EPFL

Thèse n° 7766

Quantum Monte Carlo approach to the non-equilibrium steady state of open quantum systems

Présentée le 11 juin 2020

à la Faculté des sciences de base
Laboratoire de physique théorique des nanosystèmes
Programme doctoral en physique

pour l'obtention du grade de Docteur ès Sciences

par

Alexandra NAGY

Acceptée sur proposition du jury

Prof. F. Mila, président du jury
Prof. V. Savona, directeur de thèse
Dr G. Carleo, rapporteur
Dr M. Holtzmann, rapporteur
Prof. M. Ceriotti, rapporteur

École polytechnique fédérale de Lausanne

2020

"The thing's hollow – it goes on forever –
and – oh my God! – it's full of stars!"
— Arthur C. Clarke, 2001: A Space Odyssey

A szüleimnek, Krisztinának és Jenőnek.

To my parents, Krisztina and Jenő.

Acknowledgements

This thesis presents a fairly comprehensive summary of what my last four years have been spent on, but it contains none of the manifold emotions that characterized it: the joy of understanding, the excitement of finding a new solution and the disappointment of failure. Worse yet, it can only partially acknowledge those who helped me on the way. My gratitude goes to many people whose presence was essential for the completion of this thesis, and I would like to take the chance and acknowledge them in this chapter that will probably be the most read (if not the only one) of this manuscript.

First and foremost, I would like to express my deepest gratitude to Prof. Vincenzo Savona, my thesis director. Vincenzo, I would like to thank you for trusting my skills and accepting me in your laboratory, and of course, for the great guidance throughout my work. During these years, your optimism and encouragement always gave me the motivation to overcome the inevitable difficulties that are part of research, while our discussions pushed my understanding and insight. I am especially thankful for all your support and apprehension during the most difficult months, and I thank you for your help in forming my next step after this thesis. For all this, and everything I learnt from you – I am most grateful.

I would like to offer my thanks to the members of my group: Riccardo, Kilian, Juan, Hugo, K.D, Eduardo and Tanya for the great atmosphere and the possibility to discuss scientific (and less scientific) matters. I would also like to acknowledge the relaxed environment at the ITP: for all the present and past members (you know who you are), I am thankful for our fun moments during coffee breaks and after-work hours.

I would like to express my particular appreciation to Gabriel, who also helped me with the French abstract, Patrick, Daniel, Maria and Hugo. You made my stay in Lausanne truly special, and I thank you all for the fun we shared on countless occasions and for the support in the ups and downs of everyday life. A heartfelt thanks goes to Greg and Luc, who introduced me to the Swiss life style, and made me feel at home in this country. I am grateful for the skiing lessons, the hiking and climbing trips (especially for never dropping me) and for all the adventures we had together. This acknowledgement could not be complete without expressing my gratitude towards Eya. I cannot possibly make a list of all the moments we shared, so I simply thank you for always having my back and say how grateful I am for your friendship.

I would also like to express my very special appreciation to Andrea, for all his love and support (and his help with the Italian abstract). You apprize simple sentiments, so I

just thank you for everything. You were a source of strength, even in the most difficult moments.

For my family, I would like thank in Hungarian first: *Óriási köszönettel tartozom családomnak, és különösen - amint a dedikàció szól - a szüleimnek. Ti mindig hittetek bennem, a korlátlan és feltétlen támogatásotok nélkül ez a disszertàció nem jött volna létre. Anya, engem mindig a Te leírhatatlan erőd inspiràlt.*

I feel enormously indebted to my family and especially – as the dedication reads – to my parents. You always believed in me, and I could not have done this without your limitless and unconditional support. Mom, your tremendous strength has always been my biggest inspiration.

Lausanne, June 4, 2020 Alex

Abstract

Many-body open quantum systems are exposed to an essentially uncontrollable environment that acts as a source of decoherence and dissipation. As the exact treatment of such models is generally unfeasible, it is favourable to formulate an approximate description by means of the dynamics of the reduced density operator of the system. When studying models with a weak coupling to a memoryless environment, the existence of such solution is granted and the reduced dynamics is governed by the Lindblad quantum master equation. In recent years, open quantum systems have evolved into a major field of studies. Focus of these studies are the characterization of emergent phenomena and dissipative phase transitions, as well as the ongoing debate about whether quantum computing schemes are still hard to simulate classically — and thus achieve quantum supremacy — when in presence of some degree of noise-induced decoherence. However, the challenge posed by this task lies in the complexity of the density matrix that increases exponentially with the system size, and the quest for efficient numerical methods is a research field that is still in its infancy. Here, we first develop a real-time full configuration interaction quantum Monte Carlo technique that stems from a class of methods generally known as projector Monte Carlo. The approach enables the stochastic sampling of the Lindblad time evolution of the density matrix thanks to a massively parallel algorithm, thus providing estimates of observables on the non-equilibrium steady state. We present the underlying theory and introduce an initiator technique and importance sampling to reduce statistical error. We demonstrate the efficiency of the approach by applying it to the dissipative two-dimensional XYZ spin-1/2 model on a lattice. As the importance sampling of projector approaches is often combined with variational results, we then introduce a novel method that is based on the variational Monte Carlo methods and on a neural network representation of the density matrix. Thanks to the stochastic reconfiguration scheme, the application of the variational principle is translated into the actual integration of the Lindblad quantum master equation. We test the effectiveness of the method by modeling the steady state of the dissipative two-dimensional XYZ spin-1/2 model through a dissipative phase transition, and also the real-time dynamics of the dissipative Ising model. In addition, we discuss the application of the developed methods and the open questions of the field.

Keywords: Lindblad master equation, Projector Monte Carlo, Variational Monte Carlo, Machine learning in physics, Neural network quantum states, Parallel optimization

Résumé

Les systèmes quantiques ouverts à N corps sont exposés à un environnement incontrôlable qui agit comme une source de décohérence et de dissipation. Le traitement exact de tels modèles étant généralement irréalisable, il est préférable de reformuler le problème à l'aide de l'approximation de l'opérateur de densité réduit. Pour les systèmes couplés faiblement à un environnement Markovien, l'existence de telles solutions est garantie et la dynamique de l'opérateur réduit est gouvernée par l'équation maîtresse de Linbald. Au cours des dernières années, les systèmes quantiques ouverts ont fait l'objet d'un nombre croissant de travaux de recherche. La majorité de ces études se concentrent sur la caractérisation des phénomènes émergents et des transitions de phases dissipatives ainsi que sur le débat récurrent consistant à savoir si la simulation classique des systèmes de calculs quantique reste difficile en présence de décohérence liée au bruit - et donc s'il est possible d'atteindre la suprématie quantique. Néanmoins, la complexité de la matrice densité, qui croît exponentiellement avec la taille du système, représente un défi majeur pour la résolution de ces problèmes et le développement de méthodes numériques efficaces est un domaine de recherche relativement jeune. Dans ce manuscrit, nous présentons d'abord une technique temps réel basée sur les méthodes d'interaction de configurations complète et de Monte-Carlo quantique qui trouve ses origines dans la classe de méthodes connues sous le nom de "projector Monte-Carlo". Cette approche permet l'échantillonnage stochastique de l'évolution en temps de Linbald de la matrice densité grâce à un algorithme massivement parallèle, rendant ainsi accessible l'estimation d'observables dans l'état stationnaire hors équilibre du système. Nous présentons la théorie sous-jacente et introduisons une procédure d'initialisation et une technique d'échantillonnage par importance pour réduire l'erreur statistique. Nous démontrons l'efficacité de notre approche en l'appliquant à la simulation d'un réseau bidimensionnel de spin 1/2 implémentant le modèle XYZ dissipatif. La méthode d'échantillonnage par importance des projecteurs étant généralement combinée à des résultats variationnels, nous introduisons ensuite une nouvelle technique basée sur les méthodes de Monte-Carlo variationnelles et une représentation de la matrice densité sous forme d'un réseau de neurones. Grâce à la procédure de reconfiguration stochastique, l'application du principe variationnel se traduit de facto par l'intégration de l'équation maîtresse de Linbald. Nous testons l'efficacité de cette méthode en modélisant l'état stationnaire du réseau bidimensionnel de spin 1/2 implémentant le modèle XYZ à travers une transition de phase dissipative et en simulant la dynamique en temps réel du modèle d'Ising disspatif. Finalement, nous discutons les possibilités d'application des mé-

thodes développées et les questions ouvertes de ce domaine de recherche.

Mots-clés : Equation maîtresse de Lindblad, Projector Monte-Carlo, Monte-Carlo variationnel, apprentissage automatique en physique, réseau de neurones, optimisation parallèle

Abstract

I sistemi aperti quantistici a N corpi sono esposti ad un ambiente incontrollabile che agisce come fonte di decoerenza e dissipazione. La trattazione esatta di tali modelli è però generalmente impraticabile, di conseguenza è preferibile riformularne una descrizione approssimata attraverso la dinamica della matrice densità ridotta del sistema. Per i sistemi accoppiati debolmente ad un ambiente Markoviano, l'esistenza di una soluzione è garantita e la dinamica dell'operatore ridotto è dettata dall'equazione maestra di Lindblad. Negli ultimi anni, la ricerca nei sistemi aperti quantistici è maturata in un fertile campo di studi. I temi di interesse dalla materia sono la caratterizzazione di fenomeni emergenti, lo studio transizioni di fase dissipative così come il dibattito in corso sul fatto che gli schemi di calcolo quantistico sono ancora difficili da simulare classicamente - e quindi raggiungere la supremazia quantistica - in presenza di un determinato grado di decoerenza indotta dal rumore. Tuttavia, la sfida insita in questo problema risiede nella crescita esponenziale della complessità della matrice di densità con la dimensione del sistema, laddove la ricerca di metodi numerici efficienti è un campo di ricerca che è ancora nella sua infanzia. In questa tesi, come primo risultato, descriviamo lo sviluppo di una tecnica di interazione di configurazione completa in tempo reale con quantum Monte Carlo fondata su una classe di metodi comunemente noti come metodi del proiettore. Questo approccio consente il campionamento stocastico dell'evoluzione temporale di Lindblad della matrice densità utilizzando un algoritmo parallelizzato, fornendo così stime di osservabili sullo stato stazionario fuori equilibrio. Successivamente descriviamo le fondamenta teoriche del metodo e introduciamo una tecnica di inizializzazione e campionamento ad importanza per ridurre l'errore statistico. Dimostriamo l'efficienza di questo approccio applicandolo al modello dissipativo bidimensionale XYZ per un reticolo di spin 1/2. Dato che il campionamento ad importanza su metodi di proiezione è tipicamente combinato con risultati variazionali, introduciamo un nuovo schema che si fonda su metodi di Monte Carlo variazionali e sulla rappresentazione della matrice densità come rete neurale. Attraverso l'utilizzo dello schema di riconfigurazione stocastica, l'applicazione del principio variazionale si può tradurre direttamente nell'effettiva integrazione dell'equazione maestra di Lindblad. Nell'ultima parte della tesi verifichiamo l'efficacia del nuovo metodo, modellando lo stato stazionario del modello bidimensionale XYZ di spin 1/2 attraverso una transizione di fase dissipativa, e la dinamica in tempo reale di un modello di Ising dissipativo. Inoltre, descriviamo in dettaglio l'implementazione dei metodi sviluppati e le questioni rimaste aperte nel campo.

Parole chiave: Equazione maestra di Lindblad, Metodo del proiettore in Quantum Monte Carlo, Variational Monte Carlo, Apprendimento automatico in fisica, reti neurali quantistiche, ottimizzazione parallelizzata

Contents

Introduction

The dawn of quantum mechanics, a fundamental part of modern physics, essentially stems from the declaration that the electromagnetic field is made of elementary units, photons [1, 2]. The foundation of quantum mechanics has been built on the revolutionary idea that the energy exchange between light and matter is discrete. In the early days, many of the theoretical and experimental schemes have focused on single particle quantum systems such as the Stern-Gerlach experiment, understanding the spectra of the hydrogen atom or explaining the photo-electric effect [3, 4]. This novel field of physics featured remarkable, and sometimes counter-intuitive properties which were illustrated in various thought experiments from Schrödinger's cat to Wigner's friend. Nowadays, these models are experimentally realized and consider several interacting particles, harnessing state superposition and entanglement in order to introduce a novel quantum logic that could ultimately achieve quantum supremacy.

However, while the evolution of any isolated system is generated by the Hamiltonian [3,4], the treatment of real systems must necessarily involve the description of an unavoidable environmental interaction. A particular class of this phenomenon, namely when the system is in a thermal equilibrium with its environment, has been possible to treat with the development of thermodynamics [5]. Nevertheless, there exist systems where the condition of thermal equilibrium is not satisfied, and a proper theoretical description must account for the relaxation and decoherence processes. We designate these systems as *open*. The study of the non-equilibrium dynamics of many body open quantum systems have evolved into a major research field due to the progress in several experimental areas including ultracold atomic gases, trapped ions, and superconducting circuits [6–9]. Focus of these studies include the characterization of emergent phenomena and dissipative phase transitions [6, 7, 9–18], as well as the ongoing debate about whether quantum computing schemes are still hard to simulate classically when in presence of some degree of noise-induced decoherence [19–22].

Let us note, the former definition of open quantum systems has been based on an implicit assumption, i.e. in any particular situation we can choose a partitioning of the collective degrees of freedom as a system and an environment. Typically, one thinks of the environment as a large number of degrees of freedom, e.g. a heath bath, that couple only to a small portion of other degrees of freedom. In practice, however, this division

is not always clear, and the most direct way of making the partition might not be adequate to accurately compute the influence of the environment. Instead, we provide a conceptual reasoning how to make this division by defining the system as the collection of those degrees of freedom that are "interesting" for us, and that we want to systematically understand. Although this partitioning does not always allow for the approximate treatment of the "uninteresting" environmental degrees of freedom, in this thesis we focus on systems which has a weak coupling to the environment. Assuming a memoryless, i.e. Markovian, interaction with the environment, the dynamics of these systems is governed by the quantum master equation in Lindblad form [23]. Since only a few model within this description admits an analytical solution [24, 25], the quest for efficient numerical methods is a research field that is still in its infancy.

In 1777, Compte de Buffon estimated the value of π by randomly dropping needles onto the floor, and thus introduced the first numerical approach to a complex mathematical problem [26]. Although this early example did not require a computer to solve the problem, it contains the main idea of what later becomes the foundation of computational physics: mapping a set of complex problems into an ensemble of easier actions using random numbers. Such type of algorithms are referred under the general name of Monte Carlo methods, and here we will focus on their application to quantum mechanical systems.

This is really the purpose of this thesis: to develop novel computational tools to simulate the dynamics and the asymptotic steady state of many body open quantum systems. Many recent tools have been developed following in the footsteps of well-established numerical methods for closed, Hamiltonian systems. In particular, matrix product state and tensor network schemes [27–29], a real-space renormalization approach [14, 30], cluster mean-field [18], and other *ad hoc* approximation schemes [16]. For Hamiltonian systems, various quantum Monte Carlo approaches have been the election tool to stochastically sample system properties, and two of its most commonly used classes are the versatile projector Monte Carlo techniques [31] and the variational Monte Carlo method [31]. While the projector approaches are stochastic implementation of the power method for determining expectation values of observables for the dominant eigenstate of the projector, variational Monte Carlo is based on the variational minimization of some system property. Using both of these approaches as a stepping stone in the development of novel numerical methods is the central topic of this thesis, which is organized in the following manner.

Chapter 1 covers the theoretical foundation of many body open quantum systems in the limit of weak environmental coupling. Thought experiments are introduced to interpret the resulting formulas from a physical point of view, and the general solutions of the Lindblad form are reviewed. Finally, we introduce a number of paradigmatic models that are often used as a first example to benchmark a novel numerical approach.

Chapter 2 reviews a number of existing numerical approaches for the simulation of open quantum systems with their advantages and drawbacks.

Chapter 3 covers the theory of the different projector Monte Carlo techniques in the context of closed systems, and introduces the general notion of the fermionic sign problem.

Chapter 4 reviews a particular projector approach, the full configuration interaction quantum Monte Carlo method which is later used as a stepping stone in the development of a numerical approach for open systems. We also present a study on the emergence of the fermionic sign problem in the context of this particular method.

Chapter 5 introduces a real-time full configuration interaction quantum Monte Carlo approach to model open quantum systems with Markovian system-environment coupling. We present the underlying theory and demonstrate the effectiveness of the method by applying it to a physical model.

Chapter 6 covers the theoretical foundation of variational Monte Carlo approach, and introduces both a particular optimization technique and the theory of Markov-chain sampling.

Chapter 7 provides a general overview of machine learning techniques and introduces some fundamental notions that will later be used to develop a numerical method for open systems.

Chapter 8 introduces a variational method to efficiently simulate the real-time dynamics and steady state of Markovian open systems based on a neural network representation of the density matrix. Similar approaches are also reviews, and the effectiveness of the method is once again tested on the simulation of some physical models.

A final remark is due. This thesis presents two novel numerical methods that have been developed during the last four years, and have already been published in various scientific journals [32, 33]. Thus, in some cases, a considerable amount of the description here is adapted from the corresponding publications. When this is the case, it is clearly stated in the introduction of the Chapter. To help the reader recognize Sections with only minor adjustments from already published manuscripts, we place an asterisk (*) at the end of their title.

The two novel numerical method presented in this thesis has been implemented in C++ and Python languages and can be found in the GitHub repositories:
https://github.com/EPFLLTPN/DDQMC [34]
https://github.com/EPFLLTPN/NNDM [35].

Open quantum systems

PART I

CHAPTER 1

Theory of open quantum systems

The theory of open quantum systems involves the description of models which are not isolated, so that their dynamics is eventually influenced by other quantum degrees of freedom (d.o.f). In general, the exact treatment of such systems is unfeasible, instead, it is favourable to formulate an approximate description by means of the dynamics of the reduced density operator of the system. When studying models with a weak coupling to a memoryless environment, the existence of such solution is granted and the reduced dynamics is governed by the Lindblad quantum master equation. While following such an approximation describes the average dynamics of the system, it is also possible to formulate a representation that is based on averaging over a large number of stochastically evolved quantum state trajectories. The dynamics of the individual realizations can be then associated to a stochastic Schrödinger equation.

To begin with, in this Chapter we first introduce the fundamental description of closed and open quantum systems with our emphasis on the limit of weak environmental coupling. After, we comment on the physical interpretation of the average dynamics of the reduced system and we review the possible solutions to the quantum master equation. Finally, we introduce a number of paradigmatic models that are often used as a first example to benchmark a novel numerical approach.

1.1 The definition of open quantum systems

While closed quantum systems are ideal quantum mechanical models which are perfectly isolated from the environment, open quantum systems are composed of a quantum mechanical set-up of two subsystems: a small system of interest (S) coupled to a large external environment (E) (see Fig. 1.1) where the combined total system S+E is assumed to be isolated. The interaction between the subsystems leads to such correlations that the resulting change in S can no longer be described by unitary dynamics, and only the combined evolution of S+E follows the dynamics of closed systems. The theory of open quantum systems presented in this chapter is introductory and has been largely inspired by the detailed descriptions of [23, 36–40].

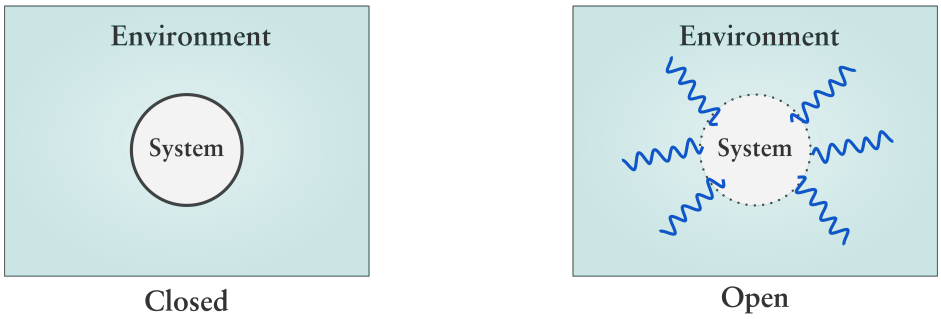

Figure 1.1 – Schematic figure of a closed versus an open quantum system.

1.1.1 The evolution of closed quantum systems

Closed quantum systems are ideal quantum mechanical models which are perfectly isolated from the environment. The postulates of quantum mechanics states that the time evolution of a state vector $|\psi\rangle$ in a Hilbert-space $\mathcal{H}$ is governed by the Schrödinger-equation

$$i\frac{d}{dt}|\psi(t)\rangle = \hat{H}(t)|\psi(t)\rangle\,, \tag{1.1}$$

where $\hat{H}$ is the Hamiltonian of the system. Here and in what follows the reduced Planck's $\hbar$ constant has been set to 1. The dynamics of the state vector can also be expressed by means of the unitary time-evolution operator $\hat{U}(t,t_0)$ which transforms the initial state $|\psi(t_0)\rangle$ at time t_0 into $|\psi(t)\rangle$ at time t

$$|\psi(t)\rangle = \hat{U}(t,t_0)|\psi(t_0)\rangle \tag{1.2}$$

with

$$\hat{U}(t,t_0) = \mathcal{T}\exp\left(-i\int_{t_0}^{t} ds\hat{H}(s)\right)\,. \tag{1.3}$$

where $\mathcal{T}$ is the time ordering. For a time-independent Hamiltonian eq. (1.3) becomes the well known expression

$$\hat{U}(t,t_0) = \exp\left(-i\hat{H}(t-t_0)\right)\,. \tag{1.4}$$

The statistical nature of quantum states can be considered using the positive semi-definite density operator $\hat{\rho}(t)$ whose time evolution derives from the Schrödinger-equation (1.1). Let us assume that at an initial time t_0 the density matrix is composed as a statistical mixture of some normalized, but not necessarily orthogonal states $\{|\psi_\alpha(t_0)\rangle\}$

$$\hat{\rho}(t_0) = \sum_\alpha p_\alpha |\psi_\alpha(t_0)\rangle\langle\psi_\alpha(t_0)| \,, \tag{1.5}$$

where $p_\alpha \in [0,1]$ and $|\psi_\alpha(t_0)\rangle$ evolve under eq. (1.1). The state of the system at some time t therefore is

$$\hat{\rho}(t) = \hat{U}(t,t_0)\hat{\rho}(t_0)\hat{U}^\dagger(t,t_0) \,, \tag{1.6}$$

and after differentiation we get the equation of motion for $\hat{\rho}$, the von-Neumann equation

$$\frac{d}{dt}\hat{\rho}(t) = -i\left[\hat{H}(t), \hat{\rho}(t)\right] . \tag{1.7}$$

1.1.2 The combined evolution of system and environment

The initial state of the total system at time $t = 0$ is described by the density matrix $\hat{\rho}_{SE}$. Let us assume that we are able to prepare $\hat{\rho}_{SE}$ as an uncorrelated product state $\hat{\rho}_{SE} = \hat{\rho}_S \otimes \hat{\rho}_E$. The total Hilbert-space is $\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_E$. The system and environment then evolve according to the unitary time evolution as

$$\hat{\rho}_{SE}(t) = \hat{U}\left(\hat{\rho}_S \otimes \hat{\rho}_E\right)\hat{U}^\dagger . \tag{1.8}$$

Since we are only interested in the evolution of system S, we need to average over the d.o.f of the environment, thus to perform a partial trace over E

$$\hat{\rho}_S(t) = \mathrm{Tr}_E\left[\hat{\rho}_{SE}(t)\right] = \mathrm{Tr}_E\left[\hat{U}\left(\hat{\rho}_S \otimes \hat{\rho}_E\right)\hat{U}^\dagger\right] . \tag{1.9}$$

Using the spectral decomposition of the density matrix $\hat{\rho}_E$ of the environment

$$\hat{\rho}_E = \sum_\alpha \eta_\alpha |\alpha\rangle\langle\alpha| \,, \tag{1.10}$$

eq. (1.9) immediately yields

$$\begin{aligned}\hat{\rho}_S(t) &= \sum_\alpha \sum_\beta \langle\alpha|\hat{U}\left(\hat{\rho}_S \otimes \eta_\beta|\beta\rangle\langle\beta|\right)\hat{U}^\dagger|\alpha\rangle \\ &= \sum_\alpha \sum_\beta \sqrt{\eta_\beta}\langle\alpha|\hat{U}|\beta\rangle\hat{\rho}_S\sqrt{\eta_\beta}\langle\beta|\hat{U}^\dagger|\alpha\rangle\end{aligned} \tag{1.11}$$

where $\hat{\rho}_S(t)$ is the reduced state of system S. As a result we obtain the operators

$$\hat{M}_{\alpha\beta} = \sqrt{\eta_\beta}\langle\alpha|\hat{U}|\beta\rangle \tag{1.12}$$

acting on $\mathcal{H}_S$ such that the matrix elements of $\hat{M}_{\alpha\beta}$ are simply $M^{i,j}_{\alpha\beta} = \langle i|\hat{M}_{\alpha\beta}|j\rangle = \langle i,\alpha|\hat{U}|j,\beta\rangle$ ($|i\rangle, |j\rangle$ are defined on $\mathcal{H}_S$). Thus the time evolution of the reduced density matrix can be represented by means of an operator sum

$$\hat{\rho}_S(t) = \mathrm{Tr}_E\left[\hat{\rho}_{SE}(t)\right] = \mathcal{M}(\hat{\rho}_S) = \sum_{\alpha,\beta}\hat{M}_{\alpha\beta}\hat{\rho}_S\hat{M}^\dagger_{\alpha\beta}\,. \tag{1.13}$$

Since the time propagator $\hat{U}$ is unitary, we have

$$\sum_{\alpha,\beta}\hat{M}^\dagger_{\alpha\beta}\hat{M}_{\alpha\beta} = \mathbb{I}_S\,. \tag{1.14}$$

The operators $\hat{M}_{\alpha\beta}$ are not unique, since the partial trace eq. (1.9) can be performed in any basis of $\mathcal{H}_E$.

1.2 The Lindblad quantum master equation

The time evolution of the reduced density matrix given in eq. (1.13) is exact, however, in practice it is not always applicable, since the complexity of the environment makes it unfeasible to follow the detailed evolution of both system and environment. A possible approach to simplify the description of the reduced dynamics is formulated by neglecting the details of the interaction between the system and the environment and only modeling the average effect of the environment on the system. Under the assumption that the system is weakly coupled to a Markovian (memoryless) environment, we obtain the evolution of the reduced density matrix in the form of the Lindblad quantum master equation. This description contains not only the unitary dynamics induced by the system Hamiltonian,

but also models dissipation processes.

1.2.1 Dynamical maps

If we assume the initial state $\hat{\rho}_{SE}$ and the time t to be fixed, then we can define a map from the Hilbert-space of system $\mathcal{H}_S$ to the space of the reduced system itself as

$$\begin{aligned} &\hat{\Phi}_t : \mathcal{H}_S \mapsto \mathcal{H}_S \\ &\qquad \hat{\rho}_S \mapsto \hat{\rho}_S(t) = \hat{\Phi}_t\left[\hat{\rho}_S\right] \equiv \mathrm{Tr}_E\left[\hat{U}\hat{\rho}_{SE}\hat{U}^\dagger\right] . \end{aligned} \tag{1.15}$$

Therefore, the map $\hat{\Phi}_t$ connects the initial density matrix to the evolved one. This action is referred to as a *dynamical map*, and can be represented by the following scheme:

$$\begin{array}{ccc} \hat{\rho}_{SE} & \xrightarrow[\hat{U}]{\text{unitary evolution}} & \hat{\rho}_{SE}(t) = \hat{U}\hat{\rho}_{SE}\hat{U}^\dagger \\ \mathrm{Tr}_E \downarrow & & \downarrow \mathrm{Tr}_E \\ \hat{\rho}_S & \xrightarrow[\hat{\Phi}_t]{\text{dynamical map}} & \hat{\rho}_S(t) = \mathrm{Tr}_E\left[\hat{U}\hat{\rho}_{SE}\hat{U}^\dagger\right] \end{array} \tag{1.16}$$

The dynamical map $\hat{\Phi}_t$ is then a *super-operator*, because it transforms an operator into another one. In order to preserve the properties of the density matrix, a dynamical map must have the following properties.

- *Linearity*: The linearity of the map is essential in order to retain the ensemble interpretation of the density matrix

$$\hat{\Phi}_t\left[\hat{\rho}_1\right] + \hat{\Phi}_t\left[\hat{\rho}_2\right] = \hat{\Phi}_t\left[\hat{\rho}_1 + \hat{\rho}_2\right] . \tag{1.17}$$

- *Continuity*: Since the propagator $\hat{U}$ is continuous, the continuity of the map is automatically preserved

$$\lim_{\tau\to 0} \hat{\Phi}_{\tau+t}\left[\hat{\rho}_S\right] - \hat{\Phi}_t\left[\hat{\rho}_S\right] = 0 . \tag{1.18}$$

- *Trace and hermiticity*: Since the diagonal terms of the density matrix describe the probabilities of finding the system in a particular state, it is fundamental that the

trace and the hermiticity is preserved

$$\begin{aligned} \mathrm{Tr}_S(\hat{\Phi}_t\left[\hat{\rho}_S\right]) &= 1 \\ \left(\hat{\Phi}_t\left[\hat{\rho}_S\right]\right)^* &= \hat{\Phi}_t\left[\hat{\rho}_S\right] . \end{aligned} \tag{1.19}$$

- *Positivity and complete positivity*: the property of positivity is such that if $\hat{\rho}_S$ is non-negative then also $\hat{\Phi}_t\left[\hat{\rho}_S\right]$ is

$$\hat{\Phi}_t\left[\hat{\rho}_S\right] \geq 0 . \tag{1.20}$$

  Although positivity is sufficient to obtain a valid density matrix, it leads to contradiction when we consider a composite system. Let us construct a proper dynamical map $\hat{\Phi}_t^A$ acting on system A. Then, considering a bipartite system, after applying the map $\hat{\Phi}_t^A \otimes \mathbb{I}_B$, we still want to obtain a valid density matrix on the total system. However, with a simply positive map, this is not always the case. We need a stronger condition on the dynamical map: $\hat{\Phi}_t$ must be completely positive such that $\hat{\Phi}_t^A \otimes \mathbb{I}_B$ is positive for any extension $\mathcal{H}_B$ of the Hilbert space $\mathcal{H}_A$.

A small technicality worth noting is that if the initial state $\hat{\rho}_S$ is factorized such that $\hat{\rho}_{SE} = \hat{\rho}_S \otimes \hat{\rho}_E$, then one obtains a linear map that is not only positive, but completely positive [41]. This feature then naturally arises in such a setting, and is indeed related to the tensor product structure of the composite system. If we allow an initial entanglement between system and environment, then the operation in eq. (1.15) in general is not a dynamical map, as it is, in general, impossible to define a linear map that can determine the subsequent evolution of S alone. The choice of separable initial state is highly compatible with the weak coupling approach, but is generally not a feasible assumption for models with strong interaction between system and environment. Despite the efforts to extend this formalism to reach beyond initially factorized states [42–44], a generally satisfactory treatment is still called for.

1.2.2 Choi-Kraus decomposition

For now we have introduced two different ways to describe the dynamics of open quantum systems. First, we considered the combined evolution of system and environment, then traced out the environment to obtain the operator-sum representation of eq. (1.13). After, we defined the properties of a dynamical map such that it transforms the initial density matrix into the evolved one. The Choi-Kraus decomposition theorem reconciles these two descriptions by stating that they are equivalent.

Theorem 1.2.1 (Choi-Kraus decomposition) *A map $\hat{\rho}_S \mapsto \hat{\Phi}_t\left[\hat{\rho}_S\right]$ from a finite dimen-*

sional Hilbert space $\mathcal{H}_S$ to a finite dimensional Hilbert space $\mathcal{H}_S$ is linear, complete positive, and trace-preserving if and only if it has a Choi-Kraus decomposition as

$$\hat{\Phi}_t\left[\hat{\rho}_S\right] = \sum_{k=1}^{K} \hat{V}_k \hat{\rho}_S \hat{V}_k^\dagger, \qquad \text{with} \qquad \sum_{k=1}^{K} \hat{V}_k^\dagger \hat{V}_k = \mathbb{I}_S\,, \tag{1.21}$$

where $K \leq N_S^2$ is the Kraus number ($\dim(\mathcal{H}_S) = N_S$).

Note the extreme representative power of the Kraus operator formalism; it is capable of compressing the infinite complexity of the environment into a set of (at most) N_S^2 operators $\hat{V}_k$. Let us remark, the form of the Kraus operators is not unique, applying any linear and unitary transformation on them leaves the dynamical map unchanged. To verify the equivalence of the two different descriptions

$$\hat{\Phi}_t\left[\hat{\rho}_S\right] = \sum_{k=1}^{K} \hat{V}_k \hat{\rho}_S \hat{V}_k^\dagger \quad \equiv \quad \mathcal{M}(\hat{\rho}_S) = \sum_{\alpha,\beta} \hat{M}_{\alpha\beta} \rho_S \hat{M}_{\alpha\beta}^\dagger\,, \tag{1.22}$$

we start by choosing a complete basis of orthonormal operators $\hat{A}_i, i = 1, \ldots, N_S^2$ in the Hilbert-Schmidt space of $\mathcal{H}_S$ such that

$$(\hat{A}_i, \hat{A}_j) = \mathrm{Tr}\left[\hat{A}_i^\dagger, \hat{A}_j\right] = \delta_{ij}\,. \tag{1.23}$$

Applying the completeness relation to each $\hat{M}_{\alpha\beta}$ defined in eq. (1.12)

$$\hat{M}_{\alpha\beta} = \sum_{i=1}^{N_S^2} \hat{A}_i (\hat{A}_i, \hat{M}_{\alpha\beta})\,, \tag{1.24}$$

eq. (1.13) transforms into

$$\mathcal{M}(\hat{\rho}_S) = \sum_{i,j=1}^{N_S^2} a_{ij} \hat{A}_i \hat{\rho}_S \hat{A}_j^\dagger \tag{1.25}$$

with

$$a_{ij} \equiv \sum_{\alpha,\beta} \left(\hat{A}_i, \hat{M}_{\alpha\beta}\right) \left(\hat{A}_j, \hat{M}_{\alpha\beta}\right)^*\,. \tag{1.26}$$

Furthermore, for any $v \in \mathbb{C}^{N_S^2}$ we have

$$\sum_{i,j=1} a_{ij} v_i^* v_j = \sum_{\alpha,\beta} \left| \left(\sum_i v_i \hat{A}_i, \hat{M}_{\alpha\beta} \right) \right|^2 \geq 0 \,, \tag{1.27}$$

which proves that $\hat{a}$ is semi-positive, hence, diagonalizable with the help of an appropriate uniform transformation $\hat{u}$

$$\hat{u}\hat{a}\hat{u}^\dagger = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & d_{N_S^2} \end{pmatrix} , \tag{1.28}$$

where d_i are non-negative. Introducing a new set of operators through

$$\hat{A}_i = \sum_{k=1}^{N_S^2} u_{ki} \hat{D}_k \,, \tag{1.29}$$

eq. (1.13) finally reads as

$$\mathcal{M}(\hat{\rho}_S) = \sum_{k=1}^{N_S^2} \left(\sqrt{d_k} \hat{D}_k \right) \hat{\rho}_S \left(\sqrt{d_k} \hat{D}_k \right)^\dagger \equiv \sum_{k=1}^{N_S^2} \hat{V}_k \hat{\rho}_S \hat{V}_k^\dagger = \hat{\Phi}_t \left[\hat{\rho}_S \right] . \tag{1.30}$$

1.2.3 The Lindblad form

In the case of closed systems, it is very convenient to characterize the dynamics with the Hamiltonian, that describes the evolution over an infinitesimal time. Now, although the Kraus representation summarizes the time evolution of the system, in practice this time-integrated equation is not very handy. Instead, we want to find a differential equation that governs the evolution of the reduced density matrix $\hat{\rho}_S$. In fact, the existence of such differential equation is not at all obvious. Indeed, such a description is only possible if the evolution of the quantum system is *Markovian*, i.e. local in time. If the dynamics is governed by a first order differential equation in time, it means that $\hat{\rho}_S(t + \delta t)$ must be completely characterized by $\hat{\rho}_S(t)$. Generally, this is not true for any open quantum system. $\hat{\rho}_S(t + \delta t)$ might depend not only on $\hat{\rho}_S(t)$, but also on $\hat{\rho}_S$ at prior times, since the environment can retain a memory of past interactions, and therefore the past can

influence the subsequent evolution of the system. In order to neglect this effect, we have to assume a clear separation between the typical correlation time of the fluctuations of the environment τ_c and the time scale of the evolution that we want to follow T_e. Solving the differential equation for $\hat{\rho}_S$ amounts to dividing T_e into small slices with duration δt and computing $\hat{\rho}_S$ at successive time steps. Mathematically, the exact evolution is then recovered in the limit when $\delta t \to 0$, however, our time slicing needs to remain coarse-grained, as the spacing of the consecutive steps must stay larger than the memory time of the environment τ_c. This can be summarized in the double inequality

$$\tau_c \ll \delta t \ll T_e \,, \tag{1.31}$$

and in what follows, by denoting $\delta t \to 0$ we refer to choosing δt in the appropriate limit of eq. (1.31).

Moreover, we need to make the hypothesis that the environment is a "large" system, i.e. its properties are not appreciably affected by its interaction with S. These two approximations together are called the *Born-Markov approximation*, and with assuming initial separability lead to the differential equation of the dynamics of open quantum systems, the *Lindblad quantum master equation*.

Let us construct the most general form of this equation starting from the evolution described by the Choi-Kraus representation (here and in the followings the S denoting the system of interest will be omitted in the expression of $\hat{\rho}_S$)

$$\hat{\rho}(t+\delta t) = \hat{\Phi}_t\left[\hat{\rho}\right] = \sum_{k=1}^{K} \hat{V}_k \hat{\rho}(t) \hat{V}_k^\dagger = \hat{\rho}(t) + \delta t \delta\hat{\rho}\,, \tag{1.32}$$

where we only kept terms up to first order in δt. We now take the limit of infinitesimal time step as $\delta t \to 0$, and without loss of generality we can organize the Kraus operators as

$$\begin{cases} \hat{V}_0 = \mathbb{I} + \delta t \hat{G} \\ \hat{V}_{k>0} = \sqrt{\delta t}\sqrt{\gamma_k}\hat{L}_k \end{cases}, \tag{1.33}$$

In order to ensure the hermiticity of $\hat{\rho}(t+\delta t)$, let us split $\hat{G}$ into $\hat{G} = \hat{J} - i\hat{H}$, where both $\hat{J} = (\hat{G} + \hat{G}^\dagger)/2$ and $\hat{H} = i(\hat{G} - \hat{G}^\dagger)/2$ are also hermitic. Then retaining once again the

terms up to linear order in δt we can write

$$\begin{aligned}\hat{\rho}(t+\delta t) &= \hat{V}_0\hat{\rho}(t)\hat{V}_0^\dagger + \sum_{k>0}\hat{V}_k\hat{\rho}(t)\hat{V}_k^\dagger \\ &= \left[\mathbb{I} + \delta t(\hat{J} - i\hat{H})\right]\hat{\rho}\left[\mathbb{I} + \delta t(\hat{J} + i\hat{H})\right] + \delta t\sum_k \gamma_k\hat{L}_k\hat{\rho}(t)\hat{L}_k^\dagger \\ &= \hat{\rho} - i\delta t\left[\hat{H},\hat{\rho}\right] + \delta t(\hat{J}\hat{\rho} + \hat{\rho}\hat{J}) + \delta t\sum_k \gamma_k\hat{L}_k\hat{\rho}(t)\hat{L}_k^\dagger\,.\end{aligned} \tag{1.34}$$

Since the operators have to respect the Kraus normalization condition, we can relate $\hat{J}$ to the other $\hat{L}_k$ as

$$\hat{J} = -\frac{1}{2}\sum_k \gamma_k\hat{L}_k\hat{L}_k^\dagger\,. \tag{1.35}$$

We can substitute $\hat{J}$ in the equation above and the take the limit $\delta t \to 0$ as $\hat{\rho}(t+\delta t) = \hat{\rho}(t) + dt\dot{\hat{\rho}}$. Finally, we obtain the *Lindblad quantum master equation* [23] in the form of

$$\begin{aligned}\frac{d\hat{\rho}}{dt} = \mathcal{L}\left[\hat{\rho}\right] &= -i\left[\hat{H},\hat{\rho}\right] + \sum_k \gamma_k\left(\hat{L}_k\hat{\rho}\hat{L}_k^\dagger - \frac{1}{2}\hat{L}_k^\dagger\hat{L}_k\hat{\rho} - \frac{1}{2}\hat{\rho}\hat{L}_k^\dagger\hat{L}_k\right) \\ &= -i\left[\hat{H},\hat{\rho}\right] + \mathcal{D}(\hat{L}_k)\left[\hat{\rho}\right]\,,\end{aligned} \tag{1.36}$$

where we have introduced the Lindblad super-operator $\mathcal{L}$ whose effect on the density matrix we will denote as $\mathcal{L}\left[\hat{\rho}\right]$ or simply $\mathcal{L}\hat{\rho}$. If the density matrix is expressed in the vectorized form $|\rho\rangle\rangle$, the Lindblad operator can also be expressed in matrix form using Kronecker products as [45]

$$\mathcal{L} = -i\left(\mathbb{I}\otimes\hat{H} + \hat{H}^T\otimes\mathbb{I}\right) - \sum_k\frac{\gamma_k}{2}\left(\mathbb{I}\otimes\hat{L}_k^\dagger\hat{L}_k + \hat{L}_k^T\hat{L}_k^*\otimes\mathbb{I} - 2\hat{L}_k^*\otimes\hat{L}_k\right)\,. \tag{1.37}$$

Note, so far we have made no assumptions about the nature of $\hat{H}$, γ_k and $\hat{L}_k$.

Coherent dynamics

In order to understand the meaning of the first term in eq. (1.36), we can use the analogy with the evolution of closed quantum systems. If we compare the Lindblad master equation (1.36) to the von-Neumann equation (1.7), one can see that $\hat{H}$ truly plays the role of the Hamiltonian of the system, and as it will be clarified in the followings $\hat{L}_k$ are called jump operators and they describe the coupling to the environment. Note, that in

general the operator $\hat{H}$ cannot be identified so easily with the Hamiltonian. Since the super-operator $\mathcal{L}$ is invariant under inhomogeneous transformations and unitary transformation of the jump operators, the form of $\hat{H}$ is not unique. To show that, in certain cases, $\hat{H}$ corresponds to the Hamiltonian of the system, one must consider the microscopic derivation of the Lindblad equation [23].

Generalized measurements and quantum jumps

After having connected the operator $\hat{H}$ to the Hamiltonian and the coherent part of the evolution, we introduce the notion of generalized measurement and we recover an analogy with the Kraus sum-representation in order to grasp the meaning of L_k.

Let us define a generalized measurement of *positive-operator valued measures (POVM)* [36] $\mathcal{M}$ by a set of non-necessarily self-adjoint operators $\hat{M}_i \in \mathcal{H}_S$ in any arbitrary number, each of them associated to an outcome m_i. When measuring the result m_i, the density matrix is projected onto the operator $\hat{M}_i$, and modified as

$$\hat{\rho}^i = \frac{\hat{M}_i \hat{\rho} \hat{M}_i^\dagger}{\mathrm{Tr}\left[\hat{M}_i \hat{\rho} \hat{M}_i^\dagger\right]} . \tag{1.38}$$

Since each of the outcomes has a probability

$$p_i = \mathrm{Tr}\left[\hat{M}_i \hat{\rho} \hat{M}_i^\dagger\right] , \tag{1.39}$$

in order to satisfy $\sum_i p_i = 1$ we must also require

$$\sum_i \hat{M}_i^\dagger \hat{M}_i = \mathbb{I} . \tag{1.40}$$

Note, when $\hat{M}_i$ are orthogonal projectors then in eq. (1.38)-(1.40) we recover the postulates of projective measurement. We want to remark, that the extension of projective measurements to POVM is necessary on both a practical and theoretical level. First of, we must realize that real life measurements are generally not projective. As a simple example, let us consider the photo-detection of a cavity field that is in the statistical mixture of $|0\rangle$ and $|1\rangle$ Fock states. This field will always end up in the vacuum state $|0\rangle$ regardless if a photon was detected or not. These measurements cannot be captured by the hermitian projectors as the corresponding measures are $\hat{M}_0 = |0\rangle\langle 0|$ and $\hat{M}_1 = |0\rangle\langle 1|$. Moreover, on a more fundamental level, generalized measurements will provide an insight into the non-unitary dynamics of open quantum systems.

In eq. (1.38), we have assumed to know the result of the measurement. However, if we

only know that the measurement took place but we are not aware of the outcome, the density matrix is in a statistical mixture of all possible outcomes as

$$\tilde{\hat{\rho}} = \sum_i p_i \hat{\rho}_i = \sum_i \hat{M}_i \hat{\rho} \hat{M}_i^\dagger \,, \tag{1.41}$$

that is, an *unread generalized measurement*.

Now we can recover the analogy with the Kraus operator-sum representation by realizing that eq. (1.21) was already cast in this form. In fact, eq. (1.21) shows that the Lindblad master equation can be interpreted as a continuous unread measurement performed on the system by identifying $\tilde{\hat{\rho}}$ with $\hat{\rho}(t+\delta t)$.

All that is left now to understand the nature of $\hat{L}_k$. For this, we devise a simple process where the coherent evolution of the system is dictated by $\hat{H}$, and the environment can be modeled with a set of $\hat{L}_k$. We continuously perform generalized measurements that are spaced with distance δt, and we apply a set of operators that have outcomes m_k associated to $\hat{M}_k$. We recall

$$\begin{cases} \hat{M}_0 = \mathbb{I} + \delta t(-i\hat{H} - \frac{1}{2}\sum_k \hat{L}_k^\dagger \hat{L}_k) \\ \hat{M}_{k\neq 0} = \sqrt{\delta t}\hat{L}_k \end{cases} . \tag{1.42}$$

We also assume that at time t the system is in a pure state $|\psi_\alpha\rangle$. This assumption is not restrictive since all states composing an arbitrary $\hat{\rho}$ can be treated separately. Then the probability of the outcome m_0 is

$$p_0 = \langle\psi_\alpha|\hat{M}_0^\dagger \hat{M}_0|\psi_\alpha\rangle = 1 - \delta t \sum_k \langle\psi_\alpha|\hat{L}_k^\dagger \hat{L}_k|\psi_\alpha\rangle \,, \tag{1.43}$$

and the state after the measurement is projected onto

$$|\psi_\alpha^0\rangle = \frac{\hat{M}_0}{\sqrt{p_0}}|\psi_\alpha\rangle = \frac{\mathbb{I} - i\hat{H}\delta t - \frac{1}{2}\sum_k \hat{L}_k^\dagger \hat{L}_k \delta t}{\sqrt{p_0}}|\psi_\alpha\rangle \,. \tag{1.44}$$

When this event occurs with probability p_0 the system undergoes an infinitesimal change that can be interpreted as a non-unitary evolution over a time step δt dictated by the effective Hamiltonian $\hat{H}_{eff} = \hat{H} - \frac{i}{2}\sum_k \hat{L}_k^\dagger \hat{L}_k$. The probability of outcome k is

$$p_k = \langle\psi_\alpha|\hat{M}_k^\dagger \hat{M}_k|\psi_\alpha\rangle = \delta t \langle\psi_\alpha|\hat{L}_k^\dagger \hat{L}_k|\psi_\alpha\rangle \,, \tag{1.45}$$

with the projected state being

$$|\psi_\alpha^k\rangle = \frac{\hat{M}_k}{\sqrt{p_k}}|\psi_\alpha\rangle = \frac{\sqrt{\delta t}\hat{L}_k}{\sqrt{p_k}}|\psi_\alpha\rangle\,. \tag{1.46}$$

If such an outcome occurs, the state undergoes a dramatic change "jumping" from state $|\psi_\alpha\rangle$ to an entirely different one $\hat{L}_k|\psi_\alpha\rangle$. Thus, L_k are called the quantum jump operators, and the typical evolution of a quantum state can be described by a slow non-unitary dissipative evolution, combined with sudden quantum jumps. In eq. (1.36), the terms $\hat{L}_k\hat{\rho}(t)\hat{L}_k^\dagger$ and $-\frac{1}{2}\left\{\hat{L}_k^\dagger\hat{L}_k,\hat{\rho}(t)\right\}$ are respectively the fluctuation and dissipation terms induced by the environment on the system. Remarkably, the evolution of the system does not follow a unitary dynamics even in the absence of a quantum jump. Rather, the dissipative evolution generated by the effective Hamiltonian is the effect of having continuously measured the system, since knowing that no jump occurred already provides us information about the state of the system.

It is important to clarify, that the choice of the jump operators L_k is not unique. The Lindblad operator can be expressed with different sets of operators, and even the number of L_k can differ in the different representations of the same dissipation process. One might try to provide a physical meaning to the simulation of a single quantum state $|\psi_\alpha\rangle$ by interpreting it as the individual evolution of an experiment in which the environment is constantly screened by generalized measurements. Nonetheless, the choice of L_k strongly affects the obtained result.

Besides an insightful interpretation of the nature of individual trajectories, eq. (1.43)-(1.46) also provide a novel tool to simulate the dynamics of open quantum systems. Eq. (1.43)-(1.46) can be cast into the so-called *counting stochastic Schrödinger equation* which describes the stochastic evolution of the system's state vectors [23,36,40]. Stochastically representing the dynamics of open systems is often referred to as *unravelling the master equation*. While one can understand the Lindblad quantum master equation as performing a continuous unread measurement on the system, by applying the stochastic Schrödinger equation one keeps track of all the measurement results. Then instead of following the time evolution dictated by the Lindbladian, the density matrix is recovered by averaging over many individual quantum trajectories. This method is the Monte Carlo Wave Function technique, and we will review it Section 2.2.

We wish to remark, that since the choice of the Kraus operators is not unique, one can obtain more than one stochastic Schrödinger equations from the Lindblad equation. These methods are not discussed in the present thesis, and for a detailed review we refer to [46,47].

1.2.4 Solutions to the master equation

To better understand the effect of dissipation, it is worthwhile to study the solutions of the Lindblad master equation (1.36) [37–39, 48, 49]. For an initial density matrix $\hat{\rho}(0)$, the formal solution of the Lindblad master equation (1.36) is of an exponential form

$$\hat{\rho}(t) = e^{\mathcal{L}t}\hat{\rho}(0)\,. \tag{1.47}$$

It has been shown [23], that the dynamical map $\hat{\Phi}_t$ fulfils the properties of contracting semi-groups, thus eq. (1.47) describes an evolution under a one-parameter semi-group generated by the Lindbladian $\mathcal{L}$. The existence and uniqueness of a non-equilibrium steady state (NESS) satisfying

$$\mathcal{L}\left[\hat{\rho}_{ss}\right] = 0 \tag{1.48}$$

has been demonstrated under quite general assumptions [38, 39], in particular for finite-size spin and boson lattices [38]. This implies that the stationary solution of eq. (1.48) corresponds to the eigenmatrix of the Lindbladian $\mathcal{L}$ with zero eigenvalue, nonetheless, the knowledge of the steady state density matrix $\hat{\rho}_{ss}$ is not sufficient to fully understand the dynamics of the system. Therefore, one must study the spectral properties of the generator $\mathcal{L}$ (Figure 1.2), whose eigenvalues and right eigenmatrices are defined as

$$\mathcal{L}\hat{\rho}_i = \lambda_i\hat{\rho}_i\,. \tag{1.49}$$

Generally, $\mathcal{L}$ has a holomorphic dependence on the continuous parameters Ω of the physical model, and since the Lindbladian is not hermitian, it is not guaranteed that it is diagonalizable for any value of the parameters. However, function theory [50] ensures that if there is a finite region of parameter space where $\mathcal{L}$ is diagonalizable, then it is diagonalizable for any Ω except for a countable number of exceptions. The eigenvalues λ_i must lie on the complex half plane $\mathrm{Re}(\lambda_i) \leq 0$, and in order to preserve the hermiticity of the eigenmatrices, the non-real eigenvalues exist in complex conjugate pairs [39]. We can recognize another relevant quantity, the spectral gap Δ, which determines the slowest non-zero rate of convergence towards the steady state. Then the real part of the eigenvalues generates the relaxation towards the steady state, and for a unique NESS, we have that

$$\hat{\rho}_{ss} = \lim_{t\to\infty}\hat{\rho}(t) = \lim_{t\to\infty} e^{\mathcal{L}t}\hat{\rho}(0)\,, \tag{1.50}$$

that is, the Lindbladian transforms any arbitrary initial $\hat{\rho}(0)$ into the non-equilibrium steady state.

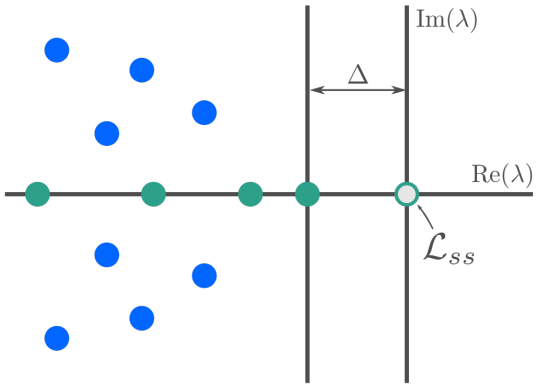

Figure 1.2 – Schematic figure of the spectrum of the Linbladian $\mathcal{L}$. The λ_i eigenvalues lie on the non-positive complex plane, and the non-real eigenvalues exist in complex conjugate pairs. The solid circles depict eigenvalues that cause a loss of proportion of the density matrix: the non-real ones as spirals, the real ones as decays. The empty circle represents the *steady state*, the eigenvalue that survives in the infinite time limit. Δ is the spectral gap, it determines the slowest non-zero rate of convergence towards the steady state.

Dissipative phase transitions

Understanding the emergence of criticality is a focal point in the research of open quantum systems. Although studying the nature of this criticality is out of the scope of the present thesis, the efficiency of each novel numerical approach is to be tested in the critical parameter region of the physical model. Therefore, we would like to give an insightful analogy between the quantum phase transitions and the emerging dissipative phase transitions of Markovian open systems, however, for a more detailed review readers are referred to [11, 37, 39, 40].

Quantum phase transitions are driven by quantum fluctuations, and they can be addressed through the competition between the non-commuting terms of the Hamiltonian. Unlike classical, or thermal, phase transitions, they occur at zero temperature, and while the thermal phase transitions induce a complete reorganization of the system itself, quantum phase transitions are characterized by an sudden change in the ground state of the system. This change is signalled by the closure of the first excitation gap at a critical value of the system parameters.

Despite the fact that the Hamiltonian and the Lindbladian are very different mathematical objects, one can draw a similarity between them in the scope of dissipative phase transitions. Dissipative phase transitions emerge due to the constant intrinsic exchange between system and environment, and there are several factors that compete in determining the NESS including non-commuting terms in the Hamiltonian and dissipation induced fluctuations. Their presence is apparent by the non-analytic behaviour of the quantum observables and the closing of the spectral gap Δ for a critical parameter value in the

thermodynamic limit of increasing system size. It has been shown [39], that while for a first order phase transition the gap only closes at the critical point, for a second order phase transition associated to a symmetry-breaking it remains closed in the whole region of broken symmetry. When the spectral gap vanishes, the lifetime of the mode with the first non-zero eigenvalue becomes infinite, and the relaxation towards the steady state is slowed down. This phenomenon is called critical slowing down and has been studied for dissipative spin [17] and bosonic [13] lattices.

1.3 Paradigmatic models

Previously we reviewed the fundamentals theory of Markovian open quantum systems, and we have also introduced a differential equation describing the dynamics of the reduced density matrix of the system, the Lindblad master equation. Within the analysis of driven-dissipative systems there are a number of paradigmatic models that are often used as a first example to benchmark a novel numerical approach. A significant amount of these models are based on spin-1/2 lattice system due to the emergence of exotic phenomena, including non-equilibrium quantum phase transitions [11, 51]. In this chapter we review three of the most prominent models that will be used throughout the thesis to compare the newly developed methods to the existing ones.

1.3.1 The dissipative XYZ model

Let us consider a spin-1/2 model on a lattice with periodic boundary condition whose Hamiltonian takes the form of a conventional Heisenberg model

$$\hat{H} = \sum_{\langle i,j \rangle} \left(J_x \hat{\sigma}_i^x \hat{\sigma}_j^x + J_y \hat{\sigma}_i^y \hat{\sigma}_j^y + J_z \hat{\sigma}_i^z \hat{\sigma}_j^z \right) , \tag{1.51}$$

where $\hat{\sigma}_j^x$, $\hat{\sigma}_j^y$, $\hat{\sigma}_j^z$ are the Pauli matrices, $\hat{\sigma}_j^{\pm} = (\hat{\sigma}_j^x \pm i\hat{\sigma}_j^y)/2$ and J_α are the coupling constants between nearest neighbour spins. This Hamiltonian is coupled to an environment that subjects each spin to a dissipation process into the state $|\sigma^z = -1\rangle$. The quantum master equation reads

$$\frac{\mathrm{d}\hat{\rho}}{\mathrm{d}t} = -i[\hat{H}, \hat{\rho}] - \frac{\gamma}{2} \sum_k \left[\left\{ \hat{\sigma}_k^+ \hat{\sigma}_k^-, \hat{\rho} \right\} - 2\hat{\sigma}_k^- \hat{\rho} \hat{\sigma}_k^+ \right] , \tag{1.52}$$

where γ is the dissipation rate. In case of an isotropic spin-spin interaction along the xy-plane, namely the XXZ model, the steady state is easily understood. For this, the

Hamiltonian can be rewritten as

$$\hat{H} = \sum_{\langle i,j \rangle} \left(2J_x(\hat{\sigma}_i^+ \hat{\sigma}_j^- + \hat{\sigma}_i^- \hat{\sigma}_j^+) + J_z \hat{\sigma}_i^z \hat{\sigma}_j^z \right) . \tag{1.53}$$

This conserves the global magnetization along the z axis , thus it does nothing to counteract the spontaneous decay. The steady state of this model is the trivial dark state, all spin polarized as $|\sigma^z = -1\rangle$.

However, for an anisotropic spin-spin coupling, the total spin is no longer conserved, a competition arises between the coherent Hamiltonian dynamics and the incoherent spin flips. The single-site Gutzwiller mean-field theory for the density matrix, which will be introduced in Section 2.1, predicts a rich phase diagram with a dissipative phase transition from a paramagnetic phase, where all the spins point along the z-axis, to a ferromagnetic one, with a finite magnetization in the xy-plane [15] (Figure 1.3). Yet, more refined calculations [14,16–18,28,52], based on numerical methods including many-body correlations showed that the phase transition does not occur for one-dimensional geometries, and the critical behaviour only emerges in two-dimension.

A viable experimental scheme were proposed in [15], and experiments were carried out using Rydberg-dressed atoms [53], Rydberg atoms [54–56], or dipolar atoms or molecules [57].

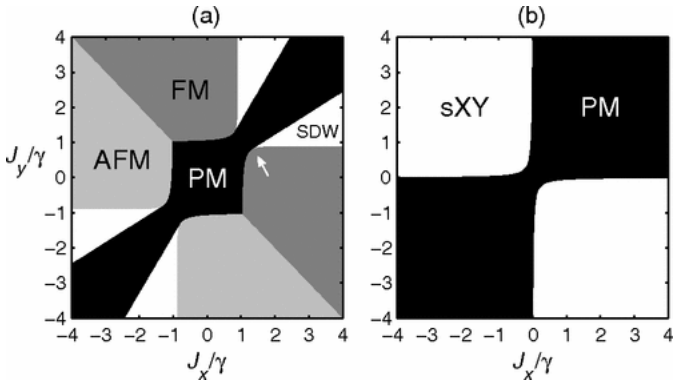

Figure 1.3 – Mean-field phase diagrams for the dissipative XYZ model with (a) $J_z/\gamma = 1$ and (b) $J_z = 0$. The different phases are paramagnetic (PM), ferromagnetic (FM), antiferromagnetic (AFM), spin-density-wave (SDW), and staggered-XY (sXY). This illustration has been presented in [15].

1.3.2 The dissipative quantum transverse Ising model

We present another theoretical model, in which we consider an array of spin-1/2 particles whose Hamiltonian is governed by

$$\hat{H} = h \sum_i \hat{\sigma}_i^x + J_z \sum_{\langle i,j \rangle} \hat{\sigma}_i^z \hat{\sigma}_j^z \,, \tag{1.54}$$

with $\hat{\sigma}_i^\alpha$ being the Pauli matrices $\{\alpha \in x, y, z\rangle\}$, h is the uniform field in transverse direction, and J_z the coupling strength between nearest neighbour spins. Once again, we consider local dissipative spin flip processes $\mathcal{D}(\hat{\sigma}_i^-)$, which fully determines the Lindblad master equation in eq. (1.36). In case there is no coupling with the environment ($\gamma = 0$), the transverse field Ising model becomes analytically solvable in one-dimension, and it features a second order phase transition from a paramagnetic phase to an antiferromagnetic one at the critical point $(h/J_z)_C = 1/2$ [58]. The phase diagram on a square lattice is qualitatively similar to the one-dimensional case, with again a paramagnetic-antiferromagnetic phase transition. The position of the critical point is known with high precision from Monte Carlo simulations $(h/J_z)_C = 1.52219(1)$ [59].

In contrast to the non-dissipative case, the phase diagram of the dissipative Ising model is still a matter of debate in the literature. Nonetheless, recent studies [60, 61] has focused on numerical models that are capable of including non-local correlations in order to determine the phase diagram in one- and two-dimensions, and successfully analysed the quantitative and qualitative changes emerging from the competition between the coherent dynamics and the dissipation. Due to the fast development of quantum simulators in recent years, this model can now be experimentally realized using Rydberg-atoms [62].

1.3.3 The dissipative synthetic Ising model

We now introduce a model that has been the focus of recent experimental studies due to an accurate mapping of multilevel Rydberg atoms on interacting spin-1/2 particles [62–66]. In the experimental setting the system is governed by the Hamiltonian

$$\hat{H} = \sum_i \left(\frac{\Omega}{2} \hat{\sigma}_i^x - \delta \hat{n}_i \right) + \frac{1}{2} \sum_{\langle i,j \rangle} U \hat{n}_i \hat{n}_j \,, \tag{1.55}$$

where the atoms are coherently coupled to the Rydberg state with a two-photon transition of Rabi frequency Ω and a detuning δ, the interaction term U arises from the interactions between the atoms, and $\hat{n}_i = (\mathbb{I} + \hat{\sigma}_i^z)/2$ is the projector on the Rydberg state for atom i. For $(\delta/U)_{TFI} = 1$ this Hamiltonian corresponds to the previously discussed transverse Ising model.

The ground state phase diagram of the model is well understood and demonstrated in

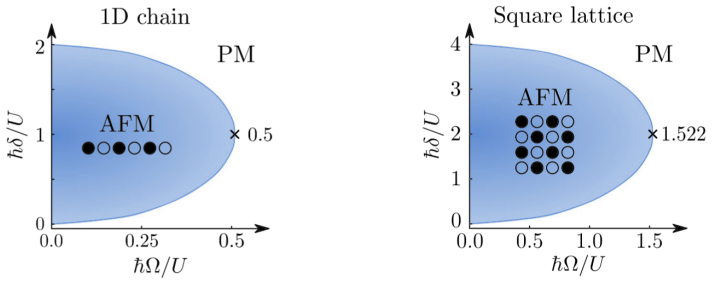

Figure 1.4 – Sketched ground state phase diagram for the synthetic Ising model in eq. (1.55) for a one-dimensional chain and a square lattice. This illustration has been presented in [62].

Figure 1.4 for both one- and two-dimensional lattices. The one dimensional array features a second order phase transition from an antiferromagnetic region to a paramagnetic one, delimited by a transition line of the $(1+1)$d universality class [58]. The phase diagram for a square lattice is qualitatively similar to the lower dimensional case with the phase boundary being in the $(2+1)$d Ising universality class [58].

Since the experimental setup has several sources of imperfection [67], describing the time evolution of the system as purely Hamiltonian is not exact. In particular, the phase noise of the excitation lasers leads to dephasing already for a single atom. This decoherence can be taken into account with a local dephasing in the scope of Markovian open quantum systems as

$$\frac{\mathrm{d}\hat{\rho}}{\mathrm{d}t} = -i[\hat{H},\hat{\rho}] - \frac{\gamma}{2}\sum_k \left[\{\hat{n}_k,\hat{\rho}\} - 2\hat{n}_k\hat{\rho}\hat{n}_k\right] . \tag{1.56}$$

This phenomenological dephasing term has been tested and found to be in great agreement with the experimental results in [62]. Note, the steady state of this model is analytically known and can be deduced as follows. The purely dephasing decoherence of eq. (1.56) yields a hermitic Lindbladian. Since the left eigenmatrix of $\mathcal{L}$ is always the identity, and the left and right eigenmatrices of hermitic super-operators must coincide, therefore eq. (1.56) describes a dynamics that converges to a fully mixed steady state. As the stationary solution is well-known, in this thesis we only study the real-time evolution of the model.

CHAPTER 2

Numerical approaches to open quantum systems

In Chapter 1 we have discussed the fundamentals of the theory of open quantum systems and, assuming Markovian coupling to the environment, we have introduced the Lindblad master equation. Only a few models within this description admit an analytical solution [24, 25], and the quest for efficient numerical methods is a research field that is still in its infancy. Although the simulation of open quantum systems is computationally much more demanding than for closed systems, many recent tools have been developed following in the footsteps of well-established numerical methods for Hamiltonian quantum systems. In general, one has two possibilities to introduce an approximate description; one can either solve the problem on a smartly truncated Hilbert space, or find an effective and efficient representation of the wave function or the density matrix. Generally, there are two different areas that are of interest when studying open quantum systems: to simulate the dynamical evolution of the system or to find the asymptotic steady state. In some cases, even when the main focus is on the properties of the non-equilibrium steady state, it is more efficient to calculate the full time evolution of the system. In this chapter we will address this distinction when discussing the most common numerical approaches.

First, we will look at mean-field methods where the master equation is decoupled into a set of effective single site equations. We then review a stochastic method called the Monte Carlo Wave Function (MCWF) in which the system-environment interaction is unravelled onto a stochastic process that adds to the unitary Hamiltonian evolution and the effective damping terms. Finally, we cover tensor network based methods and a real-space renormalization approach.

2.1 Mean-field approximations

Within the mean-field approximation to open quantum systems [68], it is possible to develop a generalized Gutzwiller approach by introducing a product ansatz for the density

matrix

$$\hat{\rho} = \bigotimes_i \hat{\rho}_i \tag{2.1}$$

with the reduced local density operators $\hat{\rho}_i = \mathrm{Tr}_{\neq i}\hat{\rho}$. By tracing out all sites except one, eq. (1.36) yields a set of single-site master equations

$$\frac{d\hat{\rho}_i}{dt} = \mathrm{Tr}_{\neq i}\left[\frac{d\hat{\rho}}{dt}\right] = -i\left[\hat{H}_i^{MF}, \hat{\rho}_i\right] + \mathcal{D}_i\left[\hat{\rho}\right] , \tag{2.2}$$

where $\hat{H}_i^{MF}$ are the mean-field Hamiltonian and $\mathcal{D}_i$ are the mean-field dissipators. For systems with translation invariance, it is sufficient to solve a lone single-site problem, otherwise eq. (2.2) is solved self-consistently.

There is a wide range of models where mean-field approximations predict the presence of bistability, including the dissipative Ising model [66, 69, 70], as well as the one-photon driven dissipative Bose-Hubbard model [71, 72]

$$\begin{gathered}\hat{H} = \sum_i \left(-\Delta\hat{a}_i^\dagger\hat{a}_i + \frac{U}{2}\hat{a}_i^\dagger\hat{a}_i^\dagger\hat{a}_i\hat{a}_i + F\hat{a}_i^\dagger + F^*\hat{a}_i\right) + J\sum_{\langle i,j\rangle}\hat{a}_i^\dagger\hat{a}_j \\ \mathcal{L}\left[\hat{\rho}\right] = -i\left[\hat{H}, \hat{\rho}\right] + \mathcal{D}(\hat{a})\left[\hat{\rho}\right] ,\end{gathered} \tag{2.3}$$

where $\hat{a}_i$ are the bosonic annihilation operators, J is the hopping strength, U is the strength of the non-linearity and F is the amplitude of the one-photon driving. This bistability is a product of the approximation, and is replaced by a first-order phase transition for both models. For the dissipative Ising model, tensor network based simulations (Sec. 2.3) has showed that while low bond dimension predicts a bistable region, it is quickly replaced with a first-order phase transition as the bond dimension is systematically increased [28]. As for the driven dissipative Bose-Hubbard model, a theoretical analysis based on Keldysh formalism confirmed the presence of a first order transition [73].

Mean-field approximation has been used to predict second-order phase transitions for the dissipative XYZ model [15], and for the dissipative Bose-Hubbard model with two-photon pumping [12]. For the dissipative XYZ model studies have showed that the phase transition only survives in two-dimensional models, and disappears for a one-dimensional lattice [14, 16–18, 28, 52]. In case of the dissipative Bose-Hubbard model, it has been shown that two-photon driving preserves a $\mathbb{Z}_2$ symmetry of the Hamiltonian which can be then spontaneously broken [39]. This gives rise to an Ising-like second order phase transition [74, 75].

As the correlations between the different lattice sites are neglected, the single-site Gutzwiller approach can only be a crude approximation of open system dynamics. The cluster

mean-field approach tries to overcome this limitation by isolating a given subset of $\mathcal{C}$ contiguous lattice sites, instead of decoupling the sites one-by-one [14]

$$\hat{\rho} = \bigotimes_i \hat{\rho}_i \longrightarrow \hat{\rho} = \bigotimes_{\mathcal{C}} \hat{\rho}_{\mathcal{C}} \,. \tag{2.4}$$

Once again, the resulting cluster mean-field equations are solved self-consistently, but within a subset $\mathcal{C}$ the short term correlations are taken into account exactly. Therefore, the accuracy of the method can systematically improved by considering clusters of increasing size. While for smaller subsets it is feasible to directly integrate the quantum master equation, for larger sizes one needs to resort to approximate solutions. The cluster mean-field method was successfully used to calculate the phase diagram of the dissipative XYZ model [14], as well as the dissipative Ising model [60].

2.2 The Monte Carlo Wave Function method

While following the Lindblad master equation describes the average dynamics of the system, it is also possible to formulate a representation that is based on averaging over a large number of stochastically evolved quantum state trajectories. At any initial time t_0 the density matrix can be decomposed as a statistical mixture $\hat{\rho}(t_0) = \sum_\alpha \eta_\alpha |\psi_\alpha(t_0)\rangle\langle\psi_\alpha(t_0)|$, and after propagating the individual pure states $|\psi_\alpha(t_0)\rangle$ to time t one can compute the observables as

$$\langle \hat{O} \rangle = \mathrm{Tr}(\hat{\rho}\hat{O}) = \sum_\alpha \eta_\alpha \langle \psi_\alpha(t)|\hat{O}|\psi_\alpha(t)\rangle \,. \tag{2.5}$$

One can either sample the probability distribution η_α with standard Monte Carlo techniques, or start from an initial pure state $|\psi_\alpha(t_0)\rangle$ and average over a large number of trajectories. Since the individual realizations are independent, this task is performed on a parallel manner resulting in reduced computational complexity. However, one needs to make sure that the individual states are propagated on such a way that the ensemble of trajectories still satisfies $\hat{\rho}_t = \sum_\alpha \eta_\alpha |\psi_\alpha(t)\rangle\langle\psi_\alpha(t)|$.

As we have discussed in Chapter 1, the Monte Carlo Wave Function technique provides a possible unravelling of the Lindblad master equation, where the individual trajectories can be seen as a photon counting process where we continuously measure the state of the system [76–79]. A recent study [80] has extensively covered the details of the approach, here we only review the fundamental principles. Recall, we can rewrite the Lindblad equation (1.36) in a more convenient form by defining the effective Hamiltonian $\hat{H}_{\mathrm{eff}}$ and the

jump super-operator $\mathcal{L}_j$

$$\begin{aligned} \hat{H}_{\text{eff}} &= \hat{H} - \frac{i}{2}\sum_k \gamma_k \hat{L}_k^\dagger \hat{L}_k \\ \mathcal{L}_j\left[\hat{\rho}\right] &= \sum_k \hat{L}_k \hat{\rho} \hat{L}_k^\dagger \\ \dot{\hat{\rho}}(t) &= -i\left[\hat{H}_{\text{eff}}, \hat{\rho}\right] + \mathcal{L}_j\left[\hat{\rho}\right] . \end{aligned} \tag{2.6}$$

The typical evolution of a quantum state then can be described by a slow non-unitary dissipative evolution, combined with sudden quantum jumps. Remarkably, the evolution of the system does not follow a unitary dynamics even in the absence of a quantum jump, since the knowledge that a jump did not occur already provides us information about the system. This split in the dynamics is then reflected in the quantum jump method such that at each time step δt we perform the following algorithm to simulate the evolution of a quantum state $|\psi_\alpha(t)\rangle$:

1. From the initial state $|\psi_\alpha(t)\rangle$ compute the probability of no-jump p_0 (eq. (1.43)), and the probabilities p_k associated to the jump operators $\hat{L}_k$ (eq. (1.45)).

2. Determine whether a jump will occur or not. If it does occur, choose randomly a jump operator L_k. In practice, both of these events are determined by drawing a random number $r \in [0,1]$. If $r < p_0$ then the system follows the non-unitary evolution dictated by $\hat{H}_{eff}$, otherwise a jump takes place and its index κ is smallest integer such that $\sum_{k=0}^{\kappa} p_k > r$.

3. If no jump occurred, compute the infinitesimal evolution generated by $\hat{H}_{eff}$, and renormalize the new state with p_0 to obtain the state $|\psi_\alpha(t+\delta t)\rangle$ (eq. (1.45)).

4. If a jump did occur, calculate the new state $|\psi_\alpha(t+\delta t)\rangle$ and renormalize it with p_κ (eq. (1.46)).

This procedure is repeated until the desired time $t_{\max}$ has been reached. The algorithm is schematically illustrated in Figure 2.1. An alternative and computationally more efficient approach is based on the so called waiting time distribution. In this case, instead of making a decision at each δt time step, the state $|\psi_\alpha\rangle$ is propagated continuously under the effective Hamiltonian, until its norm drops below a previously drawn random number $r \in [0,1]$, when a jump event is performed.

Then, a large number of individual trajectories is constructed this way, all starting from the same initial state $|\psi_\alpha(0)\rangle$. The density matrix is then recovered as the average of the

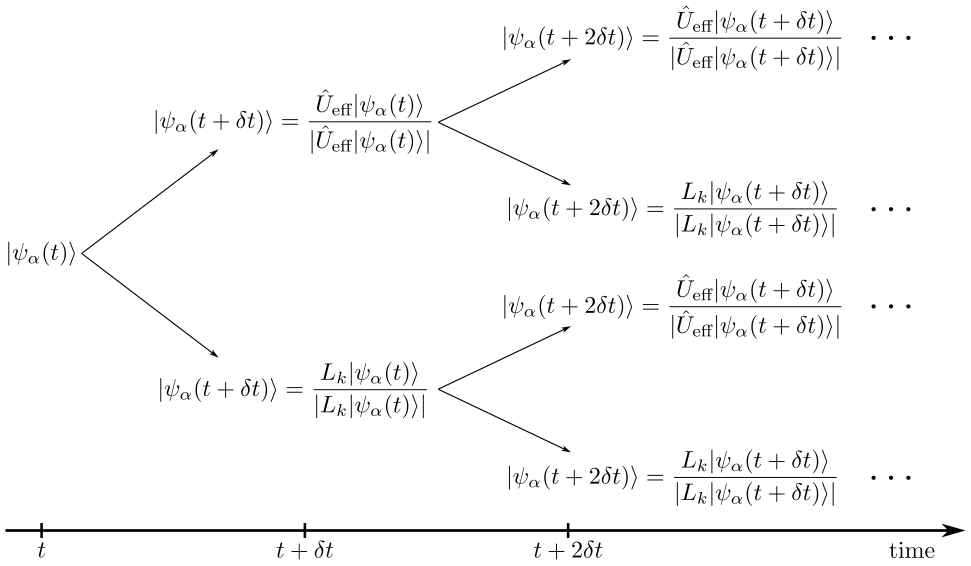

Figure 2.1 – The figure illustrates the stochastic time evolution of a single trajectory in a Monte Carlo Wave function simulation [81].

projectors

$$\hat{\rho}(t_{\rm max}) \equiv \bar{\Pi}(t_{\rm max}) = \frac{1}{N_{\rm traj}} \sum_\alpha |\psi_\alpha(t_{\rm max})\rangle\langle\psi_\alpha(t_{\rm max})| \,, \tag{2.7}$$

where $N_{\rm traj}$ is the number of individual trajectories. If the initial density matrix is a statistical mixture of quantum states, then an additional average should be performed. We can verify the consistency of eq. (2.7) by first writing $\bar{\Pi}(t+\delta t)$ after a time step δt as a mixture of the no-jump and jump evolution with weights p_k. Then by replacing the wave functions with their explicit expression (eq. (1.43)-(1.46)) and setting $d\bar{\Pi}/dt = (\bar{\Pi}(t)+\delta t) - \bar{\Pi}(t))/\delta t$ we get

$$\frac{d\bar{\Pi}}{dt} = -i\left[\hat{H}, \bar{\Pi}\right] + \sum_k \gamma_k \left(\hat{L}_k \bar{\Pi} \hat{L}_k^\dagger - \frac{1}{2}\hat{L}_k^\dagger \hat{L}_k \bar{\Pi} - \frac{1}{2}\bar{\Pi}\hat{L}_k^\dagger \hat{L}_k\right) , \tag{2.8}$$

which is the Lindblad master equation [36, 80]. An alternative and computationally more efficient approach is based on the so called waiting time distribution. In this case, instead of making a decision at each δt time step, the state $|\psi_\alpha\rangle$ is propagated continuously under the effective Hamiltonian, until its norm drops below a previously drawn random number $r \in [0, 1]$. While for most system (and Hilbert space $\mathcal{H}$) sizes quantum trajectories can be numerically computed for the whole Hilbert space, for larger sizes it has to be

combined with various approaches to reduce $\mathcal{H}$ to a physically relevant region, e.g. matrix product states (Sec. 2.3). Another possibility that allows further dimensionality reduction is provided by the Gutzwiller Monte Carlo approach [10,18,82] which employs a special ansatz for $|\psi_\alpha\rangle$ in the form of a product state over different lattice sites

$$|\psi_\alpha^G\rangle = \bigotimes_i |\phi_i\rangle\,, \tag{2.9}$$

where i denotes the different modes or lattice sites and $|\phi_i\rangle$ is a single-mode wave function on site i. Hence, while the degrees of freedom of the total wave function grow exponentially with the system size N as d^N with d being the dimension of the local Hilbert space, the dimensionality of $|\psi_\alpha^G\rangle$ grows only linearly with Nd. This method has similar computational cost as the previously introduced mean-field approximation (Sec. 2.1), however, sampling a set of trajectories in the form $|\psi_\alpha^G\rangle$ can contain more information than eq. (2.1). While neither ansatz is capable of capturing quantum entanglement between different lattice sites, an ensemble of trajectories can sustain non-trivial classical correlations whereas a product state of local density operators can not [80].

In this thesis, the MCWF calculations have been performed with the help of the QuTiP library [83,84].

2.3 Matrix product operator approach

The notion of the "physical corner of Hilbert space", namely that the physically relevant states constitute only an exceedingly small fraction of the exponentially large Hilbert space, was first introduced in the context of many-body quantum systems. For many of these models, the entanglement entropy of the ground state grows like the boundary of a system and not like its volume [85–87]. Therefore, the pure states exhibiting low entanglement in the sense of satisfying this area law constitute a very small subset of all pure states, the "corner of Hilbert space".

The motivation behind the introduction of tensor network states was to design such a variational representation $|\psi_{\text{MPS}}\rangle$ for the wave function $|\psi\rangle$ that satisfies the area law, and also capable of representing the quantum correlations of the physically relevant states efficiently and in a controllable manner. This method can be seen as the minimization of the variational energy $\langle\psi_{\text{MPS}}|\hat{H}|\psi_{\text{MPS}}\rangle$, and turned out to be a generalization of the previously well-known and widely used density matrix renormalization group (DMRG) [88,89] algorithm.

These are the so called Matrix Product States (MPS) [90–92]

$$|\psi_{\text{MPS}}\rangle = \sum_{s_1,\ldots,s_N=1}^{d-1} c_{s_1,s_2,\ldots,s_N} |s_1 s_2 \ldots s_N\rangle\,, \tag{2.10}$$

where d is the dimension of the Hilbert space and the coefficients $c_{s_1,s_2,\ldots,s_N}$ can be written as products of tensors

$$c_{s_1,s_2,\ldots,s_N} = A^{[s_1]}A^{[s_2]}\cdots A^{[s_N]}\,, \tag{2.11}$$

with $A^{[s_i]}$ being rectangular matrices that are computed with singular value decomposition (SVD). The dimensionality reduction is done by truncating the result of the SVD, and keeping only the D most significant values, with D referred to as the bond dimension. It can be shown [91], that this quantity directly controls the amount of entanglement present in the representation. The properties of the MPS ansatz are well understood, and the contraction of tensors can be done exactly and efficiently [93]. MPS based methods produced extremely accurate results for ground state calculations of closed quantum systems [94].

It is possible to extend the MPS formalism to operators, giving rise to the MPO methods, however, there are several intrinsic bottlenecks limiting their application such as positivity, hermiticity and dimensionality. An MPO representation of the density matrix for a system with N sites were formulated in [91] as

$$\hat{\rho}_{\text{MPO}} = \sum_{s_1,s_1'\ldots,s_N,s_N'=1}^{d} \left(B^{[s_1,s_1']}\cdots B^{[s_N,s_N']}\right)|s_1 s_2\ldots s_N\rangle\langle s_1' s_2'\ldots s_N'|\,, \tag{2.12}$$

where $B^{[s_k,s_k']}$ are $D_k \times D_{k+1}$ dimensional matrices that can be decomposed as

$$B^{[s_k,s_k']} = \sum_{a=1} A^{[s_k,a]}\otimes\left(A^{[s_k',a]}\right)^*\,. \tag{2.13}$$

Such a construction ensures the positivity of the density matrix, however, the required bond dimension makes it inefficient to calculate [95]. An alternative representation has been proposed in [96], where the density matrix is recast into a vector form and can be represented with an MPS as

$$|\hat{\rho}_{\text{MPS}}\rangle\rangle = \sum_{s_1,\ldots,s_N=1}^{d^2-1} c_{s_1,s_2,\ldots,s_N}|s_1\rangle\rangle\otimes|s_2\rangle\rangle\otimes\cdots\otimes|s_N\rangle\rangle\,, \tag{2.14}$$

where $|s_i\rangle\rangle$ are operators recast in vector representation forming a local basis of dimension d^2. The vectorized density matrix $|\hat{\rho}_{\text{MPS}}\rangle\rangle$ is then evolved using Time Evolving Block Decimation (TEBD) [96], assuming that the Lindblad superoperator can be decomposed as a sum of terms involving at most nearest neighbours. Although this method has proven to be efficient and simple, positivity was not guaranteed.

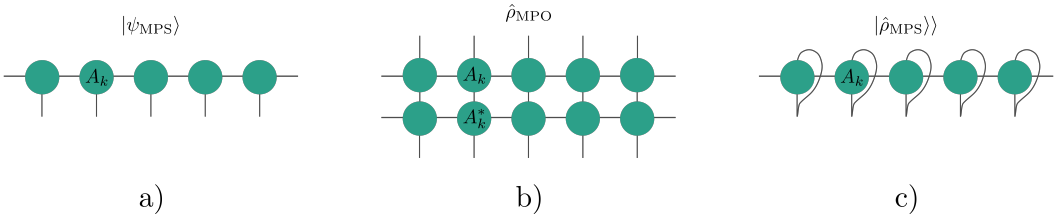

Figure 2.2 – Illustration of matrix product states. a) Writing the wave function in an MPS form. b) Diagram of the density matrix as an MPO. c) Diagram of the vectorized density matrix.

These dynamical MPO/MPS techniques employ the long-time dynamics to obtain the NESS in several different settings [97–102], however, simulating the real-time dynamics has its disadvantages. While it captures the transient behaviour of the system, it may require large bond dimensions for the intermediate states even though the final steady state is well-represented with a small bond MPO [97, 99]. Lately, this method has been improved by two independent works [27,29], where instead of the real time evolution they introduce a variational minimization for $\|\mathcal{L}|\hat{\rho}_{\mathrm{MPS}}\rangle\rangle\|$ and $\left\|\mathcal{L}^\dagger\mathcal{L}|\hat{\rho}_{\mathrm{MPS}}\rangle\rangle\right\|$ respectively.

The generalization of MPS states for higher dimensions, also known as Projected Entangled Pair states (PEPS) [103] has unfortunately many limitations, including the high programming effort and the non-trivial PEPS contractions. Nonetheless, in recent years a significant amount of studies have been focusing on developing it into a state-of-the-art tool [104–106], even though their success is mostly limited for ground state calculation of closed systems.

2.4 Corner-space renormalization

Another approach that is based on trying to find the most relevant subspace in an exponentially growing Hilbert space is the Corner-Space Renormalization (CSR) method [30]. For this, let us consider two lattices, A and B, whose density matrices are $\hat{\rho}_A$ and $\hat{\rho}_B$. We are interested in the properties of the total system that results from spatially merging A and B. The assumption of CSR is, that one can construct the "relevant" corner of the Hilbert space $\hat{H}_A \otimes \mathcal{H}_B$ by using the most probable eigenvectors of the sub-lattices. To find this corner one first diagonalizes A and B as

$$\begin{aligned}\hat{\rho}_A &= \sum_a \eta_a |\phi_a^A\rangle\langle\phi_a^A| \\ \hat{\rho}_B &= \sum_b \eta_b |\phi_b^B\rangle\langle\phi_b^B|\,,\end{aligned} \tag{2.15}$$

where $|\phi_a^A\rangle$ and $|\phi_b^B\rangle$ form an orthonormal basis for lattice A and B respectively, and η_a, η_b are the probabilities. The two lattices then spatially merged such that $\mathcal{H} = \hat{H}_A \otimes \mathcal{H}_B$, and the basis for this space is then constructed from the P most probable product state in the form $|\phi_a^A\rangle|\phi_a^B\rangle$, ranked by the joint probability $\eta_a\eta_b$. This effectively means considering the subspace spanned by

$$\mathcal{C}(P) = \{|\phi_{a_1}^A\rangle|\phi_{b_1}^A\rangle, |\phi_{a_2}^A\rangle|\phi_{b_2}^A\rangle, \ldots, |\phi_{a_P}^A\rangle|\phi_{b_P}^A\rangle\} \tag{2.16}$$

where $\eta_{a_i}\eta_{b_i} \geq \eta_{a_{i+1}}\eta_{b_{i+1}}$. Finally, the steady state of $\hat{\rho} \in \mathcal{H}$ is computed on the truncated basis. The value of the cut-off P is successively increased until convergence of the observables is achieved. Note, this algorithm becomes exact in the limit of $P = \dim(\mathcal{H})$. The CSR method is graphically illustrated in Figure 2.3.

The CSR was first introduced for the driven dissipative Bose-Hubbard model reaching lattice sizes up to 4×4 and 6×3 [30], then successfully confirmed the existence of critical behaviour in the two-photon driven dissipative Bose-Hubbard model [74]. It has also been used to study the critical properties of two-dimensional spin lattices interacting via the dissipative XYZ model [16]. The method provides a reasonably fast numerical algorithm, nonetheless, its limitations present themselves for spatially inhomogeneous systems and for highly mixed systems, where it has convergence issues even for large truncation values.

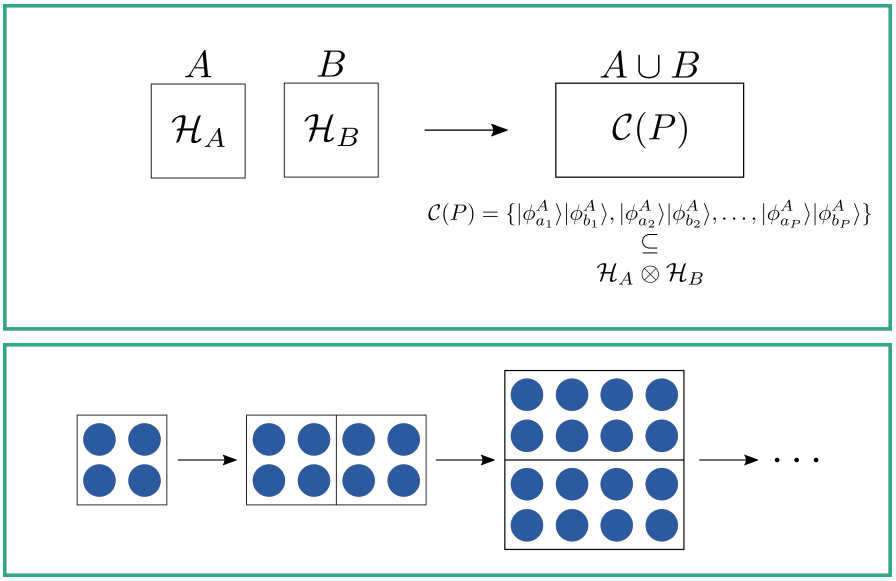

Figure 2.3 – The figure illustrates the corner-space renormalization algorithm.

Driven-dissipative quantum Monte Carlo

PART II

CHAPTER 3

Projector Monte Carlo methods

In this Part, we introduce a real-time quantum Monte Carlo approach for open quantum systems which has been developed in the footsteps of a well-established Hamiltonian method, the Full Configuration Interaction Quantum Monte Carlo (FCIQMC). FCIQMC belongs to a wider class of methods, called Projector Monte Carlo (PMC), and we start by familiarizing ourselves with the formalism and taxonomy of these approaches in the scope of closed systems.

3.1 Formalism

PMC methods are tailored to calculate the ground state properties of Hamiltonian systems, both at zero and finite temperature [31, 107, 108], and they are based on the same theoretical foundation, a stochastic implementation of the imaginary time Schrödinger equation. We start by considering the time-dependent Schrödinger equation of the Hamiltonian $\hat{H}$

$$i\frac{\partial|\psi\rangle}{\partial t} = \hat{H}|\psi\rangle\,, \tag{3.1}$$

whose formal solution reads as

$$|\psi_t\rangle = e^{-i\hat{H}t}|\psi_0\rangle\,. \tag{3.2}$$

By transforming from real-time to imaginary-time with the Wick rotation $\tau = it$, we get the imaginary-time Schrödinger equation and its solution as

$$\frac{\partial|\psi\rangle}{\partial\tau} = -\left(\hat{H} - S\mathbb{I}\right)|\psi\rangle \tag{3.3}$$

$$|\psi_\tau\rangle = e^{-(\hat{H}-S\mathbb{I})\tau}|\psi_0\rangle\,, \tag{3.4}$$

where we also introduced the methodologically crucial energy shift S. Using the spectral decomposition of the evolution operator, the infinite propagation time limit reads as

$$\lim_{\tau\to\infty}|\psi_\tau\rangle = \lim_{\tau\to\infty}\sum_i e^{-(E_i-S)\tau}|\Phi_i\rangle\langle\Phi_i|\psi_0\rangle\,, \tag{3.5}$$

with eigenvalues E_i and the eigenstates $|\Phi_i\rangle$ of the Hamiltonian $\hat{H}$. Eq. (3.5) then lets us infer the asymptotic behaviour depending on the relation between the ground state energy E_0 and the value of the energy shift S,

i if $S < E_0$, $\lim_{\tau\to\infty}|\psi_\tau\rangle = 0$, the wave function decays at an exponential rate,

ii if $S > E_0$, $\lim_{\tau\to\infty}|\psi_\tau\rangle = \infty$, the wave function grows exponentially fast,

iii if $S = E_0$, $\lim_{\tau\to\infty}|\psi_\tau\rangle \propto |\Phi_0\rangle$, the wave function converges to the ground state assuming there is a non-zero overlap between $|\psi_0\rangle$ and $|\Phi_0\rangle$.

Identifying $e^{-(\hat{H}-S\mathbb{I})\tau} = \hat{P}(\tau)$ as the projector, this behaviour lays the foundation for the projector techniques. We note here, that the word "projector" is used rather loosely in the context of PMC, a projector is any function of the Hamiltonian whose dominant eigenstate is the ground state. To obtain $|\psi_{\tau\to\infty}\rangle$, PMC methods numerically iterate the equation

$$|\psi_{\tau+\Delta\tau}\rangle = \hat{P}(\Delta\tau)|\psi_\tau\rangle\,, \tag{3.6}$$

which, for $S = E_0$, converges to the ground state, regardless of the initial state. Note, that reaching convergence by the repeated application of the projector is equivalent to the power method [109], which is used to find the largest eigenvector of a matrix. To get around the memory limitation imposed by the exponentially growing Hilbert space, PMC methods store at any instant in time only a random sample of vector or matrix elements, i.e. they perform a random walk in the space spanned by the basis elements.

3.2 Taxonomy

While all members of the PMC family rely on the same theoretical foundation, they differ in many practical aspects, that can be categorized according to the following characteristics.

- *Projector*: the choice of the projector is not unique, it can be any function of the Hamiltonian, whose dominant eigenstate is the ground state. The most common choices include the exponential projector $\hat{P}(\Delta\tau) = e^{-(\hat{H}-S\mathbb{I})\Delta\tau}$ and its first order linear approximation $\hat{P}(\Delta\tau) = \mathbb{I} - (\hat{H} - S\mathbb{I})\Delta\tau$.
- *Temperature*: the different approaches can be tailored either for finite or zero temperature calculations.
- *Basis*: the methods are distinguished by the choice of the basis states, i.e. the space where the Monte Carlo walk is performed.

Table 3.1 classifies some of the most commonly used projector techniques, such as Diffusion Monte Carlo (DMC) [110, 111] and Green's Function Monte Carlo (GFMC) [112, 113]. It also includes the recently developed FCIQMC for quantum chemistry calculations [114–117], that has been introduced for modeling both ground state [118, 119] and real-time dynamics [120]. We also present the Density Matrix Quantum Monte Carlo (DMQMC) [121–123], a finite temperature variant of FCIQMC, that has been designed to sample the space of density matrices, instead of the space of wave functions. The practical steps of the implementation, such as how to adjust the energy shift to achieve $S = E_0$, will only be introduced for the FCIQMC method in Chapter 4.

Table 3.1 – Classification of the most commonly used PMC techniques. The comment "(smp. τ)" denotes that τ is not a fixed value but sampled from a probability distribution.

Method	Projector	Temperature	Basis
DMC	$e^{-(\hat{H}-S\mathbb{I})\Delta\tau}$	zero	1st quantized
GFMC	$e^{-(\hat{H}-S\mathbb{I})\Delta\tau}$ (smp. τ)	zero	1st quantized
FCIQMC	$\mathbb{I} - (\hat{H} - S\mathbb{I})\Delta\tau$	zero	2nd quantized
DMQMC	$\mathbb{I} - (\hat{H} - S\mathbb{I})\Delta\tau$	finite	2nd quantized

3.3 The fermionic sign problem

While PMC methods are computationally effective tools to address many-body bosonic quantum systems, their success is greatly limited by the sign problem for the majority of fermionic models. The features of how the problem manifests itself differs for the various numerical approaches, and discussing these details is out of the scope of the present thesis. In general, however, it arises when the dominant eigenstate of the projector does not correspond to the fermionic ground state of the Hamiltonian. Then the repeated application of the projector induces the growth of a state relative to the ground state, which generally leads to an exponential error growth in the quantum mechanical estimators. Therefore, the severity of the sign problem can be qualified by the gap between the dominant eigenstate of the projector and the fermionic ground state. In a first quantized basis, the projector has solely non-negative off-diagonal elements and the dominant eigenstate is entirely of one sign. This state corresponds to the bosonic ground state, thus making it impossible for it to represent the fermionic ground state which is well-known to have a nodal structure. Hence, the sign problem is connected to the gap between the bosonic and fermionic ground states [110, 111]. In a second quantized basis however, the dominant eigenstate does not necessarily correspond to a physical bosonic state [31]. In this thesis, we discuss the details of the emergence of the sign problem only in the scope of FCIQMC simulations for a detailed review on other numerical approaches readers are referred to [31, 124, 125].

CHAPTER 4

Full Configuration Interaction Quantum Monte Carlo

In Chapter 3, we covered the theoretical foundation of projector Monte Carlo techniques, and we introduced the concept of the fermionic sign problem. Recently, a particular class of PMC techniques has been developed for quantum chemistry simulations, namely the Full Configuration Interaction Quantum Monte Carlo [114–117, 126]. The method has been extended to sample both ground state [118, 119] and real-time dynamics [120] of closed many-body systems, and due to a particular sampling strategy, it has proven to significantly alleviate the sign problem [115, 127].

As we expect the sign problem to be highly relevant in open many-body calculations, using FCIQMC as a stepping stone in the development of a novel open system approach holds the promise of a computationally effective tool. In this chapter, we overview the practicalities of the FCIQMC algorithm, as well as its effect in challenging the sign problem.

4.1 Overview

FCIQMC, as all other PMC techniques, relies on the formal solution of the imaginary-time Schrödinger equation by exploiting the fact, that its long-time limit evolution projects out the ground state from any arbitrary initial wave function (assuming to have non-zero overlap with it). Having introduced an additional energy shift S and a linear projector (see in Chapter 3), this solution reads as

$$|\psi_{\tau+\Delta\tau}\rangle = \hat{P}(\Delta\tau)|\psi_\tau\rangle = \mathbb{I} - \left(\hat{H} - S\mathbb{I}\right)|\psi_\tau\rangle\,. \tag{4.1}$$

For $S = E_0$ with E_0 being the ground state energy, in the asymptotic limit of $\tau \to \infty$ the wave function converges to the ground state. Since the value of E_0 is a priori unknown, the simulation starts from an arbitrary S shift which is then slowly adjusted throughout

the simulation. FCIQMC stochastically samples the first order Euler approximation of eq. (4.1) on a discrete basis set $\{|\Phi_i\rangle\}$ where the wave function reads as

$$|\psi_\tau\rangle = \sum_i c_i^\tau |\Phi_i\rangle \,. \tag{4.2}$$

From eq. (4.1), the evolution of the expansion coefficients c_i^τ then yields

$$c_i^{\tau+\Delta\tau} = \left[1 - \Delta\tau \left(H_{ii} - S\right)\right] c_i^\tau - \Delta\tau \sum_{j\neq i} H_{ij} c_j^\tau \,, \tag{4.3}$$

where $H_{ij} = \langle\Phi_i|\hat{H}|\Phi_j\rangle$. To stochastically represent the evolution in eq. (4.3), we introduce the notion of a fundamental unit, called walker. Each walker has a sign ($q = \pm 1$), and can reside on one of the basis states $|\Phi_i\rangle$, i.e. they contribute to the sampling of one of the coefficients c_i^τ. If n_i^+ is the number of positive and n_i^- is the number of negative walkers residing on $|\Phi_i\rangle$, then the amplitude of that state is proportional to the net population as $c_i^\tau \propto n_i^\tau = n_i^+ - n_i^-$. For complex coefficients, this method can be generalized by using two sets of walkers for the real and imaginary parts respectively [118]. Then starting from an initial distribution, the walkers evolve by a set of rules that represent a stochastically exact mapping of the discrete time evolution in eq. (4.3). This dynamics is repeated until convergence is reached.

Note, that the last two terms of eq. (4.3) describe the rate at which c_i^τ changes. Thus, they are sufficient to define the dynamical rules of the walker evolution. At each time step, on the entire walker population the following operations are performed:

i *Spawning*: the last term of eq. (4.3) describes at which rate the other coefficients $c_{j\neq i}^\tau$ contribute to c_i^τ. The coefficients $c_{j\neq i}^\tau$ linked with a non-zero H_{ij} contribution are called "connected" to c_i^τ. While looping through walkers residing on $|\Phi_i\rangle$, this step of the algorithm is, in fact, equivalent to computing the contributions of c_i^τ to its connected basis elements $c_{j\neq i}^\tau$ as

$$p_{\text{spawn}} = \sum_{j\neq i} p_{\text{spawn}}(j|i) = \sum_{j\neq i} -q_i \Delta\tau H_{ji} \,. \tag{4.4}$$

$p_{\text{spawn}}(j|i)$ represents the probability that a walker residing on c_i^τ (parent) spawns another signed walker on c_j^τ (child). In case of a full Hamiltonian, summing over all $p_{\text{spawn}}(j|i)$ involves $\mathcal{O}(\dim\mathcal{H} \times N_w)$ operations, where N_w is the total number of walkers $N_w = \sum_i |n_i^+ - n_i^-|$. Instead, it is computationally more efficient to sample only one of the connected states as

$$p_{\text{spawn}}(j|i) = -q_i \Delta\tau \frac{H_{ji}}{p_{\text{gen}}(j|i)} \,, \tag{4.5}$$

reducing the number of iterations to $\mathcal{O}(N_w)$ at each iteration step. The probability p_{gen} of suggesting state j from state i is arbitrarily chosen as long as $\sum_j p_{\text{gen}}(j|i) = 1$. While this sampling increases the necessary number of iteration steps until convergence is reached, the cost of each step becomes so much cheaper that the overall computational efficiency if improved [114]. If $|p_{\text{spawn}}(j|i)| > 1$ then the largest integer part, denoted as $\lfloor |p_{\text{spawn}}(j|i)| \rfloor$, is realized deterministically, and the fractional part stochastically. In practice, this means that a random number $r \in [0, 1]$ is drawn, and if $r < (|p_{\text{spawn}}(j|i) - \lfloor p_{\text{spawn}}(j|i) \rfloor|)$, then another walker is spawned.

ii *Clone/Death*: the second term in eq. (4.3) corresponds to the rate at which c_i^τ is projected onto itself as

$$p_{\text{clone}}(i) = \Delta\tau \left(H_{ii} - S\right) . \tag{4.6}$$

Numerically, this represents the probability that a walker is being copied (clone) if $p_{\text{clone}}(i) < 0$, or removed from the simulation (death) if $p_{\text{clone}}(i) > 0$. Just as before, if $|p_{\text{clone}}(i)| > 1$, the integer part $\lfloor p_{\text{clone}}(i) \rfloor$ is realized deterministically, and the fractional part stochastically.

iii *Annihilation*: if two walkers with opposite sign populate the same basis element, their contribution to the amplitude is zero, and they can be removed from the simulation. Therefore, at the end of the iteration step each $|\Phi_i\rangle$ basis state has walkers of the same sign only.

4.2 Population control and convergence criteria

As we have seen in Chapter 3, the relation between the shift S and the ground state energy E_0 determines whether the wave function converges to the ground state or infinitely diverges. In the scope of FCIQMC, keeping S constant at a value above E_0 then yields to an exponentially growing walker population, while $S < E_0$ results in a decrease. Since the value of E_0 is a priori unknown, the value of S is first kept above the ground state energy until the population reaches a desired N_w total number, then periodically adjusted to balance the population growth [111, 114] as

$$S(\tau) = S(\tau - \Delta\tau) + \frac{\delta}{\Delta\tau} \log\left(\frac{N_w(\tau)}{N_w(\tau - \Delta\tau)}\right) . \tag{4.7}$$

The damping term δ prevents large fluctuations in S, and its value is to be decided empirically. Once convergence is reached, the value of the shift provides an estimate of the ground state energy $S \approx E_0$.

We note here, that while the exponential projector will always project onto the ground state regardless of the size of the time step, this is not the case for its linear approximation.

Recall, that after n FCIQMC iterations

$$\hat{P}^n|\psi_0\rangle = \left[1 - \Delta\tau(\hat{H} - S\mathbb{I})\right]^n |\psi_0\rangle = \sum_i \left[1 - \Delta\tau(E_i - S)\right]^n v_i|\phi_i\rangle \quad \text{with } \hat{H}|\phi_i\rangle = E_i|\phi_i\rangle,\ |\psi_0\rangle = \sum_i v_i|\phi_i\rangle\,, \tag{4.8}$$

which converges to the ground state only if $|1 - \Delta\tau(E_i - S)| < 1$ for all E_i eigenvalues. Using that $S \approx E_0$, this corresponds to

$$\Delta\tau < \tau_m = \frac{2}{E_{\max} - E_0} = \frac{2}{\Delta E}\,, \tag{4.9}$$

where ΔE is the spectral width of the Hamiltonian. However, as long as the time step is below τ_m, the simulation has no time step error and converges to the exact ground state.

4.3 Stochastic estimators

The FCIQMC method provides two alternative ways to compute the value of the ground state energy. One estimate is obtained from the energy shift S, which, as explained earlier, is periodically adjusted over the course of the simulation to keep the population number more or less constant. Once this has been established, the ground state normalization is conserved, and the estimated ground state energy is equal to the energy shift.

An alternative approach is given by the projected estimator

$$\langle E\rangle_\tau = \frac{\langle\Psi^T|\hat{H}|\psi_\tau\rangle}{\langle\Psi^T|\psi_\tau\rangle} = \frac{\sum_i c_i^\tau \sum_{j\in T} H_{ij} t_j}{\sum_{i\in T} c_i^\tau t_i} = \frac{\sum_i c_i^\tau \mathcal{H}_i}{\sum_{i\in T} c_i^\tau t_i}\,, \tag{4.10}$$

where $|\Psi^T\rangle$ is the trial wave function expressed as

$$|\Psi^T\rangle = \sum_{i\in T} t_i|\Phi_i\rangle\,, \tag{4.11}$$

and T is the set of basis states that contribute to the trial function. Since $\langle E\rangle_\tau$ is a zero-bias, zero-variance estimator when the trial wave function is equal to the ground state, improving the quality of $|\Psi^T\rangle$ reduces the stochastic fluctuations. Typical trial wave functions are constructed from a single basis state, or as a snapshot of the coefficients c_i^τ at a given time step. This second implementation can be achieved with a negligible computational overhead by storing the $\mathcal{H}_i$ terms for each occupied basis state [126]. Accordingly, the expectation value of an arbitrary observable $\hat{O}$ at a given instant in time is computed

similarly

$$\langle \hat{O} \rangle_\tau = \frac{\sum_i c_i^\tau \sum_{j \in T} O_{ij} t_j}{\sum_{i \in T} c_i^\tau t_i} = \frac{\sum_i c_i^\tau \mathcal{O}_i}{\sum_{i \in T} c_i^\tau t_i} \,. \tag{4.12}$$

Once the simulation has asymptotically approached the ground state, i.e. the shift S converged to a constant value, $\langle \hat{O} \rangle_\tau$ is averaged over a sufficiently large number of iteration steps to obtain its ground state expectation value.

4.4 Algorithm

Let us now summarize the course of a typical FCIQMC simulation. We start by dispensing an initial walker distribution, and setting a constant energy shift such that $S > E_0$ to obtain a rapid population growth. Once the total number of walkers has reached a predetermined N_w value, we begin to vary the shift according to eq. (4.7). When the observables has reached a more or less constant value, the simulation continues until sufficient number of samples has been collected for the stochastic estimators.

In each time step the following steps are performed:

1. **Spawn:** for a walker residing on site i a connected site j is chosen randomly and a spawning event is made possible with a rate

$$p_{\text{spawn}}(j|i) = -q_i \Delta\tau \frac{H_{ji}}{p_{\text{gen}}(j|i)} = -q_i \Delta\tau H_{ji} N_{\text{conn}} \,, \tag{4.13}$$

   with N_{conn} being the total number of connected states [32].

2. **Clone / death:** for each walker on a given basis state i a cloning event is sampled with the stochastically realized probability

$$p_{\text{clone}}(i) = \Delta\tau \left(H_{ii} - S \right) \,. \tag{4.14}$$

3. **Annihilate:** those pair of walkers that reside on the same state with opposite sign are removed from the simulation.

4. **Update the energy shift:** either, in constant shift mode, we do nothing or, in constant N_w mode, S is adjusted as to eq. (4.7).

5. **Sample observables:** compute the observables at time τ following eq. (4.12).

4.5 Sign problem in FCIQMC and population dynamics

In Chapter 3, we have made some general comments about the nature of the fermionic sign problem, here, we discuss how it manifests itself during the course of an FCIQMC simulation. We will follow the analysis of Spencer et. al. [115] by first considering two population of non-interacting walkers, i.e. positive and negative ones. Removing the annihilation then lets us understand its crucial role in the long-time convergence, and explains the typical "plateau"-like behaviour present in FCIQMC dynamics [114].

Let us denote with n_i^+ the positive and n_i^- the negative walkers residing on state $|\Phi_i\rangle$, and define the transition matrix as

$$\begin{aligned} \frac{dc_i^\tau}{d\tau} &= \sum_j T_{ij} c_j^\tau \\ \hat{T} = -(\hat{H} - S\mathbb{I}) &= \hat{T}^+ - \hat{T}^- \,, \end{aligned} \tag{4.15}$$

where $\hat{T}^+$ contains the positive, and $\hat{T}^-$ the absolute value of the negative elements of $\hat{T}$. Then the dynamics of positive and negative walkers is governed by

$$\begin{aligned} \frac{dn_i^+}{d\tau} &= \sum_j \left(T_{ij}^+ n_j^+ + T_{ij}^- n_j^- \right) \\ \frac{dn_i^-}{d\tau} &= \sum_j \left(T_{ij}^+ n_j^- + T_{ij}^- n_j^+ \right) . \end{aligned} \tag{4.16}$$

These equations can be decoupled into the the evolution of two independent quantities

$$\begin{aligned} \frac{d(n_i^+ + n_i^-)}{d\tau} &= \sum_j \left(T_{ij}^+ + T_{ij}^- \right) \left(n_i^+ + n_i^- \right) \\ \frac{d(n_i^+ - n_i^-)}{d\tau} &= \sum_j \left(T_{ij}^+ - T_{ij}^- \right) \left(n_i^+ - n_i^- \right) , \end{aligned} \tag{4.17}$$

where we recognize the evolution of the wave function coefficients $c_i^\tau \propto n_i^+ - n_i^-$. As $\tau \to \infty$, this dynamics tends to the dominant eigenvector of $\hat{T}$, just like we expect it from eq. (3.3), the imaginary time Schrödinger equation. However, there is an addition signal $n_i^+ + n_i^-$ present in the simulation that converges to the dominant eigenstate of $\hat{T}^+ + \hat{T}^-$ instead. It has been shown [115], that the dominant eigenvalue of $\hat{T}^+ + \hat{T}^-$ will always

be greater than or equal to the ground state energy. Thus, while it is in principle still possible to obtain the ground state as $n_i^+ - n_i^-$, it will be lost due to statistical noise as $n_i^+ + n_i^-$ grows exponentially large. To study the effect of annihilation, we introduce a simple process into eq. (4.16) which cancels walkers of opposite sign at a constant κ rate [115]

$$\begin{aligned} \frac{dn_i^+}{d\tau} &= \sum_j \left(T_{ij}^+ n_j^+ + T_{ij}^- n_j^- \right) - \kappa n_i^+ n_i^- \\ \frac{dn_i^-}{d\tau} &= \sum_j \left(T_{ij}^+ n_j^- + T_{ij}^- n_j^+ \right) - \kappa n_i^+ n_i^- \,, \end{aligned} \tag{4.18}$$

where κ is a small, positive constant. Then the decoupled differential equations become

$$\begin{aligned} \frac{d(n_i^+ + n_i^-)}{d\tau} &= \sum_j \left(T_{ij}^+ + T_{ij}^- \right) \left(n_i^+ + n_i^- \right) - 2\kappa n_i^+ n_i^- \\ \frac{d(n_i^+ - n_i^-)}{d\tau} &= \sum_j \left(T_{ij}^+ - T_{ij}^- \right) \left(n_i^+ - n_i^- \right) \,. \end{aligned} \tag{4.19}$$

It is clear now how annihilation helps the emergence of the physical ground state. As the population $n_i^+ + n_i^-$ grows, the annihilation events increase with them, until the rate of spawning is balanced by the opposite walker cancellation. The physical signal $n_i^+ - n_i^-$ however, is not affected by this process. Figure 5.11 shows the crucial role of annihilation in the ground state convergence. Finally, we can conclude that the severity of the sign problem in FCIQMC is characterized by the gap between the dominant eigenstate of the physical projector $\hat{T}^+ - \hat{T}^-$, and the dominant eigenstate of $\hat{T}^+ + \hat{T}^-$, the projector with all off-diagonal elements replaced by their absolute value.

Separating the walker population into two signals $n_i^+ - n_i^-$ and $n_i^+ + n_i^-$ in eq. (4.19) also explains the typical "plateau"-like evolution of N_w in constant shift mode (Figure 4.2) [114, 115]. When S is kept constant at value that is slightly larger than the ground state, in the first phase of the simulation the annihilation cannot counteract the exponential growth of $n_i^+ + n_i^-$. The evolution then enters a plateau phase where the dynamics is driven by a balance between the walker cancellation and spawning events. The population remains roughly constant, until the steadily growing ground state signal $n_i^+ - n_i^-$ starts to dominate. From then we experience a second population growth, although now at a rate determined by $n_i^+ - n_i^-$. Allowing the walker population to grow further only serves to reduce the statistical error, thus it is this stage when the adjustment in the energy shift is introduced. The final phase of the simulation then proceeds with a more or less constant walker population and the sampling of ground state properties might begin. Here we have only given an empirical description however, extensive studies has shown [115, 127] that the intrinsic features of the physical model are reflected in the properties of the plateau, namely, that the walker population required to reach it (the height) is directly related to the severity of the sign problem.

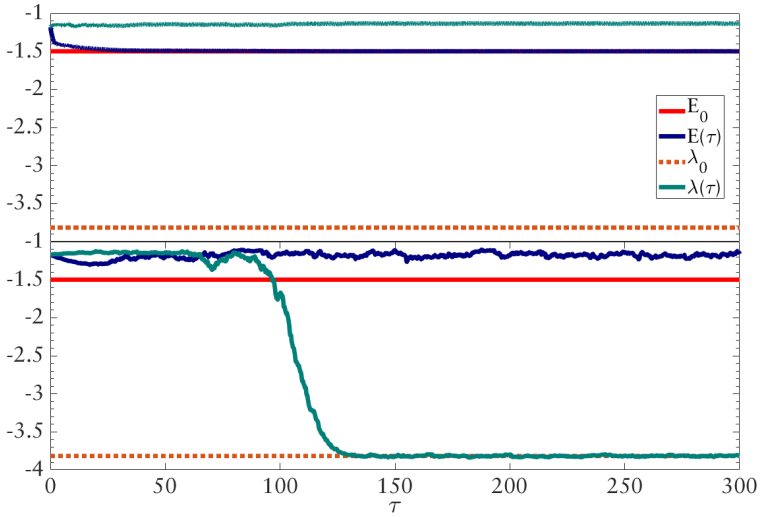

Figure 4.1 – The effect of removing the annihilation step from the FCIQMC algorithm. The method is applied to a 100×100 randomly generated, real and hermitian matrix $\hat{H}$. E_0 is the exact ground state, while $E(\tau)$ is its stochastic estimation. λ_0 and $\lambda(\tau)$ are the analogous quantities for the matrix $\hat{H}^+ + \hat{H}^-$ with $\hat{H} = \hat{H}^+ - \hat{H}^-$. While the traditional algorithm lets the ground state emerge, in the absence of annihilation the simulation converges to the ground state of $\hat{H}^+ + \hat{H}^-$ instead.

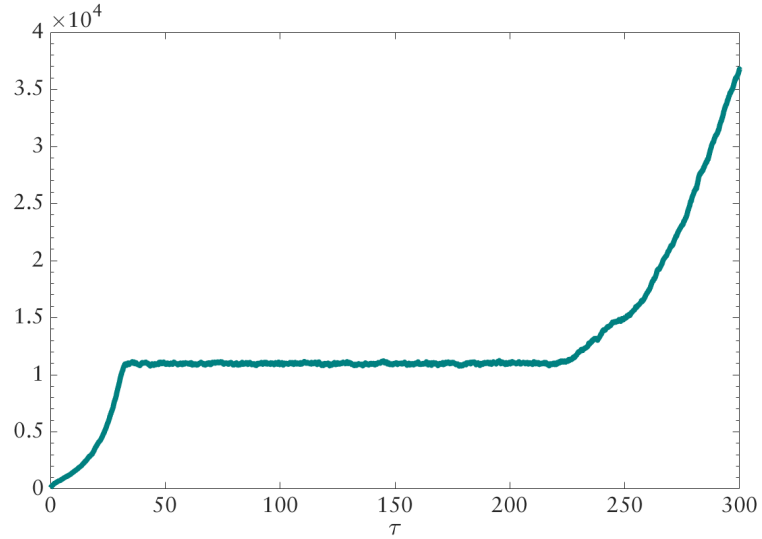

Figure 4.2 – A typical evolution of the walker population where the value of the energy shift S is kept at a constant value that is slightly larger than the ground state.

CHAPTER 5

Driven-dissipative quantum Monte Carlo

In this chapter, we introduce a real-time quantum Monte Carlo approach for open systems, which we call Driven-dissipative quantum Monte Carlo (DDQMC). DDQMC has been developed in the footsteps of the Full Configuration Interaction Quantum Monte Carlo (Chapter 4), using it to significantly alleviate the fermionic sign problem [115,127].

The development of this method was directed by Vincenzo Savona, and it was published in *Physical Review A* [32]. A considerable amount of the description presented in this thesis has been extracted from [32]. The author's contribution to the project was part of the theoretical development, the numerical implementation and the physical model simulations.

The method was implemented in C++ and the software can be accessed in the repository *https://github.com/EPFLLTPN/DDQMC* [34].

5.1 Overview*

A mutual feature of the Lindbladian dynamics and the imaginary-time Schrödinger equation is the fact that, in the long-time limit, the eigenstate with the smallest-real-part-eigenvalue will survive. In the Lindbladian case, where the eigenvalues of the Lindblad superoperator are complex valued with non-positive real part [37], this corresponds to the zero eigenvalue solution — to the NESS (Table 5.1). It is therefore natural to apply the foundation of projector Monte Carlo techniques to the simulation of the steady state.

5.1.1 Theory

DDQMC shares many of the features of FCIQMC, but while FCIQMC samples the imaginary time evolution of the wave function, DDQMC is focused on the real time dynamics of the complex-valued density matrix. The general problem we aim to solve is in the form

of eq. (1.36), the Lindblad quantum master equation. We introduce an additional shift into eq. (1.36) with

$$\frac{d\hat{\rho}(t)}{dt} = \mathcal{L}\left[\hat{\rho}(t)\right] - S\mathbb{I}\hat{\rho} \equiv \widetilde{\mathcal{L}}\left[\hat{\rho}(t)\right] , \tag{5.1}$$

which then can be sampled stochastically in a similar manner as eq. (4.3). Like for FCIQMC, we take the first order Euler approximation

$$\hat{\rho}(t + \Delta t) = \hat{\rho}(t) + \widetilde{\mathcal{L}}\left[\hat{\rho}\right]\Delta t\,, \tag{5.2}$$

and we introduce a set of walkers which now sample the amplitudes of basis operators $|\Phi_i\rangle\langle\Phi_i|$, from now on referred to as "configurations". The expectation value of an observable $\hat{O}$ at a given instant in time can be computed as

$$\langle\hat{O}\rangle_t = \frac{\mathrm{Tr}[\hat{O}\hat{\rho}(t)]}{\mathrm{Tr}[\hat{\rho}(t)]} = \frac{\sum_{i,j}\rho_{ij}(t)O_{ji}}{\sum_i \rho_{ii}(t)}\,. \tag{5.3}$$

Once the simulation has asymptotically approached the steady state – i.e. the shift S has reached the steady state eigenvalue of the Lindbladian, $S = 0$ – the numerator and the denominator can be averaged separately over a sufficiently large number of iteration steps. Not setting $\mathrm{Tr}(\hat{\rho}) = 1$ in the simulation is analogous to not setting $\langle\psi|\psi\rangle = 1$ in the standard FCIQMC scheme. Eq. (5.3) ensures that the norm is correctly accounted for in the calculation of expectation values.

Table 5.1 – A parallel is drawn between the imaginary time evolution of closed systems and the real time evolution of the quantum master equation for open quantum systems. In the case of open systems we assume, here and throughout this work, that a unique steady state exists.

	Closed system	Open system
Operator	Hamiltonian	Lindbladian
	$\hat{H}$	$\mathcal{L}$
Dynamics	Imaginary time	Real time
	$\dot{\lvert\psi_\tau\rangle} = -(\hat{H} - S\mathbb{I})\lvert\psi\rangle$	$\dot{\hat{\rho}}_t = \mathcal{L}\left[\hat{\rho}\right]$
Long time limit	Ground state	Steady state
	$e^{-(\hat{H}-S\mathbb{I})\tau} : \lvert\psi\rangle \xrightarrow{\tau\to\infty} \lvert\Phi_0\rangle$	$e^{\mathcal{L}t} : \hat{\rho} \xrightarrow{t\to\infty} \hat{\rho}_{ss}$

5.1.2 Multinomial formalism

The FCIQMC algorithm was originally performed by looping through all the walkers in the simulation [114]. Here, we introduce a computationally more efficient strategy which scales with the number of simultaneously occupied configurations, rather than the total walker number. We refer to this variant as "multinomial formalism". Let $p(j|i)$ be the probability of choosing the j-th connected state from i

$$p(j|i) = \frac{|H_{ji}|\Delta\tau}{\sum_k |H_{ki}|\Delta\tau} = \frac{|H_{ji}|\Delta\tau}{P_{tot}} \,. \tag{5.4}$$

The number of actual spawning events N_{sp}^i that take place for N_i walkers residing on site i is determined by a stochastic process following a binomial distribution

$$f(N_{sp}^i; N_i, P_{tot}) = \frac{N_i!}{N_{sp}^i!(N_i - N_{sp}^i)!} P_{tot}^{N_{sp}^i} (1 - P_{tot})^{N_i - N_{sp}^i} \,. \tag{5.5}$$

Then the N_{sp}^i walkers are divided into groups $\{M_1 \dots M_l\}$, where l is the number of states connected to the starting one by a non-zero Hamiltonian element. For each group, M_j walkers are spawned to state j-th with a sign $q_j = -\text{sign}(H_{ji})q_i$. The set of integers $\{M_j\}$ is drawn randomly following the multinomial distribution

$$\begin{aligned} &f_M(M_1 \dots M_l; N_{sp}^i, p(1|i) \dots p(l|i)) = \\ &\quad = \tfrac{N_{sp}^i!}{M_1! \cdots M_l!}\, p(1|i)^{M_1} \times \cdots \times p(l|i)^{M_l} \,. \end{aligned} \tag{5.6}$$

For systems with local interaction, the Hamiltonian can be represented as a sparse matrix, and a computationally efficient state representation makes the extra memory allocation negligible (for an occupied state i it is necessary to store all the possible connected states with the corresponding probabilities). Efficient algorithms for binomial and multinomial random number generation are also present in the literature [128–130].

Similarly, rather than looping through the individual walkers, the number of clones for a given state i can be determined at once as

$$f(N_{cl}^i; N_i, p_{\text{clone}}(i)) = \frac{N_i!}{N_{cl}^i!(N_i - N_{cl}^i)!} p_{\text{clone}}(i)^{N_{sp}^i} (1 - p_{\text{clone}}(i))^{N_i - N_{cl}^i} \,, \tag{5.7}$$

where $p_{\text{clone}}(i) = \Delta\tau|(H_{ii} - S)|$, and if $p_{\text{clone}}(i) < 0$ then the progenies have the same sign, if $p_{\text{clone}}(i) > 0$ then the opposite sign as the walkers residing on state i.

5.1.3 Algorithm

The dynamics of the walker population is determined by a set of rules designed to stochastically sample eq. (5.2). Since the density matrix is complex-valued, following [118] we introduce two sets of walkers for the real and imaginary parts respectively. If the density matrix is expressed in the vectorized form $|\rho\rangle\rangle$, the shifted Lindblad operator can also be expressed in matrix form using Kronecker products as [45]

$$\mathcal{L} = -i\left(\mathbb{I}\otimes\hat{H} + \hat{H}^T\otimes\mathbb{I}\right) - S\mathbb{I}\otimes\mathbb{I} - \sum_k \frac{\gamma_k}{2}\left(\mathbb{I}\otimes L_k^\dagger L_k + L_k^T L_k^*\otimes\mathbb{I} - 2L_k^*\otimes L_k\right) . \quad (5.8)$$

Then eq. (5.1) can be written in the form

$$\frac{d\rho_{ij}}{dt} = \widetilde{\mathcal{L}}_{ij}^{ij}\rho_{ij} + \sum_{l,m\neq i,j}\widetilde{\mathcal{L}}_{ij}^{lm}\rho_{lm} , \quad (5.9)$$

where $\widetilde{\mathcal{L}}_{ij}^{lm}$ are the matrix elements of the superoperator, and ρ_{ij}represents the now complex valued population on a configuration $|\Phi_i\rangle\langle\Phi_j|$. For the sampling protocol we apply the previously introduced multinomial formalism (Subsec. 5.1.2). Let us denote this scheme as a function, introducing a shorthand notation

$$y = \textbf{Multinomial}(x) , \quad (5.10)$$

where x takes values of the Lindbladian superoperator elements and y returns the number of spawned walkers on a given configuration as defined in Eq. (5.4)-(5.6). Figure 5.1 shows the real and imaginary walker representation of the complex-valued density matrix elements, as well as the connection between the different sets of walkers during the spawning process.

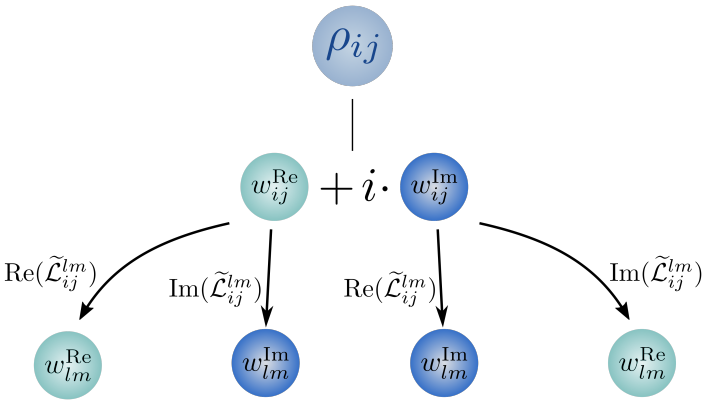

Figure 5.1 – Real and imaginary walker representation of the complex-valued density matrix elements and the spawning procedure.

At each time step, we loop over the list of simultaneously occupied configurations and perform the following steps:

i *Spawning:* consider the complex walker population residing on ρ_{ij} and perform spawning to all the connected configurations. The real (Re) and imaginary (Im) parts of $\widetilde{\mathcal{L}}_{ij}^{lm}$ are considered in turn and two spawning attempts are realized for real and imaginary parents, respectively. The number of walkers spawned to each ρ_{lm} are determined by the multinomial formalism.

For real parents

$$\begin{cases} N_{lm}^{\mathrm{Re}} = \mathbf{Multinomial}(\mathrm{Re}(\widetilde{\mathcal{L}}_{ij}^{lm})) \\ \mathrm{sign} = \mathrm{sign}(\mathrm{Re}(\rho_{ij})\mathrm{Re}(\widetilde{\mathcal{L}}_{ij}^{lm})) \end{cases} \tag{5.11}$$

$$\begin{cases} N_{lm}^{\mathrm{Im}} = \mathbf{Multinomial}(\mathrm{Im}(\widetilde{\mathcal{L}}_{ij}^{lm})) \\ \mathrm{sign} = \mathrm{sign}(\mathrm{Re}(\rho_{ij})\mathrm{Im}(\widetilde{\mathcal{L}}_{ij}^{lm})) \end{cases} \tag{5.12}$$

and for imaginary parent walkers

$$\begin{cases} N_{lm}^{\mathrm{Re}} = \mathbf{Multinomial}(\mathrm{Im}(\widetilde{\mathcal{L}}_{ij}^{lm})) \\ \mathrm{sign} = -\mathrm{sign}(\mathrm{Im}(\rho_{ij})\mathrm{Im}(\widetilde{\mathcal{L}}_{ij}^{lm})) \end{cases} \tag{5.13}$$

$$\begin{cases} N_{lm}^{\mathrm{Im}} = \mathbf{Multinomial}(\mathrm{Re}(\widetilde{\mathcal{L}}_{ij}^{lm})) \\ \mathrm{sign} = \mathrm{sign}(\mathrm{Im}(\rho_{ij})\mathrm{Re}(\widetilde{\mathcal{L}}_{ij}^{lm})) \end{cases} \tag{5.14}$$

where N_{lm}^{Re} and N_{lm}^{Im} are the number of real and imaginary walkers being spawned to configuration ρ_{lm} and 'sign' indicates the sign of the progeny.

ii *Clone / Death:* This step is required as a real (imaginary) walker can produce an imaginary (real) walker on the same configuration. The spawning occurs on-site with a population determined by the binomial distribution. The real and imaginary parts of $\widetilde{\mathcal{L}}_{ij}^{ij}$ are considered in turn and two spawning attempts are realized respectively for the real and imaginary population.

iii *Annihilation:* on a given site the real and imaginary population are considered in turn, and pairs of walkers with opposite sign are removed from the simulation.

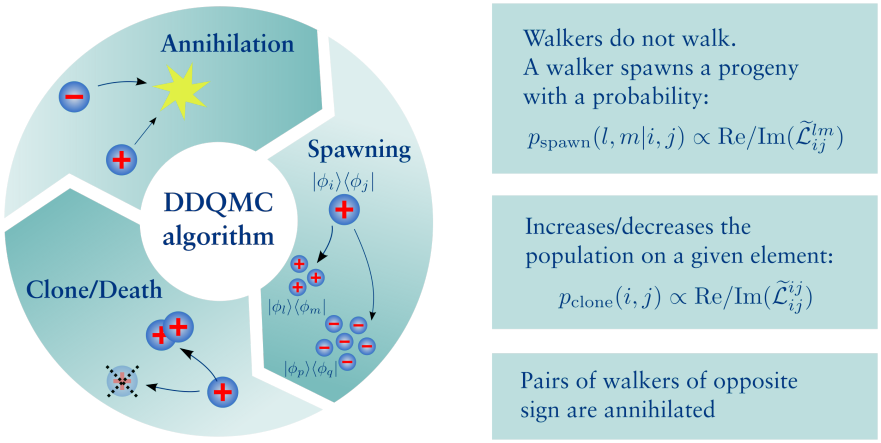

Figure 5.2 – Schematic diagram of the steps of the DDQMC algorithm.

Unlike in FCIQMC, here the exact value of the diagonal shift S is known. In the infinite time limit the master equation is drives the density matrix into the steady state, which by definition belongs to the zero eigenvalue of the Lindbladian. Therefore, the time evolution of the shift S will indicate whether convergence has been reached, as well as the quality of the simulation. As before, the shift is slowly adjusted in order to maintain a constant walker population. Since estimators for most operators only receive contributions from walkers on or near the diagonal elements, we chose to control the amount of population distributed along the diagonal of the density matrix. The value of the shift is then adjusted according to the familiar shift-update algorithm implemented in FCIQMC [114] calculations

$$S(t) = S(t - \Delta t) + \frac{\delta}{\Delta t} \log\left(\frac{N_w(t)}{N_w(t - \Delta t)}\right), \tag{5.15}$$

where δ is a damping parameter, and N_w is the total number of real walkers residing on diagonal density matrix elements. The method does not have a built-in constraint on the diagonal elements being real. The value of the imaginary part fluctuates around zero and its expectation value naturally vanishes during the simulation.

5.2 Initiator approach and importance sampling

The DDQMC algorithm allows to stochastically sample the steady state density matrix of open quantum systems whose dynamics follows the Lindblad master equation. However, with increasing system size the walker population tends to become dilute, resulting in a situation where the simulation contains only a few walkers per configuration. This leads to an increased statistical error and a dramatically reduced sampling accuracy. In order to address this issue one needs to increase the walker number in the system which – with large number of basis states – becomes computationally unfeasible.

In order to overcome this issue two different methods were introduced: initiator approach and importance sampling. Each of these techniques reduces the minimal required walker population by decreasing the number of simultaneously occupied configurations, however the strategy of selecting the configurations to be sampled is fundamentally different. The initiator approach allows the significant configurations to emerge naturally during the simulation, whilst importance sampling gives the possibility to drive the walker population to a selected subset of presumably relevant configurations. These methods can improve the sampling quality with great success, however, they have to be applied carefully, as both introduce a bias on the result.

5.2.1 Initiator approach

Cleveland et. al [117] has demonstrated that by introducing an additional criteria for the newly spawned walkers one can dramatically reduce the population required to reach convergence. We introduce the notion of 'initiators', which are configurations that have the ability to spawn progeny onto unoccupied basis states. However, walkers from the non-initiator states can only survive if they spawn to a diagonal configuration or to one that is already populated. All the diagonal configurations are initiators by definition, and during the simulation a configuration might become initiator, if its population exceeds a preset value (I_{limit}). This method results in a series of systematically improvable approximations which will tend to the original algorithm in three limits:

i with decreasing I_{limit} every basis element will become initiator. All the progeny survives regardless of the parent state, which is equivalent to the original method;

ii with increasing total population all configuration will acquire walkers, therefore, all spawned children will survive regardless of the flag of the parent state;

iii extending the initiator space by definition will result again in all the configurations becoming initiators, consequently all progeny will survive.

By setting a non-zero initiator limit one introduces a dynamical truncation on the available configurations, leading to a biased result. In order to compute the unbiased expec-

tation values, we progressively decrease the initiator limit in different simulations and fit the estimated expectation values, thus extrapolating the limit $I_{\text{limit}} \to 0$ (Figure 5.3).

5.2.2 Importance sampling

We start by identifying the basis elements whose sampling needs to be improved. Then a straightforward way to improve the accuracy is to reduce the probability of spawning out of these configurations. Walkers that do reside on unessential elements are given a correspondingly larger weight, hence the expectation values of the observables will be unchanged. We define the following simple importance sampling procedure. The evolution of the density matrix in the DDQMC formalism follows eq. (5.9). We can assign importance to the density matrix elements by multiplying them with a factor w_{ij} as

$$\rho_{ij} \quad \longrightarrow \quad \widetilde{\rho}_{ij} = w_{ij}\rho_{ij} \,. \tag{5.16}$$

The time evolution of the importance-sampled density matrix then becomes

$$\frac{d\widetilde{\rho}_{ij}}{dt} = \widetilde{\mathcal{L}}^{ij}_{ij}\widetilde{\rho}_{ij} + \sum_{l,m\neq i,j} \left(w_{ij}\widetilde{\mathcal{L}}^{lm}_{ij}\frac{1}{w_{lm}} \right) \widetilde{\rho}_{lm} \,, \tag{5.17}$$

which is fully analogous to eq. (5.9) and can be simulated by the DDQMC method. The expectation value of an observable $\hat{O}$ is thus

$$\langle \hat{O} \rangle = \frac{\sum_{ij} \frac{\widetilde{\rho}_{ij}}{w_{ij}} O_{ji}}{\sum_i \widetilde{\rho}_{ii}} \,. \tag{5.18}$$

In this work we parametrize the importance sampling with a single variable $p > 0$ by assigning all off-diagonal elements a weight $w_{ij} = e^{-p}$. Meanwhile the diagonal coefficients are not altered. This strategy focuses on sampling the diagonal density matrix elements and provides easy access to tune the effect of the importance sampling.

Figure 5.4 shows the amount of simultaneously occupied density matrix elements before and after using the initiator approach and importance sampling for the dissipative XYZ model that has been introduced in Section 1.3.1.

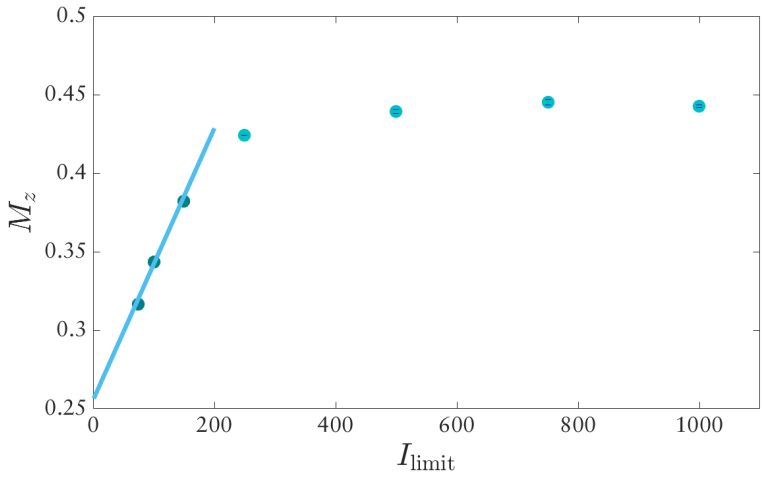

Figure 5.3 – The initiator approach used to extrapolate the magnetization M_z (in units of $\hbar$) in case of the 4×4 dissipative XYZ Heisenberg-model. Parameters of the model: $J_x/\gamma = 0.225$, $J_y/\gamma = 0.335$, $J_z/\gamma = 0.25$. Parameters of the simulation: $p = 2.5$, with a population of 10^6 walkers. The straight line is a linear extrapolation of the lowest initiator values.

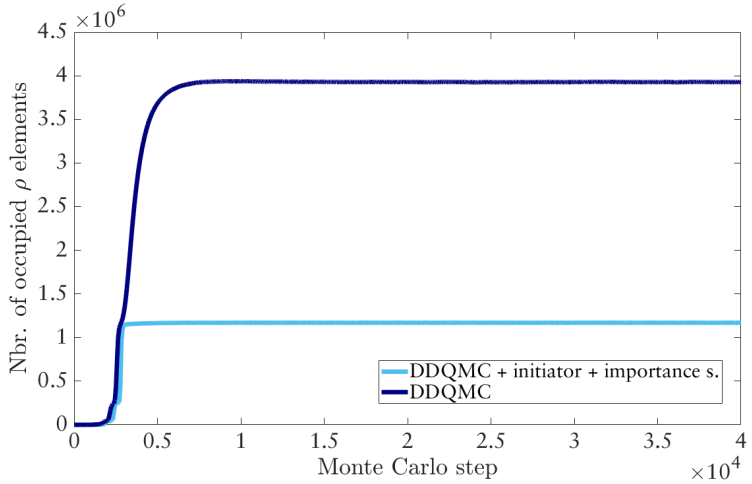

Figure 5.4 – The amount of simultaneously occupied density matrix elements, with and without using the initiator approach and importance sampling in the case of the 4×4 XYZ Heisenberg lattice. Parameters of the model: $J_x/\gamma = 0.225$, $J_y/\gamma = 0.225$, $J_z/\gamma = 0.25$. Parameters of the simulation: $p = 1.5$, $I_{limit} = 25; 75$ with a population of 10^8 walkers.

5.3 Parallel implementation

In DDQMC simulations, the sign problem is manifested as the fast growth of an unphysical state relative to the non-equilibrium steady state. The annihilation procedure can suppress this growth and allow the simulation to converge to the physical solution, but only if a minimal and system-dependent walker population is present. Therefore, building on the massively parallel nature of the method, a computationally efficient implementation can offer an insight into the study of systems with severe sign problem. In this work, we designed the algorithm in C++, and it has been organized into separate interchangeable modules, meaning that it is relatively easy to modify the model system or the dimensionality.

Parallel computing provides a tool to carry out many execution processes simultaneously, and it allows simulations on multicore computers or computer clusters of interconnected nodes. Its two main branches are based on different underlying memory architecture model, namely shared or distributed memory. The results presented in this thesis were calculated almost exclusively on computational clusters, thus we implemented a distributed memory parallelization scheme using MVAPICH2 [131], a freely available Message Passing Interface (MPI).

In principle, DDQMC is a highly parallelizable algorithm. Most of the computational effort is spent on performing the spawn and clone/death steps for every simultaneously occupied configuration. The different configurations are independent from each other, therefore this stage can be implemented in parallel. The only step that requires communication between the MPI processes is annihilation, which may induce a significant computational overhead. Finding an efficient annihilation implementation is then crucial to the scalability of the algorithm.

5.3.1 Dual hashing

Rather than keeping in memory the whole configuration space, at each iteration step we only store a random sample of simultaneously occupied basis operators. Then these configurations need to be distributed between the parallel processes, where they are organized into a list consisting not only of the numerical representation of the configurations, but also of their respective signs, flags (initiator) and properties (connected operators, spawning probabilities, observable matrix elements). Since only the instantaneously populated configurations are kept, the annihilation has to be performed explicitly. Therefore, there are two bottlenecks that need to be considered: fast access to the stored information, and a uniform and traceable configuration distribution between the processes.

The core of our numerical design is based on a dual hashing procedure [126]. A hash function is a many-to-one mapping that is used to map data of arbitrary size (numerical representation of the configurations) to a fixed-size value. A hash function can be used to index the elements of a fixed-size table, the hash table. Hashing is a computationally efficient form of data access and storage, since it provides a small and nearly constant

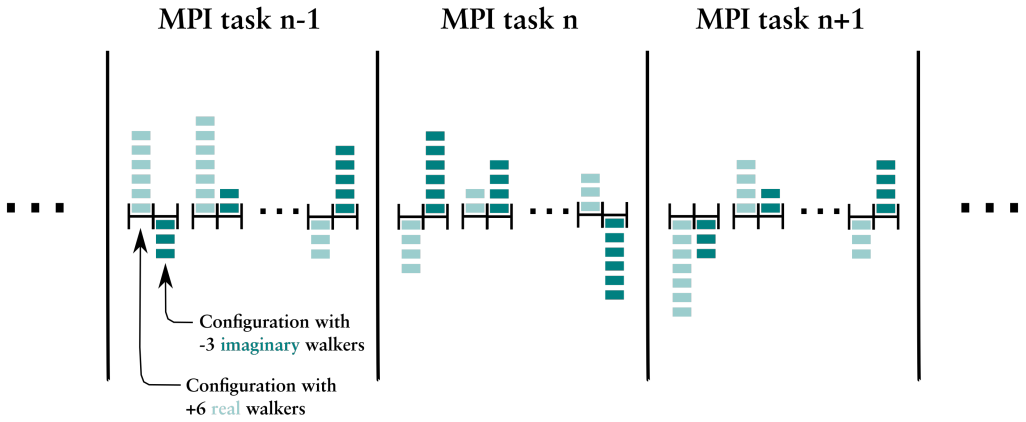

Figure 5.5 – Schematic representation of the walker distribution between the Slave processes [133].

retrieval time. We use this hashing procedure in two steps:

i on each parallel process, the instantaneously occupied configurations are stored in a hash table which provides — in average — $\mathcal{O}(1)$ complexity for search and insertion;

ii the populated configurations are distributed between the MPI processes (Figure 5.5) using a simple Merkle-Damgard type hash that has proven to be useful in mapping a bitset based basis representation (e.g. 01101... for a spin configuration) to a single integer (the process number) [126, 132].

5.3.2 Parallel program flow

Our parallel implementation is based on a "master-slave" paradigm, where the main process (the master) supervises the general program flow and gathers the statistics (e.q. total population, observables, shift, ...), while the sub-processes (the slaves) execute the DDQMC steps. The parallel program flow of the sub-processes is illustrated in Figure 5.6, and can be summarized in the following three steps:

I **DDQMC**: each process performs the spawn, clone/death steps on its list of configurations by using a sequential code;

II **Distribution**: when new walkers are spawned to a basis operator, its hash value is computed between the available sub-processes, giving its designated MPI task. Then its bitset representation with the number of spawned walkers are stored in a non-contiguous "spawned" array, separately from the main configuration list;

III **Communication**: after step II, all the MPI tasks have two lists of configurations. The first one contains the updated value of the old list after the clone/death step. These configurations are situated on the right process, since this step takes place locally. The second, "spawned" list consists of walkers that populates different configurations than their parents. These elements need to be sent to their respective MPI task, where they are sorted and merged with the main configuration table.

The sequential program flow of the master and slave processes is shown in Figure 5.7 and in Figure 5.8, and they consist of the previously introduced DDQMC steps including the initiator approach.

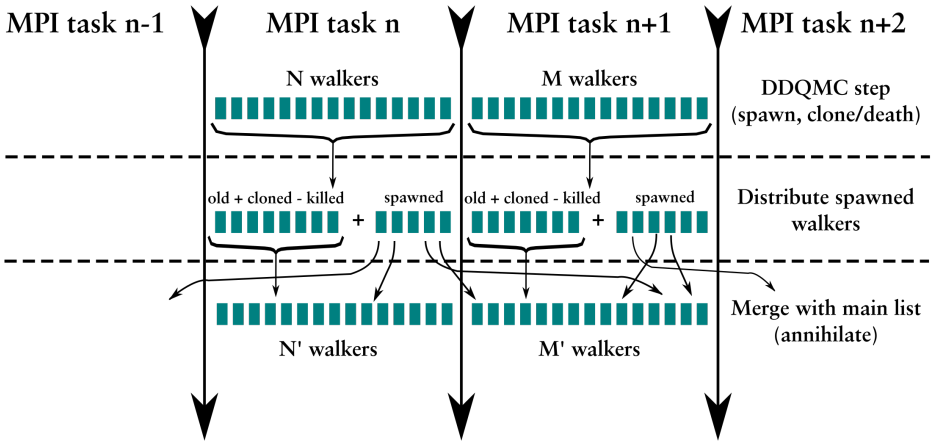

Figure 5.6 – The parallel program flow for Slave processes [133].

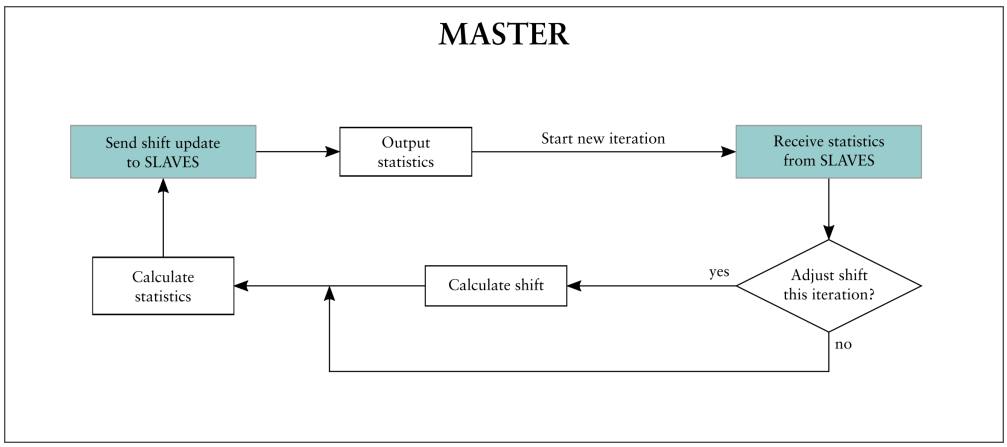

Figure 5.7 – The sequential program flow for the Master process.

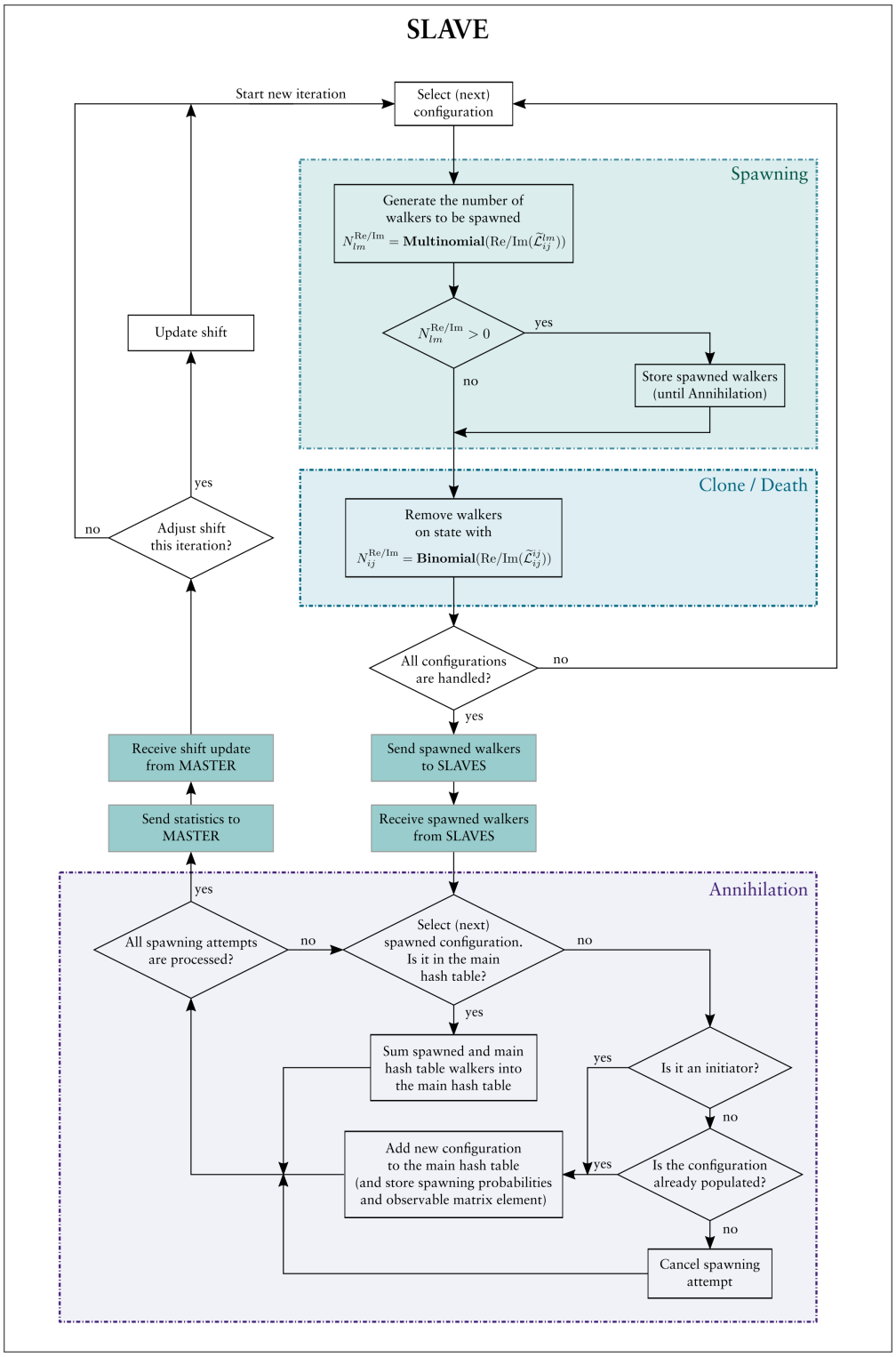

Figure 5.8 – The sequential program flow for the Slave processes.

5.3.3 Load balancing

The process of distributing the configurations such that the workload is equally balanced between the MPI processes is referred to as load balancing. Although the subprocesses only have to communicate with each other once during an iteration step, all of them need to wait for the slowest one to finish in order to sort the spawned population. This means that the computational efficiency will suffer if the configurations (and walkers) are not evenly distributed.

An advantage of implementing a hash function to distribute the configurations is the fact that as we increase the walker population, a larger part of the Hilbert space is simultaneously occupied. For a given number of parallel tasks this leads to an increased parallel performance, as the load is more evenly balanced between the processes. If we increase the system size, or tackle a model with severe sign problem, the increased parallelism of the code will grant to run more efficiently on larger computational resources [126].

Figure 5.9 is a schematic representation of the movement of configurations between the processes, and Figure 5.10 shows the parallel speed-up as the number of parallel processes is increased. The computational speed-up is defined as a number that measures the relative performance of two systems solving the same task. Here, we introduce it as the ratio between the execution times of arbitrary number of processes and a basic processing unit, 8 threads. As we expected, the scaling improves for a larger walker population, as more configurations are occupied and the load is more evenly balanced. In this case the computational effort is spent on the individual DDQMC steps and not on the parallel communication.

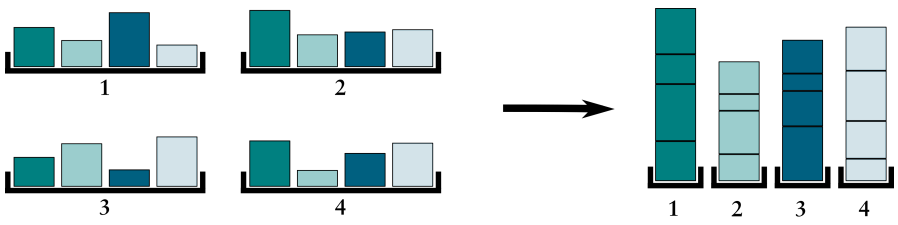

Figure 5.9 – Schematic representation of the movement of configurations between the processes.

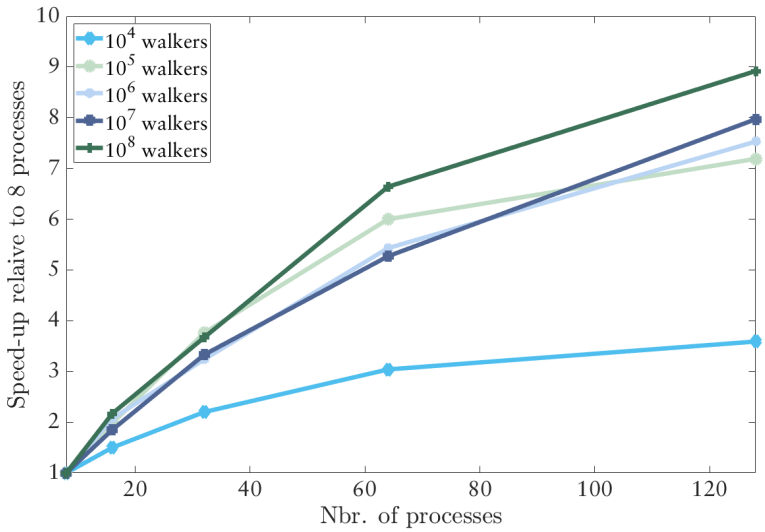

Figure 5.10 – The computational speed-up as the number of parallel processes is increased from 8 to 128. The scaling improves as the walker population is increased. The model is the 2 dimensional XYZ Heisenberg model with 16 sites.

5.4 Results*

In order to demonstrate the effectiveness of DDQMC, we simulated the two-dimensional spin-1/2 XYZ Heisenberg lattice in the presence of a dissipating channel which tends to relax each spin into the $|s_z = -1/2\rangle$ state (Chapter 1.3).

The model follows the Lindblad equation and the Hamiltonian is governed by ($\hbar = 1$)

$$\hat{H} = \sum_{\langle l,m \rangle} \left(J_x \hat{\sigma}_l^x \hat{\sigma}_m^x + J_y \hat{\sigma}_l^y \hat{\sigma}_m^y + J_z \hat{\sigma}_l^z \hat{\sigma}_m^z \right) \tag{5.19}$$

$$\frac{\mathrm{d}\hat{\rho}}{\mathrm{d}t} = -i[\hat{H}, \hat{\rho}] - \frac{\gamma}{2} \sum_k \left[\left\{ \hat{\sigma}_k^+ \hat{\sigma}_k^-, \hat{\rho} \right\} - 2\hat{\sigma}_k^- \hat{\rho} \hat{\sigma}_k^+ \right] \tag{5.20}$$

where $\hat{\sigma}_l^\alpha$ are the Pauli spin operators acting on the l-th spin, J_α are the coupling constants between nearest neighbour spins, γ is the dissipation rate, and $\hat{\sigma}_m^\pm = \hat{\sigma}_m^x \pm i\hat{\sigma}_m^y$.

In order to study the model we chose three different lattice sizes: 2×2, 3×3 and 4×4. The first two sizes are small enough to derive an exact numerical solution of the master equation in the steady state, thus allowing a direct check of the accuracy of our DDQMC results. In case of the 4×4 lattice, the magnetization is compared to those obtained by Monte Carlo wave function technique.

5.4.1 Magnetization in the steady-state

The steady-state magnetization per site is defined as

$$M_z = \frac{1}{N}\sum_{k=1}^{N} \mathrm{Tr}(\hat{\rho}\hat{\sigma}_k^z), \tag{5.21}$$

where N is the number of lattice sites. Figure 5.11(a) shows the magnetization of the 3×3 lattice as a function of the Monte Carlo iteration step with a diagonal population of 5×10^4 walkers. The exact solution obtained by directly solving the linear system is also plotted. Increasing the diagonal population to 2×10^6 reduces the statistical error as seen in the corresponding result in Fig. 5.11(b).

In Fig. 8.5, we present the magnetization per site M_z as a function of the normalized coupling constant J_y/γ for square lattices of different size. In order to assess the validity of the method we compare it with exact numerical solutions for the steady state in the case of the two smallest lattice sizes. For a 4×4 lattice, the exact numerical solution of the master equation is beyond computational reach, and we compare instead with results obtained by Monte Carlo Wave-function method [76], for which we accurately verified the convergence with respect to the number of quantum trajectories and to the time step size. Both the exact and Monte Carlo Wave-function calculations are in agreement with the one obtained by DDQMC.

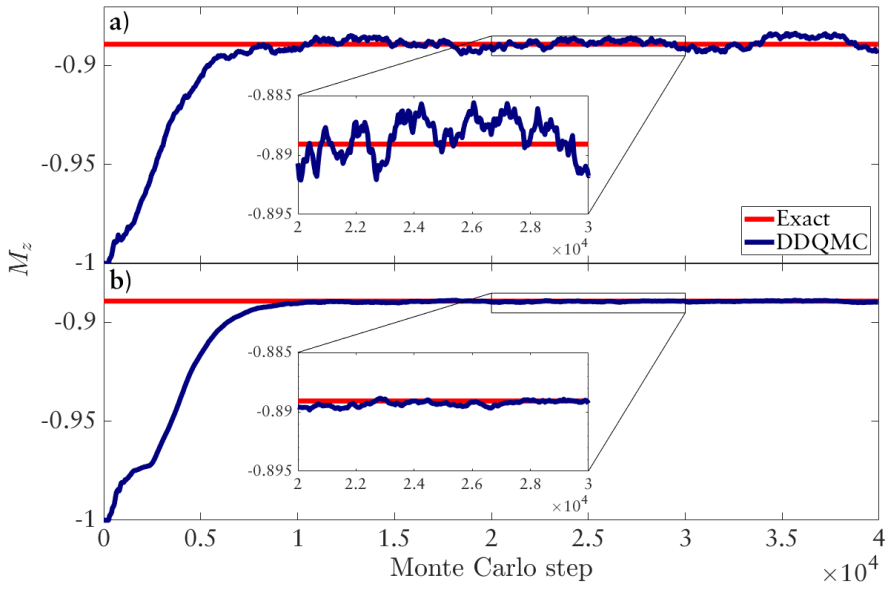

Figure 5.11 – The exact and the DDQMC magnetization values (in units of $\hbar$) for the 3×3 dissipative XYZ Heisenberg lattice with periodic boundary condition. The coupling parameters are $J_x/\gamma = 0.225$, $J_y/\gamma = 0.335$ and $J_z/\gamma = 0.25$. The diagonal population was limited to (a) 5×10^4 and (b) 2×10^6 walker.

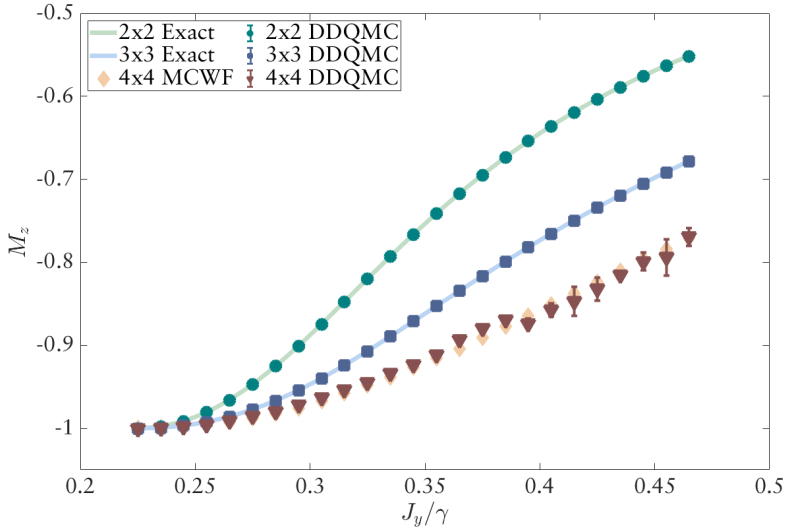

Figure 5.12 – The magnetization M_z per site (in units of $\hbar$) as a function of the normalized coupling parameter J_y/γ for different lattice sizes. The other coupling parameters are $J_x/\gamma = 0.225$ and $J_z/\gamma = 0.25$. The exact (2×2 and 3×3) and MCWF (4×4) results are plotted for comparison. Error bars, when not shown, are smaller than the symbol size.

5.4.2 Angularly-averaged susceptibility

Following the scheme presented in [16], we study the system in the presence of an applied magnetic field in the xy plane

$$\hat{H}_{ext} = \sum_i h(\cos(\theta)\hat{\sigma}_i^x + \sin(\theta)\hat{\sigma}_i^y)\,. \tag{5.22}$$

The linear response is then summarized in the 2×2 susceptibility tensor

$$\chi_{\alpha\beta} = \frac{\partial M_\alpha}{\partial h_\beta}\bigg|_{h=0}, \tag{5.23}$$

with $\alpha, \beta = x, y$. It is convenient to calculate one single quantity, the angularly-averaged susceptibility

$$\chi_{av} = \frac{1}{2\pi}\int_0^{2\pi} \mathrm{d}\theta \frac{\partial|\vec{M}(h,\theta)|}{\partial h}\bigg|_{h=0}, \tag{5.24}$$

where

$$\frac{\partial|\vec{M}(h,\theta)|}{\partial h}\bigg|_{h=0} = \left|\begin{pmatrix}\chi_{xx}\cos(\theta) + \chi_{xy}\sin(\theta)\\ \chi_{yx}\cos(\theta) + \chi_{yy}\sin(\theta)\end{pmatrix}\right|. \tag{5.25}$$

For a more complete derivation readers are referred to [16]. In Fig. 5.13, we present the angularly-averaged susceptibility χ_{av} as a function of the normalized coupling parameter J_y/γ for different lattice sizes. The magnetic susceptibility for the different lattice sizes exhibits a peak of increasing height which qualitatively corresponds to the results obtained in [16].

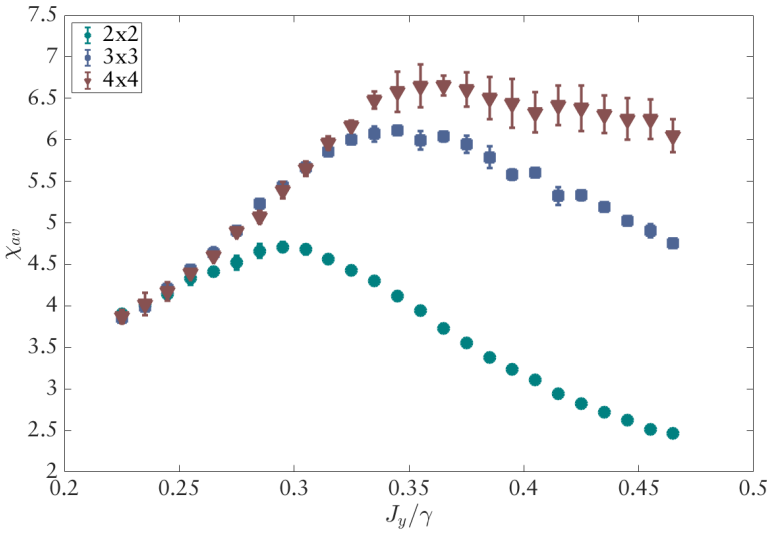

Figure 5.13 – The angle-averaged susceptibility χ_{av} per site as a function of the normalized coupling parameter J_y/γ for different lattice sizes. The other coupling parameters are $J_x/\gamma = 0.225$ and $J_z/\gamma = 0.25$. Each point on the plot was determined by 21 simulations, which corresponds to 525 calculations per lattice size. For each point, we considered $3 + 3 + 1$ values of the applied field (3 for each in-plane direction and 1 with no external field), and for each setting an extrapolation over three different initiator values was carried out.

5.5 Outlook*

The DDQMC method presented here, constitutes a basic PMC approach to the non-equilibrium steady state of open quantum systems. As such, it contains only the basic building blocks of the PMC method, and its effectiveness may be considerably improved by introducing any of the several tools that are common in other PMC schemes. Here, we describe as an outlook three such possible improvements.

The first possible improvement consists in the implementation of a mixed-estimator scheme, in analogy with the one used in projector and diffusion Monte Carlo to find the ground state of Hamiltonian systems [31, 107, 108]. Here, a possible mixed estimator strategy may consist in formally carrying out an exact real-time evolution, starting from a DDQMC sampled density matrix. More specifically, let us assume that at time t the current DDQMC sample of the density matrix is $\hat{\rho}(t)$. We can formally apply the exact time evolution for an additional time T and then evaluate the expectation value of an observable $\hat{O}$ at time $t+T$ as

$$\langle \hat{O} \rangle = \frac{\mathrm{Tr}(\hat{O} e^{\mathcal{L}T} \hat{\rho}(t))}{\mathrm{Tr}(e^{\mathcal{L}T} \hat{\rho}(t))} = \frac{\mathrm{Tr}(\hat{O} e^{\mathcal{L}T} \hat{\rho}(t))}{\mathrm{Tr}(\hat{\rho}(t))}\,, \tag{5.26}$$

where the second equality results from the trace preserving character of the time evolution. In the limit $T \to \infty$, Eq. (5.26) provides the steady-state expectation value independently of the actual value of the density matrix $\hat{\rho}(t)$, when assuming that a unique steady state exists. A mixed estimator strategy would then consist in building a "trial" observable $\hat{O}_T$ which can still be efficiently computed element-wise, and such that $\hat{O}_T \simeq \hat{O}_H(T) = \hat{O} e^{\mathcal{L}T}$. Here, $\hat{O}_H(t)$ represents the Heisenberg picture of the observable $\hat{O}$ and, for time-independent Liouvillian maps, it obeys the adjoint quantum master equation $\frac{d\hat{O}_H(t)}{dt} = \mathcal{L}^\dagger \hat{O}_H(t)$ [23]. Hence, the mixed estimator approach in the present case would require the knowledge of an approximate time dependence for $\hat{O}_H(t)$, which may be obtained, for instance, from a time-dependent variational principle [134] applied to a separable or short-range-correlated ansatz for the observable.

A second improvement would consist in using a "guiding density matrix" for the importance sampling. A natural choice for such a guiding density matrix would again be a variational ansatz, as the variational principle for the NESS is now well established, and some variational approaches have already been developed [27, 29, 135]. We wish to remark, that recently also a neural network based variational ansatz has been simultaneously introduced in [13, 33, 136, 137]. This is the second numerical approach that has been developed for the present thesis, and it will be reviewed in the following chapters.

A possible application of the method could be to sample the full time dynamics of the model, in analogy to the way in which finite temperature states are sampled in the finite-temperature density matrix formalism introduced in [121–123]. In fact, sampling the real-time dynamics would require averaging over several Monte Carlo instances in order to accumulate sufficient statistics for each given time. The steady state on the other hand can be sampled from one single Monte Carlo realization, once stationarity has been

reached in the simulation.

Finally, the present scheme is based on the Euler method (5.2) for the numerical solution of the time-dependent master equation. The Euler method is a first order method in the time step, and is only stable if Δt is chosen to be smaller than the inverse of the full spectral width of the master equation. In PMC, several approaches have been proposed to sample a higher-order discrete-timestep propagator [111], or even the exact one [138]. While a similar approach would be highly beneficial to FCIQMC and DDQMC, the question is still open, whether higher-order propagators may be efficiently sampled within the spawn-annihilation sampling protocol characterizing these Monte Carlo methods.

5.6 Conclusions*

In this chapter we have introduced a quantum Monte Carlo approach to open many-body quantum systems with Markovian system-bath coupling, called DDQMC. The method is based on the FCIQMC algorithm exploiting the analogy between the long-time dynamics of the Lindbladian master equation and the imaginary-time Schrödinger equation. DDQMC allows direct sampling of the steady state density matrices in any discrete basis set, and in all cases studied it has proven to be accurate.

DDQMC, as FCIQMC, uses an annihilation procedure which helps to alleviate the sign problem. The introduction of the initiator approach and importance sampling can lead to a significant improvement in the statistical accuracy and reduce the required walker population. The validity of the method was proven by investigating a dissipative phase transition on the two-dimensional Heisenberg-model. The defining feature of DDQMC is that it samples the whole density matrix and it does not introduce a truncation in Hilbert-space. Experience showed that the applicability of the code does not solely depend on the system size, but also on the correlations characterizing the steady state. The application presented in this work is a proof of principle, demonstrating the possibility to stochastically sample the Lindblad equation in a finite difference approximation. DDQMC holds promise as a powerful tool in the study of open quantum systems.

In the next chapters we review a second numerical approach that has been developed for the present thesis. For this, we will first review the variational quantum Monte Carlo method, and the foundations of machine learning. Finally, we introduce the concept of a neural network ansatz for the density matrix, and present a variational approach to simulate the real-time dynamics and the steady state of open quantum systems.

Variational quantum Monte Carlo

PART III

CHAPTER 6

Variational quantum Monte Carlo

Quantum Monte Carlo methods encompass a large family of numerical approaches that are versatile and effective tools to handle quantum many-body systems. The two most popular classes for ground state calculations are the Projector Monte Carlo techniques (Chapter 3) and the Variational Monte Carlo (VMC) method. VMC relies on the variational principle to approximate the ground state of quantum many-body systems by optimizing the parameters of a variational wave function which is used to sample quantum mechanical expectation values. The numerous flavours of VMC can differ in the optimization technique and the form of the variational ansatz. Here we discuss the underlying ideas of VMC and a particular optimization method called Stochastic Reconfiguration. After establishing the principles of a variational algorithm, we introduce a specially successful and versatile wave function ansatz in Chapter 8.

6.1 The foundation of Variational Monte Carlo

Let us start by assuming once again that the Hilbert space of a system is spanned by the computational basis $|\boldsymbol{\sigma}\rangle$, where $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_N)$ labels the states of N d.o.f that compose the system. Here and in what follows we will consider binary local degrees of freedom, with $\sigma_i = \{-1, 1\}$, which applies to the broad class of interacting spin-1/2 or qubit models. This is not mathematically necessary, and one can translate all equations for a continuous basis by replacing the sums with integrals. Using the completeness of the basis

$$\sum_{\boldsymbol{\sigma}} |\boldsymbol{\sigma}\rangle\langle\boldsymbol{\sigma}| = \mathbb{I}\,, \tag{6.1}$$

the wave function is expressed as

$$|\psi\rangle = \sum_{\boldsymbol{\sigma}} |\boldsymbol{\sigma}\rangle\langle\boldsymbol{\sigma}|\psi\rangle = \sum_{\boldsymbol{\sigma}} \psi(\boldsymbol{\sigma})|\boldsymbol{\sigma}\rangle\,. \tag{6.2}$$

Thus, we can recast the expectation value of any quantum mechanical observable into the form of

$$\begin{aligned}\langle\hat{\mathcal{O}}\rangle &= \frac{\langle\psi|\hat{\mathcal{O}}|\psi\rangle}{\langle\psi|\psi\rangle} = \frac{\sum_{\boldsymbol{\sigma}}\langle\psi|\boldsymbol{\sigma}\rangle\langle\boldsymbol{\sigma}|\hat{\mathcal{O}}|\psi\rangle}{\sum_{\boldsymbol{\sigma}}\langle\psi|\boldsymbol{\sigma}\rangle\langle\boldsymbol{\sigma}|\psi\rangle} \\ &= \frac{\sum_{\boldsymbol{\sigma}}|\langle\boldsymbol{\sigma}|\psi\rangle|^2\frac{\langle\boldsymbol{\sigma}|\hat{\mathcal{O}}|\psi\rangle}{\langle\boldsymbol{\sigma}|\psi\rangle}}{\sum_{\boldsymbol{\sigma}}|\langle\boldsymbol{\sigma}|\psi\rangle|^2} = \frac{\sum_{\boldsymbol{\sigma}}|\langle\boldsymbol{\sigma}|\psi\rangle|^2\mathcal{O}_{\mathrm{loc}}(\boldsymbol{\sigma})}{\sum_{\boldsymbol{\sigma}}|\langle\boldsymbol{\sigma}|\psi\rangle|^2}\,,\end{aligned} \tag{6.3}$$

where $p(\boldsymbol{\sigma}) \sim |\langle\boldsymbol{\sigma}|\psi\rangle|^2$ is the probability of finding the system in configuration $\boldsymbol{\sigma}$ for a wave function $|\psi\rangle$, and $\mathcal{O}_{\mathrm{loc}}(\boldsymbol{\sigma})$ is called the local value of the observable for state $\boldsymbol{\sigma}$. However, the number of basis elements increases exponentially with the system size, which makes it computationally impractical to calculate the exact sum in eq. (6.3). Instead, one can approximate the sum over all the basis elements with a sum over a randomly selected $\mathcal{S} = \{\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, ..., \boldsymbol{\sigma}_{N_{MH}}\}$ set of them as

$$\langle\hat{\mathcal{O}}\rangle \approx \frac{\sum_{\boldsymbol{\sigma}\in\mathcal{S}}|\langle\boldsymbol{\sigma}|\psi\rangle|^2\mathcal{O}_{\mathrm{loc}}(\boldsymbol{\sigma})}{\sum_{\boldsymbol{\sigma}\in\mathcal{S}}|\langle\boldsymbol{\sigma}|\psi\rangle|^2},. \tag{6.4}$$

In practice, this means to generate N_{MH} number of $\boldsymbol{\sigma}_i$ configurations randomly, then calculate the local value of the observable $\mathcal{O}_{\mathrm{loc}}(\boldsymbol{\sigma}_i)$ and weight it with the probability $|\langle\boldsymbol{\sigma}_i|\psi\rangle|^2$. However, this approach proves to be computationally ineffective, since there is no guarantee to find the configurations which contribute with a large weight. Rather than choosing the basis elements completely randomly, one needs to generate the configurations according to their probability and weigh them equally. From a mathematical point of view, this means to generate a set $\mathcal{S}_p = \{\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, ..., \boldsymbol{\sigma}_{N_{MH}}\}$ of N_{MH} configurations that are distributed according to $p(\boldsymbol{\sigma}_i) \sim |\langle\boldsymbol{\sigma}_i|\psi\rangle|^2$ as

$$\langle\hat{\mathcal{O}}\rangle \approx \frac{1}{N}\sum_{\boldsymbol{\sigma}\in\mathcal{S}_p}\mathcal{O}_{\mathrm{loc}}(\boldsymbol{\sigma})\,. \tag{6.5}$$

This approach is called the Markov Chain Monte Carlo or Metropolis-Hastings algorithm [139]. We will introduce this method in Section 6.2, but for now let us assume that we know how to generate $\mathcal{S}_p$.

For now, it was unclear how we can use eq. (6.3) to obtain the ground state, however, let us note, that the form of $\langle\hat{\mathcal{O}}\rangle$ is accurate for any physical observable, including the Hamiltonian $\hat{H}$. If we replace the wave function $|\psi\rangle$ with a variational ansatz such that

$$|\psi\rangle \quad\longrightarrow\quad |\psi_\chi\rangle\,, \tag{6.6}$$

and $\chi = (\chi_1, \chi_2, ..., \chi_{N_p})$ are a set of variational parameters, it is then straightforward to apply the Rayleigh-Ritz principle [140] to find the ground state of the system. The theorem states that the expectation value of the energy is bounded from below by the ground state energy. For a Hamiltonian with eigenvectors $|\phi_i\rangle$ and eigenvalues $E_0 \leq E_1 \leq ... \leq E_N$

$$E_\chi = \frac{\langle\psi_\chi|\hat{H}|\psi_\chi\rangle}{\langle\psi_\chi|\psi_\chi\rangle} = \frac{\sum_{i=0}^{N} E_i\langle\psi_\chi|\phi_i\rangle\langle\phi_i|\psi_\chi\rangle}{\langle\psi_\chi|\psi_\chi\rangle} \geq \frac{\sum_{i=0}^{N} E_0\langle\psi_\chi|\phi_i\rangle\langle\phi_i|\psi_\chi\rangle}{\langle\psi_\chi|\psi_\chi\rangle} = E_0\,, \tag{6.7}$$

so by changing the parameters χ we will always obtain a variational energy that is larger than equal to the ground state value, and there is an optimal set of parameters that minimizes the expectation value of the energy.

Now we have all the ingredients to assemble a general algorithm for the variational quantum Monte Carlo method by iterating the following steps

i *Sampling*: we apply an optimization method in order to minimize E_χ. Any quantity that is needed for the algorithm can be sampled using eq. (6.3).

ii *Feedback*: obtain a feedback from E_χ. In practice, calculate the new parameters that will provide a lower variational energy.

iii *Adjust*: adapt the value of the variational parameters to the new values.

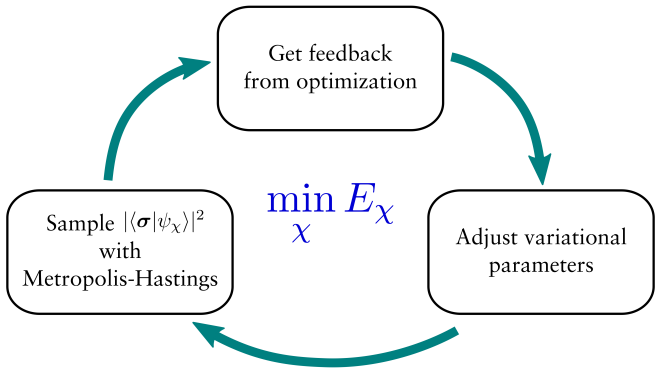

Figure 6.1 – Schematic illustration of the general VMC algorithm.

From a mathematical point of view, there are several optimization algorithms that could be used to adjust the variational parameters, however, a method called Stochastic Reconfiguration has proven to be especially successful for quantum systems. The details of this approach will be introduced in Section 6.3.

Clearly, the ground state energy estimation can only be as good as the variational ansatz, hence the success of the variational approach will highly depend on the form we choose for the wave function. This is possibly the greatest limitation of VMC, as it makes it particularly easy to miss phenomena that we have not expected. We will review an exceptionally rewarding and system independent variational ansatz in Section 8.

An important feature of variational Monte Carlo is a property called "zero variance" [141]. Let us assume that the variational wave function $|\psi_\chi\rangle$ coincides with an eigenstate $|\psi_n\rangle$ of the Hamiltonian, which is not necessarily the ground state. Then the local value of the energy reads as

$$E_{\mathrm{loc}}(\boldsymbol{\sigma}) = \frac{\langle \boldsymbol{\sigma}|\hat{H}|\psi_n\rangle}{\langle \boldsymbol{\sigma}|\psi_n\rangle} = E_n \frac{\langle \boldsymbol{\sigma}|\psi_n\rangle}{\langle \boldsymbol{\sigma}|\psi_n\rangle} = E_n \,. \tag{6.8}$$

Thus, the random variable $E_{\mathrm{loc}}(\boldsymbol{\sigma})$ is independent of the configuration $\boldsymbol{\sigma}$, which implies that its mean value is equal to E_n and that its variance is zero. One can also show [141], that the variance of $E_{\mathrm{loc}}(\boldsymbol{\sigma})$ is exactly equal to the variance of the Hamiltonian on a variational ansatz $|\psi_\chi\rangle$. Therefore, the smaller the variance of $E_{\mathrm{loc}}(\boldsymbol{\sigma})$ is, the closer the variational wave function is to the exact ground state. This is a particularly useful property, as in practice one never knows the value of the ground state energy a priori, hence the accuracy of the simulation outcome is unclear. Monitoring the variance then provides a way to measure the quality of the result.

6.2 Stochastic sampling using Markov-chains

In order to evaluate eq. (6.5), we need to generate a set of many body configurations $\mathcal{S}_p = \{x_1, x_2, ..., x_{N_{MH}}\}$ that are distributed according to $p(x) \sim |\langle x|\rho_\chi\rangle|^2$. For this we adopt the Metropolis-Hastings (or Markov Chain Monte Carlo, MCMC) algorithm [139, 142].

The method can be used to sample any distribution function $p(x)$ regardless of analytical complexity or dimension as long as it is normalizable with

$$\int_x p(x')dx' < \infty \,. \tag{6.9}$$

MCMC has been motivated by an analogy with the behaviour of systems in statistical mechanics that approach an equilibrium whose state is independent of the kinetics of the system. In the present context, by system we simply mean a point x in the configuration

space that fully describes the state of a physical model, e.g. $|x\rangle = |\boldsymbol{\sigma}\rangle$. By kinetics, we mean a stochastic process that governs the transition of the system such that

$$f(x_{n+1}, x_n) = K(x_{n+1}|x_n)p(x_n)\,, \tag{6.10}$$

where $f(x_{n+1}, x_n)$ is the probability of moving from x_n to x_{n+1} is expressed as the a priori chance of finding the system in x_n (i.e. $p(x_n)$) multiplied by the conditional probability $K(x_{n+1}|x_n)$ that it will transit to x_{n+1} from x_n. In general, the non-deterministic operation that takes the system from one state to the other is called a Markov process, and the series of states the system passes through by the repeated application of the same Markov process forms a Markov-chain. The Markov processes characterized by eq. (6.10) are memoryless in the sense that the transition probabilities only depend on the current state of the system rather than the previous moves.

Now suppose that we generate N_{MH} steps in the random walk as $x_1 \to x_2 \to \cdots \to x_{N_{MH}}$. Each of the x_n is a random variable and has an associated probability q_n, where q_n can be any distribution of x. It can be shown [142], that if the kinetics K is such that the resulting Markov-chain satisfies *detailed balance* and *ergodicity*, then q_n converges to $p(x)$ as

$$\lim_{n\to\infty} q_n(x) = p(x)\,, \tag{6.11}$$

and we are able to sample the sum (6.3) with a set of $\mathcal{S}_p = \{\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, ..., \boldsymbol{\sigma}_{N_{MH}}\}$ configurations that are distributed according to $p(x) \sim |\psi_\chi(\boldsymbol{\sigma})|^2$.

As we shall discuss in detail, in order to converge to a stationary distribution, it is sufficient, but not necessary that the random walk satisfies the so-called detailed balance. That is, quite simply, that the system is to move with the same probability from a specific point x_n into x_{n+1} as to move exactly in the reverse direction. From eq. (6.10), this corresponds to

$$K(x_n|x_{n+1})p(x_{n+1}) = K(x_{n+1}|x_n)p(x_n)\,. \tag{6.12}$$

While treating physical systems, one generally has the task to find $p(x)$ while K is known, in MCMC one needs to define a convenient kinetics that satisfies the detailed balance and equilibrates the system to the stationary distribution $p(x)$. In order to do so, Metropolis et al. [139] have suggested a simple and elegant scheme, where transitions are proposed from essentially any distribution $T(x_{n+1}|x_n)$, as long as it ensures ergodicity, and in order to define a Markov chain that yields to detailed balance they are accepted only with a

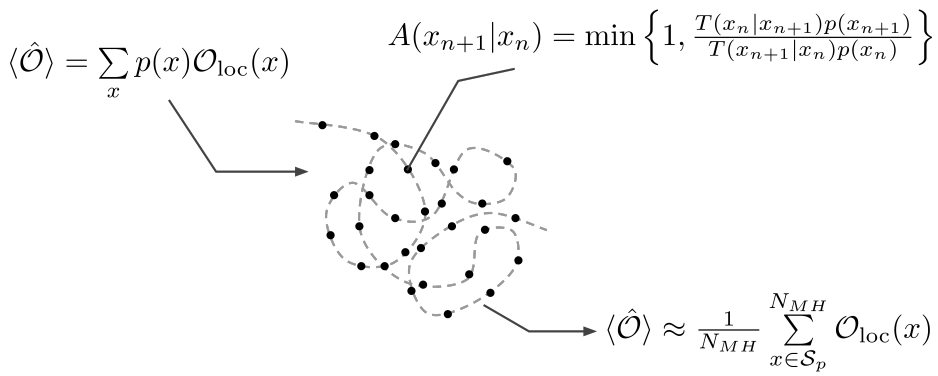

Figure 6.2 – Schematic illustration of the random walk generated by the MCMC algorithm.

probability

$$A(x_{n+1}|x_n) = \min\left\{1, \frac{T(x_n|x_{n+1})p(x_{n+1})}{T(x_{n+1}|x_n)p(x_n)}\right\}. \tag{6.13}$$

Then the resulting conditional probability reads as

$$K(x_{n+1}|x_n) = T(x_{n+1}|x_n)A(x_{n+1}|x_n)\,. \tag{6.14}$$

Although the proof will be omitted [142], we would like to mention that the Markov-chain generated by eq. (6.13) and (6.14) is truly ergodic and satisfies detailed balance, as well as that these two properties are sufficient conditions that yield to a unique limiting distribution.

Finally, if x_n is the configuration at step n, the MCMC algorithm can be summarized as follows:

i A move is proposed by generating a new configuration $\tilde{x}$ with probability $T(\tilde{x}|x_n)$.

ii The transition is accepted with the probability $A(\tilde{x}|x_n)$. In practice, this means that a random number $r \in [0,1]$ is drawn, and if $r < A(\tilde{x}|x_n)$, then the next x_{n+1} is taken to be equal to $\tilde{x}$, otherwise it is kept unchanged as $x_{n+1} = x_n$.

A significant advantage introduced by MCMC is the fact that one only needs to know the desired probability distribution $p(x)$ up to a normalization constant, since the computation of $A(x_{n+1}|x_n)$ only requires the knowledge of the ratio $p(x_{n+1})/p(x_n)$ in eq. (6.13). This allows us to avoid the computationally sometimes prohibitive calculation of the nor-

malization factor, i.e. in eq. (6.3) one can omit $\langle\psi|\psi\rangle$.

A small technicality that is worth noting, is that in practice we initialize the Markov-chain in a specific configuration y, which sets the initial distribution to $q_0(x) = \delta_{x,y}$. A practical implication of this is that one should always discard a number of iterations at the beginning of the MCMC algorithm, since the initial delta distribution needs time to converge to the stationary $p(x)$. This phase of the random walk is called equilibration period. By choosing the transition probability $T(x_{n+1}|x_n)$ such that the overlap between $p(x_{n+1})$ and $p(x_n)$ is large, one can increase the acceptance probability of consecutive moves and find an optimal equilibration period.

We need to mention however, that a having a high acceptance probability is a double edged sword. While a large $A(x_{n+1}|x_n)$ can grant a shorter equilibration phase, having similar configurations as adjacent steps also introduces a correlation between the local observable measurements. The mean value of

$$\langle\hat{\mathcal{O}}\rangle \approx \frac{1}{N_{MH}} \sum_{x\in\mathcal{S}_p} \mathcal{O}_{\mathrm{loc}}(x) \tag{6.15}$$

is equal to the expectation value of the observable, since x are distributed according to $p(x)$, but the variance $\mathrm{Var}(\langle\hat{\mathcal{O}}\rangle)$ may be underestimated [141]. This can be solved by either taking a measurement only every few Monte Carlo steps, or by using the more sophisticated technique called binning [141]. Then one divides the Markov-chain of length N_{MH} into B so called bins with N_B configurations and reorders the sum in eq. (6.15) as

$$\langle\hat{\mathcal{O}}\rangle = \frac{1}{B}\sum_{b=1}^{B}\langle\hat{\mathcal{O}}\rangle_b \qquad \text{with} \quad \langle\hat{\mathcal{O}}\rangle_b = \frac{1}{N_B}\sum_{x\in\mathcal{S}_p}^{N_B} \mathcal{O}_{\mathrm{loc}}(x)\,. \tag{6.16}$$

This operation does not alter the expectation value, however, the bin means $\langle\hat{\mathcal{O}}\rangle_b$ will no longer be correlated and one can use their variance to compute

$$\mathrm{Var}(\langle\hat{\mathcal{O}}\rangle) = \frac{1}{B(B-1)}\sum_{b=1}^{B}\left(\langle\hat{\mathcal{O}}\rangle_b - \langle\hat{\mathcal{O}}\rangle\right)^2\,. \tag{6.17}$$

Since eq. (6.17) only holds for non-correlated variables, one needs to select the bin length N_B such that the averages are effectively independent random variables. This can be done by introducing the concept of correlation time and correlation function, for which readers are referred to [141].

6.3 Minimization with Stochastic Reconfiguration

In Section 6.1 we have introduced the fundamental ideas behind the variational Monte Carlo simulations of quantum many-body systems, and we have seen that the approach relies on the Rayleigh-Ritz principle and the variational optimization of a trial wave function. In principle, any numerical optimization method could be used to find the approximate ground state, but here we review a particular approach called Stochastic Reconfiguration (SR) that has proven to be exceptionally effective in the context of VMC [141, 143, 144].

Let $|\psi_\chi\rangle$ be the variational wave function that depends on a set of N_p parameters $\chi = (\chi_1, \chi_2, ..., \chi_{N_p})$. In order to determine the correct parameter change that converges to the ground state, first we study the change in the wave function after a small perturbation in the parameters. By setting the wave function with parameters $\tilde{\chi}$ as the first order Taylor expansion of $|\psi_\chi\rangle$, it reads as

$$|\psi_{\tilde{\chi}}\rangle = |\psi_\chi\rangle + \sum_{k=1}^{N_p} \delta\chi_k \frac{\partial|\psi_\chi\rangle}{\partial\chi_k}\,. \tag{6.18}$$

Eq. (6.18) can be rewritten using the completeness of the basis $|\boldsymbol{\sigma}\rangle$ such that

$$\begin{aligned} |\psi_{\tilde{\chi}}\rangle &= |\psi_\chi\rangle + \sum_{k=1}^{N_p} \delta\chi_k \frac{\partial}{\partial\chi_k} \sum_{\boldsymbol{\sigma}} |\boldsymbol{\sigma}\rangle\langle\boldsymbol{\sigma}|\psi_\chi\rangle \\ &= |\psi_\chi\rangle + \sum_{k=1}^{N_p} \delta\chi_k \sum_{\boldsymbol{\sigma}} \frac{\partial\langle\boldsymbol{\sigma}|\psi_\chi\rangle}{\partial\chi_k}|\boldsymbol{\sigma}\rangle \\ &= |\psi_\chi\rangle + \sum_{k=1}^{N_p} \delta\chi_k \sum_{\boldsymbol{\sigma}} \frac{\partial \ln\langle\boldsymbol{\sigma}|\psi_\chi\rangle}{\partial\chi_k}|\boldsymbol{\sigma}\rangle\langle\boldsymbol{\sigma}|\psi_\chi\rangle\,. \end{aligned} \tag{6.19}$$

We can define the logarithmic derivatives o_k of the wave function for any configuration $|\boldsymbol{\sigma}\rangle$ as the variational derivatives with respect to the k-th parameter

$$o_k(\boldsymbol{\sigma}) = \frac{1}{\psi_\chi(\boldsymbol{\sigma})}\frac{\partial\psi_\chi(\boldsymbol{\sigma})}{\partial\chi_k}\,, \tag{6.20}$$

which form the diagonal operators

$$O_k = \begin{cases} \mathbb{I} & \text{for } k = 0 \\ \sum_{\boldsymbol{\sigma}} o_k(\boldsymbol{\sigma})|\boldsymbol{\sigma}\rangle\langle\boldsymbol{\sigma}| & \text{for } k \neq 0\,. \end{cases} \tag{6.21}$$

In linear approximation the new wave function can then be written in the compact form of

$$|\psi_{\tilde{\chi}}\rangle = \sum_{k=0}^{N_p} \delta\chi_k O_k |\psi_\chi\rangle \,, \tag{6.22}$$

which can be considered as an expansion over the basis spanned by $\{O_k|\psi_\chi\rangle\}$. Note, for the equation above we used $\delta\chi_0 = 1$, which is not necessarily true, and in that case the variation of the parameters must be rescaled as

$$\delta\chi_k \longrightarrow \frac{\delta\chi_k}{\delta\chi_0} \,. \tag{6.23}$$

Now our purpose is to exploit the linear approximation in eq. (6.22) by designing an iterative scheme such that changing the parameters from χ to $\tilde{\chi}$ yields to a lower energy as $E_{\tilde{\chi}} < E_\chi$ and a wave function that approaches the true ground state.

One possibility to find a state $|\psi'\rangle$ with an energy lower than $|\psi_\chi\rangle$ is to follow the Hamiltonian dynamics by applying the projector $(\Lambda - \hat{H})$ as

$$|\psi'\rangle = (\Lambda - \hat{H})|\psi_\chi\rangle \,, \tag{6.24}$$

where Λ is a large enough energy shift which ensures convergence. In the repeated application of this projector one can recognize the power method [109]. While for a continuous system, where the energy is unbounded from above, the value of Λ should be infinite, in a lattice system with a finite basis also Λ will be limited, as the Hamiltonian is bounded from above. Ensuring that $|\psi_{\tilde{\chi}}\rangle$ has lower energy than $|\psi_\chi\rangle$ corresponds to equating eq. (6.22) and (6.24) as $|\psi'\rangle = |\psi_{\tilde{\chi}}\rangle$ in a space spanned by $\{O_k|\psi_\chi\rangle\}$. Projecting onto the k'-th element we get

$$\langle\psi_\chi|O_{k'}(\Lambda - \hat{H})|\psi_\chi\rangle = \sum_{k=0}^{N_p} \delta\chi_k \langle\psi_\chi|O_{k'}O_k|\psi_\chi\rangle \,. \tag{6.25}$$

Following eq. (6.25), the quantities $\delta\chi_k$ correspond to the variation of the wave function parameters that lower the energy E_χ. In order to rewrite eq. (6.25) in terms of expectation values, we divide both side by $\langle\psi_\chi|\psi_\chi\rangle$

$$\Lambda\langle O_{k'}\rangle - \langle O_{k'}\hat{H}\rangle = \sum_{k=0}^{N_p} \delta\chi_k \langle O_{k'}O_k\rangle \,. \tag{6.26}$$

This is a system of $(N_p + 1)$ linear equations that can be reduced to N_p. Indeed, solving eq. (6.26) for $k' = 0$ and using that $O_o = \mathbb{I}$ we get

$$\delta\chi_0 = \Lambda - \langle \hat{H} \rangle - \sum_{k=1}^{N_p} \delta\chi_k \langle O_k \rangle \,. \tag{6.27}$$

Substituting eq. (6.27) into (6.26) we obtain the final equations that determine the variations $\delta\chi_k$ that lower the energy of the wave function as

$$\langle O_{k'} \rangle \langle \hat{H} \rangle - \langle O_{k'} \hat{H} \rangle = \sum_{k=0}^{N_p} \delta\chi_k \left(\langle O_{k'} O_k \rangle - \langle O_{k'} \rangle \langle O_k \rangle \right) \,. \tag{6.28}$$

This can be written in a convenient matrix form as

$$\vec{F} = \hat{S} \delta\vec{\chi} \,, \tag{6.29}$$

where we introduced the generalized forces F_k and the covariance matrix $S_{kk'}$ for a given set of variational parameters such that

$$\begin{aligned} F_k &= \langle O_k \rangle \langle \hat{H} \rangle - \langle O_k \hat{H} \rangle \\ S_{kk'} &= \langle O_k O_{k'} \rangle - \langle O_k \rangle \langle O_{k'} \rangle \,. \end{aligned} \tag{6.30}$$

After solving eq. (6.29), the variational parameters are updated following

$$\tilde{\chi}_k = \chi_k + \frac{\delta\chi_k}{\delta\chi_0} = \chi_k + \nu \delta\chi_k \,, \tag{6.31}$$

where ν is small enough to guarantee convergence. Note, that the variation of the parameters depends on Λ only indirectly, through the variable $\delta\chi_0/\nu$. It has been shown [141], that the choice of ν can be controlled during the simulation so that it ensures steady and stable convergence.

6.4 Steepest descent vs Stochastic Reconfiguration

In a standard steepest descent (SD) method the expectation value of the energy E_χ is minimized by changing the parameters χ according to the variational derivatives of E_χ. Notice, that the generalized forces F_k correspond to the same quantity such that

$$F_k = -\frac{\partial E_\chi}{\partial \chi_k} = -\frac{\partial}{\partial \chi_k}\frac{\langle \psi_\chi | \hat{H} | \psi_\chi \rangle}{\langle \psi_\chi | \psi_\chi \rangle}\,. \tag{6.32}$$

Therefore, the steepest descent method is recovered from the stochastic reconfiguration, whenever one replaces the covariance matrix with the identity. A basic requirement for every stable iterative method is that the updated parameters $\tilde{\chi}$ are close to the previous χ ones according to some metric. SR has proven to be particularly successful in variational Monte Carlo due to a fundamental difference that is related to the definition of this metric [145]. While for SD the new parameters are close to the old ones corresponding to the Cartesian distance

$$\Delta_C = \sum_k |\tilde{\chi}_k - \chi_k|^2\,, \tag{6.33}$$

SR works ideally in the Hilbert space metric of the wave function $|\psi_\chi\rangle$ such that

$$\Delta_H = \sum_{k,k'} S_{k,k'}(\tilde{\chi}_k - \chi_k)(\tilde{\chi}_{k'} - \chi_{k'}) = \frac{||\psi_{\tilde{\chi}}\rangle - |\psi_\chi\rangle|^2}{||\psi_\chi\rangle|^2}\,. \tag{6.34}$$

Indeed, the diagonal elements of the covariance matrix $\hat{S}$ give direct information about the fluctuations of the variational derivatives O_k. Multiplying the forces with the inverse of the fluctuations allows the optimization to move along a line where the variance of the corresponding operators O_k are small. Moreover, the presence of the off-diagonal parameters take into account the correlations between the parameters χ. Therefore, the advantage of SR compared to SD comes from the fact, that a small change in the variational parameters can correspond to a big change in the wave function.

CHAPTER 7

Fundamentals of machine learning

The rise of machine learning (ML) techniques in industrial applications has quickly been followed by a significant interest in physics and in fundamental research in general. A groundbreaking progress in the numerical simulation of both the ground state and the dynamics of closed quantum systems has recently been made with the introduction of the neural network variational ansatz [146–151], which efficiently represents highly correlated quantum states and whose parameters are easily optimized by means of the variational Monte Carlo method. In Chapter 8 we use this technique as a stepping stone to introduce a novel approach for the simulation of open quantum systems, and for the purpose of this thesis we now briefly explain some fundamental concepts in machine learning. However, a full account of the theoretical foundation and broad range of applicability falls beyond the scope of this thesis, and for further reading we recommend [152–157].

Machine learning has been loosely defined by Samuel [158] as the field of study that gives computers the ability to learn without being explicitly programmed. It can be classified into broad categories based on whether they learn with or without human supervision, whether they can learn on the fly, or whether they are data or model based. In the following sections we will review the first type of classification and introduce some fundamental concepts of machine learning.

The machine learning algorithms fall roughly into three basic paradigms: supervised, unsupervised and reinforcement learning. Supervised learning requires a feedback whether a prediction is right or wrong, whereas an unsupervised algorithm does not need a user input, it tries to categorize the data based on its underlying structure. Reinforcement learning lies somewhere in between, in the sense that it does involve a feedback but not for each input or state. Figure 7.1 shows a schematic representation of the ML classification with their respective application areas.

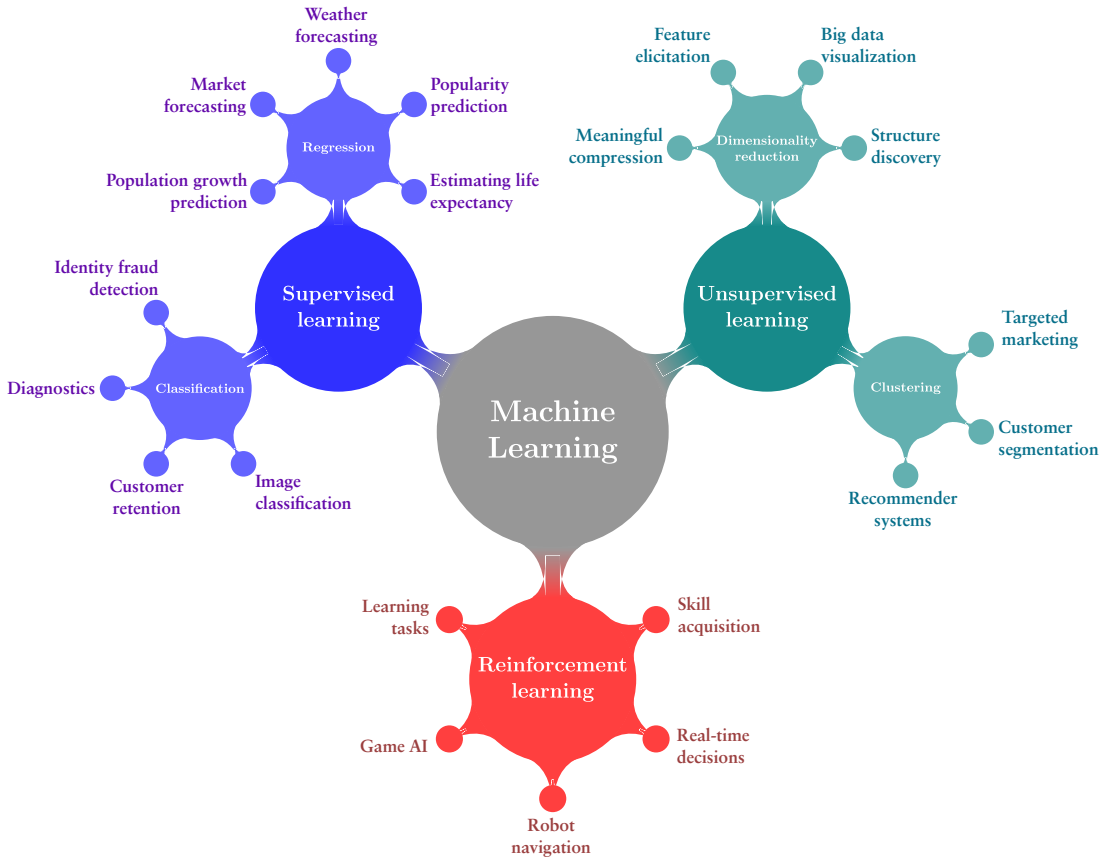

Figure 7.1 – Schematic illustration of the three basic paradigms of machine learning with their respective field of application.

7.1 Supervised machine learning

Supervised machine learning is the simplest paradigmatic model to understand, and it entails making predictions on unseen data based on existing observations. There are two levels of understanding of this model. First, we can consider it as fitting a function $y = f(\mathbf{X})$ for inter- or extrapolation, or second, from a probabilistic point of view, the purpose of the model is to learn the conditional probability $p(y|\mathbf{X})$. We call $\mathcal{D} = (\mathbf{X}, y)$ the dataset, with $\mathbf{X}$ being the data, and y the label. The dataset consists of N samples such that $\mathbf{X} = \{\mathbf{x}_1, ..., \mathbf{x}_N\}$ and $y = \{y_1, ..., y_N\}$, where $\mathbf{x}_\mu \in \mathbb{R}^d$. For a continuous label the task is called regression, while for binary labels it is classification. A sample $\mathbf{x}_\mu$ is for instance the pixels of an image reshaped into a vector and y_μ determines what is seen on the picture. Typical applications of supervised learning algorithms include – among many – image classification, weather forecasting and identity fraud detection.

The function f is expressed in terms of a set of parameters θ, leading to f_θ, and the accuracy of the model is quantified using a cost function $\mathcal{C}_\theta$. The probabilistic interpre-

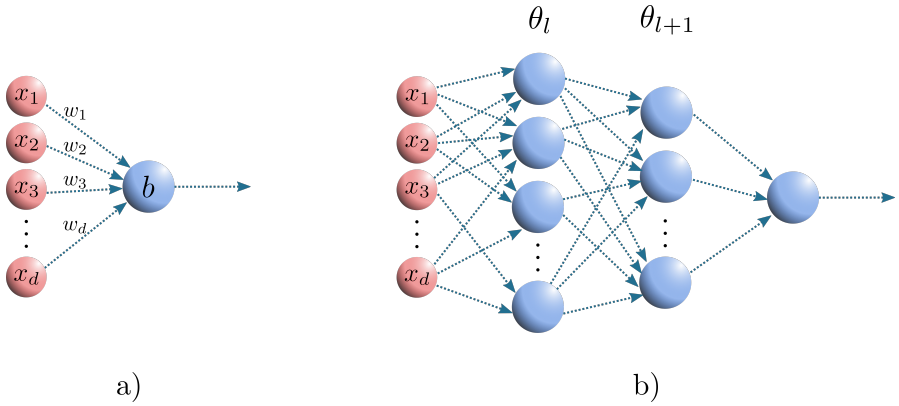

Figure 7.2 – Schematic illustration of a) an artificial neuron and b) a feed-forward neural network.

tation of supervised learning provides a possibility to design the cost function, since f_θ parametrizes a conditional probability distribution for a predicted label. Our goal is then to minimize the negative log-likelihood averaged over the dataset as

$$\mathcal{C}_\theta = -\frac{1}{N} \sum_{(\mathbf{x}_\mu, y_\mu \in \mathcal{D})} \ln \left[p_\theta(y_\mu | \mathbf{x}_\mu) \right], . \tag{7.1}$$

During the training procedure one needs to perform an optimization over the cost function in order to learn the conditional probability $p(y|\mathbf{X})$. There are many optimization techniques that can be used like evolution strategies, simulated annealing or random guessing. Stochastic gradient descent, one of the most commonly used one, relies on the gradient information of the loss function with respect to the parameters θ.

Artificial neural networks

There are numerous variants of supervised learning, the most basic one being a simple linear regression. Other classical learning tools include Gaussian process regression, kernel regression and decision trees, however, the sole class of methods, that stands behind the machine learning revolution of recent years, is based on the so called artificial neural networks. The building block of a neural network is a biologically inspired artificial neuron, which is illustrated in Figure 7.2. The neuron computes the weighted sum $\Sigma(\vec{\mathcal{W}}, b)$ as

$$\Sigma(\vec{\mathcal{W}}, b) = b + \sum_{i=1}^{d} w_i x_i \,, \tag{7.2}$$

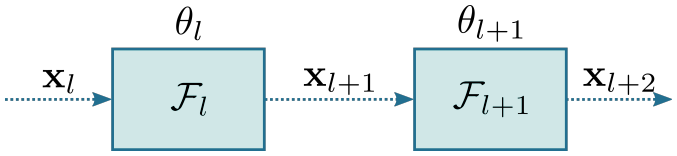

Figure 7.3 – Schematic illustration of the information flow through a feed-forward network. Each layer is represented ny a function $\mathcal{F}_l(\theta_l, \mathbf{x}_l)$.

where $\vec{\mathcal{W}} = \{w_1, w_2, \ldots, w_d\}$ are the weights, and b is the bias. This quantity is then passed through a non-linear transformation, the so called activation function $\mathcal{A}(\vec{\mathcal{W}}, b)$. Popular activation functions include the Sigmoid, the ReLU and the Softplus functions [155]. One can think of a neuron as a switch that depending on the weighted sum of the input elements activates or not. A neural network is built by arranging the neurons into layers, and connecting them. Figure 7.2 represents a particular class of neural network, the so called feed-forward network, since the connections have a direction and the information flows from left to right until the final layer. For each layer, the collection of variables θ_l consist of the weights ($\vec{\mathcal{W}}_n^l$) and biases (b_n^l) of the individual neurons as $\theta_l = (\{\vec{\mathcal{W}}_n^l\}, \{b_n^l\})$, where l denotes a specific layer and n represents a given neuron. Figure 7.3 shows the information flow in a feed-forward neural network denoting each layer with a function $\mathcal{F}_l(\theta_l, \mathbf{x}_l)$ such that

$$\mathbf{x}_{l+1} = \mathcal{F}_l(\theta_l, \mathbf{x}_l) = \begin{pmatrix} \mathcal{A}_1^l(\vec{\mathcal{W}}_1^l, b_1^l) \\ \mathcal{A}_2^l(\vec{\mathcal{W}}_2^l, b_2^l) \\ \vdots \\ \mathcal{A}_{N_n}^l(\vec{\mathcal{W}}_{N_n}^l, b_{N_n}^l) \end{pmatrix}, \tag{7.3}$$

where N_n is the number of neurons in layer l, $\mathbf{x}_l$ denotes a single sample $\mathbf{x}_\mu$ after the transformation of layer l and $\mathcal{A}_n(\vec{\mathcal{W}}_n^l, b_n^l)$ is the activation function. Neural networks are universal function approximators in the sense, that using only a single layer they can approximate any function with an arbitrary accuracy by increasing the number of neurons [159,160]. Practice shows that it is more rewarding to increase the depth, i.e. the number of layers, of the networks, rather than increasing the number of nodes in a single one, hence the name "deep learning".

7.2 Unsupervised learning

As for supervised learning, the input for unsupervised learning is the dataset $\mathcal{D} = (\mathbf{X})$, however, there are no labels y available. The goal of these techniques is to recover and understand the underlying data structure, without any a priori assumptions. Its typical

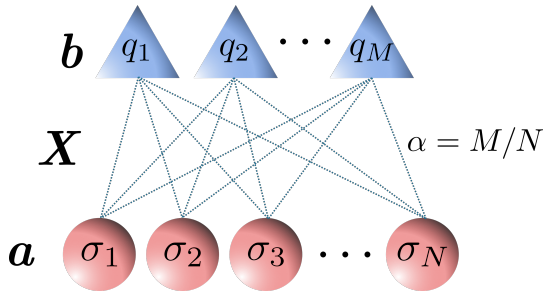

Figure 7.4 – Schematic representation of a Restricted Boltzmann Machine.

applications are clustering algorithms, structure discovery and recommending systems, and its basic methods include principal component analysis or k-means clustering. While the previously introduced discriminative learning techniques model the conditional probability $p(y|\mathbf{X})$, unsupervised learning often aims at learning about the joint probability distribution of $p(\mathbf{X}, y)$. This is referred to as generative modeling [161]. With a generative model one can also support a discriminative task using the Bayes formula

$$p(y|\mathbf{X}) = \frac{p(\mathbf{X}, y)}{p(\mathbf{X})}\,, \tag{7.4}$$

where $p(\mathbf{X}) = \sum_y p(\mathbf{X}, y)$, thus, it also supports reinforcement or semi-supervised learning tasks. Furthermore, one can produce new samples that are conditioned on the label such that $p(\mathbf{X}|y) = p(\mathbf{X}, y)/p(y)$, hence the name generative modeling. Summarizing, the purpose of unsupervised learning is to learn and sample a high dimensional distribution function.

Restricted Boltzmann machines

The key idea behind generative modeling lies in imposing certain structures on the form of the probability distribution. Therefore, several successful models have been developed for the various tasks including variational autoencoders [162], normalizing flow [163], autoregressive models [164] and generative adversarial networks [165]. However, in the context of quantum many body system a rewarding example is the Restricted Boltzmann machine (RBM), which is a particular case of general Boltzmann machines (BM) [155, 166].

Boltzmann machines are methods inspired by statistical physics that model the probability as a Boltzmann distribution

$$p(\mathbf{x}) = \frac{e^{-E(\mathbf{x})}}{\mathcal{Z}}\,, \tag{7.5}$$

where $\mathcal{Z}$ is the normalizing partition function, $E(\mathbf{X})$ is an energy function and we considered a single sample $\mathbf{x} \in \mathbf{X}$. BM-s are practically inverse Ising models where the dataset $\mathcal{D}$ contains samples of the Boltzmann distribution of an Ising model that is connected via pair-wise interactions. The goal is then to learn the parameters of the energy function $E(\mathbf{x})$ such that the probability of the observed samples is high. In order to increase the representative power of the model, we introduce hidden variables in $E(\mathbf{x})$ which then need to be marginalized in order to obtain the probability distribution

$$p(\mathbf{x}) = \frac{1}{\mathcal{Z}} \sum_{\mathbf{q}} e^{-E(\mathbf{x},\mathbf{q})} \,. \tag{7.6}$$

Restricted Boltzmann machines are a particular class of BM-s that combine learn-ability and expressibility [155]. A RBM is composed of two layers of binary valued nodes (see Fig. 7.4): a visible layer for the input vector and a hidden layer for auxiliary variables. Each neuron is associated with a bias (a and b parameters) and the nodes between different layers are connected via a set of weighted edges (X parameters). The energy function reads

$$E(\mathbf{x}, \mathbf{q}) = -\sum_i a_i x_i - \sum_m b_m q_m - \sum_{m,i} q_m x_i X_{mi} \,. \tag{7.7}$$

While in theory a RBM can represent any probability distribution given a sufficiently large number of hidden nodes, this number can grow exponentially. To further improve their efficiency one can introduce the deep Boltzmann machines (DBM) that have more than one layer of hidden neurons. It has been shown that although DBM is capable of representing data that is computationally prohibitive using an RBM [167], it is computationally harder to train and sample them due to the interactions between the hidden units [168].

7.3 Reinforcement learning

Situated between supervised and unsupervised learning, reinforcement learning is focused on sequential decision making problems where there is limited feedback available. The basic paradigm is generally modeled as a Markov decision process [154, 169, 170] with the introduction of an agent that needs to learn how to behave in an environment where the only feedback is a scalar valued reward sign. The goal of the agent is to perform actions on the environment such that it maximizes the reward sign in the long run. Formally one can model the paradigm of a Markov decision process with the following elements.

- *States*: a state is a unique characterization of the environment. The set of environmental states is defined as the finite set $\mathcal{S} = \{s_1, ..., s_N\}$.

- *Actions*: the actions of the agent are defined with another finite set $\mathcal{A} = \{a_1, ..., a_M\}$. Actions are used to control the state of the environment.

- *Transition*: by applying an action $a \in \mathcal{A}$ to a state $s \in \mathcal{S}$ the environment changes state from s to some s'. If the environment is assumed to be Markovian then the probability of transition from state s to s' under action a reads as

$$T(s, a, s') = P(s'|s, a)\,. \tag{7.8}$$

- *Reward*: the reward function determines the reward for being in a state s or performing an action a on state s. The state reward function

$$R_s : \mathcal{S} \longrightarrow \mathbb{R} \tag{7.9}$$

specifies the reward for being in a state. Let us remark, that one can also define the functions

$$\begin{aligned} R_{sa} &: \mathcal{S} \times \mathcal{A} \longrightarrow \mathbb{R} \\ R_{sas} &: \mathcal{S} \times \mathcal{A} \times \mathcal{S} \longrightarrow \mathbb{R}\,, \end{aligned} \tag{7.10}$$

which determine the reward for performing an action on a state (R_{as}) and a particular combination between states (R_{sas}). In this thesis, we will only consider the state reward function R_s.

Figure 7.5 shows the schematic illustration of a general reinforcement learning algorithm. One of the biggest success of reinforcement learning is the AlphaZero computer program that can play the traditional board game Go on a "superhuman" level. Other typical applications include robot navigation, real-time decision making and learning tasks.

Note, the link between the paradigm of reinforcement learning (Fig. 7.5) and the variational algorithm (Fig. 6.1) implies that VMC can also be interpreted as a reinforcement learning scheme. This connection has been first established by Carleo et al. in [146], and inspired the introduction of neural networks into VMC calculations. In the next chapter, we combine the reinforcement scheme with the restricted Boltzmann machines into a variational approach to obtain the non-equilibrium steady state of open quantum systems.

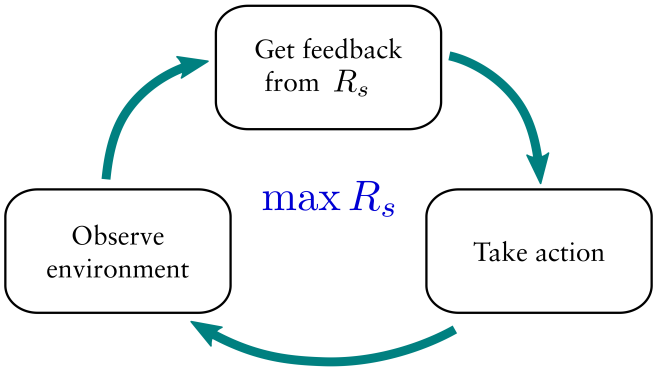

Figure 7.5 – Schematic illustration of the paradigm of reinforcement learning.

CHAPTER 8

Variational quantum Monte Carlo with neural network ansatz

In Chapter 6 we have introduced the fundamentals of the VMC method, which relies on the variational principle to approximate the ground state of many-body quantum systems. In recent years, groundbreaking progress has been made by combining VMC with machine learning tools. Carleo et al. [146] have been the first to establish the link between the variational algorithm (Fig. 6.1) and the paradigm of reinforcement learning (Fig. 7.5), which inspired the introduction of the restricted Boltzmann machine as a variational ansatz. The neural network wave function has proven to efficiently represent highly correlated quantum states, whose parameters can be optimized by means of the variational Monte Carlo method.

In this chapter we introduce the concept of a neural network ansatz for the density matrix, and present a VMC approach to simulate the non-equilibrium steady state of open quantum systems. The development of this method was directed by Vincenzo Savona, and it was published in *Physical Review Letters* [33]. A considerable amount of the description presented in this chapter has been extracted from [33]. The author's contribution to the project was part of the theoretical development, the numerical implementation and the physical model simulations. While developing the present result, we became aware of three related independent works that have been carried out in parallel [61, 136, 137].

The method was implemented in Python and the software can be accessed in the repository *https://github.com/EPFLLTPN/NNDM* [35].

8.1 Variational principle for open quantum systems

The dynamics of the density matrix $\hat{\rho}$ of an open quantum system is governed by the quantum master equation which – in case of Markovian coupling to the environment –

takes the Lindblad form

$$\frac{d\hat{\rho}}{dt} = -i[\hat{H}, \hat{\rho}] - \sum_k \frac{\gamma_k}{2} \left[\left\{ \hat{L}_k^\dagger \hat{L}_k, \hat{\rho} \right\} - 2\hat{L}_k \hat{\rho} \hat{L}_k^\dagger \right] , \tag{8.1}$$

where the curly brackets denote the anti-commutator. A promising route to the numerical computation of the NESS is provided by the variational principle [27, 29, 135]. In cases where a unique steady state exists [38], the NESS corresponds to the eigenmatrix of the Lindblad super-operator $\mathcal{L}$ with zero eigenvalue [39] as

$$\mathcal{L}\left[\hat{\rho}_{ss}\right] = 0\,. \tag{8.2}$$

As all other eigenvalues have strictly negative real part, the NESS can be formally derived as the matrix that maximizes the real part of the expectation value (computed in matrix space) of the Lindbladian.

8.2 Neural network density matrix*

As we have done in Chapter 6, we assume that the Hilbert space of the system is spanned by the computational basis $|\boldsymbol{\sigma}\rangle$, where $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_N)$ labels the states of N degrees of freedom that compose the system, and we consider binary local degrees of freedom, with $\sigma_i = \{-1, 1\}$. The density matrix in this basis is formally expressed as $\rho(\boldsymbol{\sigma}, \boldsymbol{\eta}) = \langle \boldsymbol{\sigma} | \hat{\rho} | \boldsymbol{\eta} \rangle$ in terms of the density operator $\hat{\rho}$. We denote a specific variational ansatz for the density matrix as $\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})$, where $\chi = (\chi_1, \chi_2, \ldots, \chi_{N_p})$ is a set of variational parameters.

A neural network ansatz for a self-adjoint, positive semi-definite density matrix was recently introduced [171] in the specific form of a restricted Boltzmann machine (RBM). In a variational approach, RBMs present the significant advantage that the sum over the hidden-spin configurations can be carried out analytically, and the logarithmic derivatives with respect to the variational parameters admit simple expressions [146]. Here we describe how this ansatz can be derived from simple considerations on the density matrix. A self-adjoint, positive semi-definite expression for the density matrix is

$$\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta}) = \sum_{j=1}^{J} p_j(\chi) \cdot \psi_j(\boldsymbol{\sigma}, \chi) \psi_j^*(\boldsymbol{\eta}, \chi) \tag{8.3}$$

The states $\psi_j(\boldsymbol{\sigma}, \chi)$ are not necessarily mutually orthogonal and the sum extends over J states, with $J \leq d$ and $d = 2^N$ is the dimension of the Hilbert space under study.

We start by introducing a RBM ansatz for each state $\psi_j(\boldsymbol{\sigma}, \chi)$ entering expression (8.3)

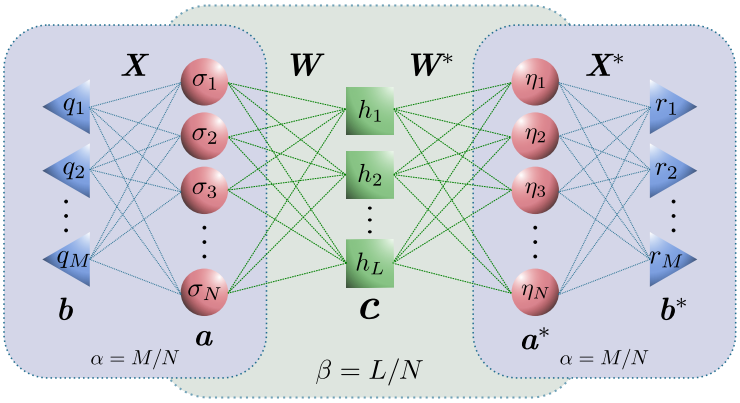

Figure 8.1 – Graphical representation of the neural network ansatz for the density matrix. The input states $|\boldsymbol{\sigma}\rangle, |\boldsymbol{\eta}\rangle$ are encoded in the visible layer, represented by circles. The hidden spins in the triangles encode the correlation between the physical spins in each state of the statistical mixture, while the hidden spins in the squares encode the mixture between the states. This structure is easily seen to coincide with a RBM, where the hidden layer is composed by the triangle and square nodes.

such that

$$\psi(\boldsymbol{\sigma}, \chi) = \sum_{\{q\}} \exp\left(\sum_i a_i \sigma_i + \sum_m b_m q_m + \sum_{m,i} q_m \sigma_i X_{mi}\right) \tag{8.4}$$

where $q_m = \pm 1$ are a set of $M = \alpha \times N$ hidden spin variables and the leftmost sum runs over all possible hidden spin configurations $\{q\}$. A RBM is composed of two layers of binary valued nodes (see Fig. 8.1): a visible layer for encoding the physical state and a hidden layer. Each node is associated with a bias (a- and b-parameters) and nodes within the same layer are connected via a set of weighted edges (X-parameters).

In order to express the mixed structure in eq. (8.3) as a single RBM, we embed an intermediate set of L hidden nodes that are used to express the probabilities $p_j(\chi)$ in RBM form as $p_j(\chi) = \exp(\sum_l c_l h_l)$, with $h_l = \pm 1$ and $c_l \in \mathbb{R}$. To index the different states in the mixture accordingly, this new set of hidden nodes must also enter the (8.4) RBM expression of the wave functions. The full RBM form of the density matrix that we propose is

$$\begin{aligned}\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta}) = &\sum_{\{h\}} \sum_{\{q\}} \sum_{\{r\}} \exp\left(\sum_l c_l h_l\right) \\ &\times \exp\left(\sum_i a_i \sigma_i + \sum_m b_m q_m + \sum_{m,i} q_m \sigma_i X_{mi} + \sum_{l,i} h_l \sigma_i W_{li}\right) \\ &\times \exp\left(\sum_i a_i^* \eta_i + \sum_n b_n^* r_n + \sum_{n,i} r_n \eta_i X_{ni}^* + \sum_{l,i} h_l \eta_i W_{li}^*\right) .\end{aligned} \tag{8.5}$$

Since no intra-layer connection is allowed, the hidden variables can be explicitly traced out and the neural network density matrix reads as

$$\begin{aligned}\rho_\chi(\boldsymbol{\sigma},\boldsymbol{\eta}) = \exp\left(\sum_i a_i\sigma_i\right)\exp\left(\sum_i a_i^*\eta_i\right) \\ \times\prod_{l=1}^{L}\cosh\left(c_l+\sum_i W_{li}\sigma_i+\sum_i W_{li}^*\eta_i\right) \\ \times\prod_{m=1}^{M}\cosh\left(b_m+\sum_i X_{mi}\sigma_i\right) \\ \times\prod_{n=1}^{M}\cosh\left(b_n^*+\sum_i X_{ni}^*\eta_i\right)\,.\end{aligned} \tag{8.6}$$

The RBM is sketched in Fig. 8.1, and $\chi = \{a_i, b_m, X_{mi}, c_l, W_{li}\}$ is the final set of parameters, which are assumed as complex valued with the exception of c_l that must take real values. The representational power of the RBM is determined by the number of hidden nodes [172]. Here we set the densities of hidden nodes through the parameters $\alpha = M/N$, $\beta = L/N$, which measure the representational power of the RBM ansatz independently of the size of the spin lattice. When separately accounting for the real and imaginary parts of complex-valued parameters, the total number of computational parameters in the RBM ansatz is $N_p = N[(\alpha+\beta)(2N+1)+\alpha+2]$. In what follows, we will always assume $\alpha = \beta$ for simplicity.

On the representative power of RBMs

A research field that received a lot of attention recent years is the application of neural networks, and in particular, RBMs as variational ansatz. Although for a large number of hidden nodes, RBMs are known to describe quantum correlations efficiently [147, 149], a deep understanding of their efficiency and representative power is still a matter of debate. Indeed, unlike tensor network states, we cannot identify a quantity that is in direct and quantitative connection with the entanglement properties of the physical systems. Recent studies [149, 150] demonstrated that there is a strong relation between some particular type of RBMs and subclasses of tensor network states, providing additional insight into their efficiency. However, while the structure of tensor network states is intended to model quantum states whose entanglement entropy follows area laws, finding which physical states the geometry of the general RBMs describe is still an open question.

Let us remark, that RBM is not the only neural network model that has been successfully applied for many-body quantum systems [146, 149, 150, 173]. As it has been shown by Gao and Duan [151], there are ground states of Hamiltonians that cannot be efficiently represented with an RBM without invoking an exponential growth in the number of parameters. Instead, one might resort to deep Boltzmann machines [147] and convolutional

networks [148].

8.3 Stochastic reconfiguration for open quantum systems

It is convenient to rewrite eq. (8.2) in a vectorized form by reshaping $\hat{\rho}$ into a column vector $|\rho\rangle$. Following [45], $\mathcal{L}$ takes matrix form and the steady state density matrix fulfils $\langle\rho|\mathcal{L}|\rho\rangle = 0$. Therefore the expectation value over the variational density matrix $\langle\langle\mathcal{L}\rangle\rangle = \langle\rho_\chi|\mathcal{L}|\rho_\chi\rangle/\langle\rho_\chi|\rho_\chi\rangle$ is a function of the variational parameters χ. Now and in what follows we denote the normalized expectation value taken over the variational density matrix $|\rho_\chi\rangle$ with $\langle\langle\cdot\rangle\rangle$. The parameter values that best approximate $\langle\langle\mathcal{L}\rangle\rangle = 0$ can be found by means of various optimization procedures [146, 150, 171]. In this thesis, we choose to adopt the Stochastic Reconfiguration (SR) scheme by Sorella et al. [143] which we extend to open quantum systems.

Similarly to eq. (6.22), after an infinitesimal time step δt the variational density matrix reads

$$\begin{aligned} |\rho_{\tilde{\chi}}\rangle &= |\rho_\chi\rangle + \delta t \sum_{k=1}^{N_p} \delta\dot{\chi}_k \hat{O}_k |\rho_\chi\rangle \\ &= \delta t \sum_{k=0}^{N_p} \delta\dot{\chi}_k \hat{O}_k |\rho_\chi\rangle \,. \end{aligned} \tag{8.7}$$

We can define now the logarithmic derivatives o_k of the density matrix for any configuration $|\boldsymbol{\sigma}, \boldsymbol{\eta}\rangle$ as the variational derivatives with respect to the k-th parameter

$$o_k(\boldsymbol{\sigma}, \boldsymbol{\eta}) = \frac{1}{\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})} \cdot \frac{\partial \rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})}{\partial \chi_k} \,. \tag{8.8}$$

which form the diagonal operators

$$O_k = \begin{cases} \mathbb{I} & \text{for } \; k = 0 \\ \sum\limits_{\boldsymbol{\sigma},\boldsymbol{\eta}} o_k(\boldsymbol{\sigma}, \boldsymbol{\eta}) |\boldsymbol{\sigma}, \boldsymbol{\eta}\rangle\langle\boldsymbol{\sigma}\, \boldsymbol{\eta}| & \text{for } \; k \neq 0 \,. \end{cases} \tag{8.9}$$

Open quantum systems evolve under a one-parameters semigroup dictated by the Lindblad super-operator which for an infinitesimal time step reads as

$$|\rho_\chi^{\text{ex}}\rangle = e^{\mathcal{L}\delta t} |\rho_\chi\rangle \simeq (\mathbb{I} + \mathcal{L}\delta t) |\rho_\chi\rangle \,. \tag{8.10}$$

Under general assumptions, this dynamics asymptotically converges to the non-equilibrium steady state. Once again, ensuring that $|\rho_{\tilde{\chi}}\rangle$ follows the evolution of $|\rho_\chi^{\text{ex}}\rangle$ corresponds to equating eq. (8.7) and (8.10) in the subspace spanned by the vectors $\{O_k|\rho_\chi\rangle\}$. Doing so

we obtain

$$\delta\dot{\chi}_k = \sum_{k'} S^{-1}_{kk'} F_{k'} \,, \tag{8.11}$$

where we introduced the generalized forces and the covariance matrix for a given set of variational parameters as

$$\begin{aligned} F_k &= \langle\langle O^*_k \mathcal{L}\rangle\rangle - \langle\langle \mathcal{L}\rangle\rangle\langle\langle O^*_k\rangle\rangle \\ S_{kk'} &= \langle\langle O^*_k O_{k'}\rangle\rangle - \langle\langle O^*_k\rangle\rangle\langle\langle O_{k'}\rangle\rangle \,, \end{aligned} \tag{8.12}$$

with $k, k' = 1, 2, \ldots, N_p$. After solving eq. (8.11), the variational parameters are updated following

$$\tilde{\chi}_k = \chi_k + \delta t \delta\dot{\chi}_k \,, \tag{8.13}$$

where δt is small enough to guarantee convergence. We point out that, while the expression for S in (8.12) results in the VMC iterations following the real time evolution, minimization can be achieved by using any positive semi-definite covariance matrix. In particular, setting S as the identity results in the steepest descent procedure. Since S can be non-invertible, we apply an explicit regularization scheme, as introduced in [146]:

$$S^{\text{reg}}_{kk'} = S_{kk'} + \lambda(n)\delta_{k,k'} S_{kk'} \,. \tag{8.14}$$

where $\lambda(n) = \max(\lambda_0 b^n, \lambda_{\min})$ with n being the iteration number. For the present calculations, they were set to $\lambda_0 = 100$, $b = 0.998$ and $\lambda_{\min} = 10^{-2}$. Note that, to obtain the correct real-time dynamics the diagonal regularization cannot be applied and S^{-1} denotes the Moore-Penrose inverse [146].

8.4 Observables*

Once the optimal parameter values have been determined, the expectation value of any quantum mechanical observable $\hat{\mathcal{O}}$ over the steady state can be expressed as

$$\langle\hat{\mathcal{O}}\rangle = \mathrm{Tr}(\hat{\mathcal{O}}\hat{\rho}_\chi) = \sum_{\boldsymbol{\sigma},\boldsymbol{\eta}} |\rho_\chi(\boldsymbol{\sigma},\boldsymbol{\eta})|^2 \cdot \frac{\mathcal{O}(\boldsymbol{\eta},\boldsymbol{\sigma})}{\rho_\chi(\boldsymbol{\sigma},\boldsymbol{\eta})^*} \,, \tag{8.15}$$

which can also be evaluated using the Metropolis-Hastings algorithm (see Sec. 6.2). For all the quantities considered here, the expectation values were additionally averaged over 100 sets of parameter values chosen in the asymptotic region of the SR iteration, in order to improve the statistical accuracy. The overall error in the sampled observables has, in addition to the contribution from the Metropolis-Hastings algorithms, a contribution from the SR scheme and a systematic contribution related to the representational power of the RBM ansatz, as measured by the α and β parameters.

8.5 Comment on optimization techniques

We wish to remark, the previously introduced SR method that we have also adopted in [33] is one among several optimization techniques. Here, we want to illustrate the various, and conceptually different approaches that have been introduced for the variational study of open quantum systems [13, 136, 137]. Some of these also provide insight into the real-time dynamics of the density matrix, others can only approximate the steady state one. In order to review these approaches, let us first introduce a shorthand notation for

i the approximate variational density matrix

$$|\rho_\chi^{\delta\chi}\rangle \equiv \delta t \sum_{k=0}^{N_p} \delta\dot{\chi}_k O_k |\rho_\chi\rangle \tag{8.16}$$

ii the real-time evolved density matrix

$$|\rho_\chi^{\delta t}\rangle \equiv e^{\mathcal{L}\delta t}|\rho_\chi\rangle \simeq (\mathbb{I} + \mathcal{L}\delta t)|\rho_\chi\rangle \tag{8.17}$$

iii the density matrix evolved by $\mathcal{L}^\dagger\mathcal{L}$

$$|\rho_\chi^{\delta\tau}\rangle \equiv e^{-\mathcal{L}^\dagger\mathcal{L}\delta\tau}|\rho_\chi\rangle \simeq (\mathbb{I} - \mathcal{L}^\dagger\mathcal{L}\delta\tau)|\rho_\chi\rangle\,. \tag{8.18}$$

Finally, we introduce the five distinct approaches that has been presented in [33, 61, 136, 137].

Approach A:

For studying: **Real-time dynamics / NESS**
ODE given by:

$$\langle\rho_\chi|O_l^*|\rho_\chi^{\delta\chi}\rangle = \langle\rho_\chi|O_l^*|\rho_\chi^{\delta t}\rangle$$

This approach has been presented in [33]. The method relies on equating the approximate variational density matrix $|\rho_\chi^{\delta\chi}\rangle$ with the exact infinitesimal time evolution $|\rho_\chi^{\delta t}\rangle$. This leads to the system of linear equations in (8.11), where the generalized forces correspond to the variational derivatives of $\langle\langle\mathcal{L}\rangle\rangle$, in analogy to eq. (6.32).

Approach B:

For studying: **Real-time dynamics / NESS**
ODE given by:

$$\delta = \left\| |\rho_\chi^{\delta\chi}\rangle - |\rho_\chi^{\delta t}\rangle \right\|_2^2$$

$$\frac{\partial \delta}{\partial \dot{\chi}_k} = 0$$

This approach has been presented in [136], and it minimizes the 2-norm distance between the variationally evolved state $|\rho_\chi^{\delta\chi}\rangle$ and the the exactly evolved one $|\rho_\chi^{\delta t}\rangle$.

Approach C:

For studying: **Real-time dynamics / NESS**
ODE given by:

$$\mathcal{D} = \mathrm{dist}_{FS}(|\rho_\chi^{\delta\chi}\rangle, |\rho_\chi^{\delta t}\rangle) = \arccos \sqrt{\frac{|\langle \rho_\chi^{\delta\chi} | \rho_\chi^{\delta t}\rangle|^2}{\langle \rho_\chi^{\delta\chi} | \rho_\chi^{\delta\chi}\rangle \langle \rho_\chi^{\delta t} | \rho_\chi^{\delta t}\rangle}}$$

$$\frac{\partial \mathcal{D}}{\partial \dot{\chi}_k} = 0$$

This approach has been presented in [136]. The method minimizes the Hilbert space distance, which is given by the Fubini-Study norm, between the variationally evolved state $|\rho_\chi^{\delta\chi}\rangle$ and the the exactly evolved one $|\rho_\chi^{\delta t}\rangle$.

Approach D:

For studying: **NESS**
ODE given by:

$$\langle \rho_\chi | O_l^* | \rho_\chi^{\delta\chi}\rangle = \langle \rho_\chi | O_l^* | \rho_\chi^{\delta\tau}\rangle$$

This approach has been presented in [137]. The technique relies on equating the approximate variational density matrix $|\rho_\chi^{\delta\chi}\rangle$ with $|\rho_\chi^{\delta\tau}\rangle$, generated by $\mathcal{L}^\dagger\mathcal{L}$. $\delta\dot{\chi}_k$ are given by a system of linear equations similar to eq. (8.11), where the generalized forces will now correspond to the variational derivatives of $\langle\langle \mathcal{L}^\dagger\mathcal{L}\rangle\rangle$. The method yields a purely dissipative dynamics using the super-operator $\mathcal{L}^\dagger\mathcal{L}$ as a generator. In the case of a unique steady state, this super-operator is self-adjoint and positive semi-defined, with the only null eigenvalue being associated to the steady-state solution. We argue that this effective dynamics should be more robust to the choice of the time step. The super-operator $\mathcal{L}^\dagger\mathcal{L}$ is however less sparse than $\mathcal{L}$ on the computational basis, calling for an efficient sampling scheme.

Approach E:

For studying: **NESS**
ODE given by:

$$\mathcal{C} = \frac{\left|\frac{d\hat{\rho}}{dt}\right|_2^2}{|\hat{\rho}|_2^2} = \frac{\mathrm{Tr}\left[\mathcal{L}^\dagger\left[\hat{\rho}^\dagger\right]\mathcal{L}\left[\hat{\rho}\right]\right]}{\mathrm{Tr}\left[\hat{\rho}^\dagger\hat{\rho}\right]}$$

$$-\frac{\partial \mathcal{C}}{\partial \dot{\chi}_k} = 0$$

This approach has been presented in [61]. The method minimizes the 2-norm of the time derivative of the density matrix, where the normalization factor is the purity. Although it cannot describe real-time evolution, the steady state density matrix can be obtained since $\dot{\hat{\rho}} = 0$ if and only if $\hat{\rho}_\chi = \hat{\rho}_{SS}$.

8.6 Stochastic sampling*

The various expectation values in eq. (8.12) must be evaluated at each iteration step. Similarly to eq. (6.3), we cast the general forces and the covariance matrix into a suitable form as

$$\begin{aligned}
F_k(n) &= \sum_x |\langle x|\rho_\chi\rangle|^2 \cdot \left(\frac{\partial \ln\langle x|\rho_\chi\rangle}{\partial \chi_k}\right)^* \frac{\langle x|\mathcal{L}|\rho_\chi\rangle}{\langle x|\rho_\chi\rangle} \\
&\quad - \sum_x |\langle x|\rho_\chi\rangle|^2 \cdot \frac{\langle x|\mathcal{L}|\rho_\chi\rangle}{\langle x|\rho_\chi\rangle} \sum_{x'} |\langle x'|\rho_\chi\rangle|^2 \cdot \left(\frac{\partial \ln\langle x'|\rho_\chi\rangle}{\partial \chi_k}\right)^* \\
S_{kk'}(n) &= \sum_x |\langle x|\rho_\chi\rangle|^2 \cdot \left(\frac{\partial \ln\langle x|\rho_\chi\rangle}{\partial \chi_k}\right)^* \left(\frac{\partial \ln\langle x|\rho_\chi\rangle}{\partial \chi_{k'}}\right) \\
&\quad - \sum_x |\langle x|\rho_\chi\rangle|^2 \cdot \left(\frac{\partial \ln\langle x|\rho_\chi\rangle}{\partial \chi_k}\right)^* \sum_{x'} |\langle x'|\rho_\chi\rangle|^2 \cdot \left(\frac{\partial \ln\langle x'|\rho_\chi\rangle}{\partial \chi_{k'}}\right),
\end{aligned} \tag{8.19}$$

where $|x\rangle = |\boldsymbol{\sigma}, \boldsymbol{\eta}\rangle$ are the basis elements spanning the space of the vectorized density matrix, and $|\langle x|\rho_\chi\rangle|^2 = |\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})|^2$ plays the role of a probability distribution. We then compute these quantities stochastically over a Markov-chain of N_{MH} configurations $(\boldsymbol{\sigma}, \boldsymbol{\eta})$ sampling the square modulus of the density matrix $|\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})|^2$. For this we adopt the MCMC algorithm, which we have introduced in Section 6.2. In the limit of $N_{MH} \to \infty$, the statistical error decays as $1/\sqrt{N_{MH}}$. Choosing an appropriate set of rules for the random walk is key to an efficient Monte Carlo sampling, and here we introduce a particular technique which is based on the transitions dictated by the Lindbladian. We will review this move generation protocol in Section 8.7.

8.7 Implementation

The neural network based VMC algorithm was implemented in Python3.6 and has been designed to contain several interchangeable modules that provide extensive functionality:

- *Physical models*: both in one- and two-dimensions
  - dissipative XYZ model
  - transverse Ising model
  - synthetic Ising model
- *Real-time dynamics* (**Approach A**)
  - first order Euler
  - second order Runge-Kutta
- *Non-equilibrium steady state* (**Approach D**)
  - full sampling of $\langle\langle \mathcal{L}^\dagger \mathcal{L} \rangle\rangle$
  - partial sampling of $\langle\langle \mathcal{L}^\dagger \mathcal{L} \rangle\rangle$ (see Sec. 8.7.3)

The number of floating point operations to evaluate eq. (8.11) scales as N_p^3, if we assume that the number of Metropolis-Hastings steps N_{MH} is set to roughly the number of parameters N_p, as in Ref. [146]. The MCMC procedure also scales with the number of connected states $N_{c_{av}}$, i.e. with the average number of non-zero elements in a column of the Lindbladian matrix. Therefore, the efficiency of the whole procedure scales as $O(N_p^3 + N_p N_{c_{av}})$. In order to improve the computational efficiency, we have optimized the stochastic sampling, installed a parallel scheme by splitting the MCMC chain into several independent threads and introduced GPU-accelerated linear algebra calculations.

8.7.1 Optimal calculus

In each optimization step, one needs to compute the generalized forces F and the covariance matrix S in order to evaluate eq. (8.11). This task can be achieved using the previously introduced MCMC technique, in which at each step of the Markov-chain we need to be able to efficiently calculate

- the Metropolis-Hastings acceptance ratio

$$A(\boldsymbol{\sigma}', \boldsymbol{\eta}' | \boldsymbol{\sigma}, \boldsymbol{\eta}) \propto \left| \frac{\rho_\chi(\boldsymbol{\sigma}', \boldsymbol{\eta}')}{\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})} \right|^2 \tag{8.20}$$

- the local value of the Lindbladian

$$\frac{\langle \boldsymbol{\sigma}, \boldsymbol{\eta} | \mathcal{L} | \rho_\chi \rangle}{\langle \boldsymbol{\sigma}, \boldsymbol{\eta} | \rho_\chi \rangle} = \sum_{|\boldsymbol{\alpha}, \boldsymbol{\beta}\rangle}^{N_c} c_{\alpha,\beta} \frac{\langle \boldsymbol{\alpha}, \boldsymbol{\beta} | \rho_\chi \rangle}{\langle \boldsymbol{\sigma}, \boldsymbol{\eta} | \rho_\chi \rangle} \,, \tag{8.21}$$

  where N_c is the number of states connected to $\langle \boldsymbol{\sigma}, \boldsymbol{\eta}|$ through the Lindbladian $\mathcal{L}$.

- and the variational derivatives

$$\frac{1}{\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})} \cdot \frac{\partial \rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta})}{\partial \chi_k} \,. \tag{8.22}$$

For this, we follow [146] by introducing the so-called effective angles

$$\begin{aligned} \tilde{\theta}_l(\boldsymbol{\sigma}, \boldsymbol{\eta}) &= c_l + \sum_i W_{li} \sigma_i + \sum_i W^*_{li} \eta_i \\ \theta_m(\boldsymbol{\sigma}) &= b_m + \sum_i X_{mi} \sigma_i \,, \end{aligned} \tag{8.23}$$

that also allows us to write the density matrix ansatz in a condensed form as

$$\rho_\chi(\boldsymbol{\sigma}, \boldsymbol{\eta}) = e^{(\sum_i a_i \sigma_i)} e^{(\sum_i a^*_i \eta_i)} \prod_{l=1}^{L} \prod_{m=1}^{M} \prod_{n=1}^{M} \cosh \tilde{\theta}_l(\boldsymbol{\sigma}, \boldsymbol{\eta}) \cosh \theta_m(\boldsymbol{\sigma}) \cosh \theta^*_n(\boldsymbol{\eta}) \,. \tag{8.24}$$

In order to efficiently calculate the ratios in eq. (8.20) and (8.21) as well as the variational derivatives, it is useful to keep in memory the effective angles and update them at each step of the random walk according to

$$\begin{aligned} \tilde{\theta}'_l(\boldsymbol{\sigma}', \boldsymbol{\eta}') &= \tilde{\theta}_l(\boldsymbol{\sigma}, \boldsymbol{\eta}) + \sum_i W_{li} S^\sigma_i + \sum_i W^*_{li} S^\eta_i \\ \theta'_m(\boldsymbol{\sigma}') &= \theta_m(\boldsymbol{\sigma}) + \sum_i X_{mi} S^\sigma_i \,, \end{aligned} \tag{8.25}$$

where

$$S^\sigma_i = \begin{cases} 0 & \text{if } \sigma'_i = \sigma_i \\ -2\sigma_i & \text{if } \sigma'_i \neq \sigma_i \end{cases} \quad \text{and} \quad S^\eta_i = \begin{cases} 0 & \text{if } \eta'_i = \eta_i \\ -2\eta_i & \text{if } \eta'_i \neq \eta_i \,. \end{cases} \tag{8.26}$$

Moreover, given the expression (8.24), the variational derivatives can also be written in a

rather compact form as

$$\begin{aligned}
\frac{1}{\rho_\chi(\boldsymbol{\sigma},\boldsymbol{\eta})}\cdot\frac{\partial\rho_\chi(\boldsymbol{\sigma},\boldsymbol{\eta})}{\partial\mathrm{Re}(a_k)} &= \sigma_k+\eta_k \\
\frac{1}{\rho_\chi(\boldsymbol{\sigma},\boldsymbol{\eta})}\cdot\frac{\partial\rho_\chi(\boldsymbol{\sigma},\boldsymbol{\eta})}{\partial\mathrm{Im}(a_k)} &= i(\sigma_k-\eta_k) \\
\frac{1}{\rho_\chi(\boldsymbol{\sigma},\boldsymbol{\eta})}\cdot\frac{\partial\rho_\chi(\boldsymbol{\sigma},\boldsymbol{\eta})}{\partial\mathrm{Re}(b_k)} &= \tanh\left(\theta_k(\boldsymbol{\sigma})\right)+\tanh\left(\theta_k^*(\boldsymbol{\eta})\right) \\
\frac{1}{\rho_\chi(\boldsymbol{\sigma},\boldsymbol{\eta})}\cdot\frac{\partial\rho_\chi(\boldsymbol{\sigma},\boldsymbol{\eta})}{\partial\mathrm{Im}(b_k)} &= i\left[\tanh\left(\theta_k(\boldsymbol{\sigma})\right)-\tanh\left(\theta_k^*(\boldsymbol{\eta})\right)\right] \\
\frac{1}{\rho_\chi(\boldsymbol{\sigma},\boldsymbol{\eta})}\cdot\frac{\partial\rho_\chi(\boldsymbol{\sigma},\boldsymbol{\eta})}{\partial\mathrm{Re}(X_{kl})} &= \tanh\left(\theta_k(\boldsymbol{\sigma})\right)\sigma_l+\tanh\left(\theta_k^*(\boldsymbol{\eta})\right)\eta_l \\
\frac{1}{\rho_\chi(\boldsymbol{\sigma},\boldsymbol{\eta})}\cdot\frac{\partial\rho_\chi(\boldsymbol{\sigma},\boldsymbol{\eta})}{\partial\mathrm{Im}(X_{kl})} &= i\left[\tanh\left(\theta_k(\boldsymbol{\sigma})\right)\sigma_l-\tanh\left(\theta_k^*(\boldsymbol{\eta})\right)\eta_l\right] \\
\frac{1}{\rho_\chi(\boldsymbol{\sigma},\boldsymbol{\eta})}\cdot\frac{\partial\rho_\chi(\boldsymbol{\sigma},\boldsymbol{\eta})}{\partial c_k} &= \tanh\left(\tilde{\theta}_l(\boldsymbol{\sigma},\boldsymbol{\eta})\right) \\
\frac{1}{\rho_\chi(\boldsymbol{\sigma},\boldsymbol{\eta})}\cdot\frac{\partial\rho_\chi(\boldsymbol{\sigma},\boldsymbol{\eta})}{\partial\mathrm{Re}(W_{kl})} &= \tanh\left(\tilde{\theta}_l(\boldsymbol{\sigma},\boldsymbol{\eta})\right)(\sigma_l+\eta_l) \\
\frac{1}{\rho_\chi(\boldsymbol{\sigma},\boldsymbol{\eta})}\cdot\frac{\partial\rho_\chi(\boldsymbol{\sigma},\boldsymbol{\eta})}{\partial\mathrm{Im}(W_{kl})} &= i\tanh\left(\tilde{\theta}_l(\boldsymbol{\sigma},\boldsymbol{\eta})\right)(\sigma_l-\eta_l)\,.
\end{aligned} \tag{8.27}$$

8.7.2 Optimal stochastic sampling*

In order to perform the stochastic sampling proposed in Sec. 8.6, we apply a standard MCMC protocol where the configurations are generated based on a proposition scheme, and accepted according to

$$A(\boldsymbol{\sigma}',\boldsymbol{\eta}'|\boldsymbol{\sigma},\boldsymbol{\eta}) = \min\left(1,\frac{T(\boldsymbol{\sigma},\boldsymbol{\eta}|\boldsymbol{\sigma}',\boldsymbol{\eta}')|\rho_\chi(\boldsymbol{\sigma}',\boldsymbol{\eta}')|^2}{T(\boldsymbol{\sigma}',\boldsymbol{\eta}'|\boldsymbol{\sigma},\boldsymbol{\eta})|\rho_\chi(\boldsymbol{\sigma},\boldsymbol{\eta})|^2}\right). \tag{8.28}$$

Our move generation is based on the transitions dictated by the Lindblad super-operator which occur by applying $\mathcal{L}|\boldsymbol{\sigma},\boldsymbol{\eta}\rangle$. Doing so decreases the required number of thermalization and sampling steps by promoting, in the random walk, the configurations with the highest importance. However, in case of a close-to-pure steady-state, a handful of density matrix elements have magnitude dominating the others. This causes the random walk to perform poorly, as some regions in configuration space become scarcely accessible. Hence, we introduced additional moves which occur with a low probability and reach a configuration by randomly flipping one spin in both $\boldsymbol{\sigma}$ and $\boldsymbol{\eta}$. The propositions always need to obey detailed balance.
For sampling $|\rho_\chi(\boldsymbol{\sigma},\boldsymbol{\eta})|^2$ in the case of a two-dimensional spin lattice model, we introduce the following possible moves on the configuration space spanned by $|\boldsymbol{\sigma},\boldsymbol{\eta}\rangle$

1. *Column-hopping in* $\boldsymbol{\sigma}$: a site σ_j and its right neighbour are flipped. This move accounts for single spin hopping as well as double excitation or loss.

2. *Row-hopping in* $\boldsymbol{\sigma}$: a site σ_j and its down neighbour are flipped.

3. *Column-hopping in* $\boldsymbol{\eta}$: a site η_j and its right neighbour are flipped.

4. *Row-hopping in* $\boldsymbol{\eta}$: a site η_j and its down neighbour are flipped.

5. *Field in* $\boldsymbol{\sigma}$: a site σ_j is flipped. This move corresponds to the effect of having an external field in the direction perpendicular to the quantization axis.

6. *Field in* $\boldsymbol{\eta}$: a site η_j is flipped.

7. *Dissipator*: sites $\sigma_j = \eta_j$ are flipped with an asymmetrical transition ratio. The dissipative moves are always proposed, while excitations are only proposed with ten-percent probability.

8. *Jumper:* two randomly chosen spins σ_l, η_m are flipped.

For the move proposition we choose one of the actions with uniform probability. In Figure 8.2 we present a comparison between the efficiency of using random move proposition, where at each MCMC step a random spin is flipped in $\boldsymbol{\sigma}$, $\boldsymbol{\eta}$, and applying the Lindbladian based proposition scheme. The graph shows the relative error of $\langle\langle\mathcal{L}\rangle\rangle$ for an increasing number of sampling steps N_{MH}. Indeed, using the Lindbladian proposition

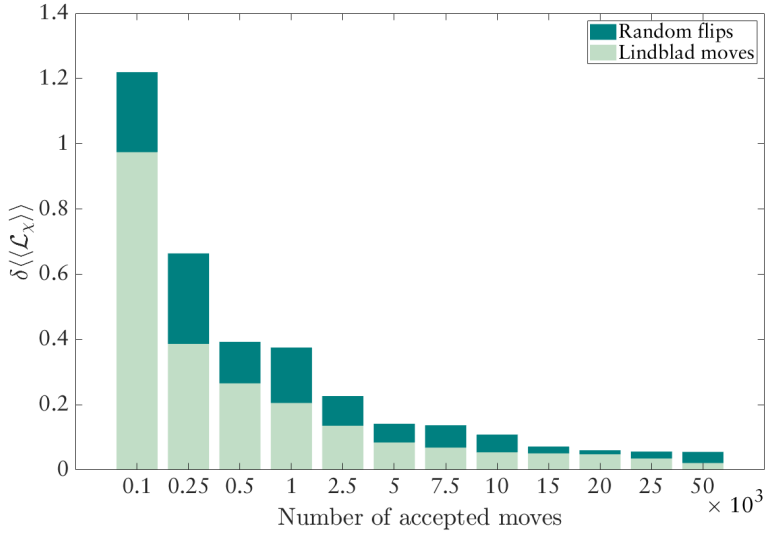

Figure 8.2 – Comparing the efficiency of using random move proposition versus the Lindbladian scheme. The graph shows the relative error of $\langle\langle\mathcal{L}\rangle\rangle$ as function of the number of accepted MCMC moves. This result has been obtained for a 3×3 lattice XYZ Heisenberg model with parameters $J_x/\gamma = 0.9$, $J_y/\gamma = 1.2$, $J_z/\gamma = 1$ and $\alpha = \beta = 4$.

scheme provides in average 10% better accuracy than using random spin flips. We wish to note here, that we set the number of *accepted* moves instead of the total number of moves. This is necessary as, for some choices of the physical parameters, the density matrix is highly concentrated on a few matrix elements and consequently MCMC moves are seldom accepted.

8.7.3 Sampling $\mathcal{L}^\dagger\mathcal{L}$

As we have shown in Section 8.5, one does not necessarily have to follow the real-time dynamics towards the steady state, it is also possible to study the purely dissipative dynamics using the super-operator $\mathcal{L}^\dagger\mathcal{L}$ as generator, whose zero eigenvalue solution is associated to the NESS. Although this choice can be more robust against the choice of the time step, the generator $\mathcal{L}^\dagger\mathcal{L}$ is however less sparse, calling for an effective sampling scheme of what now plays the role of the local variational energy

$$\langle\langle\mathcal{L}^\dagger\mathcal{L}\rangle\rangle_{\mathrm{loc}} = \frac{\langle\boldsymbol{\sigma},\boldsymbol{\eta}|\mathcal{L}^\dagger\mathcal{L}|\rho_\chi\rangle}{\langle\boldsymbol{\sigma},\boldsymbol{\eta}|\rho_\chi\rangle} = \sum_{|\boldsymbol{\alpha},\boldsymbol{\beta}\rangle}^{N_c^{\sigma,\eta}} c_{\alpha,\beta}\frac{\langle\boldsymbol{\alpha},\boldsymbol{\beta}|\mathcal{L}|\rho_\chi\rangle}{\langle\boldsymbol{\sigma},\boldsymbol{\eta}|\rho_\chi\rangle} = \sum_{|\boldsymbol{\alpha},\boldsymbol{\beta}\rangle}^{N_c^{\sigma,\eta}} c_{\alpha,\beta}\sum_{|\boldsymbol{\gamma},\boldsymbol{\delta}\rangle}^{N_c^{\alpha,\beta}} d_{\gamma,\delta}\frac{\langle\boldsymbol{\gamma},\boldsymbol{\delta}|\rho_\chi\rangle}{\langle\boldsymbol{\sigma},\boldsymbol{\eta}|\rho_\chi\rangle}\,, \tag{8.29}$$

where $N_c^{\sigma,\eta}$ is the number of states connected to $\langle\boldsymbol{\sigma},\boldsymbol{\eta}|$ through $\mathcal{L}^\dagger$ and $N_c^{\alpha,\beta}$ is the number of states connected to $\langle\boldsymbol{\alpha},\boldsymbol{\beta}|$ through $\mathcal{L}$. Thus, the complexity of the sum (8.29) scales with the square of the average connectivity of the Lindbladian. In our implementation one has two choices to compute eq. (8.29): we can either perform the sum explicitly which yields to a computationally slower MCMC step, or can perform a stochastic sampling of

$$\sum_{|\boldsymbol{\gamma},\boldsymbol{\delta}\rangle}^{N_c^{\alpha,\beta}} d_{\gamma,\delta}\frac{\langle\boldsymbol{\gamma},\boldsymbol{\delta}|\rho_\chi\rangle}{\langle\boldsymbol{\sigma},\boldsymbol{\eta}|\rho_\chi\rangle}\,. \tag{8.30}$$

In practice, this means that for every Markov-chain we need to store a running average of $\langle\langle\mathcal{L}^\dagger\mathcal{L}\rangle\rangle_{\mathrm{loc}}$ for each visited state $\langle\boldsymbol{\sigma},\boldsymbol{\eta}|$. We have also introduced a parameter N_{draw} that is the number of states $\langle\boldsymbol{\gamma},\boldsymbol{\delta}|$ that will be randomly chosen and taken into account every time the random walk visits $\langle\boldsymbol{\sigma},\boldsymbol{\eta}|$. This procedure reduces the number of operations to $\mathcal{O}(N_c N_{\mathrm{draw}})$ for every MCMC step. Even though the number of steps before the random walk converges will increase, the cost of each move will become so much cheaper that the overall numerical efficiency is improved. The parameter N_{draw} can be used to strike a balance between the convergence time of the Markov-chain and the computational cost of each move. In Section 8.8.3 we show an example how this latter sampling techniques allows us to tackle larger system sizes.

8.7.4 Parallel computing

We implemented the MCMC algorithm in a parallel fashion by splitting the Markov-chain into several independent threads that are run in parallel using MVAPICH2 [131]. Both a CPU and a GPU version of the code has been developed. GPU calculations allowed to significantly speed up the update of the O_k expectation values, as well as the iterative solution of the linear system in the SR, using the MINRES-QLP algorithm [174,175], which can correctly handle the case of a singular matrix. The GPU version brought considerable advantage over CPU when the number of variational parameters was larger than 1000. Figure 8.3 shows a schematic representation of the GPU accelerated parallel program flow.

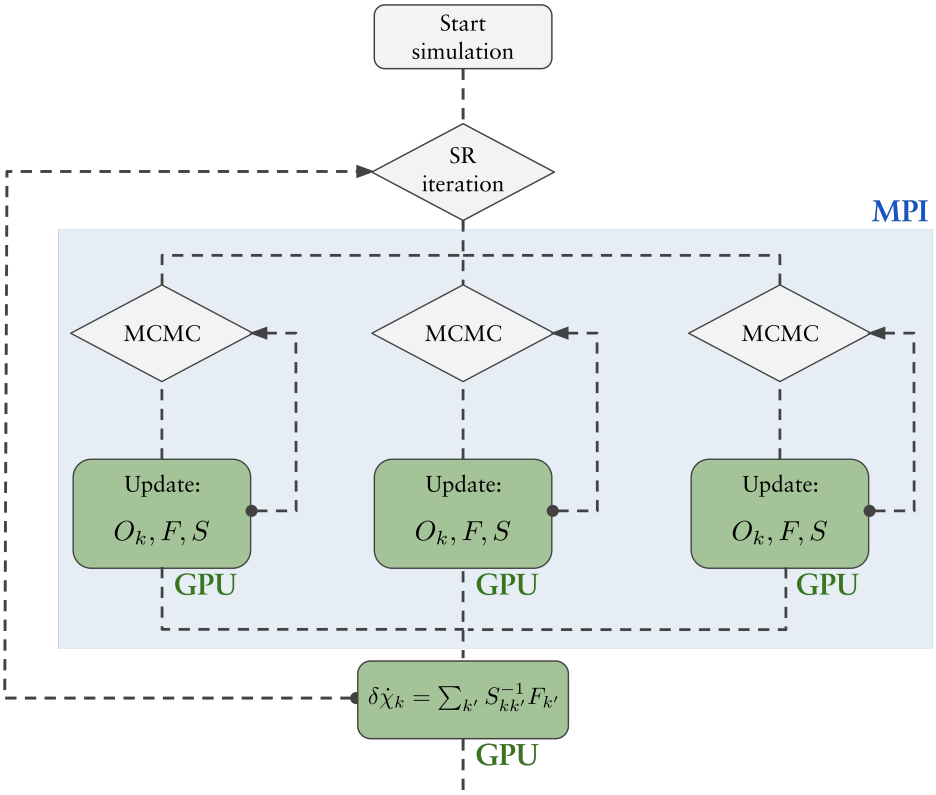

Figure 8.3 – Schematic illustration of the parallel program flow.

8.8 Results

To assess the accuracy and effectiveness of the neural network variational ansatz for the density matrix we present three studies; the first one models the NESS of the XYZ Heisenberg model through a dissipative phase transition, in the second we observe the NESS of the dissipative Ising model, and finally we present a real-time dynamics study of the synthetic Ising model.

8.8.1 The dissipative XYZ Heisenberg model*

We study a spin-1/2 XYZ model on a two-dimensional lattice with periodic boundary condition. As we have discussed in Chapter 1.3, the Hamiltonian and the quantum master equation read

$$\hat{H} = \sum_{\langle i,j \rangle} \left(J_x \hat{\sigma}_i^x \hat{\sigma}_j^x + J_y \hat{\sigma}_i^y \hat{\sigma}_j^y + J_z \hat{\sigma}_i^z \hat{\sigma}_j^z \right) \tag{8.31}$$

$$\frac{\mathrm{d}\hat{\rho}}{\mathrm{d}t} = -i[\hat{H}, \hat{\rho}] - \frac{\gamma}{2} \sum_k \left[\left\{ \hat{\sigma}_k^+ \hat{\sigma}_k^-, \hat{\rho} \right\} - 2\hat{\sigma}_k^- \hat{\rho} \hat{\sigma}_k^+ \right] \tag{8.32}$$

where $\hat{\sigma}_j^x$, $\hat{\sigma}_j^y$, $\hat{\sigma}_j^z$ are the Pauli matrices, $\hat{\sigma}_j^{\pm} = (\hat{\sigma}_j^x \pm i\hat{\sigma}_j^y)/2$, J_α are the coupling constants between nearest neighbour spins and γ is the dissipation rate. The excitations in the system – induced by the anisotropic spin coupling – compete with the isotropic dissipative process, and this competition is at the origin of the dissipative phase transition [14, 15, 17, 17, 18]. The effectiveness of the neural network ansatz is demonstrated by studying the system observables across a phase boundary, and we apply the SR presented for Approach A (see Sec. 8.5)

In addition to the expectation value $\langle\langle \mathcal{L} \rangle\rangle$, we study the local magnetization

$$M_z = \frac{1}{N} \sum_{i=1}^{N} \mathrm{Tr}(\hat{\rho} \hat{\sigma}_i^z) \,, \tag{8.33}$$

and the steady-state structure factor

$$S_{ss}^{xx}(\mathbf{k}) = \frac{1}{N(N-1)} \sum_{\mathbf{j} \neq \mathbf{l}} e^{-i\mathbf{k}(\mathbf{j}-\mathbf{l})} \langle \hat{\sigma}_{\mathbf{j}}^x \hat{\sigma}_{\mathbf{l}}^x \rangle \,, \tag{8.34}$$

computed for the asymptotic steady state. Fig. 8.4 shows the convergence of $S_{ss}^{xx}(\mathbf{k} = \mathbf{0})$ and $S_{ss}^{xx}(\mathbf{k} = (2\pi/3, 0))$ to the exact result for a 3×3 lattice, as $\alpha = \beta$ are increased. The parameters χ are initialized randomly and updated at each VMC step according to the SR scheme. The parameters of the model are chosen to lie in the vicinity of the dissipative phase transition, i.e. $J_x/\gamma = 0.9$, $J_y/\gamma = 1.2$, $J_z/\gamma = 1.0$. A clear convergence towards

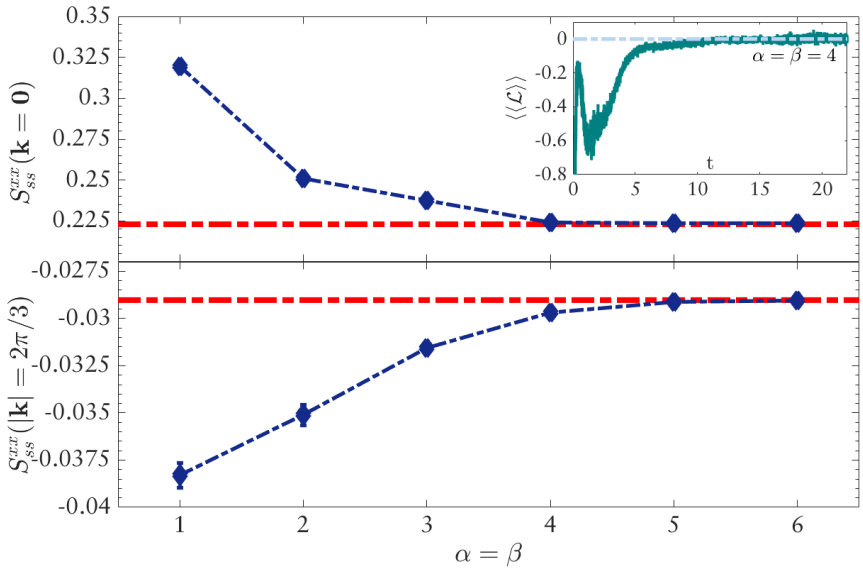

Figure 8.4 – The steady-state spin structure factor $S^{xx}_{ss}(\mathbf{k})$ computed as a function of $\alpha=\beta$ for a 3×3 lattice and $\mathbf{k}=\mathbf{0}$ (upper panel) and $\mathbf{k}=(2\pi/3,0)$ (lower panel). The red dot-dashed line represents in both panels the exact result. The inset shows the evolution of $\langle\langle\mathcal{L}\rangle\rangle$ over the VMC run. Parameters: $J_x/\gamma=0.9$, $J_y/\gamma=1.2$, $J_z/\gamma=1.0$.

the exact value upon increasing $\alpha=\beta$ is found. The inset in Fig. 8.4 shows the SR evolution of $\mathrm{Re}(\langle\langle\mathcal{L}_\chi\rangle\rangle)$ over a typical VMC run. The oscillations at early times are a feature of the unitary part of the dynamics in the quantum master equation.

In Fig. 8.5 we display the magnetization as computed for different lattice sizes and as a function of the coupling parameter J_y/γ. For this choice of parameters, a para-to-ferromagnetic phase transition is expected to occur when increasing the coupling through the value $J_y \gtrapprox 1.04$ [16, 17], while a second phase boundary between a ferromagnetic and a paramagnetic region has been predicted by cluster mean-field calculations at around $J_y \gtrapprox 1.4$. For 2×2 and 3×3 lattices the VMC result agrees well with the exact calculation for a large enough number of variational parameters. In Fig. 8.6 we display the spin structure factor $S^{xx}_{ss}(\mathbf{k}=\mathbf{0})$ for the same parameters as in Fig. 8.5. The quantity $S^{xx}_{ss}(\mathbf{k}=\mathbf{0})$ vanishes when in a paramagnetic phase, while it takes a finite value in the ferromagnetic region of the phase diagram. This behaviour is displayed both by the exact calculation for small lattices, and by the VMC data, in the vicinity of the phase boundary at $J_y \gtrapprox 1.04$. For values $J_y > 1.4$ the system should become again paramagnetic in the thermodynamic limit of large lattices, but this feature was not displayed by the present data up to the largest lattice under study, in agreement also with recent stochastic Gutzwiller calculations [18].

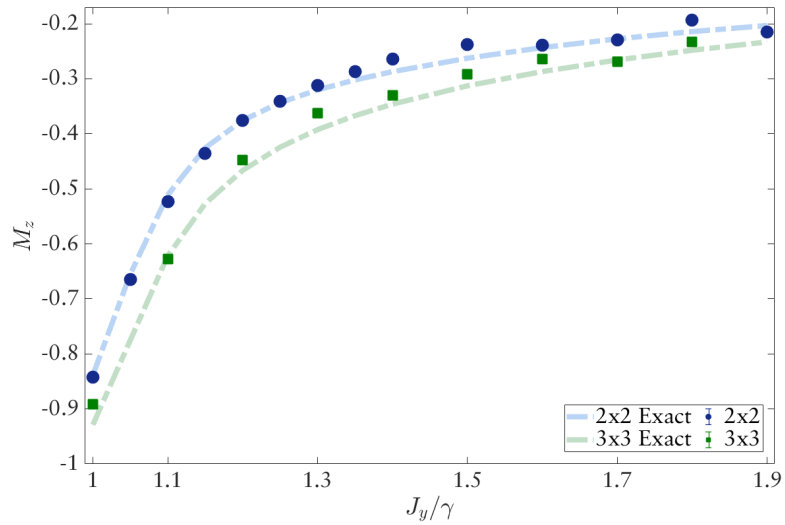

Figure 8.5 – The magnetization M_z computed as a function of the coupling J_y/γ. VMC and exact values are compared. Error bars, when not shown, are smaller than the symbol. Other parameters: $J_x/\gamma = 0.9$, $J_z/\gamma = 1.0$, $\alpha = \beta = 3$.

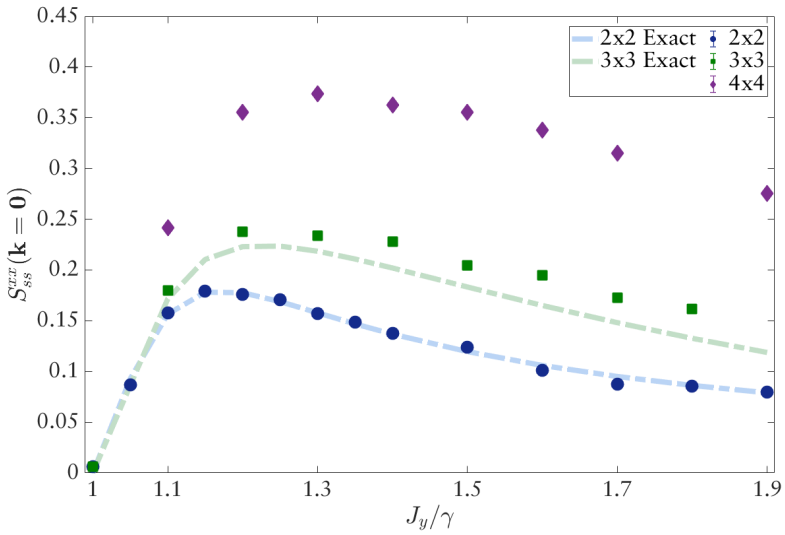

Figure 8.6 – The steady-state spin structure factor $S^{xx}_{ss}(\mathbf{k} = \mathbf{0})$ computed as a function of the coupling J_y/γ. VMC and exact values are compared. Other parameters: $J_x/\gamma = 0.9$, $J_z/\gamma = 1.0$, $\alpha = \beta = 3$.

8.8.2 Time evolution of the dissipative synthetic Ising model

To benchmark the accuracy of the time evolution for $\hat{\rho}$ as obtained from eq. (8.11), we consider the dissipative synthetic Ising model that has been in the focus of extensive experimental research [62, 64–66]. Using a mapping on multilevel Rydberg atoms [63], the experimental setup can efficiently simulate the effective Ising model that is governed by

$$\hat{H} = \sum_i \left(\frac{\Omega}{2} \hat{\sigma}_i^x - \delta \hat{n}_i \right) + \frac{1}{2} \sum_{\langle i,j \rangle} U \hat{n}_i \hat{n}_j \tag{8.35}$$

$$\frac{\mathrm{d}\hat{\rho}}{\mathrm{d}t} = -i[\hat{H}, \hat{\rho}] - \frac{\gamma}{2} \sum_k \left[\{\hat{n}_k, \hat{\rho}\} - 2\hat{n}_k \hat{\rho} \hat{n}_k \right] , \tag{8.36}$$

where the atoms are coherently coupled to the Rydberg state with a two-photon transition of Rabi frequency Ω and a detuning δ, and $\hat{n}_i = (\mathbb{I} + \hat{\sigma}_i^z)/2$ is the projector on the Rydberg state for atom i. As we have previously derived, the NESS is well-known to be fully mixed when the dissipative channel is pure dephasing. However, the short time dynamics of the system is subject to many experimental studies. Therefore, we chose to reproduce the experimental parameters presented in [62] and study how the number of excited Rydberg atoms n evolves in time with

$$n = \frac{1}{N} \sum_{i=1}^{N} \mathrm{Tr}(\hat{\rho} \hat{n}_i) \, . \tag{8.37}$$

Figure 8.7 shows the results for a chain of 6 spins with open boundary conditions. The dynamics generated by the SR algorithm provides a quantitatively accurate approximation of the dynamics already for a hidden node density $\alpha = \beta = 1$.

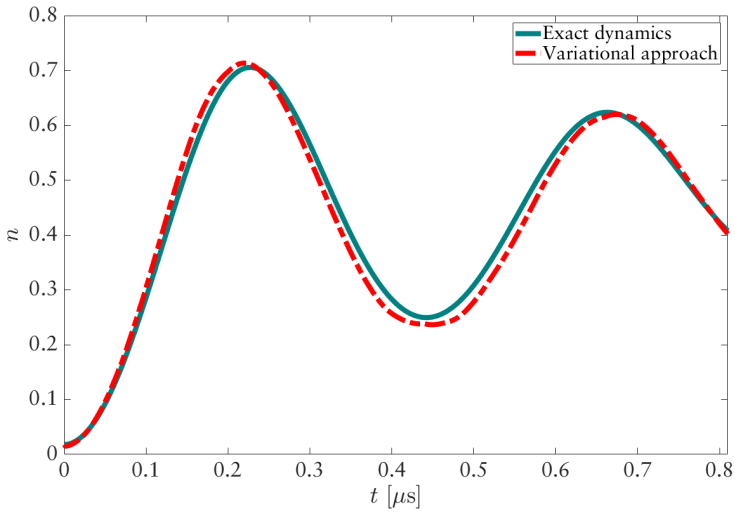

Figure 8.7 – The real-time dynamics of the number of excited Rydberg atoms n with $\Omega = 3.6\pi$, $U = 5.4\pi$, $\delta = 0$, $\alpha = \beta = 1$.

8.8.3 The steady state for the dissipative Ising model

Besides following the real-time evolution, we have also implemented a direct search for the NESS as in Approach D of Section 8.5. Adjusting the variational parameters to follow the dynamics dictated by $\mathcal{L}^\dagger\mathcal{L}$ leads to the same system of equations as given in eq. (8.11), but where the generalized forces and the covariance matrix reads

$$\begin{aligned} F_k &= \langle\langle \mathcal{L} \rangle\rangle \langle\langle O_k^* \rangle\rangle - \langle\langle O_k^* \mathcal{L} \rangle\rangle \\ S_{kk'} &= \langle\langle O_k^* O_{k'} \rangle\rangle - \langle\langle O_k^* \rangle\rangle \langle\langle O_{k'} \rangle\rangle \,. \end{aligned} \tag{8.38}$$

In order to benchmark the method and test its scalability, we consider the one-dimensional dissipative quantum transverse Ising model with periodic boundary conditions, whose Hamiltonian and master equation follows

$$\hat{H} = h \sum_i \hat{\sigma}_i^x + J_z \sum_{\langle i,j \rangle} \hat{\sigma}_i^z \hat{\sigma}_j^z \tag{8.39}$$

$$\frac{\mathrm{d}\hat{\rho}}{\mathrm{d}t} = -i[\hat{H}, \hat{\rho}] - \frac{\gamma}{2} \sum_k \left[\left\{ \hat{\sigma}_k^+ \hat{\sigma}_k^-, \hat{\rho} \right\} - 2\hat{\sigma}_k^- \hat{\rho} \hat{\sigma}_k^+ \right] . \tag{8.40}$$

As an example of convergence, in Figure 8.8 we report a typical evolution of $\langle\langle \mathcal{L}^\dagger \mathcal{L} \rangle\rangle$ in the SR minimization procedure for a fixed set of parameters, that shows a good convergence towards the global minimum for 5 and 25 spins respectively. To obtain the results we have applied the partial sampling strategy introduced in Sec. 8.7 and a hidden node density $\alpha = \beta = 1$.

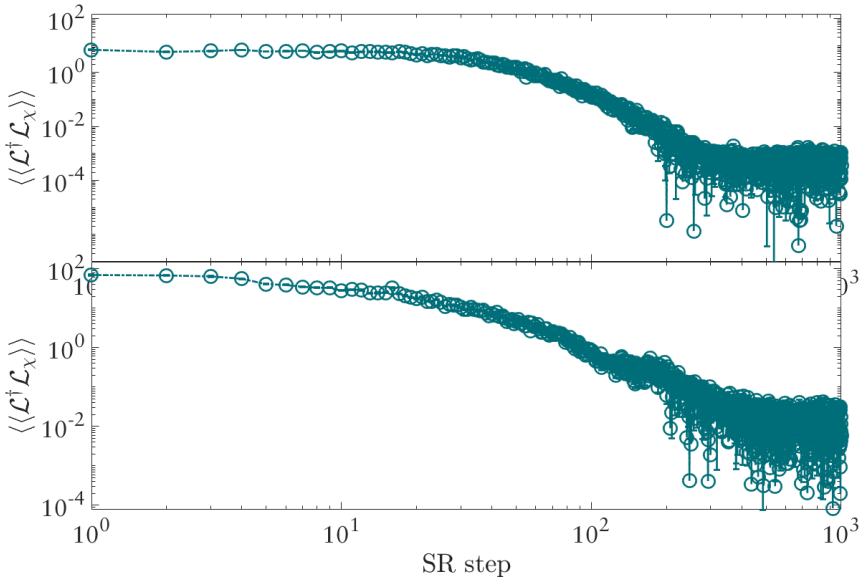

Figure 8.8 – SR evolution of $\langle\langle \mathcal{L}^\dagger \mathcal{L}_\chi \rangle\rangle$ for the 1D dissipative Ising model for 5 (upper panel) and 25 (lower panel) spins with $J_z/\gamma = 0.5$, $h = 0.5$, $\alpha = \beta = 1$.

8.9 Conclusion

Existing numerical approaches to the simulation of the steady state of a Markovian open quantum system either require the full representation of the Hilbert space into memory, or rely on a properly chosen truncation of the Hilbert space to a relevant subspace. The present VMC approach is free of these two limitations, thanks to the stochastic evaluation of expectation values by means of the MCMC algorithm. The neural network ansatz in terms of a RBM is highly representative of quantum correlated statistical mixtures, while being simple to handle numerically. In cases with very strong quantum correlations, this ansatz could be extended to deep network representations, as was recently done in the case of Hamiltonian problems [147,148,151,176]. For some of these networks [176], the hidden degrees of freedom can still be summed analytically, as for RBMs. Neural network representations are not restricted to spin degrees of freedom and have been successfully adopted to represent bosonic many-body states efficiently [177]. For these reasons, the present VMC approach may emerge as the election tool to numerically model open quantum systems, with considerable impact on the study of fundamental physics and on the modeling of near-term, noisy quantum information platforms [22].

Summary

PART IV

CHAPTER 9

Conclusion and outlook

In this thesis we have developed two complementary numerical approaches to simulate the dynamics and the asymptotic steady state resulting from the Lindblad quantum master equation. The possibility to simulate the properties of many-body open quantum systems with a large number of degrees of freedom is the premise to the solution of several outstanding problems in quantum science and quantum information. However, as within the Markovian description only few model admits analytical solution, the quest for finding efficient numerical methods is still ongoing.

In Part I, we explored the fundamental principles of the theory of many-body open quantum systems, we have presented several paradigmatic models that are often used as a first example to benchmark a novel technique, and we reviewed the most commonly used computational approaches with a particular focus on their advantages and disadvantages.

Then, in Part II, we have introduced a novel numerical method for quantum systems with Markovian system-environment coupling, called DDQMC. The approach stems from the FCIQMC algorithm and exploits the analogy between the long-time dynamics of the Lindblad master equation and the imaginary-time Schrödinger equation. DDQMC allows the direct sampling of the density matrix in any discrete basis set, and the introduction of initiator approach and importance sampling has lead to a significant improvement in the statistical accuracy. We have demonstrated the validity of the method by investigating a dissipative phase transition on the two-dimensional XYZ Heisenberg model. The defining feature of DDQMC is that it samples the whole density matrix and it does not introduce a truncation in Hilbert space. The biggest limitation of the algorithm emerged with increasing system size, as the absence of a guiding wave function resulted in an exponentially increasing statistical error on the observables.

Finally, in Part III, we have presented in detail a variational approach to efficiently simulate both the dynamics and the non-equilibrium steady state of Markovian open quantum systems. The density matrix is parametrized using a neural network ansatz, and the parameters are varied using an extension of the stochastic reconfiguration method, which was shown to approximate the real-time dynamics of the system. As the derived SR is one among several possible optimization techniques, we also illustrated the other conceptually different approaches that have been introduced for the variational study of open

quantum systems. We have demonstrated a proof-of-principle of the theoretical scheme by successfully benchmarking both dynamical and steady state results for a set of spins described by dissipative quantum spin models.

The neural network ansatz in terms of a RBM is highly representative of quantum correlated mixtures, while being simple to handle numerically. However, it has been shown that there are certain classes of states that cannot be accurately simulated without invoking an exponential growth in the number of parameters. We wish to remark, that the presented variational approach does not depend on the specific network topology, hence in general, it can be applied to other neural networks. Nevertheless, while a RBM ansatz can be extended to deep network representations for the wave function, for open systems the problem does not only arise from an increased difficulty in their sampling, but also from defining a model that preserves positivity for the density matrix.

However, since the problem of positivity does not emerge for wave function representations and the ansatz does not depend on the topology of the physical model, combining the neural network ansatz with most existing methods could stand as a computationally more efficient alternative to tensor network states. Nevertheless, if we consider the combination with Monte Carlo Wave Function simulations, their application could be limited to systems with slow dissipation rates, as is yet to be found a compact way to apply the quantum jump operators to the neural network state, without having to train a new network to represent the effect of a jump.

Instead, a promising approach, that could solve the scaling problem for both developed methods, would consist of applying DDQMC and the variational approach as complementary techniques. This means, that one could first find an approximate representation of the steady state density matrix with the variational RBM ansatz, then use it as a guiding function for the importance sampling in DDQMC. Contrary to existing numerical approaches which either require the full representation of the Hilbert space in memory, or rely on a properly chosen truncation of it, this combination would stochastically sample the full space of the density matrix.

We believe that this thesis has introduced novel methods that hold promise as a powerful tool in the study of open quantum systems, since both of them can be straightforwardly adopted to also represent bosonic lattice problems, and none of them depends on the spatial arrangement of the physical models. Applying them independently, we experienced a limitation with increasing system size and correlation, however, there are many future developments on the horizon, and their complementary application might break free of these problems.

Bibliography

[1] M. Planck. Zur Theorie des Gesetzes der Energieverteilung im Normalspectrum. *Verhandlungen der Deutschen Physikalischen Gesellschaft*, (2):237, 1900.

[2] A. Einstein. Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt. *AnP*, 322(6):132–148, 1905.

[3] C. Cohen-Tannoudji, B. Diu, and F. Laloe. *Quantum Mechanics, Volume 1*, volume 2. Wiley, New York, 1991.

[4] L. Landau and E. Lifshitz. *Quantum Mechanic: Non-relativistic theory*, volume 3. Elsevier Science, 3 edition, 1981.

[5] L. Landau and E. Lifshitz. *Statistical Physics*, volume 5. Elsevier Science, 3 edition, 2013.

[6] Iacopo Carusotto and Cristiano Ciuti. Quantum fluids of light. *Rev. Mod. Phys.*, 85(1):299–366, February 2013.

[7] Michael J. Hartmann. Quantum simulation with interacting photons. *J. Opt.*, 18(10), 2016.

[8] Karyn Le Hur, Loïc Henriet, Alexandru Petrescu, Kirill Plekhanov, Guillaume Roux, and Marco Schiró. Many-body quantum electrodynamics networks: Non-equilibrium condensed matter physics with light. *Comptes Rendus Physique*, 17(8):808–835, October 2016.

[9] Changsuk Noh and Dimitris G. Angelakis. Quantum simulations and many-body physics with light. *Rep. Prog. Phys.*, 80(1):016401, November 2016.

[10] Wim Casteels and Michiel Wouters. Optically bistable driven-dissipative Bose-Hubbard dimer: Gutzwiller approaches and entanglement. *Phys. Rev. A*, 95(4):043833, April 2017.

[11] E. M. Kessler, G. Giedke, A. Imamoglu, S. F. Yelin, M. D. Lukin, and J. I. Cirac. Dissipative phase transition in a central spin system. *Phys. Rev. A*, 86(1):012116, July 2012.

[12] Vincenzo Savona. Spontaneous symmetry breaking in a quadratically driven nonlinear photonic lattice. *Phys. Rev. A*, 96(3):033826, September 2017.

[13] Filippo Vicentini, Fabrizio Minganti, Riccardo Rota, Giuliano Orso, and Cristiano Ciuti. Critical slowing down in driven-dissipative Bose-Hubbard lattices. *Phys. Rev. A*, 97(1):013853, January 2018.

[14] Jiasen Jin, Alberto Biella, Oscar Viyuela, Leonardo Mazza, Jonathan Keeling, Rosario Fazio, and Davide Rossini. Cluster Mean-Field Approach to the Steady-State Phase Diagram of Dissipative Spin Systems. *Phys. Rev. X*, 6(3):031011, July 2016.

[15] Tony E. Lee, Sarang Gopalakrishnan, and Mikhail D. Lukin. Unconventional Magnetism via Optical Pumping of Interacting Spin Systems. *Phys. Rev. Lett.*, 110(25):257204, June 2013.

[16] R. Rota, F. Storme, N. Bartolo, R. Fazio, and C. Ciuti. Critical behavior of dissipative two-dimensional spin lattices. *Phys. Rev. B*, 95(13):134431, April 2017.

[17] R. Rota, F. Minganti, A. Biella, and C. Ciuti. Dynamical properties of dissipative XYZ Heisenberg lattices. *New J. Phys.*, 20(4):045003, April 2018.

[18] Wim Casteels, Ryan M. Wilson, and Michiel Wouters. Gutzwiller Monte Carlo approach for a critical dissipative spin model. *Phys. Rev. A*, 97(6):062107, June 2018.

[19] Michael J. Bremner, Ashley Montanaro, and Dan J. Shepherd. Achieving quantum supremacy with sparse and noisy commuting quantum computations. *Quantum*, 1:8, April 2017.

[20] Xun Gao and Luming Duan. Efficient classical simulation of noisy quantum computation. *arXiv:1810.03176 [quant-ph]*, October 2018. arXiv: 1810.03176.

[21] Aram W. Harrow and Ashley Montanaro. Quantum computational supremacy. *Nature*, 549(7671):203–209, September 2017.

[22] John Preskill. Quantum Computing in the NISQ era and beyond. *Quantum*, 2:79, August 2018.

[23] Heinz-Peter Breuer and Francesco Petruccione. *The Theory of Open Quantum Systems*. Oxford University Press, January 2007.

[24] Tomaž Prosen. Exact Nonequilibrium Steady State of a Strongly Driven Open XXZ Chain. *Phys. Rev. Lett.*, 107(13):137201, September 2011.

[25] Tomaž Prosen. Exact Nonequilibrium Steady State of an Open Hubbard Chain. *Phys. Rev. Lett.*, 112(3):030603, January 2014.

[26] Lee Badger. Lazzarini's Lucky Approximation of π. *Mathematics Magazine*, 67(2):83–91, 1994.

[27] Jian Cui, J. Ignacio Cirac, and Mari Carmen Bañuls. Variational Matrix Product Operators for the Steady State of Dissipative Quantum Systems. *Phys. Rev. Lett.*, 114(22):220601, June 2015.

[28] Augustine Kshetrimayum, Hendrik Weimer, and Román Orús. A simple tensor network algorithm for two-dimensional steady states. *Nat Commun*, 8(1):1–7, November 2017.

[29] Eduardo Mascarenhas, Hugo Flayac, and Vincenzo Savona. Matrix-product-operator approach to the nonequilibrium steady state of driven-dissipative quantum arrays. *Phys. Rev. A*, 92(2):022116, August 2015.

[30] S. Finazzi, A. Le Boité, F. Storme, A. Baksic, and C. Ciuti. Corner-Space Renormalization Method for Driven-Dissipative Two-Dimensional Correlated Systems. *Phys. Rev. Lett.*, 115(8):080604, August 2015.

[31] C. J. Umrigar. Observations on variational and projector Monte Carlo methods. *J. Chem. Phys.*, 143(16):164105, October 2015.

[32] Alexandra Nagy and Vincenzo Savona. Driven-dissipative quantum Monte Carlo method for open quantum systems. *Phys. Rev. A*, 97(5):052129, May 2018.

[33] Alexandra Nagy and Vincenzo Savona. Variational Quantum Monte Carlo Method with a Neural-Network Ansatz for Open Quantum Systems. *Phys. Rev. Lett.*, 122(25):250501, June 2019.

[34] Alexandra Nagy. Epflltpn/DDQMC.

[35] Alexandra Nagy. Epflltpn/NNDM.

[36] Serge Haroche and Jean-Michel Raimond. *Exploring the Quantum: Atoms, Cavities, and Photons*. Oxford Graduate Texts. Oxford University Press, Oxford, New York, April 2013.

[37] Victor V. Albert and Liang Jiang. Symmetries and conserved quantities in Lindblad master equations. *Phys. Rev. A*, 89(2):022118, February 2014.

[38] Davide Nigro. On the uniqueness of the steady-state solution of the Lindblad–Gorini–Kossakowski–Sudarshan equation. *J. Stat. Mech.*, 2019(4):043202, April 2019.

[39] Fabrizio Minganti, Alberto Biella, Nicola Bartolo, and Cristiano Ciuti. Spectral theory of Liouvillians for dissipative phase transitions. *Phys. Rev. A*, 98(4):042118, October 2018.

[40] Fabrizio Minganti. Out-of-Equilibrium Phase Transitions in Nonlinear Optical Systems. page 173, 2019.

[41] Michael A. Nielsen and Isaac L. Chuang. *Quantum computation and quantum information*. Cambridge University Press, Cambridge ; New York, 10th anniversary ed edition, 2010.

[42] Bassano Vacchini and Giulio Amato. Reduced dynamical maps in the presence of initial correlations. *Scientific Reports*, 6:37328, November 2016.

[43] A. R. Usha Devi, A. K. Rajagopal, and Sudha. Open-system quantum dynamics with correlated initial states, not completely positive maps, and non-Markovianity. *Phys. Rev. A*, 83(2):022109, February 2011.

[44] Kavan Modi. Operational approach to open dynamics and quantifying initial correlations. *Scientific Reports*, 2:581, August 2012.

[45] Matthias Jakob and Stig Stenholm. Variational functions in driven open quantum systems. *Phys. Rev. A*, 67(3):032111, March 2003.

[46] N. Gisin and I. C. Percival. The quantum-state diffusion model applied to open systems. *J. Phys. A: Math. Gen.*, 25(21):5677–5691, November 1992.

[47] Hendrik Weimer, Augustine Kshetrimayum, and Román Orús. Simulation methods for open quantum many-body systems. *arXiv:1907.07079 [cond-mat, physics:physics, physics:quant-ph]*, July 2019. arXiv: 1907.07079.

[48] Bernhard Baumgartner and Heide Narnhofer. Analysis of quantum semigroups with GKS–Lindblad generators: II. General. *J. Phys. A: Math. Theor.*, 41(39):395303, September 2008.

[49] Marko Žnidari□. Relaxation times of dissipative many-body quantum systems. *Phys. Rev. E*, 92(4):042143, October 2015.

[50] Tosio Kato. *Perturbation Theory for Linear Operators*. Classics in Mathematics. Springer-Verlag, Berlin Heidelberg, 2 edition, 1995.

[51] J. Eisert and T. Prosen. Noise-driven quantum criticality. *arXiv:1012.5013 [cond-mat, physics:quant-ph]*, December 2010. arXiv: 1012.5013.

[52] Alberto Biella, Jiasen Jin, Oscar Viyuela, Cristiano Ciuti, Rosario Fazio, and Davide Rossini. Linked cluster expansions for open quantum systems on a lattice. *Phys. Rev. B*, 97(3):035103, January 2018.

[53] N. Henkel, R. Nath, and T. Pohl. Three-Dimensional Roton Excitations and Supersolid Formation in Rydberg-Excited Bose-Einstein Condensates. *Phys. Rev. Lett.*, 104(19):195302, May 2010.

[54] Isabelle Bouchoule and Klaus Mølmer. Spin squeezing of atoms by the dipole interaction in virtually excited Rydberg states. *Phys. Rev. A*, 65(4):041803, April 2002.

[55] M. D. Lukin, M. Fleischhauer, R. Cote, L. M. Duan, D. Jaksch, J. I. Cirac, and P. Zoller. Dipole Blockade and Quantum Information Processing in Mesoscopic Atomic Ensembles. *Phys. Rev. Lett.*, 87(3):037901, June 2001.

[56] M. Saffman, T. G. Walker, and K. Mølmer. Quantum information with Rydberg atoms. *Rev. Mod. Phys.*, 82(3):2313–2363, August 2010.

[57] T. Lahaye, C. Menotti, L. Santos, M. Lewenstein, and T. Pfau. The physics of dipolar bosonic quantum gases. *Rep. Prog. Phys.*, 72(12):126401, November 2009.

[58] Sachdev, Subir. *Quantum phase transitions 2nd edition | Condensed matter physics, nanoscience and mesoscopic physics*. Cambridge University Press, 2 edition, 2011.

[59] Henk W. J. Blöte and Youjin Deng. Cluster Monte Carlo simulation of the transverse Ising model. *Phys. Rev. E*, 66(6):066110, December 2002.

[60] Jiasen Jin, Alberto Biella, Oscar Viyuela, Cristiano Ciuti, Rosario Fazio, and Davide Rossini. Phase diagram of the dissipative quantum Ising model on a square lattice. *Phys. Rev. B*, 98(24):241108, December 2018.

[61] Filippo Vicentini, Alberto Biella, Nicolas Regnault, and Cristiano Ciuti. Variational Neural-Network Ansatz for Steady States in Open Quantum Systems. *Phys. Rev. Lett.*, 122(25):250503, June 2019.

[62] Vincent Lienhard, Sylvain de Léséleuc, Daniel Barredo, Thierry Lahaye, Antoine Browaeys, Michael Schuler, Louis-Paul Henry, and Andreas M. Läuchli. Observing the Space- and Time-Dependent Growth of Correlations in Dynamically Tuned Synthetic Ising Models with Antiferromagnetic Interactions. *Phys. Rev. X*, 8(2):021070, June 2018.

[63] Antoine Browaeys, Daniel Barredo, and Thierry Lahaye. Experimental investigations of dipole–dipole interactions between a few Rydberg atoms. *J. Phys. B: At. Mol. Opt. Phys.*, 49(15):152001, June 2016.

[64] Henning Labuhn, Daniel Barredo, Sylvain Ravets, Sylvain de Léséleuc, Tommaso Macrì, Thierry Lahaye, and Antoine Browaeys. Tunable two-dimensional arrays of single Rydberg atoms for realizing quantum Ising models. *Nature*, 534(7609):667–670, June 2016.

[65] Sylvain de Léséleuc, Sebastian Weber, Vincent Lienhard, Daniel Barredo, Hans Peter Büchler, Thierry Lahaye, and Antoine Browaeys. Accurate Mapping of Multilevel Rydberg Atoms on Interacting Spin-1/2 Particles for the Quantum Simulation of Ising Models. *Phys. Rev. Lett.*, 120(11):113602, March 2018.

[66] Matteo Marcuzzi, Emanuele Levi, Sebastian Diehl, Juan P. Garrahan, and Igor Lesanovsky. Universal Nonequilibrium Properties of Dissipative Rydberg Gases. *Phys. Rev. Lett.*, 113(21):210401, November 2014.

[67] Sylvain de Léséleuc, Daniel Barredo, Vincent Lienhard, Antoine Browaeys, and Thierry Lahaye. Analysis of imperfections in the coherent optical excitation of single atoms to Rydberg states. *Phys. Rev. A*, 97(5):053803, May 2018.

[68] Sebastian Diehl, Andrea Tomadin, Andrea Micheli, Rosario Fazio, and Peter Zoller. Dynamical Phase Transitions and Instabilities in Open Atomic Many-Body Systems. *Phys. Rev. Lett.*, 105(1):015702, July 2010.

[69] Tony E. Lee and M. C. Cross. Spatiotemporal dynamics of quantum jumps with Rydberg atoms. *Phys. Rev. A*, 85(6):063822, June 2012.

[70] Tony E. Lee, H. Häffner, and M. C. Cross. Antiferromagnetic phase transition in a nonequilibrium lattice of Rydberg atoms. *Phys. Rev. A*, 84(3):031402, September 2011.

[71] Jiasen Jin, Davide Rossini, Rosario Fazio, Martin Leib, and Michael J. Hartmann. Photon Solid Phases in Driven Arrays of Nonlinearly Coupled Cavities. *Phys. Rev. Lett.*, 110(16):163605, April 2013.

[72] Alexandre Le Boité, Giuliano Orso, and Cristiano Ciuti. Steady-State Phases and Tunneling-Induced Instabilities in the Driven Dissipative Bose-Hubbard Model. *Phys. Rev. Lett.*, 110(23):233601, June 2013.

[73] Mohammad F. Maghrebi and Alexey V. Gorshkov. Nonequilibrium many-body steady states via Keldysh formalism. *Phys. Rev. B*, 93(1):014307, January 2016.

[74] Riccardo Rota, Fabrizio Minganti, Cristiano Ciuti, and Vincenzo Savona. Quantum Critical Regime in a Quadratically Driven Nonlinear Photonic Lattice. *Phys. Rev. Lett.*, 122(11):110405, March 2019.

[75] Amit Dutta, Gabriel Aeppli, Bikas K. Chakrabarti, Uma Divakaran, Thomas F. Rosenbaum, and Diptiman Sen. Quantum phase transitions in transverse field spin models: from statistical physics to quantum information. *arXiv:1012.0653 [cond-mat, physics:quant-ph]*, December 2010. arXiv: 1012.0653.

[76] C. W. Gardiner, A. S. Parkins, and P. Zoller. Wave-function quantum stochastic differential equations and quantum-jump simulation methods. *Phys. Rev. A*, 46(7):4363–4381, October 1992.

[77] Jean Dalibard, Yvan Castin, and Klaus Mølmer. Wave-function approach to dissipative processes in quantum optics. *Phys. Rev. Lett.*, 68(5):580–583, February 1992.

[78] R. Dum, P. Zoller, and H. Ritsch. Monte Carlo simulation of the atomic master equation for spontaneous emission. *Phys. Rev. A*, 45(7):4879–4887, April 1992.

[79] H. J. Carmichael. Quantum trajectory theory for cascaded open systems. *Phys. Rev. Lett.*, 70(15):2273–2276, April 1993.

[80] Andrew J. Daley. Quantum trajectories and open many-body quantum systems. *Advances in Physics*, 63(2):77–149, March 2014.

[81] Veronica Øverbye. Optical Cavities as Single Photon Sources. page 133, 2012.

[82] D. Huybrechts and M. Wouters. Cluster methods for the description of a driven-dissipative spin model. *Phys. Rev. A*, 99(4):043841, April 2019.

[83] J. R. Johansson, P. D. Nation, and Franco Nori. QuTiP: An open-source Python framework for the dynamics of open quantum systems. *Computer Physics Communications*, 183(8):1760–1772, August 2012.

[84] J. R. Johansson, P. D. Nation, and Franco Nori. QuTiP 2: A Python framework for the dynamics of open quantum systems. *Computer Physics Communications*, 184(4):1234–1240, April 2013.

[85] J. Eisert, M. Cramer, and M. B. Plenio. Colloquium: Area laws for the entanglement entropy. *Rev. Mod. Phys.*, 82(1):277–306, February 2010.

[86] J. Eisert. Entanglement and tensor network states. *arXiv:1308.3318 [cond-mat, physics:quant-ph]*, September 2013. arXiv: 1308.3318.

[87] Yimin Ge and Jens Eisert. Area laws and efficient descriptions of quantum many-body states. *New J. Phys.*, 18(8):083026, August 2016.

[88] Steven R. White. Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.*, 69(19):2863–2866, November 1992.

[89] Steven R. White. Density-matrix algorithms for quantum renormalization groups. *Phys. Rev. B*, 48(14):10345–10356, October 1993.

[90] F. Verstraete, V. Murg, and J. I. Cirac. Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems. *Advances in Physics*, 57(2):143–224, March 2008.

[91] F. Verstraete, D. Porras, and J. I. Cirac. Density Matrix Renormalization Group and Periodic Boundary Conditions: A Quantum Information Perspective. *Phys. Rev. Lett.*, 93(22):227205, November 2004.

[92] Guifré Vidal. Efficient Classical Simulation of Slightly Entangled Quantum Computations. *Phys. Rev. Lett.*, 91(14):147902, October 2003.

[93] Norbert Schuch, Michael M. Wolf, Frank Verstraete, and J. Ignacio Cirac. Computational Complexity of Projected Entangled Pair States. *Phys. Rev. Lett.*, 98(14):140506, April 2007.

[94] U. Schollwöck. The density-matrix renormalization group. *Rev. Mod. Phys.*, 77(1):259–315, April 2005.

[95] Gemma De las Cuevas, Norbert Schuch, David Pérez-García, and J. Ignacio Cirac. Purifications of multipartite states: limitations and constructive methods. *New J. Phys.*, 15(12):123021, December 2013.

[96] 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, November 2004.

[97] Zi Cai and Thomas Barthel. Algebraic versus Exponential Decoherence in Dissipative Many-Particle Systems. *Phys. Rev. Lett.*, 111(15):150403, October 2013.

[98] Chaitanya Joshi, Felix Nissen, and Jonathan Keeling. Quantum correlations in the one-dimensional driven dissipative XY model. *Phys. Rev. A*, 88(6):063835, December 2013.

[99] Lars Bonnes, Daniel Charrier, and Andreas M. Läuchli. Dynamical and steady-state properties of a Bose-Hubbard chain with bond dissipation: A study based on matrix product operators. *Phys. Rev. A*, 90(3):033612, September 2014.

[100] J. J. Mendoza-Arenas, T. Grujic, D. Jaksch, and S. R. Clark. Dephasing enhanced transport in nonequilibrium strongly correlated quantum systems. *Phys. Rev. B*, 87(23):235130, June 2013.

[101] Tomaž Prosen and Marko Žnidari. Matrix product simulations of non-equilibrium steady states of quantum spin chains. *J. Stat. Mech.*, 2009(02):P02035, February 2009.

[102] Michael J. Hartmann. Polariton Crystallization in Driven Arrays of Lossy Nonlinear Resonators. *Phys. Rev. Lett.*, 104(11):113601, March 2010.

[103] J. Ignacio Cirac, José Garre-Rubio, and David Pérez-García. Mathematical open problems in Projected Entangled Pair States. *arXiv:1903.09439 [cond-mat, physics:math-ph, physics:quant-ph]*, March 2019. arXiv: 1903.09439.

[104] Philippe Corboz and Frédéric Mila. Crystals of Bound States in the Magnetization Plateaus of the Shastry-Sutherland Model. *Phys. Rev. Lett.*, 112(14):147203, April 2014.

[105] H. J. Liao, Z. Y. Xie, J. Chen, Z. Y. Liu, H. D. Xie, R. Z. Huang, B. Normand, and T. Xiang. Gapless Spin-Liquid Ground State in the $S=1/2$ Kagome Antiferromagnet. *Phys. Rev. Lett.*, 118(13):137202, March 2017.

[106] Y. H. Matsuda, N. Abe, S. Takeyama, H. Kageyama, P. Corboz, A. Honecker, S. R. Manmana, G. R. Foltin, K. P. Schmidt, and F. Mila. Magnetization of SrCu2 BO32 in Ultrahigh Magnetic Fields up to 118 T. *Phys. Rev. Lett.*, 111(13):137204, September 2013.

[107] Jindrich Kolorenc and Lubos Mitas. Applications of quantum Monte Carlo methods in condensed systems. *Rep. Prog. Phys.*, 74(2):026502, January 2011.

[108] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal. Quantum Monte Carlo simulations of solids. *Rev. Mod. Phys.*, 73(1):33–83, January 2001.

[109] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes in C (2Nd Ed.): The Art of Scientific Computing*. Cambridge University Press, New York, NY, USA, 1992.

[110] Ioan Kosztin, Byron Faber, and Klaus Schulten. Introduction to the diffusion Monte Carlo method. *American Journal of Physics*, 64(5):633–644, May 1996.

[111] C. J. Umrigar, M. P. Nightingale, and K. J. Runge. A diffusion Monte Carlo algorithm with very small time□step errors. *J. Chem. Phys.*, 99(4):2865–2890, August 1993.

[112] M. H. Kalos. Monte Carlo Calculations of the Ground State of Three- and Four-Body Nuclei. *Phys. Rev.*, 128(4):1791–1795, November 1962.

[113] D. M. Ceperley and M. H. Kalos. Quantum Many-Body Problems. In Kurt Binder, editor, *Monte Carlo Methods in Statistical Physics*, Topics in Current Physics, pages 145–194. Springer, Berlin, Heidelberg, 1986.

[114] George H. Booth, Alex J. W. Thom, and Ali Alavi. Fermion Monte Carlo without fixed nodes: A game of life, death, and annihilation in Slater determinant space. *J. Chem. Phys.*, 131(5):054106, August 2009.

[115] J. S. Spencer, N. S. Blunt, and W. M.c. Foulkes. The sign problem and population dynamics in the full configuration interaction quantum Monte Carlo method. *J. Chem. Phys.*, 136(5):054110, February 2012.

[116] George H. Booth and Ali Alavi. Approaching chemical accuracy using full configuration-interaction quantum Monte Carlo: A study of ionization potentials. *J. Chem. Phys.*, 132(17):174104, May 2010.

[117] Deidre Cleland, George H. Booth, and Ali Alavi. Communications: Survival of the fittest: Accelerating convergence in full configuration-interaction quantum Monte Carlo. *J. Chem. Phys.*, 132(4):041103, January 2010.

[118] George H. Booth, Andreas Grüneis, Georg Kresse, and Ali Alavi. Towards an exact description of electronic wavefunctions in real solids. *Nature*, 493(7432):365–370, January 2013.

[119] F. R. Petruzielo, A. A. Holmes, Hitesh J. Changlani, M. P. Nightingale, and C. J. Umrigar. Semistochastic Projector Monte Carlo Method. *Phys. Rev. Lett.*, 109(23):230201, December 2012.

[120] Jarrod R. McClean and Alán Aspuru-Guzik. Clock quantum Monte Carlo technique: An imaginary-time method for real-time quantum dynamics. *Phys. Rev. A*, 91(1):012311, January 2015.

[121] N. S. Blunt, T. W. Rogers, J. S. Spencer, and W. M. C. Foulkes. Density-matrix quantum Monte Carlo method. *Phys. Rev. B*, 89(24):245124, June 2014.

[122] Fionn D. Malone, N. S. Blunt, Ethan W. Brown, D. K. K. Lee, J. S. Spencer, W. M. C. Foulkes, and James J. Shepherd. Accurate Exchange-Correlation Energies for the Warm Dense Electron Gas. *Phys. Rev. Lett.*, 117(11):115701, September 2016.

[123] Fionn D. Malone, N. S. Blunt, James J. Shepherd, D. K. K. Lee, J. S. Spencer, and W. M. C. Foulkes. Interaction picture density matrix quantum Monte Carlo. *J. Chem. Phys.*, 143(4):044116, July 2015.

[124] Zi-Xiang Li and Hong Yao. Sign-Problem-Free Fermionic Quantum Monte Carlo: Developments and Applications. *Annu. Rev. Condens. Matter Phys.*, 10(1):337–356, March 2019.

[125] Mauro Iazzi, Alexey A. Soluyanov, and Matthias Troyer. Topological origin of the fermion sign problem. *Phys. Rev. B*, 93(11):115102, March 2016.

[126] George H. Booth, Simon D. Smart, and Ali Alavi. Linear-scaling and parallelisable algorithms for stochastic quantum chemistry. *Molecular Physics*, 112(14):1855–1869, July 2014.

[127] James J. Shepherd, Gustavo E. Scuseria, and James S. Spencer. Sign problem in full configuration interaction quantum Monte Carlo: Linear and sublinear representation regimes for the exact wave function. *Phys. Rev. B*, 90(15):155130, October 2014.

[128] Voratas Kachitvichyanukul and Bruce W. Schmeiser. Binomial Random Variate Generation. *Commun. ACM*, 31(2):216–222, February 1988.

[129] Charles S. Davis. The computer generation of multinomial random variates. *Computational Statistics & Data Analysis*, 16(2):205–217, August 1993.

[130] Wolfgang Hörmann. The generation of binomial random variates. *Journal of Statistical Computation and Simulation*, 46(1-2):101–110, April 1993.

[131] MVAPICH :: http://mvapich.cse.ohio-state.edu/.

[132] D.C. Merkle. *Advances in Cryptology - CRYPTO '89: Proceedings*. Lecture Notes in Computer Science. Springer-Verlag, New York, 1990.

[133] Karl Roald Leikanger. Full Configuration Interaction Monte Carlo Studies of Quantum Dots. 2013.

[134] Jutho Haegeman, J. Ignacio Cirac, Tobias J. Osborne, Iztok Pižorn, Henri Verschelde, and Frank Verstraete. Time-Dependent Variational Principle for Quantum Lattices. *Phys. Rev. Lett.*, 107(7):070601, August 2011.

[135] Hendrik Weimer. Variational Principle for Steady States of Dissipative Quantum Many-Body Systems. *Phys. Rev. Lett.*, 114(4):040402, January 2015.

[136] Michael J. Hartmann and Giuseppe Carleo. Neural-Network Approach to Dissipative Quantum Many-Body Dynamics. *Phys. Rev. Lett.*, 122(25):250502, June 2019.

[137] Nobuyuki Yoshioka and Ryusuke Hamazaki. Constructing neural stationary states for open quantum many-body systems. *Phys. Rev. B*, 99(21):214306, June 2019.

[138] K. E. Schmidt, Parhat Niyaz, A. Vaught, and Michael A. Lee. Green's function Monte Carlo method with exact imaginary-time propagation. *Phys. Rev. E*, 71(1):016707, January 2005.

[139] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of State Calculations by Fast Computing Machines. *J. Chem. Phys.*, 21(6):1087–1092, June 1953.

[140] Walter Ritz. Über eine neue Methode zur Lösung gewisser Variationsprobleme der mathematischen Physik. *Journal für die reine und angewandte Mathematik (Crelles Journal)*, 1909(135):1–61, 1909.

[141] Sandro Sorella and Federico Becca. SISSA Lecture notes on Numerical methods for strongly correlated electrons. 2016.

[142] Malvin H. Kalos and Paula A. Whitlock. *Monte Carlo Methods*. Wiley, New York, NY, USA, 2008.

[143] Sandro Sorella, Michele Casula, and Dario Rocca. Weak binding between two aromatic rings: Feeling the van der Waals attraction by quantum Monte Carlo methods. *J. Chem. Phys.*, 127(1):014105, July 2007.

[144] Robert Rüger. Implementation of the Variational Monte Carlo method for the Hubbard model.

[145] Sandro Sorella. Generalized Lanczos algorithm for variational quantum Monte Carlo. *Phys. Rev. B*, 64(2):024512, June 2001.

[146] Giuseppe Carleo and Matthias Troyer. Solving the quantum many-body problem with artificial neural networks. *Science*, 355(6325):602–606, February 2017.

[147] Giuseppe Carleo, Yusuke Nomura, and Masatoshi Imada. Constructing exact representations of quantum many-body systems with deep neural networks. *Nat Commun*, 9(1):1–11, December 2018.

[148] Zi Cai and Jinguo Liu. Approximating quantum many-body wave functions using artificial neural networks. *Phys. Rev. B*, 97(3):035116, January 2018.

[149] Ivan Glasser, Nicola Pancotti, Moritz August, Ivan D. Rodriguez, and J. Ignacio Cirac. Neural-Network Quantum States, String-Bond States, and Chiral Topological States. *Phys. Rev. X*, 8(1):011006, January 2018.

[150] Yusuke Nomura, Andrew S. Darmawan, Youhei Yamaji, and Masatoshi Imada. Restricted Boltzmann machine learning for solving strongly correlated quantum systems. *Phys. Rev. B*, 96(20):205152, November 2017.

[151] Xun Gao and Lu-Ming Duan. Efficient representation of quantum many-body states with deep neural networks. *Nat Commun*, 8(1):1–6, September 2017.

[152] Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar. *Foundations of machine learning*. Adaptive computation and machine learning series. MIT Press, Cambridge, MA, 2012.

[153] Shai Shalev-Shwartz and Shai Ben-David. *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press, Cambridge, 2014.

[154] Martijn van Otterlo and Marco Wiering. Reinforcement Learning and Markov Decision Processes. In Marco Wiering and Martijn van Otterlo, editors, *Reinforcement Learning: State-of-the-Art*, Adaptation, Learning, and Optimization, pages 3–42. Springer, Berlin, Heidelberg, 2012.

[155] Jin-Guo Liu, Shuo-Hui Li, and Lei Wang. Lecture Note on Deep Learning and Quantum Many-Body Computation. page 58.

[156] Giuseppe Carleo, Ignacio Cirac, Kyle Cranmer, Laurent Daudet, Maria Schuld, Naftali Tishby, Leslie Vogt-Maranto, and Lenka Zdeborová. Machine learning and the physical sciences. *arXiv:1903.10563 [astro-ph, physics:cond-mat, physics:hep-th, physics:physics, physics:quant-ph]*, March 2019. arXiv: 1903.10563.

[157] Pecht, G. Michael and Kang, Myeongsu. *Prognostics and Health Management of Electronics: Fundamentals, Machine Learning, and the Internet of Things*. Wiley, 2019.

[158] A L Samuel. Some Studies in Machine Learning Using the Gameof Checkers. II-Recent Progress. *MACHINE LEARNING*.

[159] G. Cybenko. Approximation by superpositions of a sigmoidal function. *Math. Control Signal Systems*, 2(4):303–314, December 1989.

[160] Kurt Hornik. Approximation capabilities of multilayer feedforward networks. *Neural Networks*, 4(2):251–257, January 1991.

[161] Karpathy, Andrej, Abbeel, Pieter, Brockman, Greg, Chen, Peter, Cheung, Vicki, Duan, Rocky, GoodFellow, Ian, and Kingma, Durk. Generative Models, June 2016.

[162] Diederik P. Kingma and Max Welling. Auto-Encoding Variational Bayes. *arXiv:1312.6114 [cs, stat]*, May 2014. arXiv: 1312.6114.

[163] Laurent Dinh, David Krueger, and Yoshua Bengio. NICE: Non-linear Independent Components Estimation. *arXiv:1410.8516 [cs]*, April 2015. arXiv: 1410.8516.

[164] Diederik P. Kingma, Tim Salimans, Rafal Jozefowicz, Xi Chen, Ilya Sutskever, and Max Welling. Improving Variational Inference with Inverse Autoregressive Flow. *arXiv:1606.04934 [cs, stat]*, January 2017. arXiv: 1606.04934.

[165] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative Adversarial Networks. *arXiv:1406.2661 [cs, stat]*, June 2014. arXiv: 1406.2661.

[166] Geoffrey E. Hinton. Training Products of Experts by Minimizing Contrastive Divergence. *Neural Comput.*, 14(8):1771–1800, August 2002.

[167] Jing Chen, Song Cheng, Haidong Xie, Lei Wang, and Tao Xiang. Equivalence of restricted Boltzmann machines and tensor network states. *Phys. Rev. B*, 97(8):085104, February 2018. arXiv: 1701.04831.

[168] Ruslan Salakhutdinov. Learning Deep Generative Models. *Annu. Rev. Stat. Appl.*, 2(1):361–385, April 2015.

[169] Richard S Sutton and Andrew G Barto. Reinforcement Learning: An Introduction.

[170] Puterman,Martin L. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley, 2005.

[171] Giacomo Torlai and Roger G. Melko. Latent Space Purification via Neural Density Operators. *Phys. Rev. Lett.*, 120(24):240503, June 2018.

[172] Nicolas Le Roux and Yoshua Bengio. Representational Power of Restricted Boltzmann Machines and Deep Belief Networks. *Neural Computation*, 20(6):1631–1649, February 2008.

[173] Nahuel Freitas, Giovanna Morigi, and Vedran Dunjko. Neural network operations and Susuki–Trotter evolution of neural network states. *Int. J. Quantum Inform.*, 16(08):1840008, November 2018.

[174] Sou-Cheng T. Choi, Christopher C. Paige, and Michael A. Saunders. MINRES-QLP: A Krylov Subspace Method for Indefinite or Singular Symmetric Systems. *SIAM J. Sci. Comput.*, 33(4):1810–1836, January 2011.

[175] Y. Liu. syangliu/MINRES-QLP.

[176] Kenny Choo, Giuseppe Carleo, Nicolas Regnault, and Titus Neupert. Symmetries and Many-Body Excitations with Neural-Network Quantum States. *Phys. Rev. Lett.*, 121(16):167204, October 2018.

[177] Hiroki Saito and Masaya Kato. Machine Learning Technique to Find Quantum Many-Body Ground States of Bosons on a Lattice. *J. Phys. Soc. Jpn.*, 87(1):014001, December 2017.