symmetric exclusion process. We find that we can rephrase the stochastic master equation as a quantum equation that fully reproduces the classical dynamics. The quantum perturbations modify the steady state behavior and allow for two different types of currents, which, each on their own, can vary as a function of the site. Numerical simulations of the full master equation indicate, that the underlying classical phase-diagram of the stochastic process is respected. The steady state responds to driving due to the boundary terms with a different behavior in current and density. A further step would be to investigate the current-current correlation function of the SEP, to see whether the quantum perturbation has an effect to the current fluctuations. Furthermore, other, more complex models with an interplay between stochastic and coherent dynamics can be investigated along these lines.

Bibliography

- [1] P. A. M. Dirac. Quantum mechanics of many-electron systems. *Proc. R. Soc. London, Ser. A.*, 123(714), 1929.
- [2] L. D. Landau and E. M. Lifshitz. *Quantum Mechanics: Nonrelativistic Theory*. Pergamon Press, 1997.
- [3] M. M. Wolf D. Perez-Garcia, F. Verstraete and J. I. Cirac. Matrix product state representations. *Quantum Inf. Comput.*, 7:401, 2007.
- [4] V. Murg F. Verstraete and J. I. Cirac. Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems. *Adv. Phys.*, 57:143, 2008.
- [5] Stefan Rommer and Stellan Östlund. Class of ansatz wave functions for one-dimensional spin systems and their relation to the density matrix renormalization group. *Phys. Rev. B*, 55(4):2164–2181, Jan 1997.
- [6] M. Fannes, B. Nachtergaele, and R. F. Werner. Finitely correlated states on quantum spin chains. *Commun. Math. Phys.*, 144:443, 1992.
- [7] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculation by fast computing machines. *J. Chem. Phys.*, 21:1087, 1953.
- [8] S. Durr, Z. Fodor, J. Frison, C. Hoelbling, R. Hoffmann, S. D. Katz, S. Krieg, T. Kurth, L. Lellouch, T. Lippert, K. K. Szabo, and G. Vulvert. Ab Initio Determination of Light Hadron Masses. *Science*, 322(5905):1224–1227, 2008.
- [9] D. Geman S. Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intell.*, 6:721, 1984.
- [10] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by Simulated Annealing. *Science*, 220(4598):671–680, 1983.
- [11] ed. M Suzuki. *Quantum Monte Carlo Methods in Equilibrium and Nonequilibrium Systems*. Vol. 74 of Springer Series in Solid-State Sciences (Springer-Verlag), 1988.

-
- [12] Richard Feynman. Simulating physics with computers. *International Journal of Theoretical Physics*, 21:467–488, 1982. 10.1007/BF02650179.
- [13] Seth Lloyd. Universal Quantum Simulators. *Science*, 273(5278):1073–1078, 1996.
- [14] F. Verstraete, M. M. Wolf, and J. I. Cirac. Quantum computation and quantum-state engineering driven by dissipation. *Nature Physics*, 5:633 – 636, 2009.
- [15] I. L. Chuang M. A. Nielsen. *Quantum computation and quantum information*. Cambridge University Press, 2000.
- [16] R. Bhatia. *Matrix Analysis*. Springer-Verlag, New York, 1997.
- [17] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 2007.
- [18] E. Seneta. *Non-negative matrices and Markov chains*. Springer, 1981.
- [19] D. W. Strook. *An Introduction to Markov Processes*. Springer, 2005.
- [20] H. Hinrichsen. Non-equilibrium critical phenomena and phase transitions into absorbing states. *Adv. in. Phys*, 49(7):815 – 958, 2000.
- [21] R. Stinchcombe. Stochastic non-equilibrium systems. *Adv. in. Phys*, 50(5):431–469, 2001.
- [22] E. Wilmer D. Levin, Y. Peres. *Markov Chains and Mixing Times*. American Mathematical Society, 2008.
- [23] B. A. Berg. *Markov Chain Monte Carlo Simulations and Their Statistical Analysis*. Singapore, World Scientific, 2004.
- [24] F. Barahona. On the computational complexity of ising spin glass models. *J. Phys. A: Math. Gen.*, 15:3241 – 3253, 1982.
- [25] M. Ligget. *Interacting Particle Systems*. Springer-Verlag, New York, 1985, 2005.
- [26] C. W. Gardiner. *Handbook of Stochastic Methods: for Physics, Chemistry and the Natural Sciences*. Springer, 2004.
- [27] Man-Duen Choi. Completely positive linear maps on complex matrices. *Linear Algebra and its Applications*, 10(3):285 – 290, 1975.
- [28] M. Wolf. Lecture notes on quantum channels (held '08/'09).
- [29] J. Schwinger. Unitary operator bases. *Proc. Natl. Acad. Sci. U S A*, 46:570 – 579, 1960.
- [30] D. E. Evans and R. Høgh Krohn. Spectral properties of positive maps on c^* -algebras. *J. London Math. Soc.*, 17:345 – 355, 1978.

- [31] M. M. Wolf, M. Sanz, D. Perez-Garcia and J. I. Cirac. A quantum version of Wielandt's inequality. *Information Theory, IEEE Transactions on*, 56(9):4668–4673, 2010.
- [32] G. Lindblad. On the generators of quantum dynamical semigroups. *Communications in Mathematical Physics*, 48:119–130, 1976. 10.1007/BF01608499.
- [33] V. Gorini, A. Kossakowski, and E. C. G. Sudarshan. Completely positive dynamical semigroups of n -level systems. *J. Math. Phys.*, 17:821, 1976.
- [34] P. Zoller, W. Gardiner. *Quantum Noise*. Springer, Berlin, 2004.
- [35] Maximilian Schlosshauer. *Decoherence and the Quantum-To-Classical Transition*. Springer Berlin Heidelberg, 2007.
- [36] H. P. Breuer and F. Petruccione. *The theory of open quantum systems*. Oxford University Press, 2002.
- [37] Sadao Nakajima. On quantum theory of transport phenomena. *Progress of Theoretical Physics*, 20(6):948–959, 1958.
- [38] Robert Zwanzig. Ensemble method in the theory of irreversibility. *The Journal of Chemical Physics*, 33(5):1338–1341, 1960.
- [39] Steven R. White. Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.*, 69(19):2863–2866, Nov 1992.
- [40] U. Schollwöck. The density-matrix renormalization group. *Rev. Mod. Phys.*, 77(1):259–315, Apr 2005.
- [41] F. Verstraete and J. I. Cirac. Matrix product states represent ground states faithfully. *Phys. Rev. B*, 73(9):094423, Mar 2006.
- [42] Guifré Vidal. Efficient classical simulation of slightly entangled quantum computations. *Phys. Rev. Lett.*, 91(14):147902, Oct 2003.
- [43] L. P. Kadanoff. Scaling laws for Ising models near t_c . *Physics*, 2:263, 1966.
- [44] F. Verstraete, J. I. Cirac, J. I. Latorre, E. Rico, and M. M. Wolf. Renormalization-group transformations on quantum states. *Phys. Rev. Lett.*, 94(14):140601, Apr 2005.
- [45] C. Schön, E. Solano, F. Verstraete, J. I. Cirac, and M. M. Wolf. Sequential generation of entangled multiqubit states. *Phys. Rev. Lett.*, 95(11):110503, Sep 2005.
- [46] Diaconis and D. Stroock. Geometric bounds on eigenvalues of Markov chains. *The Annals of Applied Probability*, 1(1):36–61, 1991.

- [47] A. Fill. Eigenvalue bounds on convergence to stationarity for nonreversible markov chains with an application to the exclusion process. *The Annals of Applied Probability*, 1(1):62–87, 1991.
- [48] M. Mihail. *Combinatorial Aspects of Expanders*. PhD thesis, Dept. Computer Science, Harvard University, 1989.
- [49] K. Pearson. On the criterion that a given system of deviations from the probable in the case of correlated system of variables is such that it can be reasonable supposed to have arisen from random sampling. *Philos. Mag.*, 50:157–172, 1900.
- [50] K. Temme, T.J. Osborne, K.G. Vollbrecht, D. Poulin, and F. Verstraete. Quantum metropolis sampling. *arXiv:0911.3635*, 2009.
- [51] E. A. Morozova and N. N. Chentsov. Markov invariant geometry on the state manifolds (russian). *Itoji Nauki i Tekhniki*, 36:69 – 102, 1990.
- [52] D. Petz. Monotone metrics on matrix spaces. *Linear Algebr. Appl.*, 244:81 – 96, 1996.
- [53] D. Petz and Cs. Sudar. Geometries of quantum states. *J. Math. Phys.*, 37:2662 – 2673, 1996.
- [54] D. Petz and M. B. Ruskai. Contraction of generalized relative entropy under stochastic mappings on matrices. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.*, 1:83 – 90, 1998.
- [55] H. Araki. Relative entropy of state of von neumann algebras. *Publ. RIMS Kyoto Univ.*, 9:809 – 833, 1976.
- [56] H. Araki. Recent progress in entropy and relative entropy. In *Proceedings of the VIIIth International Congress on Mathematical Physics*, pages 354 – 365. World Scientific, Singapore, 1987.
- [57] A. Amari and H. Nagaoka. Methods of information geometry. *Translations of Mathematical Monographs*, 191, 2000.
- [58] E. Lieb. Convex trace functions and the wigner-yanase-dyson conjecture. *Adv. Math.*, 11:267–288, 1973.
- [59] A. Lesniewski and M. B. Ruskai. Monotone riemannian metrics and relative entropy on noncommutative probability spaces. *J. Math. Phys.*, 40:5702 – 5724, 1999.
- [60] A. Jencova and M.B. Ruskai. A unified treatment of convexity of relative entropy and related trace functions, with conditions for equality. *arXiv:0903.2895*, 2009.

- [61] H. Hasegawa and D. Petz. On the riemannian metric of α -entropies of density matrices. *Lett. Math. Phys.*, 38:221 – 225, 1996.
- [62] F. Hansen. Metric adjusted skew information. *Proc. National Academy of Science of the USA*, 105:9909 – 9916, 2008.
- [63] D. Petz. Quasi-entropies for finite quantum systems'. *Reports on Math. Phys.*, 23:57 – 65, 1986.
- [64] D. Petz and C. Ghinea. Introduction to quantum fisher information. *arXiv:1008.2417*, 2010.
- [65] D. Petz. Quasi-entropies for finite quantum systems. *Rep. Math. Phys.*, 23:57 – 65, 1986.
- [66] D. Petz. Quasi-entropies for states of a von neumann algebra. *Publ. RIMS. Kyoto Univ.*, 21:781–800, 1985.
- [67] M. B. Ruskai and F. H. Stillinger. Convexity inequalities for estimating free energy and relative entropy. *J. Phys. A*, 23:2421 – 2437, 1990.
- [68] David Perez-Garcia, Michael M. Wolf, Denes Petz, and Mary Beth Ruskai. Contractivity of positive and trace-preserving maps under l_p norms. *Journal of Mathematical Physics*, 47(8):083506, 2006.
- [69] R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 2006.
- [70] M. B. Ruskai. Beyond strong subadditivity? improved bounds on the contraction of generalized relative entropy. *Rev. Math. Phys.*, 6:1147–1161, 1994.
- [71] Alberto Frigerio. Simulated annealing and quantum detailed balance. *Journal of Statistical Physics*, 58:325–354, 1990. 10.1007/BF01020296.
- [72] W. A. Majewski. The detailed balance condition in quantum statistical mechanics. *J. Math. Phys.*, 25:614, 1984.
- [73] R. F. Streater A. Majewski. Detailed balance and quantum dynamical maps. *J. Phys. A* 31, pages 7981–7995, 1998.
- [74] Robert Alicki. On the detailed balance condition for non-hamiltonian systems. *Reports on Mathematical Physics*, 10(2):249 – 258, 1976.
- [75] M B Hastings. An area law for one-dimensional quantum systems. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(08):P08024, 2007.

- [76] Daniel S. Abrams and Seth Lloyd. Quantum algorithm providing exponential speed increase for finding eigenvalues and eigenvectors. *Phys. Rev. Lett.*, 83(24):5162–5165, Dec 1999.
- [77] Alán Aspuru-Guzik, Anthony D. Dutoi, Peter J. Love, and Martin Head-Gordon. Simulated Quantum Computation of Molecular Energies. *Science*, 309(5741):1704–1707, 2005.
- [78] N. Schuch and F. Verstraete. Computational complexity of interacting electrons and fundamental limitations of density functional theory. *Nature Physics*, 5:732 – 735, 2009.
- [79] Edward Farhi, Jeffrey Goldstone, Sam Gutmann, Joshua Lapan, Andrew Lundgren, and Daniel Preda. A Quantum Adiabatic Evolution Algorithm Applied to Random Instances of an NP-Complete Problem. *Science*, 292(5516):472–475, 2001.
- [80] Barbara M. Terhal and David P. DiVincenzo. Problem of equilibration and the computation of correlation functions on a quantum computer. *Phys. Rev. A*, 61(2):022301, Jan 2000.
- [81] K. Binder. *Monte Carlo and molecular dynamics simulations in polymer science*. Oxford University Press, 1995.
- [82] Jiwen Liu and Erik Luijten. Rejection-free geometric cluster algorithm for complex fluids. *Phys. Rev. Lett.*, 92(3):035504, Jan 2004.
- [83] Robert H. Swendsen and Jian-Sheng Wang. Nonuniversal critical dynamics in monte carlo simulations. *Phys. Rev. Lett.*, 58(2):86–88, Jan 1987.
- [84] H. G. Evertz. The loop algorithm. *Adv. Phys.*, 52, 2003.
- [85] A. Y. Kitaev. Quantum measurements and the abelian stabilizer problem. *arXiv: quant-ph/9511026*, 1995.
- [86] R. Cleve, A. Ekert, C. Macchiavello, and M. Mosca. Quantum algorithms revisited. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454(1969):339–354, 1998.
- [87] Dominic Berry, Graeme Ahokas, Richard Cleve, and Barry Sanders. Efficient quantum algorithms for simulating sparse hamiltonians. *Communications in Mathematical Physics*, 270:359–371, 2007. 10.1007/s00220-006-0150-x.
- [88] W. H. Zurek W. K. Wootters. A single quantum cannot be cloned. *Nature*, 299:802 – 803, 1982.
- [89] C. Marriott and J. Watrous. Quantum arthur-merlin games. *Computational Complexity*, 14(2):122 – 152, 2005.

- [90] A. M. Childs. *Quantum information processing in continuous time*. PhD thesis, Massachusetts Institute of Technology, 2004.
- [91] Yong Zhang Daniel Nagaj, Pawel Wocjan. Fast amplification of qma. *Quantum Inf. Comput.*, 9(11):1053 – 1068, 2009.
- [92] K. Temme, M. J. Kastoryano, M. B. Ruskai, M. M. Wolf, and F. Verstraete. The χ^2 - divergence and mixing times of quantum markov processes. *arXiv: 1005.2358v2*, 2010.
- [93] A Yu Kitaev. Quantum computations: algorithms and error correction. *Russian Mathematical Surveys*, 52(6):1191, 1997.
- [94] David Poulin and Pawel Wocjan. Sampling from the thermal quantum gibbs state and evaluating partition functions with a quantum computer. *Phys. Rev. Lett.*, 103(22):220502, Nov 2009.
- [95] E. Seneta. Explicit forms for ergodicity coefficients and spectrum localization. *Linear Algebra and its Applications*, 60:187 – 197, 1984.
- [96] A Yu. Mitrophanov. Sensitivity and convergence of uniformly ergodic markov chains. *J. Appl. Probab.*, 42(4):1003–1014, 2005.
- [97] G. J. Milburn. Quantum optical fredkin gate. *Phys. Rev. Lett.*, 62(18):2124–2127, May 1989.
- [98] T. Monz, K. Kim, W. Hansel, M. Riebe, A. S. Villar, P. Schindler, M. Chwalla, M. Henrich, and R. Blatt. Realization of the quantum toffoli gate with trapped ions. *Phys. Rev. Lett.*, 102(4):040501, Jan 2009.
- [99] B. Derrida, E. Domany, and D. Mukamel. An exact solution of a one-dimensional asymmetric exclusion model with open boundaries. *Journal of Statistical Physics*, 69:667–687, 1992. 10.1007/BF01050430.
- [100] B. Derrida and M.R. Evans. Exact correlation functions in an asymmetric exclusion model with open boundaries. *J. Phys. I France*, 3(2):311–322, 1993.
- [101] B. U. Felderhof. Note on spin relaxation of the ising chain. *Reports on Mathematical Physics*, 2(2):151 – 152, 1971.
- [102] R A Blythe and M R Evans. Nonequilibrium steady states of matrix-product form: a solver’s guide. *Journal of Physics A: Mathematical and Theoretical*, 40(46):R333, 2007.
- [103] J. Cardy P. Calabrese and B. Doyon ed. Special issue: Entanglement entropy in extended quantum systems. *J. Phys. A: Math. Theor.*, 42(50), 2009.

- [104] A. Winter. Secret, public and quantum correlation cost of triples of random variables. In *Proceedings of the 2005 IEEE International Symposium on Information Theory*, pages 2270 – 2274. Adelaide, Australia (IEEE, Piscataway, NJ, 2005), 2005.
- [105] Christoph Holzhey, Finn Larsen, and Frank Wilczek. Geometric and renormalized entropy in conformal field theory. *Nuclear Physics B*, 424(3):443 – 467, 1994.
- [106] G. Vidal, J. I. Latorre, E. Rico, and A. Kitaev. Entanglement in quantum critical phenomena. *Phys. Rev. Lett.*, 90(22):227902, Jun 2003.
- [107] Pasquale Calabrese and John Cardy. Entanglement entropy and quantum field theory. *Journal of Statistical Mechanics: Theory and Experiment*, 2004(06):P06002, 2004.
- [108] J. Eisert, M. Cramer, and M. B. Plenio. Colloquium: Area laws for the entanglement entropy. *Rev. Mod. Phys.*, 82(1):277–306, Feb 2010.
- [109] F. Verstraete, M. M. Wolf, D. Perez-Garcia, and J. I. Cirac. Criticality, the area law, and the computational power of projected entangled pair states. *Phys. Rev. Lett.*, 96(22):220601, Jun 2006.
- [110] A. M. Polyakov. Conformal symmetry of critical fluctuations. *Sov. Phys. - JETP Lett.*, 12:381–383, 1970.
- [111] J. Cardy. Conformal field theory and statistical mechanics, 2008. arXiv:0807.3472.
- [112] Barbara M. Terhal, Michał Horodecki, Debbie W. Leung, and David P. DiVincenzo. The entanglement of purification. *Journal of Mathematical Physics*, 43(9):4286–4298, 2002.
- [113] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley, 1991.
- [114] A. Wyner. The common information of two dependent random variables. *Information Theory, IEEE Transactions on*, 21(2):163 – 179, March 1975.
- [115] F. C. Alcaraz, V. Rittenberg, and G. Sierra. Shared information in stationary states of stochastic processes. *Phys. Rev. E*, 80(3):030102, Sep 2009.
- [116] B Derrida, M R Evans, V Hakim, and V Pasquier. Exact solution of a 1d asymmetric exclusion model using a matrix formulation. *Journal of Physics A: Mathematical and General*, 26(7):1493, 1993.
- [117] M. Uchiyama and M. Wadati. Correlation function of asymmetric simple exclusion process with open boundaries. *JNMP*, 12:676 – 688, 2005.
- [118] B. Schmittmann. Fixed-point hamiltonian for a randomly driven diffusive system. *EPL (Europhysics Letters)*, 24(2):109, 1993.

- [119] Kevin E. Bassler and Zoltán Rácz. Bicritical point and crossover in a two-temperature, diffusive kinetic ising model. *Phys. Rev. Lett.*, 73(10):1320–1323, Sep 1994.
- [120] M B Plenio and S F Huelga. Dephasing-assisted transport: quantum networks and biomolecules. *New Journal of Physics*, 10(11):113019, 2008.
- [121] Patrick Rebentrost, Masoud Mohseni, Ivan Kassal, Seth Lloyd, and Alan Aspuru-Guzik. Environment-assisted quantum transport. *New Journal of Physics*, 11(3):033003, 2009.
- [122] R Augusiak, F M Cucchietti, F Haake, and M Lewenstein. Quantum kinetic ising models. *New Journal of Physics*, 12(2):025021, 2010.
- [123] P. Jordan and E. Wigner. Über das paulische äquivalenzverbot. *Zeitschrift für Physik A Hadrons and Nuclei*, 47:631–651, 1928. 10.1007/BF01331938.
- [124] Tomaž Prosen and Iztok Pižorn. Quantum phase transition in a far-from-equilibrium steady state of an xy spin chain. *Phys. Rev. Lett.*, 101(10):105701, Sep 2008.
- [125] M. Znidaric. Exact solution for a diffusive nonequilibrium steady state of an open quantum chain. *Journal of Statistical Mechanics: Theory and Experiment*, 2010(05):L05002, 2010.
- [126] F. Verstraete, J. J. Garcia-Ripoll, and J. I. Cirac. Matrix product density operators: Simulation of finite-temperature and dissipative systems. *Phys. Rev. Lett.*, 93(20):207204, Nov 2004.
- [127] Michael Zwolak and Guifré Vidal. Mixed-state dynamics in one-dimensional quantum lattice systems: A time-dependent superoperator renormalization algorithm. *Phys. Rev. Lett.*, 93(20):207205, Nov 2004.
- [128] B. M. Terhal R. Oliveira. The complexity of quantum spin systems on a two-dimensional square lattice. *Quant. Inf. Comp.*, 8(10):0900 – 0924, 2008.
- [129] Dorit Aharonov, Daniel Gottesman, Sandy Irani, and Julia Kempe. The power of quantum systems on a line. *Communications in Mathematical Physics*, 287:41–65, 2009. 10.1007/s00220-008-0710-3.
- [130] Daniel S. Abrams and Seth Lloyd. Simulation of many-body fermi systems on a universal quantum computer. *Phys. Rev. Lett.*, 79(13):2586–2589, Sep 1997.
- [131] Mario Szegedy. Quantum speed-up of markov chain based algorithms. *Foundations of Computer Science, Annual IEEE Symposium on*, 0:32–41, 2004.
- [132] R. D. Somma, S. Boixo, H. Barnum, and E. Knill. Quantum simulations of classical annealing processes. *Phys. Rev. Lett.*, 101(13):130504, Sep 2008.

-
- [133] G. Ortiz, J. E. Gubernatis, E. Knill, and R. Laflamme. Quantum algorithms for fermionic simulations. *Phys. Rev. A*, 64(2):022319, Jul 2001.
- [134] Rolando D. Somma, Gerardo Ortiz, Emanuel H. Knill, and James Gubernatis. Quantum simulations of physics problems. *Quantum Information and Computation*, 5105(1):96–103, 2003.
- [135] J. Hubbard. Electron Correlations in Narrow Energy Bands. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 276(1365):238–257, 1963.
- [136] Ivan Kassal and Alan Aspuru-Guzik. Quantum algorithm for molecular properties and geometry optimization. *The Journal of Chemical Physics*, 131(22):224102, 2009.
- [137] B. P. Lanyon, J. D. Whitfield, G. G. Gillett, M. E. Goggin, M. P. Almeida, I. Kassal, J. D. Biamonte, M. Mohseni, B. J. Powell, M. Barbieri, A. Aspuru-Guzik, and A. G. White. Towards quantum chemistry on a quantum computer. *Nature Chemistry*, 2:106 – 111, 2010.
- [138] Kenneth G. Wilson. Confinement of quarks. *Phys. Rev. D*, 10(8):2445–2459, Oct 1974.
- [139] John Kogut and Leonard Susskind. Hamiltonian formulation of wilson’s lattice gauge theories. *Phys. Rev. D*, 11(2):395–408, Jan 1975.
- [140] D. Horn. Finite matrix models with continuous local gauge invariance. *Physics Letters B*, 100(2):149 – 151, 1981.
- [141] R. Brower, S. Chandrasekharan, and U.-J. Wiese. Qcd as a quantum link model. *Phys. Rev. D*, 60(9):094502, Sep 1999.
- [142] S. Chandrasekharan and U. J. Wiese. Quantum link models: A discrete approach to gauge theories. *Nuclear Physics B*, 492(1-2):455 – 471, 1997.

Lebenslauf

Persönliche Daten

Name: Paul Kristan Temme
Geburtsdatum: 25. Februar 1980
Geburtsort: Bad Schwalbach, Deutschland
Nationalität: deutsch

Ausbildung und Tätigkeiten

07 / 2000 Fachhochschulreife an der Technischen Fachoberschule in München
08 / 2000 – 09 / 2001 Zivildienstleistender in der Individuellen Schwerstbehindertenbetreuung
10 / 2001 – 09 / 2003 Studium und Vordiplom in Physik an der Universität Duisburg-Essen
10 / 2003 – 11 / 2006 Studium und Diplom in Physik an der Universität Heidelberg
01 / 2007 – 06 / 2007 Forschungsaufenthalt an der Universität Leeds
10 / 2007 – 12 / 2010 Dissertation an der Universität Wien

Publikationsliste

- K. Temme and T. Gasenzer,
Nonequilibrium dynamics of condensates in a lattice with the two-particle-irreducible effective action in the $1/N$ expansion, *Phys. Rev. A.* **74**(5), 053603, (2006)
- J. K. Pachos, M. Stone and K. Temme,
Graphene with geometrically induced vorticity, *Phys. Rev. Lett.* **100**(15), 156806 (2008)
- K. Temme, T.J. Osborne, K.G. Vollbrecht, D. Poulin and F. Verstraete,
Quantum Metropolis sampling, *Nature* (accepted), *e-print arXiv: 0911.3635*, (2009)
- K. Temme, M. M. Wolf and F. Verstraete,
Stochastic exclusion processes versus coherent transport, *e-print arXiv: 0912.0858*, (2009)
- K. Temme and F. Verstraete,
Stochastic Matrix Product States, *Phys. Rev. Lett.* **104**(21), 210502 (2010)
- K. Temme, M. J. Kastoryano, M. B. Ruskai, M. M. Wolf and F. Verstraete,
The χ^2 -divergence and mixing times of quantum Markov processes,
J. Math. Phys. (accepted), *e-print arXiv: 1005.2358*, (2010)