SISSA

International School for Advanced Studies

---

PhD course in Statistical Physics

VII cycle

Academic Year 2017/2018

# Simulating quantum annealing via projective quantum Monte Carlo algorithms


Thesis submitted for the degree of

*Doctor Philosophiae*

*Candidate:*

Inack Estelle Maéva

*Supervisors:*

Sebastiano Pilati and Giuseppe E. Santoro

Acquista verità e non la vendere, acquista sapienza,
istruzione e intelligenza.
— Proverbi 23:23

# Abstract

In this thesis, we implement projective quantum Monte Carlo (PQMC) methods to simulate quantum annealing on classical computers. A detailed derivation of the algorithm is presented. We show that in the regime where the systematic errors are well controlled, PQMC algorithms are indeed capable of simulating the imaginary-time dynamics of the Schroedinger equation both on continuous space model and discrete basis systems. We also demonstrate that the tunneling time of the PQMC method is quadratically faster than the one of incoherent quantum tunneling. It also shows remarkable stability when applied to frustrated systems compared to the path integral Monte Carlo algorithm, the algorithm mostly chosen to do comparisons with quantum annealers. However, a major drawback of the method comes from the finite number of random walkers needed to implement the simulations. It grows exponentially with the system size when no or poor guiding wave-functions are utilized. Nevertheless, we demonstrated that when good enough guiding wave-functions are used – in our case we choose artificial neural networks – the computational complexity seems to go from exponential to polynomial in the system size. We advocate for a search of more efficient guiding wave functions since they determine when and how feasible the PQMC simulations are on classical computers, a question closely related to a provable need or speed-up of a quantum computer.

# Contents

# Introduction

Half a century ago, the advent of integrated circuits led to a complete change of the design of computers, steering up the birth of the computer revolution. From that moment, the size of computers started to considerably reduce, from gigantic computers in labs, to personal computers in homes, to mobiles phones in pockets. However, though the number of transistors have been doubling every two years, according to Moore's law, it is now very clear to physicists and engineers that it is about to reach its limit. This phenomena is due to increasing quantum effects in the chip.

An alternative that researchers have been pursuing to palliate to the possible death of Moore's law is to build up computers that function exclusively under the laws of quantum mechanics. A traditional computer (which we will refer to as classical computer) as we know it, stores information in bits strings, where each bit can be either in the state $0$ or $1$. Whereas in a *quantum computer*, information is stored in a qubit, the quantum variant of the bit which can be simultaneously both in $0$ and $1$ states. This is an inherently quantum phenomenon called superposition, which together with other quantum properties such as coherence and entanglement is believed to provide an exponential speed-up in certain calculations compared with classical computers.

However, the road to a quantum computer revolution is full of pitfalls, the major one being the extreme sensibility of a quantum state to external perturbation, leading to its collapse into a classical state. This effect is called decoherence, the nightmare of many quantum physicists. Nevertheless, researchers have been searching for potential applications where quantum computers may have a definite superiority compared to their classical counterpart even when using "noisy" qubits. One of those applications is solving optimization problem through adiabatic quantum computation sometimes referred to as quantum annealing (QA) [1, 2].

Currently, the state of art in quantum annealing technology is owned by DWave, with their $2000Q$ system, a quantum processor chip with $2000$ superconducting qubits. Very recently also, Google announced to have built an $4$-leaf Shamrock [3] superconducting quantum annealer. The usual way in which quantum annealers (QAs) are tested to inspect if they exploit quantum effects is by performing quantum Monte Carlo (QMC) simulations on classical computers [4, 5].

QMC methods are stochastic methods used to simulate properties of quantum systems at finite or zero temperature [6, 7, 8, 9, 10]. They have also been used as heuristic optimization methods in various ways: to find the minimal configurations of Lennard-Jones potentials [11], to solve combinatorial optimization problems [12, 13, 14, 15, 16, 17], to compare with the performance and ascertain the quantum nature of quantum annealers [18, 19, 20]. Very recently, they were successfully used to assess the tunneling time of adiabatic quantum computations on quantum Ising models [21] (considered to be systems effectively described by a double well structure) and on continuous space models [21, 22].

It is important to mention that among the plethora of QMC methods available, there is a particular one that has been standing out as a term of comparison for the behavior of QAs, namely path integral Monte Carlo (PIMC). PIMC is a QMC method that simulates quantum many-body systems at finite temperature. When used to simulate QA, it implements a stochastic Monte Carlo dynamics to explore the configuration space, sometimes without succeeding. While this stochastic dynamics might effectively mimic the (real time) behavior of a quantum annealer, it is in principle distinct from the real-time Schroedinger dynamics.

In this thesis, we investigate the use of another type of QMC method to simulate quantum annealing namely the projective quantum Monte Carlo (PQMC) algorithm. It is a QMC method that emulates the dynamics of a quantum system in imaginary-time. One of the motivations is the theory of Refs. [23, 1], showing that in the adiabatic perturbative regime, imaginary-time Schroedinger dynamics is expected to simulate (in long annealing times) the behavior of quantum systems in real-time. On the contrary, when perturbative terms are present, imaginary-time annealing has been proved to provide quadratic speed-up with respect to its real-time counterpart [24].

Therefore, one of the goals is to compare the performance of PQMC with respect to PIMC. Hence, it is also essential to identify, understand, and learn how to control all the possible flaws and biases of PQMC simulations. We also explore a

route to boost the performance of the PQMC method using guiding wave functions based on artificial neural networks states [25].

# GUIDE TO THE THESIS

**Chapter 1:** We describe in details the PQMC method for continuous space models and for discrete basis models with and without importance sampling. In particular, in the prospect of quantum annealing simulations, we derive a generic equation for the PQMC algorithm that has a time-dependent guiding wave-function. In the last section of the chapter, we introduce a novel way of recovering pure estimators from PQMC simulations without importance sampling taking as a test bed the ferromagnetic quantum Ising chain.

**Chapter 2:** In this chapter, we use the diffusion Monte Carlo (DMC) to simulate quantum annealing problems on continuous space models. Comparative studies of DMC vs PIMC on double-well models are shown. We also report results of performing simulated quantum annealing (SQA) with DMC in comparison with classical annealing on multiwell potentials. The content of the chapter has been published in:
E. M. Inack and S. Pilati, Phys. Rev. E **92**, 053304 (2015).

**Chapter 3:** Here, comparisons are made between the tunneling dynamics of simple DMC simulations, of finite temperature PIMC and of incoherent quantum annealing on the Ising chain and a frustrated spin model. We also study the computational complexity of the simple DMC algorithm, in the context of finding ground state properties of local Hamiltonians with a given precision. The test bed is the quantum Ising chain (QIC). The results of the chapter are taken from the recent publication
E. M. Inack, G. Giudici, T. Parolini, G. Santoro and S. Pilati, Phys. Rev. A **97**, 032307 (2018).

**Chapter 4:** We introduce importance sampling making use of the unrestricted Boltzmann machine (uRBM) as a trial wave function to study the computational complexity of the QIC. Prior to that, we compare the variational ground-state energy predictions for the 1D ferromagnetic quantum Ising models, obtained using the uRBM introduced here, with the RBM recently introduced by Carleo and Troyer [25]. The results presented here come from a recent preprint
E. M. Inack, G. Santoro, L. Dell'Anna and S. Pilati, arXiv:1809.03562v1

**Chapter 5:** SQA using PMQC with and without importance sampling is performed on the QIC. Results are compared with both the real and imaginary time dynamics of the Schroedinger equation. We also test our protocol of recovering pure estimators in PQMC, this time at the end of annealing simulations.

# 1 Fundamentals of projective quantum Monte Carlo algorithms

In this chapter we present projective Quantum Monte Carlo (PQMC) methods. They allow one to extract certain ground-state properties of quantum many-body systems. We will first discuss the main idea behind their ability to sample the ground state wave-function of Hamiltonians that are not affected by the sign problem by giving a brief overview of PQMC methods.

We then move on to give a detailed description of the algorithm, for both continuous space and discrete basis models. We particularly highlight and analyze the possible sources of systematic errors in the method, which are the time step error and the one due to a finite population of random-walkers. Understanding those errors is very important because it can be a determining factor when using the PQMC algorithm to inspect for potential quantum speed-up in using adiabatic quantum computers instead of stochastic simulations performed on classical computers. We describe the importance sampling technique and its relevance in improving the efficiency of PQMC simulations, also in the context of solving optimization problems. Thus, a new general formalism to simulate PQMC with time-dependent trial wave-functions is introduced, including terms that had not been considered previously.

Different ways of approximating the propagator of the imaginary-time Schroedinger equation are also described, which are the linear Taylor approximation, the non-symmetric and the symmetric Trotter approximations. These are all affected by a time step error. Also, a continuous-time exact approximation which does not have a time step error is presented. For the Trotter approximation, on discrete models, we write the complete derivation that leads to a binomial sampling protocol.

Exactly solvable models are used to benchmark the algorithm. For continuous space models, we consider the Harmonic oscillator, for discrete basis models the ferromagnetic quantum Ising model. For the last model, equilibrium dynamics of the PQMC method is studied close to the quantum critical point.

Lastly, we introduce a novel way to extract pure estimators from PQMC simulations which, by default give mixed estimators. We believe that this is a promising route in the context of quantum annealing simulations

## 1.1 Projective quantum Monte Carlo in a nutshell

Let's consider the Dirac representation of the Schrödinger equation written in imaginary-time ($\tau = \frac{it}{\hbar}$)

$$-\frac{\partial}{\partial\tau}|\Psi(\tau)\rangle = (\hat{H} - E_{\mathrm{ref}})|\Psi(\tau)\rangle. \tag{1.1}$$

$\hbar$ is the reduced Planck constant. It is set to $\hbar = 1$. $|\Psi(\tau)\rangle$ is the wave-function of a quantum system at imaginary-time $\tau$. $\hat{H}$ denotes a generic quantum many-body Hamiltonian with eigenvectors $\{|\Psi_{\mathrm{n}}\rangle\}$ and corresponding eigenvalues $\{E_{\mathrm{n}}\}$. $\mathrm{n} = 0, 1, \ldots$ labels the eigenstates of the Hamiltonian. $E_{\mathrm{ref}}$ is a reference energy which, when equal to the average energy $\langle\hat{H}\rangle$, takes care of the normalization of the wave-function. In practice, it is kept close to that value in order to maintain the stability of the simulations, as it will be explained latter. $\langle\hat{H}\rangle$ stands for the expectation value of the Hamiltonian in the ground state. PQMC methods could also be used for excited states in some cases [26], but in this thesis we focus only on their use for the ground state.

At $\tau = 0$, the system is initialized in a state $|\psi_{\mathrm{ini}}\rangle$. Expanding this state on the eigenvectors of the Hamiltonian $\hat{H}$ yields

$$|\psi_{\mathrm{ini}}\rangle = \sum_{\mathrm{n}} c_{\mathrm{n}}|\Psi_{\mathrm{n}}\rangle.$$

$|\psi_{\mathrm{ini}}\rangle$ can be any arbitrary state with the only requirement that it has a non-zero overlap with the ground state, i.e $\langle\Psi_0|\psi_{\mathrm{ini}}\rangle \neq 0$. After a propagation time $\tau$, the

state of the system is given by:

$$\begin{aligned} |\Psi(\tau)\rangle =& e^{-(\hat{H}-E_{\rm ref})\tau}|\psi_{\rm ini}\rangle \\ =& \sum_{\rm n} c_{\rm n} e^{-(E_{\rm n}-E_{\rm ref})\tau}|\Psi_{\rm n}\rangle. \end{aligned} \tag{1.2}$$

As $\tau$ increases, higher energy states are exponentially suppressed if $E_{\rm ref} \simeq E_0$. At very long times, in the limit when $\tau \to \infty$, one projects out the ground state.

$$|\Psi(\tau)\rangle \xrightarrow{\tau\to\infty} c_0 e^{-(E_0-E_{\rm ref})\tau}|\Psi_0\rangle.$$

The algorithm thus described recovers both the ground state energy and the ground state wave-function of $\hat{H}$ at zero temperature. It is important to note from Eq. (1.2) that slower convergence of the algorithm is expected if the energy difference $\Delta = E_1 - E_0$ between the first excited state $E_1$ and the ground state $E_0$ (assuming that $E_{\rm ref} \simeq E_0$) is small. This happens, for e.g., in double well-models that will be considered in latter chapters in the context of optimization problems. Other factors such as the choices of the initial configuration and the guiding wave-function (discussed latter) influence the convergence rate of the algorithm. When the above algorithm is implemented via a stochastic dynamics, it is dubbed projective quantum Monte Carlo (PQMC) method.

In the following section we will discuss its practical implementation in Monte Carlo simulations.

## 1.2 Simple projective quantum Monte Carlo method

The simple projective quantum Monte Carlo method is an algorithm that evolves the Schrödinger equation (1.1) in imaginary time. It is called "simple" because it doesn't make use of the so-called importance sampling technique, well known to increase the efficiency of PQMC simulations. In the next section, we will describe how importance sampling is performed in PQMC simulations, whereas in this section, we will focus on the description of the simple PQMC method.

The Schrödinger equation in Eq. (1.1) can be solved by iteratively applying the following equation:

$$\Psi(\boldsymbol{x}, \tau + \Delta\tau) = \sum_{\boldsymbol{x}'} G(\boldsymbol{x}, \boldsymbol{x}', \Delta\tau)\Psi(\boldsymbol{x}', \tau), \tag{1.3}$$

where $\Psi(\boldsymbol{x},\tau) = \langle \boldsymbol{x}|\Psi(\tau)\rangle$ denotes the wave function at the imaginary time $\tau$.

$$G(\boldsymbol{x},\boldsymbol{x}',\Delta\tau) = \langle \boldsymbol{x}|e^{-\Delta\tau(\hat{H}-E_{\mathrm{ref}})}|\boldsymbol{x}'\rangle, \tag{1.4}$$

is the Green's function of Eq. (1.1) for a short time step $\Delta\tau$. Therefore, in order to perform simulations for long times $\tau$, one has to repeat many steps with the short time Green's function. Thus, in the $\tau \to \infty$ limit, configurations $\boldsymbol{x}$ with a probability density proportional to the ground state wave function $\Psi_0(\boldsymbol{x})$ (assumed to be real and non negative) can be sampled. This is under the assumption that the Hamiltonian $\hat{H}$ is not affected by the sign problem.

One should notice that the Green's function $G(\boldsymbol{x},\boldsymbol{x}',\Delta\tau)$ does not always define a stochastic matrix [1]. Therefore, it cannot be utilized to define the transition matrix of a conventional Markov chain Monte Carlo simulation. This problem can be circumvented by rewriting the Green's function as

$$G(\boldsymbol{x},\boldsymbol{x}',\Delta\tau) = G_{\mathrm{T}}(\boldsymbol{x},\boldsymbol{x}',\Delta\tau)b_{\boldsymbol{x}'}, \tag{1.5}$$

where $G_{\mathrm{T}}(\boldsymbol{x},\boldsymbol{x}',\Delta\tau)$ is by definition stochastic, hence it defines a transition probability to go from $\boldsymbol{x}'$ to $\boldsymbol{x}$. The normalization factor is

$$b_{\boldsymbol{x}'} = \sum_{\boldsymbol{x}} G(\boldsymbol{x},\boldsymbol{x}',\Delta\tau), \tag{1.6}$$

the weight on the configuration $\boldsymbol{x}'$. A stochastic process can then be implemented, where a large population of equivalent copies of the system, in jargon called *walkers*, is evolved. Each walker represents one possible configuration of the system $\boldsymbol{x}'_n$ (the index $n$ labels different walkers), and is gradually modified by performing configuration updates according to $G_{\mathrm{T}}(\boldsymbol{x}_n,\boldsymbol{x}'_n,\Delta\tau)$. Thereafter, their (relative) weights $w_n$ are accumulated according to the rule $w_n \to w_n b_{\boldsymbol{x}'_n}$, starting with equal initial weights $w_n = 1$ for all the walkers in the initial population.

It is important to remark that the matrix elements of the Green's function $G(\boldsymbol{x},\boldsymbol{x}',\Delta\tau)$ need to be nonnegative otherwise the Monte Carlo simulations would be impracticable. This is the "infamous" sign-problem. While this implementation is in principle correct, it is known to lead on a practical side to

[1] A stochastic matrix $A$ is a square matrix whose nonnegative real column entries sum up to unity, i.e., $\sum_{\boldsymbol{x}} A(\boldsymbol{x},\boldsymbol{x}') = 1$. Each entry $A(\boldsymbol{x},\boldsymbol{x}')$ would then define the transition probability to move from configuration $\boldsymbol{x}'$ to $\boldsymbol{x}$. Stochastic matrices are used to construct Markov processes in Monte Carlo simulations.

an exponentially fast signal loss as the number of Monte Carlo steps increases. This is due to the fact that the relative weight of few walkers quickly becomes dominant, while most other walkers give a negligible contribution to the signal.

An effective remedy consists in introducing a birth-death process called *branching*, where each walker is replicated (or annihilated) a number of times corresponding, on average, to the weight $w_n$. Thus, at each imaginary-time $\tau$, the distribution of the walkers population represents the (unnormalized) ground state wave function $\Psi_0(\boldsymbol{x})$.

The simplest correct rule consists in generating, for each walker in the population at a certain imaginary time $\tau$, a number of descendants $n_\mathrm{d}$ in the population at imaginary time $\tau + \Delta\tau$. $n_\mathrm{d}$ is defined as $\mathrm{int}\,[w_n + \eta]$, where $\eta \in [0, 1]$ is a uniform random number, and the function $\mathrm{int}\,[]$ gives the integer part of the argument [10]. This scheme allows one to take full consideration of the information contained in the weight $w_n$. Thus, if $n_\mathrm{d} > 1$, $n_\mathrm{d} - 1$ identical copies of the walker are created and added in the current population of walkers. Whereas, if $n_\mathrm{d} = 0$, the corresponding walker is annihilated from the population. After branching has been performed, all walkers have the same weight $w_n = 1$.

Therefore, the number of walkers in the population fluctuates at each PQMC iteration and can be kept close to a target value $N_\mathrm{w}$ by adjusting the reference energy $E_\mathrm{ref}$. The parameter $E_\mathrm{ref}$ regulates the variations of the walkers population by preventing them to increase or decrease in a rapid manner. There are different ways to maintain the walkers population around a given value. One simple possibility is to compute the reference energy at time $\tau + \Delta\tau$ using the formula

$$E_\mathrm{ref} = E + \mu \log \frac{N_\mathrm{w}}{N_\mathrm{cur}}. \tag{1.7}$$

Here, $E = \langle \hat{H} \rangle$ denotes the energy averaged over the walkers population $N_\mathrm{cur}$ at the previous time step. $N_\mathrm{cur}$ refers to the size that the replicas population had at time $\tau$. $N_\mathrm{w}$ is an algorithm parameter chosen by the user. It affects the accuracy of the results, as explained latter. It is kept fixed all through the simulations. $\mu$ is a small positive parameter which is used to reduce the fluctuations of the walkers population. The logarithmic function serves to regulate the rate of variations in the number of walkers.

Introducing the branching process provides one with a feasible, possibly efficient algorithm. However, such a process might actually introduce a systematic bias

if the average population size $N_w$ is not large enough. The bias originates from the spurious correlations among walkers generated from the same ancestor [27]. This effect becomes negligible in the $N_w \to \infty$ limit, but might be sizable for finite $N_w$. It is known to be the most relevant and subtle possible source of systematic errors in PQMC algorithms [28, 29, 30].

Another potential source of systematic errors in PQMC originates from the time step $\Delta\tau$. In order to illustrate that, consider a generic Hamiltonian

$$\hat{H} = \hat{H}_{\text{cl}} + \hat{H}_{\text{kin}},$$

where $\hat{H}_{\text{cl}}$ is the potential energy designating the classical part of the Hamiltonian and, $\hat{H}_{\text{kin}}$ the kinetic part responsible for quantum fluctuations. Employing a primitive Trotter approximation [10] to separate the contributions from the two non-commuting terms yields:

$$e^{-\Delta\tau(\hat{H}-E_{\text{ref}})} = e^{-\Delta\tau\hat{H}_{\text{kin}}} e^{-\Delta\tau(\hat{H}_{\text{cl}}-E_{\text{ref}})} + \mathcal{O}(\Delta\tau^2). \tag{1.8}$$

Such error is dealt with either by running a single (expensive) simulation with a very small value of $\Delta\tau$ or, by running simulations at different values of $\Delta\tau$ and extrapolating ground-state energies linearly in the limit where $\Delta\tau \to 0$.

The left hand side of Equation (1.8) can also be approximated using a Taylor expansion. It is written as (for e.g. see [31, 32, 33])

$$e^{-\Delta\tau(\hat{H}-E_{\text{ref}})} = 1 - \Delta\tau\hat{H}_{\text{kin}} - \Delta\tau(\hat{H}_{\text{cl}} - E_{\text{ref}}) + \mathcal{O}(\Delta\tau^2). \tag{1.9}$$

It is a different way to explicitly express the Green's function, using a Taylor expansion instead of a Trotter decomposition.

A more accurate estimate of the propagator, that has a higher order error in $\Delta\tau$ is given by the symmetrized version of the Trotter approximation [10]

$$e^{-\Delta\tau(\hat{H}-E_{\text{ref}})} = e^{-\frac{\Delta\tau}{2}(\hat{H}_{\text{cl}}-E_{\text{ref}})} e^{-\Delta\tau\hat{H}_{\text{kin}}} e^{-\frac{\Delta\tau}{2}(\hat{H}_{\text{cl}}-E_{\text{ref}})} + \mathcal{O}(\Delta\tau^3). \tag{1.10}$$

By neglecting the $\mathcal{O}(\Delta\tau^3)$ terms in the Green's function, one obtains a quadratic convergence of ground-state properties in the $\Delta\tau \to 0$ limit [34]. Higher order expansions of the imaginary-time propagator exist but require a more complex implementation [35]. In all our computations (unless it is otherwise stated), we implement the primitive approximation in Eq. (1.10).

It is important to mention that the systematic error coming from the time step $\Delta\tau$ can be completely eliminated in the simulations as we will discuss latter in the context of continuous-time PQMC methods.

## 1.3 Projective Quantum Monte Carlo method with importance sampling

The simple PQMC has the advantage that no prior knowledge of the wave-function is needed. However, rapidly varying potential energies can seriously affect the branching step, even to the point of rendering the calculations intractable. A promising strategy to circumvent the aforementioned problem is to introduce the so-called importance sampling technique. This is indeed a well established approach to boost the efficiency of PQMC simulations (see, e.g, Ref. [36]) because it has the potential to reduce the number of walkers needed to attain a given accuracy [27]. It consists in evolving a function

$$f(\boldsymbol{x}, \tau) = \Psi(\boldsymbol{x}, \tau)\psi_T(\boldsymbol{x}, \tau), \tag{1.11}$$

via a modified imaginary-time Schrödinger equation. $\psi_T(\boldsymbol{x}, \tau)$ is a guiding function designed to accurately approximate the ground-state wave function. Its role is to favor the sampling of configurations with high probability amplitude.

Usually, in PMQC simulations using importance sampling, the guiding function is considered to be time-independent. A novelty that we introduce in the following derivation is to have a trial wave function which explicitly depends on time. The motivation for this consideration stems from the desire to use PQMC as a quantum-inspired classical optimization method to simulate combinatorial optimization problems in an annealing setting. As a matter of fact, in a quantum annealing simulation, the kinetic part of the Hamiltonian is varied with time, and so is its spectrum. Therefore, it is expected that a trial function that accurately describes the time-dependent ground state wave function during annealing needs to be time dependent.

In order to derive how the imaginary-time Schrödinger equation changes, we

rewrite Equation (1.1) as

$$-\frac{\partial}{\partial\tau}\Psi(\boldsymbol{x},\tau)=\sum_{\boldsymbol{x}'}\Big[H_{\boldsymbol{x},\boldsymbol{x}'}-E_{\mathrm{ref}}\delta_{\boldsymbol{x},\boldsymbol{x}'}\Big]\Psi(\boldsymbol{x}',\tau), \tag{1.12}$$

where $H_{\boldsymbol{x},\boldsymbol{x}'}=\langle\boldsymbol{x}|\hat{H}|\boldsymbol{x}'\rangle$ and $\delta$ stands for the Kronecker delta function (assuming a discrete basis). Using the expression of the wave-function $\Psi(\boldsymbol{x},\tau)$ in Equation 1.11 yields:

$$\frac{\partial}{\partial\tau}\Psi(\boldsymbol{x},\tau)=\frac{\partial_\tau f(\boldsymbol{x},\tau)}{\psi_T(\boldsymbol{x},\tau)}-\frac{f(\boldsymbol{x},\tau)}{\psi_T(\boldsymbol{x},\tau)}\frac{\partial_\tau\psi_T(\boldsymbol{x},\tau)}{\psi_T(\boldsymbol{x},\tau)}.$$

Defining a fictitious time-dependent potential energy as

$$\tilde{V}(\boldsymbol{x},\tau)=-\frac{\partial_\tau\psi_T(\boldsymbol{x},\tau)}{\psi_T(\boldsymbol{x},\tau)}, \tag{1.13}$$

we obtain from Eq. (1.12) that the modified imaginary-time Schrödinger equation becomes:

$$-\frac{\partial}{\partial\tau}f(\boldsymbol{x},\tau)=\sum_{\boldsymbol{x}'}\Big[H_{\boldsymbol{x},\boldsymbol{x}'}-E_{\mathrm{ref}}\delta_{\boldsymbol{x},\boldsymbol{x}'}+\tilde{V}(\boldsymbol{x}',\tau)\delta_{\boldsymbol{x},\boldsymbol{x}'}\Big]\frac{\psi_T(\boldsymbol{x},\tau)}{\psi_T(\boldsymbol{x}',\tau)}f(\boldsymbol{x}',\tau). \tag{1.14}$$

In this thesis, an explicit time-dependence of the guiding function is only considered in the last chapter. Therefore, for the time being, the guiding function will be referred to without a dependence on $\tau$, i.e $\psi_T(\boldsymbol{x},\tau)\equiv\psi_T(\boldsymbol{x})$. Equivalently, this would imply a vanishing time-dependent potential $\tilde{V}(\boldsymbol{x},\tau)$. Nonetheless, remark that its effect could be trivially added in the PQMC algorithm given that it is diagonal in the basis state.

In the long imaginary-time limit, the walkers sample spin configurations with a probability distribution proportional to $f(\boldsymbol{x},\tau\to\infty)=\Psi_0(\boldsymbol{x})\psi_T(\boldsymbol{x})$. If $\psi_T(\boldsymbol{x})$ is a good approximation of the ground-state wave function, this distribution closely approximates the quantum-mechanical probability of finding the system in the configuration $\boldsymbol{x}$. The expectation value of the energy in PQMC is computed using

the mixed estimator:

$$\begin{aligned}
\langle \hat{H} \rangle_{mixed} &= \frac{\langle \Psi_0 | \hat{H} | \psi_T \rangle}{\langle \Psi_0 | \psi_T \rangle} \\
&= \frac{\sum_{\boldsymbol{x}} \langle \Psi_0 | \boldsymbol{x} \rangle \langle \boldsymbol{x} | \hat{H} | \psi_T \rangle}{\sum_{\boldsymbol{x}} \langle \Psi_0 | \boldsymbol{x} \rangle \langle \boldsymbol{x} | \psi_T \rangle} \\
&= \frac{\sum_{\boldsymbol{x}} \Psi_0(\boldsymbol{x}) \left[ \frac{\langle \boldsymbol{x} | \hat{H} | \psi_T \rangle}{\langle \boldsymbol{x} | \psi_T \rangle} \right] \langle \boldsymbol{x} | \psi_T \rangle}{\sum_{\boldsymbol{x}} \Psi_0(\boldsymbol{x}) \psi_T(\boldsymbol{x})} \\
&= \frac{\sum_{\boldsymbol{x}} \Psi_0(\boldsymbol{x}) E_{\mathrm{loc}}(\boldsymbol{x}) \psi_T(\boldsymbol{x})}{\sum_{\boldsymbol{x}} \Psi_0(\boldsymbol{x}) \psi_T(\boldsymbol{x})}.
\end{aligned} \tag{1.15}$$

The quantity $E_{\mathrm{loc}}(\boldsymbol{x})$, dubbed *local energy* which is defined as

$$E_{\mathrm{loc}}(\boldsymbol{x}) = \frac{\langle \boldsymbol{x} | \hat{H} | \psi_T \rangle}{\langle \boldsymbol{x} | \psi_T \rangle}, \tag{1.16}$$

plays a fundamental role for the stabilization of the simulations. If it undergoes large fluctuations then the (current) number of walkers will also fluctuate a lot.

It is important to notice that if our guiding wave function was exact, i.e., if $\psi_T(\boldsymbol{x}) = \Psi_0(\boldsymbol{x})$, then the local energy would be a constant function, $E_{\mathrm{loc}}(\boldsymbol{x}) = E_0$. This would completely suppress the fluctuations of the number of walkers (this is a consequence of the zero variance property [27]), therefore eliminating the bias due to the finite walkers population $N_w$. If $\psi_T(\boldsymbol{x})$ is, albeit not exact, a good approximation of $\Psi_0(\boldsymbol{x})$, the fluctuations of the number of walkers are still reduced compared to the case of the simple PQMC algorithm giving a faster convergence to the exact $N_w \to \infty$ limit, thus reducing the computational cost, in particular for large systems.

The ground state energy $E$ of the system is computed as

$$E = \lim_{N_c \to \infty} \frac{1}{N_c} \sum_{i=1}^{N_c} E_{\mathrm{loc}}(\boldsymbol{x}_i), \tag{1.17}$$

where $N_c$ is a number of uncorrelated configurations $\boldsymbol{x}_i$ sampled from the proba-

bility distribution

$$p(\boldsymbol{x}) = \frac{\Psi_0(\boldsymbol{x})\psi_T(\boldsymbol{x})}{\sum_{\boldsymbol{x}} \Psi_0(\boldsymbol{x})\psi_T(\boldsymbol{x})}. \tag{1.18}$$

Note that in the case of the Hamiltonian operator, and of other operators that commute with the Hamiltonian, the mixed estimator defined above provides exact estimates (within statistical error bar) of the ground state expectation value. In general, the mixed estimator is accurate only if the trial function $\psi_T(\boldsymbol{x})$ is a good approximation of the ground state wave-function. Pure estimators for operators that do not commute with the Hamiltonian can be derived too, as explained in section 1.4.

### 1.3.1 Continuous space models

In this section, we consider continuous-space models in one spatial dimension. A generic form of the Hamiltonian is:

$$\hat{H} = -D\nabla_{\boldsymbol{x}}^2 + \hat{V}(\boldsymbol{x}),$$

where $D = (2m)^{-1}$, with $m$ being the particles mass (assumed to be equal for all particles). Recall that the reduced Planck constant has been set to $\hbar = 1$ all through the manuscript. $\boldsymbol{x} = (x_1, \ldots, x_N)$ denotes the particles configuration. $x_i$ is the position of the particle $i$ (with $i = 1, \ldots, N$), and $N$ is the particle number. $\hat{V}(\boldsymbol{x})$ stands for the potential energy operator.

Introducing importance sampling with a time independent trial function $\psi_T(\boldsymbol{x})$ transforms Eq. (1.14) to:

$$\begin{aligned}
-\frac{\partial}{\partial\tau} f(\boldsymbol{x},\tau) &= \int \mathrm{d}\boldsymbol{x}' \Big[H_{\boldsymbol{x},\boldsymbol{x}'} - E_{\mathrm{ref}}\delta(\boldsymbol{x}-\boldsymbol{x}')\Big]\frac{\psi_T(\boldsymbol{x})}{\psi_T(\boldsymbol{x}')} f(\boldsymbol{x}',\tau) \\
&= \int \mathrm{d}\boldsymbol{x}' \Big[\Big(\delta(\boldsymbol{x}-\boldsymbol{x}')[-D\nabla_{\boldsymbol{x}'}^2 + V(\boldsymbol{x}')]\Big) - E_{\mathrm{ref}}\delta(\boldsymbol{x}-\boldsymbol{x}')\Big]\frac{\psi_T(\boldsymbol{x})}{\psi_T(\boldsymbol{x}')} f(\boldsymbol{x}',\tau) \\
&= \int \mathrm{d}\boldsymbol{x}' \delta(\boldsymbol{x}-\boldsymbol{x}')\Big[ - D\underbrace{\nabla_{\boldsymbol{x}'}^2\Big(\frac{\psi_T(\boldsymbol{x})f(\boldsymbol{x}',\tau)}{\psi_T(\boldsymbol{x}')}\Big)}_{T} + [V(\boldsymbol{x}') - E_{\mathrm{ref}}]\frac{\psi_T(\boldsymbol{x})}{\psi_T(\boldsymbol{x}')} f(\boldsymbol{x}',\tau)\Big].
\end{aligned} \tag{1.19}$$

The term $T$ in the kinetic energy is expanded as:

$$\begin{aligned} T &= \nabla^2_{\boldsymbol{x}'}\Big[\frac{\psi_T(\boldsymbol{x})f(\boldsymbol{x}',\tau)}{\psi_T(\boldsymbol{x}')}\Big] = \psi_T(\boldsymbol{x})\nabla^2_{\boldsymbol{x}'}\Big[\frac{f(\boldsymbol{x}',\tau)}{\psi_T(\boldsymbol{x}')}\Big] \\ &= \psi_T(\boldsymbol{x})\nabla_{\boldsymbol{x}'}\Big[\frac{\nabla_{\boldsymbol{x}'}f(\boldsymbol{x}',\tau)}{\psi_T(\boldsymbol{x}')} - \frac{f(\boldsymbol{x}',\tau)}{\psi_T(\boldsymbol{x}')}F(\boldsymbol{x}')\Big], \end{aligned}$$

where we have defined a quantum drift force $F(\boldsymbol{x})$ as

$$F(\boldsymbol{x}) = \frac{\nabla_{\boldsymbol{x}}\psi_T(\boldsymbol{x})}{\psi_T(\boldsymbol{x})}. \tag{1.20}$$

Hence, we obtain

$$\begin{aligned} T &= \psi_T(\boldsymbol{x})\Big[\frac{\nabla^2_{\boldsymbol{x}'}f(\boldsymbol{x}',\tau)}{\psi_T(\boldsymbol{x}')} - \frac{\nabla_{\boldsymbol{x}'}f(\boldsymbol{x}',\tau)}{\psi_T(\boldsymbol{x}')}F(\boldsymbol{x}') - \nabla_{\boldsymbol{x}'}\Big[\frac{f(\boldsymbol{x}',\tau)}{\psi_T(\boldsymbol{x}')}\Big]F(\boldsymbol{x}') - \frac{f(\boldsymbol{x}',\tau)}{\psi_T(\boldsymbol{x}')}\nabla_{\boldsymbol{x}'}F(\boldsymbol{x}')\Big] \\ &= \psi_T(\boldsymbol{x})\Big[\nabla^2_{\boldsymbol{x}'}f(\boldsymbol{x}',\tau) - 2F(\boldsymbol{x}')\nabla_{\boldsymbol{x}'}f(\boldsymbol{x}',\tau) + F^2(\boldsymbol{x}')f(\boldsymbol{x}',\tau) - \nabla_{\boldsymbol{x}'}[F(\boldsymbol{x}')]f(\boldsymbol{x}',\tau)\Big]\frac{1}{\psi_T(\boldsymbol{x}')}. \end{aligned}$$

One can check that

$$F^2(\boldsymbol{x}') = \frac{\nabla^2_{\boldsymbol{x}'}\psi_T(\boldsymbol{x}')}{\psi_T(\boldsymbol{x}')} - \nabla_{\boldsymbol{x}'}F(\boldsymbol{x}').$$

Therefore, after rearrangement of terms, we obtain

$$T = \psi_T(\boldsymbol{x})\Big[\nabla^2_{\boldsymbol{x}'}f(\boldsymbol{x}',\tau) - 2\nabla_{\boldsymbol{x}'}[F(\boldsymbol{x}')f(\boldsymbol{x}',\tau)] + \frac{\nabla^2_{\boldsymbol{x}'}\psi_T(\boldsymbol{x}')}{\psi_T(\boldsymbol{x}')}f(\boldsymbol{x}',\tau)\Big]\frac{1}{\psi_T(\boldsymbol{x}')}.$$

Inserting this term in Equation (1.19) yields

$$\begin{aligned} -\frac{\partial}{\partial\tau}f(\boldsymbol{x},\tau) &= \int \mathrm{d}\boldsymbol{x}'\delta(\boldsymbol{x}-\boldsymbol{x}')\psi_T(\boldsymbol{x})\Big[-D\nabla^2_{\boldsymbol{x}'}f(\boldsymbol{x}',\tau) + 2D\nabla_{\boldsymbol{x}'}[F(\boldsymbol{x}')f(\boldsymbol{x}',\tau)] \\ &+ [E_{\mathrm{loc}}(\boldsymbol{x}') - E_{\mathrm{ref}}]f(\boldsymbol{x}',\tau)\Big]\frac{1}{\psi_T(\boldsymbol{x}')}, \end{aligned}$$

where we have made used of the definition of the local energy in Equation (1.16),

which in coordinate space takes the form:

$$E_{\mathrm{loc}}(\boldsymbol{x}) = -D\frac{\nabla_{\boldsymbol{x}}^2\psi_T(\boldsymbol{x})}{\psi_T(\boldsymbol{x})} + V(\boldsymbol{x}). \tag{1.21}$$

The modified Schrödinger Equation thus assumes the following expression for continuous space-models

$$-\frac{\partial}{\partial\tau}f(\boldsymbol{x},\tau) = -D\nabla_{\boldsymbol{x}}^2 f(\boldsymbol{x},\tau) + 2D\nabla_{\boldsymbol{x}}[F(\boldsymbol{x})f(\boldsymbol{x},\tau)] + [E_{\mathrm{loc}}(\boldsymbol{x}) - E_{\mathrm{ref}}]f(\boldsymbol{x},\tau). \tag{1.22}$$

This equation displays a twofold effect due to the importance sampling. First, through the drift force $F(\boldsymbol{x})$, walkers are conducted to regions of the configuration space that are the most relevant for sampling. Second, the branching step is smoothed with the use of the local energy $E_{\mathrm{loc}}(\boldsymbol{x})$ instead of the bare potential.

Note that when the guiding function is a constant $\psi_T(\boldsymbol{x}) = 1$, in order words, importance sampling is not utilized, as expected, the imaginary-time Schrödinger equation is recovered.

$$-\frac{\partial}{\partial\tau}\Psi(\boldsymbol{x},\tau) = -D\nabla_{\boldsymbol{x}}^2\Psi(\boldsymbol{x},\tau) + [V(\boldsymbol{x}) - E_{\mathrm{ref}}]\Psi(\boldsymbol{x},\tau). \tag{1.23}$$

$D$ can be viewed as an effective diffusion constant. This is because the above equation is similar to a diffusion equation.

From Equation (1.22), we identify an effective Hamiltonian given by

$$\begin{aligned}\tilde{H} &= -D\nabla_{\boldsymbol{x}}^2 + 2D\nabla_{\boldsymbol{x}}[F(\boldsymbol{x})] + E_{\mathrm{loc}}(\boldsymbol{x}) \\ &= D\hat{\boldsymbol{p}}^2 + 2iD\hat{\boldsymbol{p}}[F(\boldsymbol{x})] + E_{\mathrm{loc}}(\boldsymbol{x}),\end{aligned}$$

such that the modified Schrödinger equation can be re-written as:

$$-\frac{\partial}{\partial\tau}f(\boldsymbol{x},\tau) = (\tilde{H} - E_{\mathrm{ref}})f(\boldsymbol{x},\tau). \tag{1.24}$$

Therefore, in the same spirit as what was done in Equation (1.3), the modified imaginary-time Schrödinger is solved by applying iteratively the integral equation:

$$f(\boldsymbol{x},\tau+\Delta\tau) = \int \mathrm{d}\boldsymbol{x}'\tilde{G}(\boldsymbol{x},\boldsymbol{x}',\Delta\tau)f(\boldsymbol{x}',\tau), \tag{1.25}$$

where $\tilde{G}(\boldsymbol{x},\boldsymbol{x}',\Delta\tau)$ is a suitable approximation for the Green's function of Eq. (1.22) and is given by the expression

$$\tilde{G}(\boldsymbol{x},\boldsymbol{x}',\Delta\tau) = \langle\boldsymbol{x}|e^{-\Delta\tau(\tilde{H}-E_{\mathrm{ref}})}|\boldsymbol{x}'\rangle. \tag{1.26}$$

Using the primitive approximation of Eq. (1.10) gives

$$\tilde{G}(\boldsymbol{x},\boldsymbol{x}',\Delta\tau) \approx G_d(\boldsymbol{x},\boldsymbol{x}',\Delta\tau)G_b(\boldsymbol{x},\boldsymbol{x}',\Delta\tau), \tag{1.27}$$

where

$$G_b(\boldsymbol{x},\boldsymbol{x}',\Delta\tau) = \exp\left[-\Delta\tau\left(\frac{E_{\mathrm{loc}}(\boldsymbol{x})+E_{\mathrm{loc}}(\boldsymbol{x}')}{2} - E_{\mathrm{ref}}\right)\right], \tag{1.28}$$

since the local energy is diagonal in the basis state. Whereas the off-diagonal contribution (coming from the kinetic energy) to the Green's function is

$$G_d(\boldsymbol{x},\boldsymbol{x}',\Delta\tau) \quad = \quad \langle\boldsymbol{x}|e^{-\Delta\tau D\hat{\boldsymbol{p}}^2 - 2i\Delta\tau D\hat{\boldsymbol{p}}[F(\boldsymbol{x})]}|\boldsymbol{x}'\rangle.$$

Using the completeness relation

$$\int \mathrm{d}\boldsymbol{p}|\boldsymbol{p}\rangle\langle\boldsymbol{p}| = 1,$$

and the expression for the single particle wave-function in coordinate space

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi}}e^{ipx},$$

we obtain

$$\begin{aligned} G_d(\boldsymbol{x},\boldsymbol{x}',\Delta\tau) &= \int \frac{\mathrm{d}\boldsymbol{p}}{(2\pi)^{N/2}} \frac{\mathrm{d}\boldsymbol{p}'}{(2\pi)^{N/2}} e^{i(\boldsymbol{p}\boldsymbol{x}-\boldsymbol{p}'\boldsymbol{x}')} e^{-\Delta\tau D\boldsymbol{p}'^2 - 2i\Delta\tau D\boldsymbol{p}'[F(\boldsymbol{x}')]} \delta(\boldsymbol{p}-\boldsymbol{p}') \\ &= \int \frac{\mathrm{d}\boldsymbol{p}}{(2\pi)^{N}} e^{i\boldsymbol{p}(\boldsymbol{x}-\boldsymbol{x}')} e^{-\Delta\tau D\boldsymbol{p}^2 - 2i\Delta\tau D\boldsymbol{p}[F(\boldsymbol{x}')]}. \end{aligned}$$

From the following integral relation

$$\int dp e^{-\frac{a}{2}p^2} e^{bp} = \sqrt{\frac{2\pi}{a}} e^{\frac{b^2}{2a}},$$

we obtain the following expression for $G_d(\boldsymbol{x},\boldsymbol{x}',\Delta\tau)$:

$$G_d(\boldsymbol{x},\boldsymbol{x}',\Delta\tau) = (4\pi\Delta\tau D)^{-N/2} \exp\left[-\frac{(\boldsymbol{x}-\boldsymbol{x}'-2\Delta\tau D F(\boldsymbol{x}'))^2}{4\Delta\tau D}\right]. \tag{1.29}$$

This Green's function is normalized and defines a standard drift-diffusion process. Remark that in the absence of a drift force, it would characterize simple diffusion processes.

Eq. (1.25) could be interpreted as the definition of a Markov chain with transition matrix equal to the (positive-definite) Green's function $\tilde{G}(\boldsymbol{x},\boldsymbol{x}',\Delta\tau)$. However, the second term $G_b(\boldsymbol{x},\boldsymbol{x}',\Delta\tau)$ in Equation (1.27) is not normalized. The Markov chain can still be defined in an extended configuration space. One has to evolve a (large) ensemble of copies of the system according to the drift-Gaussian process, with an additional branching (or killing) process where walker replicas are generated (or annihilated) proportionally to $G_b(\boldsymbol{x},\boldsymbol{x}',\Delta\tau)$. This branching process, performed as described above, takes into account the lack of normalization of the Green's function, and causes fluctuations in the random-walker number as previously discussed in Sec. (1.2).

### 1.3.2 Test case: The Harmonic oscillator

As a test bed, we present results obtained by implementing simple PQMC to simulate the harmonic potential in one dimension. The objective here is twofold: first, to show that with such a toy model the algorithm does recover ground state properties. Second, we show how a smartly chosen guiding function affects how

accurately one can predict the ground state energy of a quantum system.

The Hamiltonian of the system is given by

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(x),$$

with the harmonic potential being:

$$V(x) = \frac{1}{2} m\omega^2 x^2.$$

$m$ stands for the mass and $\omega$, the angular frequency of the oscillator, is here chosen to be the unit of energy. The value for the ground state energy is $E_0 = 0.5$.

The simple PQMC is implemented as described in Section (1.2). Walkers are initialized randomly, then evolved through a *free particle diffusion* process and branching. The diffusion part comes as a result of the Green's function $G_{\mathrm{T}}(x, x', \Delta\tau)$ in Equation (1.5) being Gaussian. Thus, during the Markov process, each walker diffuses through a random Gaussian displacement of typical length $\sqrt{\frac{\Delta\tau}{m}}$. At imaginary-time $\tau$, each walker positioned at $x'$ is updated as follows:

$$x = x' + \eta\sqrt{\frac{\Delta\tau}{m}},$$

where $\eta$ is a Gaussian random variable with variance $1$. The branching step is performed with a weight

$$w = \exp\Big[ - \Delta\tau\Big(\frac{V(x') + V(x)}{2} - E_{\mathrm{ref}}\Big)\Big].$$

Figure 1.1(a) displays how the algorithm is able to recover the ground state wave-function of the harmonic oscillator. Initially, at $\tau = 0$, walkers have a flat distribution in space. Thereafter, they are evolved in imaginary time. Different imaginary times $\tau$ ($\equiv$ MCs $\times\Delta\tau$), represent the running time at which the simulations are being performed. MCs designates the number of Monte Carlo steps. When $\tau = 225$, a converged distribution of the population is obtained. It corresponds to $\Psi_0(x)$, given by:

$$\Psi_0(x) = \Big(\frac{m\omega}{\pi}\Big)^{\frac{1}{4}} \exp\Big[-\frac{m\omega}{2}x^2\Big].$$

Next, we move on to introduce importance sampling in the algorithm. The implementation follows the description given in Section (1.3). We choose the following variational wave-function $\psi_T(x)$ as a guiding function for PQMC simulations:

$$\psi_T(x) = \Big(\frac{m\omega}{\pi}\kappa\Big)^{\frac{1}{4}} \exp\Big[-\frac{m\omega}{2}\kappa x^2\Big], \tag{1.30}$$

where $\kappa$ is a variational parameter. $\kappa = 1$ corresponds to the case where the guiding function coincides exactly with the exact ground state wave function, i.e $\psi_T(x) = \Psi_0(x)$. The diffusion process is similar to the one in simple PQMC except that an additional update which consists of a drift in the walker coordinate is added

$$x = x' + \frac{\Delta\tau}{m}F(x') + \eta\sqrt{\frac{\Delta\tau}{m}}.$$

This is a discrete *Langevin* equation with a Gaussian noise and a drift force (see Eq. 1.20),

$$F(x) = -m\omega\kappa x.$$

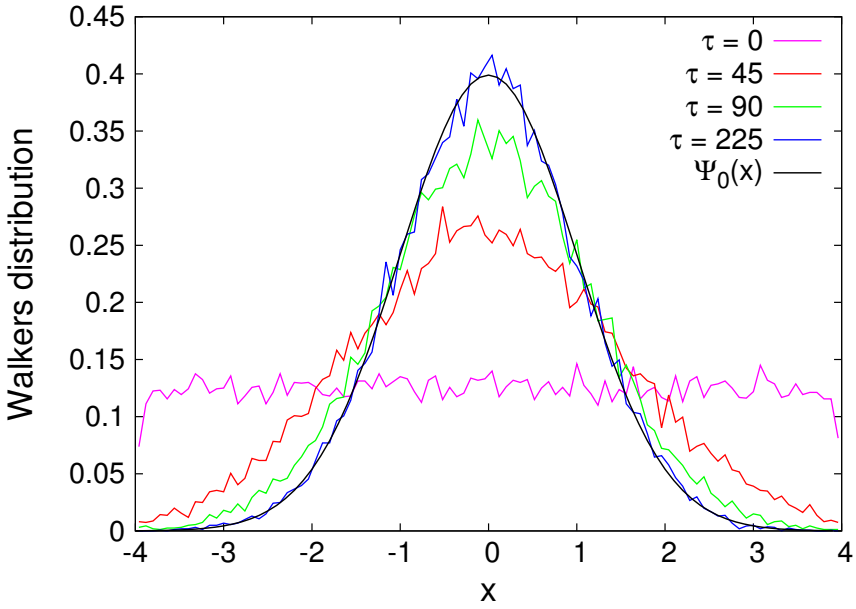
Branching with importance sampling is done using the local energy computed using Equation (1.21),

$$E_{\text{loc}}(x) = \frac{\omega\kappa}{2}\Big[1 - m\omega\kappa x^2\Big] + V(x),$$

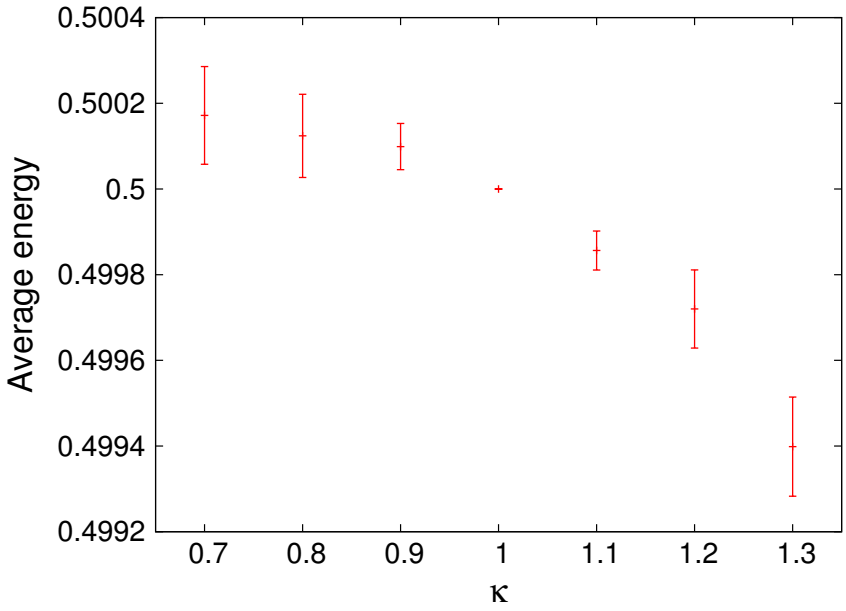
instead of the bare potential $V(x)$. The expectation value $E$ of the Hamiltonian is computed using the average value of the local energy taken on a large population of walkers.

In Figure 1.1(b), the average energy $E = \langle\hat{H}\rangle$ is plotted for different values of $\kappa$. Note that when $\kappa = 1$, the statistical fluctuations of the average energy are completely suppressed and the correct value for the ground state energy is found. However, the farther the guiding function is to the exact ground state, the less accurate are the results and the greater are the statistical fluctuations. This is because the local energy is no longer constant and this generate more fluctuations in the walkers population as the error bars of $E$ demonstrate.

Note however that other systematic errors can also become larger, for example the one due to the finite time step $\Delta\tau$. At large values of the variational parameter, convergence to the ground state energy (with the time step) is obtained from




(a)




(b)

Figure 1.1: (a) Histograms of the density distribution of the population of walkers in the harmonic potential at different imaginary times $\tau$. For a sufficiently long $\tau$, the walkers distribution converges towards the ground state wave-function of the harmonic potential $\Psi_0(x)$. (b) Average ground state energy of the harmonic oscillator computed using PQMC with $\psi_T(x)$ as a guiding function (see Eq. 1.30) as a function of the variational parameter $\kappa$. The exact value for ground state energy of the harmonic oscillator is $0.5$. Energies are given in unit of $\omega$, the harmonic oscillator strength.

below (data not shown). This explains why the data in Fig. 1.1(b) at large $\kappa$ are slightly below the exact result.

### 1.3.3 Discrete basis models

In this section, we present the implementation of the simple PQMC algorithm, i.e without importance sampling, for transverse-field Ising models. The formalism could equivalently be used for particles hopping on a lattice, but we choose to focus on spin based models which are more relevant for quantum computing, in particular for studies on solving combinatorial optimization problems via adiabatic quantum optimization. Initially, a description of simple PQMC following the theoretical formalism sketched in Ref. [32] will be given. We explain it in details here because it will latter on serve as a basis for the work presented in Chapter 3. Subsequently, we present the continuous-time Green's function Monte Carlo (CTGFMC) algorithm introduced in [37]. Its choice is motivated by the fact that it completely removes the time step error inherent in the Trotter approximation of the Green's function. It has been used in the results that will be discussed in Chapters 4 and 5.

Consider a generic Ising spin Hamiltonian defined as:

$$\hat{H} = -\sum_{i,j} J_{ij}\sigma_i^z\sigma_j^z - \Gamma\sum_{i=1}^{N}\sigma_i^x, \tag{1.31}$$

where $J_{ij}$ is the interaction strength between the spins $i$ and $j$ placed on the $N$ nodes of a graph. $\sigma_i^z$ and $\sigma_i^x$ are Pauli matrices acting at site $i$. Each spin experiences a transverse field of strength $\Gamma$, which introduces quantum fluctuations. We do not explicitly consider longitudinal fields; however, their effect could be trivially included in the algorithm.

Given $|x_i\rangle$ an eigenstate of the Pauli matrix $\sigma_i^z$ at site $i$ with eigenvalue $x_i = 1$ when $|x\rangle = |\uparrow\rangle$ and $x_i = -1$ when $|x\rangle = |\downarrow\rangle$, the quantum state of $N$ spins in the system is indicated by $|\boldsymbol{x}\rangle = |x_1x_2...x_N\rangle$. The ensemble of $2^N$ states $\{|\boldsymbol{x}\rangle\}$ is chosen as computational basis.

The Schrödinger equation (1.1) is solved by applying iteratively Equation (1.3). Recall that $G(\boldsymbol{x}, \boldsymbol{x}', \Delta\tau)$ is the Green's function of Eq. (1.1). For reasons that will

be explained latter on, we employ the primitive Trotter approximation given by:

$$G(\boldsymbol{x}, \boldsymbol{x}', \Delta\tau) = \langle \boldsymbol{x}| e^{-\Delta\tau \hat{H}_{\rm kin}} e^{-\Delta\tau(\hat{H}_{\rm cl} - E_{\rm ref})} |\boldsymbol{x}'\rangle + \mathcal{O}(\Delta\tau^2),$$

where

$$\hat{H}_{\rm cl} = -\sum_{i,j} J_{ij} \sigma_i^z \sigma_j^z, \tag{1.32}$$

and

$$\hat{H}_{\rm kin} = -\Gamma \sum_{i=1}^{N} \sigma_i^x. \tag{1.33}$$

The potential energy being diagonal on the basis state $|\boldsymbol{x}\rangle$, the function $G(\boldsymbol{x}, \boldsymbol{x}', \Delta\tau)$ transforms to:

$$G(\boldsymbol{x}, \boldsymbol{x}', \Delta\tau) \simeq \underbrace{\langle \boldsymbol{x}| e^{-\Delta\tau \hat{H}_{\rm kin}} |\boldsymbol{x}'\rangle}_{I} e^{-\Delta\tau[E_{\rm cl}(\boldsymbol{x}') - E_{\rm ref}]},$$

with $E_{\rm cl}(\boldsymbol{x}) = \langle \boldsymbol{x}| \hat{H}_{\rm cl} |\boldsymbol{x}\rangle$. Given that Pauli matrices at different sites commutes, we have the following result:

$$\begin{aligned} I &= \langle x_1 x_2 ... x_N | e^{\Delta\tau\Gamma \sum_i \hat{\sigma}_i^x} | x'_1 x'_2 ... x'_N \rangle \\ &= \langle x_1 | e^{\Delta\tau\Gamma\hat{\sigma}_1^x} | {x'}_1 \rangle \langle x_2 | e^{\Delta\tau\Gamma\hat{\sigma}_2^x} | {x'}_2 \rangle ... \langle x_N | e^{\Delta\tau\Gamma\hat{\sigma}_N^x} | {x'}_N \rangle \\ &= \prod_{i=1}^{N} \langle x_i | e^{\Delta\tau\Gamma\hat{\sigma}_i^x} | {x'}_i \rangle. \end{aligned}$$

One could verify that

$$\langle x | e^{a\hat{\sigma}^x} | x' \rangle = c e^{J_\perp x x'}, \tag{1.34}$$

where $c = \sqrt{\frac{\sinh 2a}{2}}$ and $J_\perp = -\frac{1}{2} \log \tanh a$.

The formula in Eq. (1.34) is generally used to write $I$ in terms of Ising-like interactions between same lattice site spins in the imaginary-time domain. Here we refer to the mapping of a quantum system in $D$ dimension to a classical system

in $D+1$ dimension. Thus, we obtain that

$$I = c^N e^{J_\perp \sum_{i=1}^N x_i x_i'}.$$

Now, we desire to engineer a diffusion mechanism that encodes *multi-spin flips*. To do that, we suppose that during a small time $\Delta\tau$, when passing from a spin configuration $|\boldsymbol{x}'\rangle$ to a spin configuration $|\boldsymbol{x}\rangle$, a certain number of $\delta$ spins are flipped i.e

$$\begin{cases} x_i x'_i = -1 & \text{for } \delta \text{ spins} \\ x_i x'_i = 1 & \text{for } (N-\delta) \text{ spins.} \end{cases}$$

The Green's function becomes

$$\begin{aligned} G(\boldsymbol{x},\boldsymbol{x}',\Delta\tau) &\simeq c^N e^{-\Delta\tau J_\perp \delta} e^{\Delta\tau J_\perp (N-\delta)} e^{-\Delta\tau[E_{\rm cl}(\boldsymbol{x}')-E_{\rm ref}]} \\ &= c^N e^{\Delta\tau J_\perp N} e^{-2\Delta\tau J_\perp \delta} e^{-\Delta\tau[E_{\rm cl}(\boldsymbol{x}')-E_{\rm ref}]} \\ &= \left[\frac{\sinh 2\Delta\tau\Gamma}{2}\right]^{N/2} e^{-\frac{N}{2}\log(\tanh\Delta\tau\Gamma)} e^{\delta\Big(\log(\tanh\Delta\tau\Gamma)\Big)} e^{-\Delta\tau[E_{\rm cl}(\boldsymbol{x}')-E_{\rm ref}]} \\ &= \left[\frac{2\sinh\Delta\tau\Gamma\cosh\Delta\tau\Gamma}{2}\right]^{N/2}\left[\frac{1}{\tanh\Delta\tau\Gamma}\right]^{N/2}\left[\tanh\Delta\tau\Gamma\right]^{\delta} e^{-\Delta\tau[E_{\rm cl}(\boldsymbol{x}')-E_{\rm ref}]} \\ &= \left[\cosh\Delta\tau\Gamma\right]^{N}\left[\tanh\Delta\tau\Gamma\right]^{\delta} e^{-\Delta\tau[E_{\rm cl}(\boldsymbol{x}')-E_{\rm ref}]}. \end{aligned} \tag{1.35}$$

Note that the above Green's function is not normalized, hence it cannot be used as a transition probability of a Markov chain. Following the same procedure as in continuous space models, the Green's function is rewritten as in Equation (1.27). The part of the Green's function responsible for branching is then be given by

$$\begin{aligned} G_b(\boldsymbol{x},\boldsymbol{x}',\Delta\tau) &= \sum_{\boldsymbol{x}} G(\boldsymbol{x},\boldsymbol{x}',\Delta\tau) \\ &= e^{-\Delta\tau[E_{\rm cl}(\boldsymbol{x}')-E_{\rm ref}]}\sum_{\boldsymbol{x}}\left[\cosh\Delta\tau\Gamma\right]^{N}\left[\tanh\Delta\tau\Gamma\right]^{\delta}. \end{aligned} \tag{1.36}$$

Note that if the symmetrized version of the Trotter approximation was used, it wouldn't have been possible to perform the sum (as we will assume latter on) in $\boldsymbol{x}$ because the term $E_{\rm cl}(\boldsymbol{x})$.

The Green's function portraying the diffusion of spin configurations is:

$$\begin{aligned} G_d(\boldsymbol{x}, \boldsymbol{x}', \Delta\tau) &= \frac{G(\boldsymbol{x}, \boldsymbol{x}', \Delta\tau)}{\sum_{\boldsymbol{x}} G(\boldsymbol{x}, \boldsymbol{x}', \Delta\tau)} \\ &= \frac{\Big[\cosh \Delta\tau\Gamma\Big]^N \Big[\tanh \Delta\tau\Gamma\Big]^\delta}{\underbrace{\sum_{\boldsymbol{x}} \Big[\cosh \Delta\tau\Gamma\Big]^N \Big[\tanh \Delta\tau\Gamma\Big]^\delta}_{J}}, \end{aligned}$$

where,

$$\begin{aligned} J &= \sum_{\boldsymbol{x}} \left[\frac{e^{\Delta\tau\Gamma} + e^{-\Delta\tau\Gamma}}{2}\right]^N \left[\frac{e^{\Delta\tau\Gamma} - e^{-\Delta\tau\Gamma}}{e^{\Delta\tau\Gamma} + e^{-\Delta\tau\Gamma}}\right]^\delta \\ &= \frac{1}{2^N} \sum_{\boldsymbol{x}} \Big(\underbrace{e^{\Delta\tau\Gamma} - e^{-\Delta\tau\Gamma}}_{A}\Big)^\delta \Big(\underbrace{e^{\Delta\tau\Gamma} + e^{-\Delta\tau\Gamma}}_{B}\Big)^{N-\delta}. \end{aligned}$$

Let $n_p$ be the number of spin configurations that have $p$ spins overturned when $|\boldsymbol{x}'\rangle$ goes to $|\boldsymbol{x}\rangle$. Then,

$$\begin{aligned} J &= \frac{1}{2^N}\Big(n_0 A^0 B^N + n_1 A^1 B^{N-1} + n_2 A^2 B^{N-2} + ... + n_N A^N B^0\Big) \\ &= \frac{1}{2^N} \sum_{p=0}^{N} n_p A^p B^{N-p}, \end{aligned}$$

where $\sum_{p=0}^{N} n_p = 2^N$.

$n_p$ encodes information about the number of possible configurations that one obtains after flipping $p$ spins from a chain of $N$ spins. This is so because each spin configuration with $p$ overturned spins contribute equally to the above sum. It is a valid approach when assuming a uniform proposal probability of the spins to be flipped (no longer valid in an importance sampling scheme). Therefore,

$n_p = \binom{N}{p}$. We obtain a *Binomial identity*

$$\begin{aligned} J &= \frac{1}{2^N}\sum_{p=0}^{N}\binom{N}{p}A^p B^{N-p} \\ &= \frac{1}{2^N}(A+B)^N \\ &= e^{\Delta\tau\Gamma N}. \end{aligned}$$

Hence,

$$\begin{aligned} G_d(\boldsymbol{x},\boldsymbol{x}',\Delta\tau) &= \Big[\cosh\Delta\tau\Gamma\Big]^N\Big[\tanh\Delta\tau\Gamma\Big]^\delta e^{-\Delta\tau\Gamma N} \\ &= \left[\frac{\cosh\Delta\tau\Gamma}{e^{\Delta\tau\Gamma}}\tanh\Delta\tau\Gamma\right]^\delta\left[\frac{\cosh\Delta\tau\Gamma}{e^{\Delta\tau\Gamma}}\right]^{(N-\delta)} \\ &= P_F^\delta(1-P_F)^{N-\delta}, \end{aligned} \tag{1.37}$$

with $P_F = \frac{\sinh\Delta\tau\Gamma}{e^{\Delta\tau\Gamma}}$ being the probability for *single-spin flip*.

Using Eq. 1.36, we write the branching part of the Green's function as

$$G_b(\boldsymbol{x},\boldsymbol{x}',\Delta\tau) = e^{-\Delta\tau[E_{\rm cl}(\boldsymbol{x}')-E_{\rm ref}]}e^{\Delta\tau\Gamma N} = e^{-\Delta\tau[E_{\rm loc}(\boldsymbol{x}')-E_{\rm ref}]}, \tag{1.38}$$

where the local energy is given by:

$$E_{\rm loc}(\boldsymbol{x}) = E_{\rm cl}(\boldsymbol{x}) - \Gamma N. \tag{1.39}$$

The propagator in Eq. (1.37) defines a positive-definite and column-normalized matrix. Hence, it can be used to define a conventional Markov chain. Specifically, at each iteration every spin is addressed and flipped with a probability $P_F$. Alternatively, one samples the number $\delta$ of spins to be reversed from a binomial probability distribution, and then randomly selects which spins to flip, uniformly. In practice, the two algorithms turn out to have comparable computational costs.

The quantity in Eq. (1.38) can be written in a symmetrized version so as to reduce

the systematic bias coming from the time step, i.e:

$$G_b(\boldsymbol{x}, \boldsymbol{x}', \Delta\tau) = e^{-\Delta\tau[\frac{E_{\rm loc}(\boldsymbol{x})+E_{\rm loc}(\boldsymbol{x}')}{2} - E_{\rm ref}]}. \tag{1.40}$$

We have checked on the quantum Ising chain that using $G_b(\boldsymbol{x}, \boldsymbol{x}', \Delta\tau)$ in the above form does give correct results of the ground state energy (as shown latter on), with a relative error that grows quadratically with the time step in the limit where $\Delta\tau \to 0$. The ground state energy of the Hamiltonian in Eq. (1.31) is calculated using the expression of the local energy given by the formula in Equation (1.39).

We now move to the incorporation of importance sampling in the PQMC method, where we use a Taylor expansion to approximate the Green's function instead of the Trotter approximation. The modified imaginary-time Schrödinger equation is solved via a Markov process defined by the following equation:

$$f(\boldsymbol{x}, \tau + \Delta\tau) = \sum_{\boldsymbol{x}'} \tilde{G}(\boldsymbol{x}, \boldsymbol{x}', \Delta\tau) f(\boldsymbol{x}', \tau), \tag{1.41}$$

where the modified Green's function is given by

$$\tilde{G}(\boldsymbol{x}, \boldsymbol{x}', \Delta\tau) = G(\boldsymbol{x}, \boldsymbol{x}', \Delta\tau) \frac{\psi_T(\boldsymbol{x})}{\psi_T(\boldsymbol{x}')}. \tag{1.42}$$

$G(\boldsymbol{x}, \boldsymbol{x}', \Delta\tau)$ is defined as in Eq. (1.4). A suitable approximation for the modified Green's function can be obtained by dividing the time step $\Delta\tau$ into $M$ shorter time steps

$$\delta\tau = \Delta\tau / M.$$

If $\delta\tau$ is sufficiently short, one can employ a Taylor expansion truncated at the linear term,

$$\tilde{G}(\boldsymbol{x}, \boldsymbol{x}', \Delta\tau) \cong [\tilde{g}(\boldsymbol{x}, \boldsymbol{x}', \delta\tau)]^M ,$$

where:

$$\tilde{g}(\boldsymbol{x}, \boldsymbol{x}', \delta\tau) = \left[\delta_{\boldsymbol{x},\boldsymbol{x}'} - \delta\tau(H_{\boldsymbol{x},\boldsymbol{x}'} - E_{\rm ref}\delta_{\boldsymbol{x},\boldsymbol{x}'}) \frac{\psi_T(\boldsymbol{x})}{\psi_T(\boldsymbol{x}')}\right] . \tag{1.43}$$

With this approximation, Eq. (1.41) defines a stochastic implementation of the power method of linear algebra. Convergence to the exact ground state is

guaranteed as long as $\delta\tau$ is smaller than a finite value, sufficiently small to ensure that all matrix elements of $\tilde{g}(\boldsymbol{x}, \boldsymbol{x}', \delta\tau)$ are not negative [38].

For a time step $\delta\tau$, following the same framework as in Equation (1.5), $\tilde{g}(\boldsymbol{x}, \boldsymbol{x}', \delta\tau)$ can be rewritten as

$$\tilde{g}(\boldsymbol{x}, \boldsymbol{x}', \delta\tau) = p_{\boldsymbol{x},\boldsymbol{x}'} \sum_{\boldsymbol{x}} \tilde{g}(\boldsymbol{x}, \boldsymbol{x}', \delta\tau'), \tag{1.44}$$

where $p_{\boldsymbol{x},\boldsymbol{x}'}$ is the stochastic matrix used to build to Markov chain. It is given by

$$p_{\boldsymbol{x},\boldsymbol{x}'} = \frac{\left[\delta_{\boldsymbol{x},\boldsymbol{x}'} - \delta\tau(H_{\boldsymbol{x},\boldsymbol{x}'} - E_{\mathrm{ref}}\delta_{\boldsymbol{x},\boldsymbol{x}'})\frac{\psi_T(\boldsymbol{x})}{\psi_T(\boldsymbol{x}')}\right]}{1 - \delta\tau(E_{\mathrm{loc}}(\boldsymbol{x}') - E_{\mathrm{ref}})}. \tag{1.45}$$

Given that the time step $\delta\tau$ is very small, most of the tentative to update a spin configuration from $\boldsymbol{x}'$ to $\boldsymbol{x}$ will fail. This is because in this regime the identity operator dominates ($p_{\boldsymbol{x},\boldsymbol{x}'} \approx \delta_{\boldsymbol{x},\boldsymbol{x}'}$) hence, the result will be extremely long autocorrelation times leading to pathologically inefficient simulations.

This problem can be solved by adopting the continuous-time Green's function Monte Carlo (CTGFMC) algorithm which was first introduced in [39]. An interested reader may also refer to Ref. [27] for more details of the algorithm.

The idea is to formally take the $M \to \infty$ limit, and determine the (stochastic) time interval $\delta\tau'$ that passes before the next configuration update occurs. It is convenient to bookkeep the remaining time $\delta\tau_t$ left to complete a total interval of time $\Delta\tau$. This is to ensure that each iteration of the PQMC simulation corresponds to a time step of duration $\Delta\tau$. The time interval $\delta\tau'$ is sampled using the formula

$$\delta\tau' = \mathrm{Min}\Big(\delta\tau_t, \frac{\ln(1-\xi)}{E_{\mathrm{loc}}(\boldsymbol{x}') - E_{\mathrm{cl}}(\boldsymbol{x}')}\Big),$$

with $\xi \in (0, 1)$ being a uniform random number.

With the use of importance sampling, the expression local energy is modified. Its

new expression is computed using Equation (1.16).

$$\begin{aligned} E_{\mathrm{loc}}(\boldsymbol{x}) &= \frac{\langle \boldsymbol{x}|\hat{H}|\psi_T\rangle}{\langle \boldsymbol{x}|\psi_T\rangle} \\ &= \frac{\langle \boldsymbol{x}|\hat{H}_{\mathrm{cl}}|\psi_T\rangle}{\langle \boldsymbol{x}|\psi_T\rangle} + \frac{\langle \boldsymbol{x}|\hat{H}_{\mathrm{kin}}|\psi_T\rangle}{\langle \boldsymbol{x}|\psi_T\rangle} \\ &= E_{\mathrm{cl}}(\boldsymbol{x}) + \frac{\Big[-\Gamma\Big(\langle \tilde{\boldsymbol{x}}_1| + \langle \tilde{\boldsymbol{x}}_2| + \cdots + \langle \tilde{\boldsymbol{x}}_N|\Big)|\psi_T\rangle\Big]}{\langle \boldsymbol{x}|\psi_T\rangle} \\ &= E_{\mathrm{cl}}(\boldsymbol{x}) - \Gamma\frac{\sum_{i=1}^{N}\psi_T(\tilde{\boldsymbol{x}}_i)}{\psi_T(\boldsymbol{x})}. \end{aligned} \tag{1.46}$$

$|\tilde{\boldsymbol{x}}_i\rangle$ is equivalent to the spin configuration $|\boldsymbol{x}\rangle$ where the spin $i$ has been overturned.

Remark that, in the case of simple PQMC for which the trial function is a constant, say $\psi_T(\boldsymbol{x}) = 1$, the form of the local energy in Eq. (1.39) is recovered. It is interesting to note that, for a constant trial function, since $\hat{H}_{\mathrm{kin}}$ has a non-zero action on the basis state $|\boldsymbol{x}\rangle$, the local energy has a term which depends on the kinetic energy. This is in contrast to an equivalent situation for continuous space models where the kinetic energy operator will always annihilate a constant guiding function (refer to Eq. 1.21).

The spin-configuration update $\boldsymbol{x}' \to \boldsymbol{x}$ (with $\boldsymbol{x}' \neq \boldsymbol{x}$) is then randomly selected from the probability distribution

$$\begin{aligned} t_{\boldsymbol{x},\boldsymbol{x}'} &= \frac{p_{\boldsymbol{x},\boldsymbol{x}'}}{\sum_{\boldsymbol{x}\neq\boldsymbol{x}'} p_{\boldsymbol{x},\boldsymbol{x}'}} \\ &= \begin{cases} \frac{\psi_T(\boldsymbol{x}_i)}{\sum_{i=1}^{N}\psi_T(\tilde{\boldsymbol{x}}_i)}, & \text{if } \boldsymbol{x}' \text{ is different from } \boldsymbol{x} \text{ of a single spin flip} \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \tag{1.47}$$

Notice that for a constant guiding function, the probability to propose an update of the spin configuration is uniform,

$$t_{\boldsymbol{x},\boldsymbol{x}'} = \begin{cases} \frac{1}{N}, & \text{if } \boldsymbol{x}' \text{ is different from } \boldsymbol{x} \text{ of a single spin flip} \\ 0, & \text{otherwise.} \end{cases}$$

Thus, Equation (1.47) portrays the facet of the trial function driving the configurations towards more favorable regions in the configuration space, assuming of

course that $\psi_T(x)$ is appropriately chosen.

The weight-update factor for the branching process takes the exponential form

$$b_{\boldsymbol{x}'} \;= \sum_{\boldsymbol{x}} \tilde{g}(\boldsymbol{x}, \boldsymbol{x}', \delta\tau') \;= e^{-\delta\tau'[E_{\mathrm{loc}}(\boldsymbol{x}')-E_{\mathrm{ref}}]},$$

with the local energy given by Equation (1.46).

In summary, the CTGFMC algorithm requires to perform, for each walker $n$ in the population, the following steps:

**i)** initialize the time interval $\delta\tau_t = \Delta\tau$, and the weight factor $w_n = 1$;

**ii)** sample the time $\delta\tau'$ at which the the configuration update $\boldsymbol{x}' \to \boldsymbol{x}$ might occur;

**iii)** if $\delta\tau' < \delta\tau_t$, update $\boldsymbol{x}'$ with a transition probability $t_{\boldsymbol{x},\boldsymbol{x}'}$ in Eq. (1.47), else set $\delta\tau' = \delta\tau_t$;

**iv)** accumulate the weight factor according to the rule $w_n \to w_n b_{\boldsymbol{x}'}$ and set $\delta\tau_t \to \delta\tau_t - \delta\tau'$;

**v)** Go back to step **ii)** until $\delta\tau_t = 0$;

**vi)** finally, perform branching according to the total accumulated weight factor $w_n$.

This continuous-time algorithm implicitly implements the exact imaginary-time modified Green's function $\tilde{G}(\boldsymbol{x}, \boldsymbol{x}', \Delta\tau)$.

### 1.3.4 Test case: The Ferromagnetic quantum Ising chain

In order to show that the PQMC algorithm on discrete models does recover ground-state properties of quantum many-body systems, we use as a test case the quantum Ising chain (QIC) with nearest neighbors interactions. Setting a uniform interaction strength ($J_{ij} \equiv J$) between spins in the Hamiltonian in

Equation (1.31) yields

$$\hat{H} = -J \sum_{i=1}^{N} \sigma_i^z \sigma_{i+1}^z - \Gamma \sum_{i=1}^{N} \sigma_i^x, \tag{1.48}$$

where the interaction energy $J > 0$ sets the energy scale. In the following, we set $J = 1$. All energy scales are henceforth expressed in units of $J$. Periodic boundary conditions are considered, i.e., $\sigma_{N+1}^a = \sigma_1^a$ where $a = x, y, z$.

At zero temperature this model undergoes a quantum phase transition from a paramagnetic phase at $\Gamma > 1$ to a ferromagnetic phase at $\Gamma < 1$. Its ground-state energy $E_{\rm JW}$ can be exactly determined by performing the Jordan–Wigner transformation, followed by a Fourier and the Bogoliubov transformations [40].

$$E_{\rm JW} = -\sum_{m=0}^{N-1} \left[ \Gamma^2 - 2\Gamma \cos \frac{2\pi}{N}(m + \frac{1}{2}) + 1 \right]^{1/2}. \tag{1.49}$$

Initially, we study the equilibration dynamics of the local energy (averaged over the current walkers population) computed at different Monte Carlo steps (MCs), using the formula in Eq. (1.39). It is worth mentioning that the equilibration time (informally, the number of PQMC steps required to reach a configuration which is statistically uncorrelated with the initial one) might increase with the system size or with other Hamiltonian parameters. Also such an effect could affect the scaling of the computational cost required to determine with the PQMC algorithm the ground state energy within a target accuracy.

We investigate (using simple PQMC) this possibility by analyzing the scaling of the equilibration time $\tau_{\rm eq}$ with the system size $N$ close to the ferromagnetic transition of the QIC. We set $\Gamma = 0.95$. For concreteness, we define the equilibration time as the number of PQMC iterations, times the time step $\Delta\tau$, required to reach a random-walker population with an average energy within $10\%$ of the equilibrium (i.e., long PQMC time) value. The walkers are initialized with random spin configurations, more precisely, for each replica in the population, a value between $-1$ and $1$ is assigned to each of its spin with a $50\%$ probability. Their number is scaled so that the systematic bias due to the finite population size is $1\%$ for all system sizes (as explained in a subsequent chapter).

Figure (1.2)(a) shows the histogram of the equilibration time (using about $100$ different runs) for different system sizes. The probability distributions of

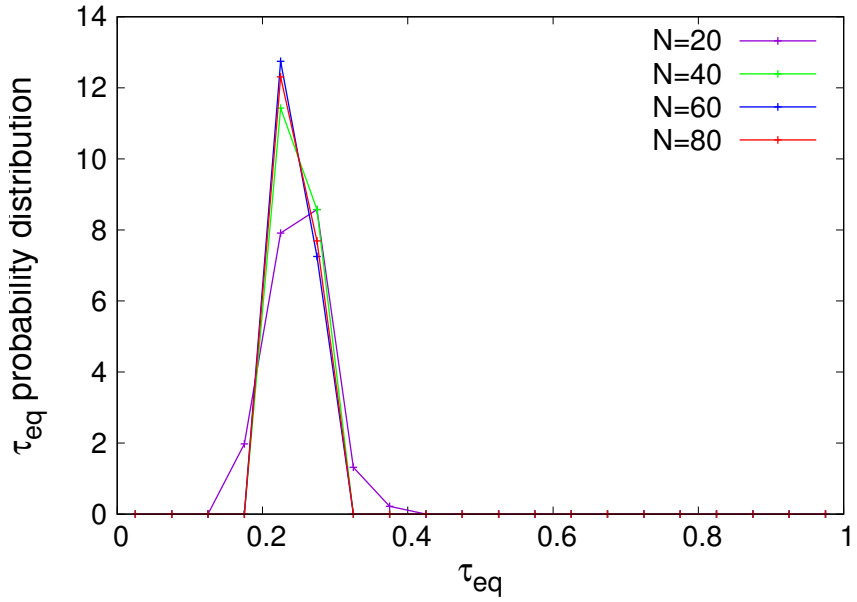
$\tau_{\rm eq}$ at different $N$ is centered around the same time interval. This suggests that $\tau_{\rm eq}$ doesn't scale up with the system size. Furthermore, as the results in Figure (1.2)(b) show, the average equilibration time is constant — within $10\%$ — when the system size varies from $N = 20$ to $N = 80$. This (admittedly non-exhaustive) analysis indicates that the finite number of random walkers is the most relevant source of systematic bias in simple PQMC simulations, as we will demonstrate latter on. A thorough analysis of the equilibration time in more complex (e.g., disordered) models would be quite useful, but is beyond the scope of this work.

We also report on the effect of the systematic error due to the time step $\Delta\tau$. This is evaluated by performing simulations with two different branching protocols. One stemming from the non symmetrized primitive approximation to the Green's function defined as in Equation (1.38). The other one from a symmetrized primitive approximation using Eq. (1.40). We quantify the accuracy by computing $e_{\rm rel}$, the relative error between the ground state energy $E$ obtained with PQMC simulations and $E_{\rm JW}$ defined in Eq. (1.49).

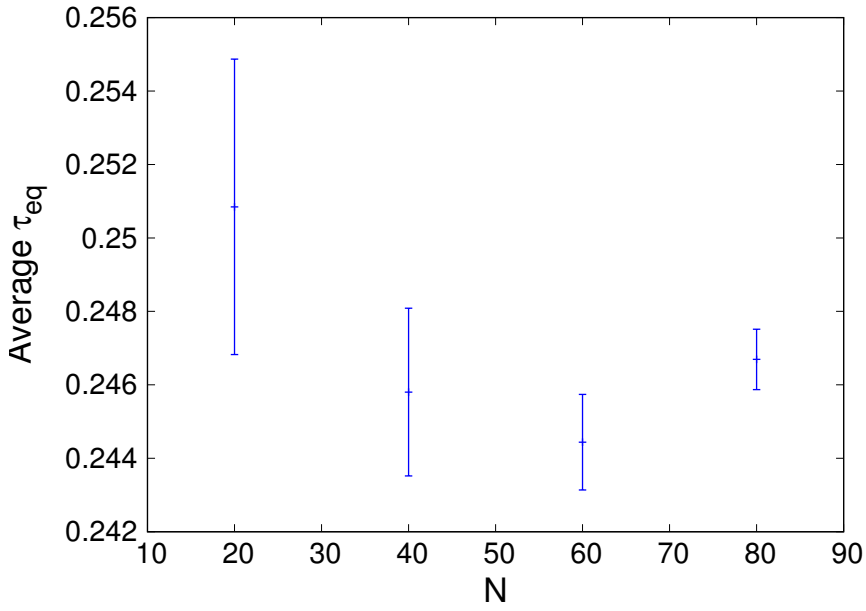
$$e_{\rm rel} = \frac{|E - E_{\rm JW}|}{|E_{\rm JW}|}. \tag{1.50}$$

Plots in Figure (1.3) clearly portray the difference between the two approximations. As expected a linear dependence on $e_{\rm rel}$ is found when using a non symmetrized Trotter approximation of the imaginary-time evolution operator. While, a quadratic dependence on $\Delta\tau$ is found when a symmetric Trotter approximation is used. For both cases, very small values of the time step is sufficient to find correct estimates, but at the expenses of running very long simulations. All through the thesis (except for CTGFMC), we do branching with Eq. (1.40) with a (small) fixed value of the time step, $\Delta\tau$ between $0.001$ and $0.01$. The different choices of $\Gamma$ in the figure are motivated by the need to be in a regime where the systematic bias due to the finite number of walkers $N_w$ does not wash out the scaling of $\Delta\tau$. The effect of $N_w$, which is arguably the most severe source of systematic errors in PQMC methods will be discussed more exhaustively in subsequent chapters.

In Figure (1.4), we plot the zero temperature equation of state of the QIC using the simple PQMC. Good agreement is found with the exact values of the ground state energy. Hence, we can conclude that, when keeping under control all the systematic errors of the algorithm, PQMC simulations are indeed powerful

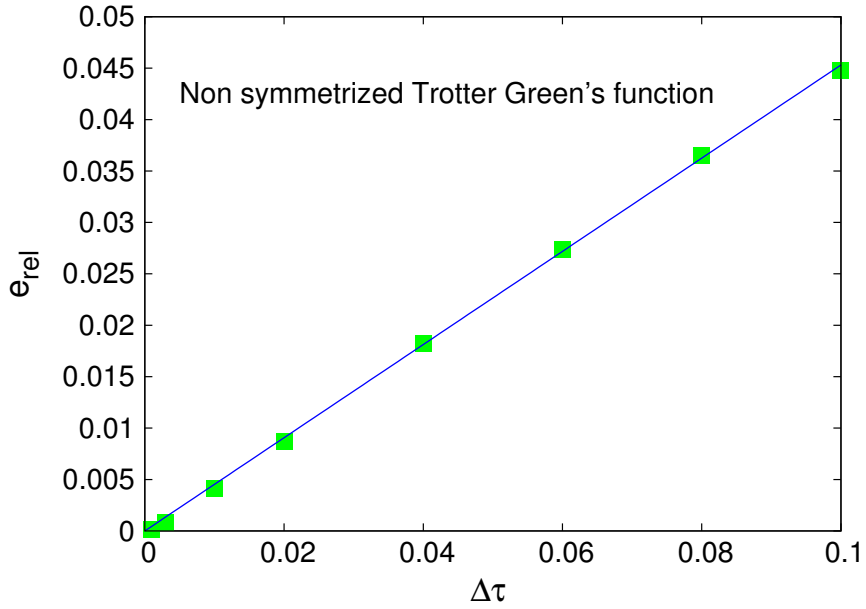



(a)

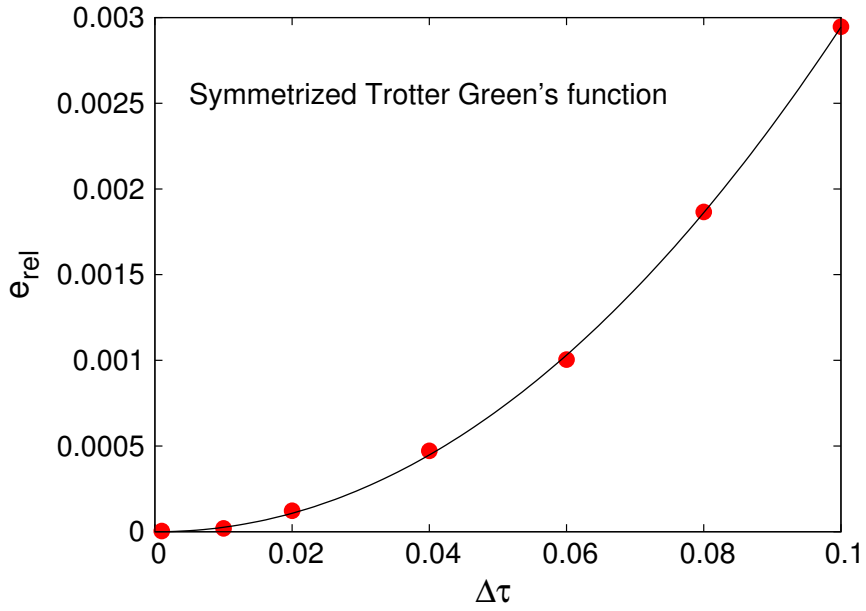



(b)

Figure 1.2: (a) Plot of the probability distribution of the equilibration time $\tau_{\rm eq}$ for distinct runs of the PQMC simulations at different system sizes $N$. The definition of the equilibration time is given in the text. (b) Average value of $\tau_{\rm eq}$ for different sizes $N$ of the spin chain. Simulations are performed for $\Gamma = 0.95$ in both (a) and (b).




(a)




(b)

Figure 1.3: Relative error as a function of the time step $\Delta\tau$. The values of $\Gamma$ are different in both plots and are chosen in a regime where the bias due to the final walker population is minimal. $N_w = 20,000$. Data are computed for a chain with $N = 20$ spins. (a) In this plot, branching is done with the Green's function in Equation (1.38) which corresponds to a non symmetrized version of the Suziki-Trotter approximation (see Equation 1.8). The intensity of the transverse field is $\Gamma = 0.9$. The blue line is a linear fit to the data. (b) The red circles data are obtained with a branching step performed using Eq. (1.40), the symmetrized version of the Trotter step. The black curve is a quadratic fit to the data. $\Gamma = 1.6$ is utilized. Error bars are smaller than the symbol size.

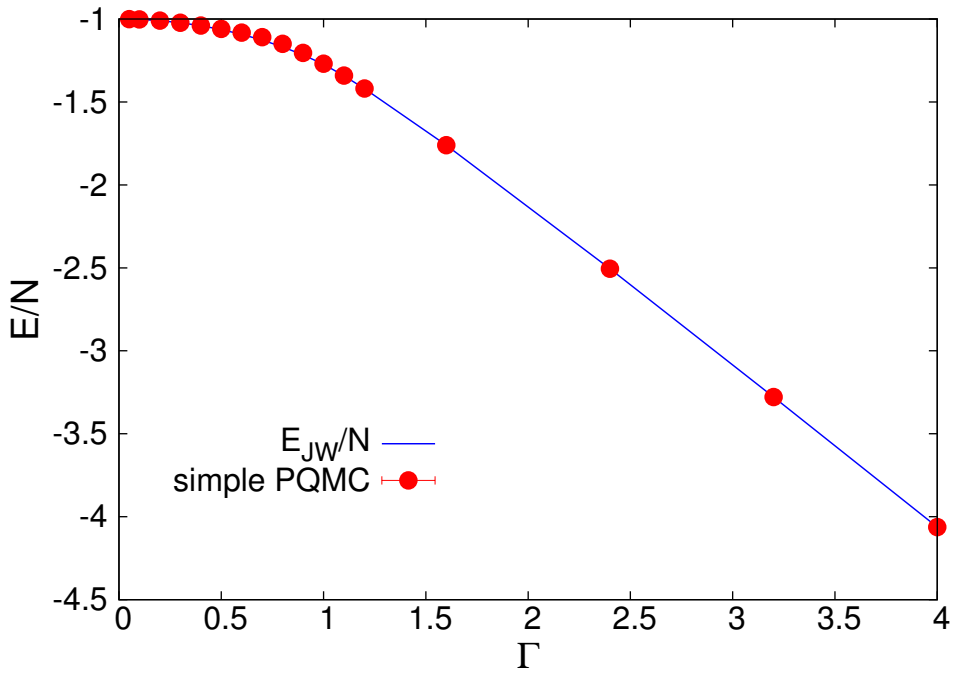



Figure 1.4: Ground state energy per spin of the ferromagnetic quantum Ising chain as a function of the transverse field intensity $\Gamma$. The red circles are data obtained with the simple PQMC method. The blue line corresponds to values of the exact ground-state energy $E_{\mathrm{JW}}$ per spin obtained using the formula in Equation (1.49). The average number of random walkers is $N_w = 20,000$ and the time step is $\Delta\tau = 10^{-3}$. The size of the spin chain is $N = 50$. Error bars are smaller than the symbol size.

methods to find ground state properties of quantum many-body systems.

## 1.4 Estimating observables

In this section, we propose a new way to evaluate pure estimators on the fly in simple PQMC simulations. The method at the present stage doesn't seem to be effective when computing equilibrium ground state properties of large system sizes. However, it is quite promising in the prospect of recovering pure estimators at the end of a quantum annealing simulations when using the simple PQMC method, even for large system sizes.

In simulations performed with PQMC methods, the ground state expectation value of a local operator $\hat{O}$ is computed with the probability distribution $p(\boldsymbol{x})$

defined in Equation (1.18). In this subsection, we will denote it as, $\langle\hat{O}\rangle_{mixed}$, the *mixed estimator* of the operator $\hat{O}$,

$$\langle\hat{O}\rangle_{mixed} = \frac{\langle\Psi_0|\hat{O}|\psi_T\rangle}{\langle\Psi_0|\psi_T\rangle}.$$

While, we will call $\langle\hat{O}\rangle_{pure}$ the *pure estimator* obtained using the quantum mechanical probability distribution:

$$\langle\hat{O}\rangle_{pure} = \frac{\langle\Psi_0|\hat{O}|\Psi_0\rangle}{\langle\Psi_0|\Psi_0\rangle}.$$

When $\hat{O}$ commutes with the Hamiltonian of the system, the mixed estimator, which is the quantity that can be directly computed in PQMC, has the same value (apart statistical errors) as the pure estimator i.e, $\langle\hat{O}\rangle_{pure} \approx \langle\hat{O}\rangle_{mixed}$. However, when $\hat{O}$ does not commute with the Hamiltonian, the mixed estimator does not provide the same result as the pure estimator.

It is worth noting that usually in PQMC simulations, the pure estimator is approximated using the Ceperley estimator [41]:

$$\langle\hat{O}\rangle_{pure} \approx 2\langle\hat{O}\rangle_{mixed} - \frac{\langle\psi_T|\hat{O}|\psi_T\rangle}{\langle\psi_T|\psi_T\rangle},$$

where the last term of the above equation is obtained from another set of simulations performed with Variational Monte Carlo. The accuracy of the Ceperley estimator depends on how accurate the guiding wave function is to the exact ground state wave-function.

Here, we test a different way of approximating pure estimators (on the fly) in projective Monte Carlo methods close in spirit to the forward walking technique [42]. The pure estimator is computed as follows (assuming $\hat{O}$ diagonal in the basis

state):

$$\begin{aligned}
\langle \hat{O} \rangle_{pure} &= \frac{\langle \Psi_0 | \hat{O} | \Psi_0 \rangle}{\langle \Psi_0 | \Psi_0 \rangle} \\
&= \frac{\sum_{\boldsymbol{x}} \langle \Psi_0 | \boldsymbol{x} \rangle \langle \boldsymbol{x} | \hat{O} | \Psi_0 \rangle}{\sum_{\boldsymbol{x}} \langle \Psi_0 | \boldsymbol{x} \rangle \langle \boldsymbol{x} | \Psi_0 \rangle} = \frac{\sum_{\boldsymbol{x}} \Psi_0^*(\boldsymbol{x}) O(\boldsymbol{x}) \frac{\psi_T(\boldsymbol{x})}{\psi_T(\boldsymbol{x})} \Psi_0(\boldsymbol{x})}{\sum_{\boldsymbol{x}} \Psi_0^*(\boldsymbol{x}) \frac{\psi_T(\boldsymbol{x})}{\psi_T(\boldsymbol{x})} \Psi_0(\boldsymbol{x})} \\
&= \frac{\sum_{\boldsymbol{x}} \Psi_0^*(\boldsymbol{x}) \left[ O(\boldsymbol{x}) \frac{\Psi_0(\boldsymbol{x})}{\psi_T(\boldsymbol{x})} \right] \psi_T(\boldsymbol{x})}{\sum_{\boldsymbol{x}} \Psi_0^*(\boldsymbol{x}) \left[ \frac{\Psi_0(\boldsymbol{x})}{\psi_T(\boldsymbol{x})} \right] \psi_T(\boldsymbol{x})} = \frac{\langle \Psi_0 | O(\boldsymbol{x}) \frac{\Psi_0(\boldsymbol{x})}{\psi_T(\boldsymbol{x})} | \psi_T \rangle}{\langle \Psi_0 | \frac{\Psi_0(\boldsymbol{x})}{\psi_T(\boldsymbol{x})} | \psi_T \rangle} \\
&= \frac{\langle O(\boldsymbol{x}) \frac{\Psi_0(\boldsymbol{x})}{\psi_T(\boldsymbol{x})} \rangle_{mixed}}{\langle \frac{\Psi_0(\boldsymbol{x})}{\psi_T(\boldsymbol{x})} \rangle_{mixed}}.
\end{aligned} \tag{1.51}$$

In the forward walking technique, the term $\frac{\Psi_0(\boldsymbol{x})}{\psi_T(\boldsymbol{x})}$ is estimated bookkeeping (on the fly) the asymptotic offspring of the descendants of the walker $\boldsymbol{x}$ (for details check [27]). Furthermore, this is done in a context of time-independent Hamiltonians. However, in the prospect of estimating pure averages in quantum annealing simulations, in which the Hamiltonian is time dependent, another alternative to estimate the ratio $\frac{\Psi_0(\boldsymbol{x})}{\psi_T(\boldsymbol{x})}$ is needed. This is because when the (annealing) simulation is stopped, forward walking can no longer be performed.

We propose to estimate the ratio in Eq. (1.51) using the simple PQMC method, thus we set the trial function to a constant. Suppose that at a given Monte Carlo sweep $i$, the number of walkers in the population is $N_{\text{cur}}$. Let $\bar{O}_i$ be the average value of the operator $\hat{O}$ taken over the current population of walkers,

$$\bar{O}_i = \frac{1}{N_{\text{cur}}} \sum_{n=1}^{N_{\text{cur}}} O_n^i.$$

$O_n^i$ designates the value that the local operator has on the $n^{th}$ walker at time $i$. Let $N_{\text{cur}}(\boldsymbol{x}_i^{(n)})$ be the number of walkers having the same spin configuration $\boldsymbol{x}_i^{(n)}$ and, $\chi_{\text{cur}}$ the total number of distinct spin configurations in a population $N_{\text{cur}}$. $\boldsymbol{x}_i^{(n)}$ stands for the configuration of the walker $n$ at time $i$. The amplitude of the ground state wave-function on that spin configuration can be estimated to be

equivalent to the density of walkers in that configuration; i.e.

$$\Psi_0(\boldsymbol{x}_i^{(n)}) \approx \frac{N_{\rm cur}(\boldsymbol{x}_i^{(n)})}{N_{\rm cur}}. \tag{1.52}$$

The above equation is correct provided that $N_{\rm cur}$ is (very) large, and representative enough of the configurations that have dominant weights. Hence, we obtain that:

$$\begin{aligned}
\bar{O}_i &= \frac{1}{N_{\rm cur}} \sum_{n=1}^{N_{\rm cur}} O_n^i \\
&= \frac{1}{N_{\rm cur}} \Big[ O_1^i + O_2^i + ... + O_{N_{\rm cur}}^i ) \Big] \\
&= \frac{1}{N_{\rm cur}} \Big[ O_1^i N_{\rm cur}(\boldsymbol{x}_i^{(1)}) + O_2^i N_{\rm cur}(\boldsymbol{x}_i^{(2)}) + ... + O_{\chi_{\rm cur}}^i N_{\rm cur}(\boldsymbol{x}_i^{(\chi_{\rm cur})}) \Big] \\
&= \sum_{m=1}^{\chi_{\rm cur}} O_m^i \frac{N_{\rm cur}(\boldsymbol{x}_i^{(m)})}{N_{\rm cur}} \\
&\approx \sum_{m=1}^{\chi_{\rm cur}} O_m^i \Psi_0(\boldsymbol{x}_i^{(m)}).
\end{aligned} \tag{1.53}$$

Therefore, Eq. (1.51) becomes

$$\begin{aligned}
\langle \hat{O} \rangle_{pure} &= \frac{\langle O(\boldsymbol{x}) \Psi_0(\boldsymbol{x}) \rangle_{mixed}}{\langle \Psi_0(\boldsymbol{x}) \rangle_{mixed}} \\
&= \frac{\overline{O(\boldsymbol{x}_i) \Psi_0(\boldsymbol{x}_i)}}{\overline{\Psi_0(\boldsymbol{x}_i)}} \\
&\approx \frac{\Big[ \sum_{m=1}^{\chi_{\rm cur}} O_m^i \Psi_0^2(\boldsymbol{x}_i^{(m)}) \Big]}{\Big[ \sum_{m=1}^{\chi_{\rm cur}} \Psi_0^2(\boldsymbol{x}_i^{(m)}) \Big]}.
\end{aligned} \tag{1.54}$$

We compute the estimate of square of the wave-function in two ways. The first one by simply squaring the value in Equation (1.52),

$$\Psi_0^2(\boldsymbol{x}_i^{(n)}) \approx \Big( \frac{N_{\rm cur}(\boldsymbol{x}_i^{(n)})}{N_{\rm cur}} \Big)^2, \tag{1.55}$$

We stress on the fact that this procedure is not statistically correct, but it is however used as a simple approximation that might be useful in comparisons

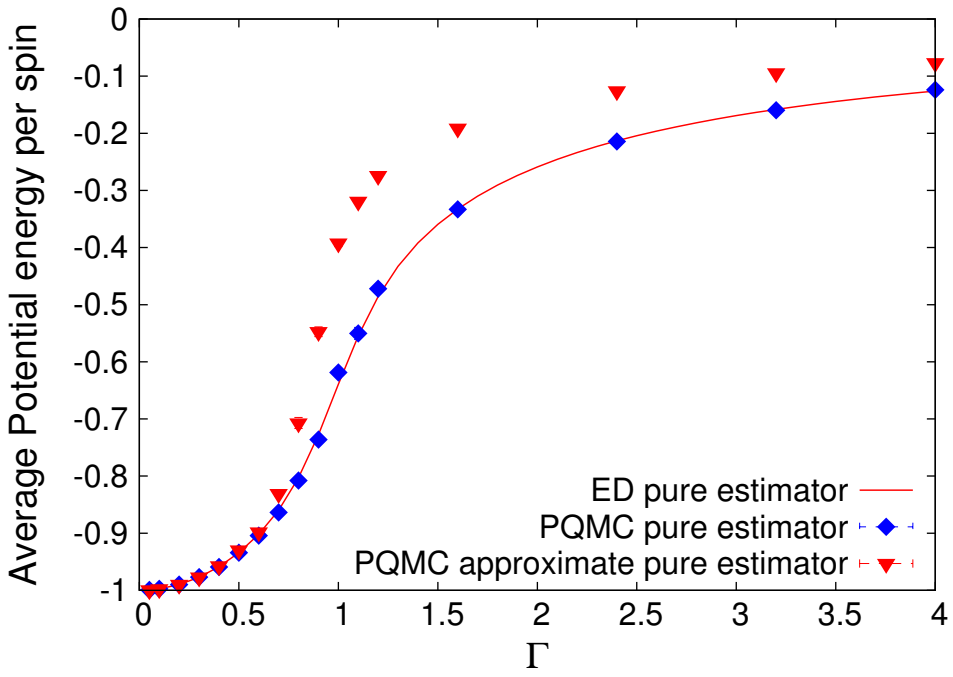



Figure 1.5: Comparison of the pure estimator of the ground state potential energy per spin obtained using respectively the almost pure estimator (see Eq. (1.55)), and the pure estimator (see Eq. (1.56)). The test bed is the quantum Ising chain. The potential energy is plotted with respect to the transverse field $\Gamma$. The red and blue symbols are data obtained with the simple PQMC with respectively the approximate estimator and the pure estimator. The solid lines are exact diagonalization (ED) data for the pure estimator. The number of walkers is $N_w = 20,000$ and the time step is $\Delta\tau = 0.01$. We considered a chain with $N = 20$ spins. Error bars are smaller than the symbol size.

between PQMC simulations and adiabatic quantum computers.

The second way, which is the statistically correct procedure, is implemented evolving in parallel two distinct replicas population $A$ and $B$ such that:

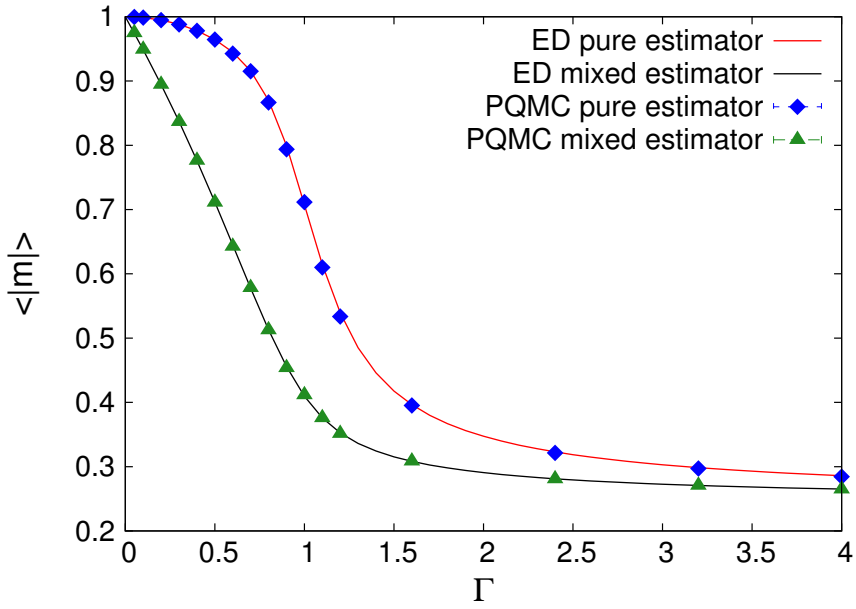
$$\Psi_0^2(\boldsymbol{x}_i^{(n)}) \approx \left(\frac{N_{\rm cur}(\boldsymbol{x}_i^{(n)})}{N_{\rm cur}}\right)^A \times \left(\frac{N_{\rm cur}(\boldsymbol{x}_i^{(n)})}{N_{\rm cur}}\right)^B . \tag{1.56}$$

In Figure (1.5), we test both estimators by computing the expectation value of the potential energy in the ground state of the QIC for different transverse field intensities. The simulations are carried out on a chain of 20 spins, without using importance sampling. Comparisons are made using exact diagonalization data (ED) obtained with the pure estimator. It is clearly seen that approximating the

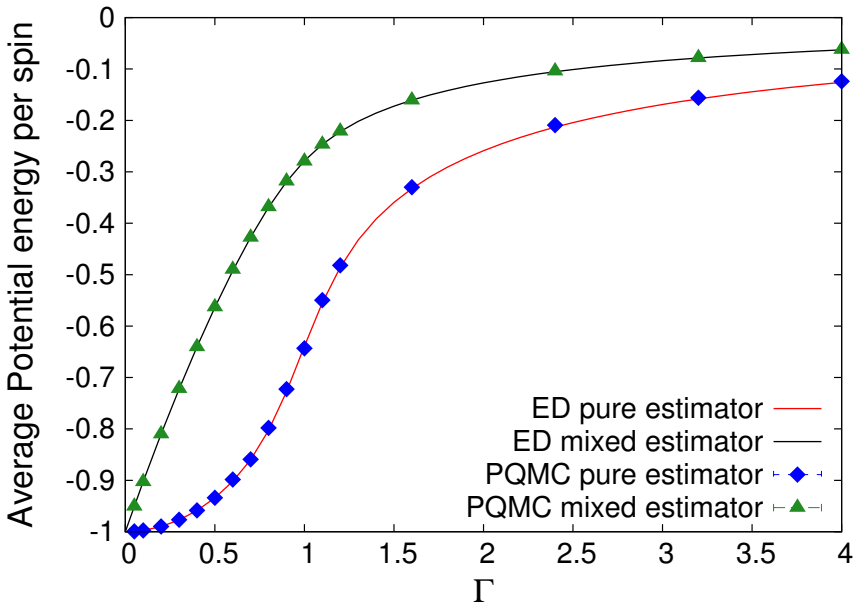
ground state wave-function with Equation (1.55) what we call "approximate estimator" is not always correct. Whereas, using Equation (1.56) renders satisfactory results. Nevertheless, the "approximate" estimator turns out to be quite accurate at small $\Gamma$. This is the interesting regime for quantum annealing simulations, since in that case, one makes measurements at the end of the annealing, where $\Gamma$ is $0$.

In Figure (1.6), we plot also the average of the absolute value of the magnetization per spin $m = \frac{1}{N}\sum_{i=1}^{N}\sigma_i^z$ of the QIC. The blue diamond data are obtained using simple PQMC with the pure estimator, which we showed to be the "correct" procedure. Comparisons are also made with ED, taking into account corresponding estimators. The mixed estimates of both the magnetization and the potential energy give wrong results as expected (check the green triangles data). However, it is interesting to notice that when $\Gamma \approx 0$, the mixed and pure estimators data converge to the same value. Henceforth, it could then be possible that estimating with a mixed estimator say, the average potential energy at the end of annealing (when $\Gamma \to 0$), that one does obtains correct values.

In conclusion to this section, we have to say that there is a price to pay in estimating the ground state wave-function as Eq. (1.56). One should expect that the number of walkers needed to have a non-zero signal in the approximation $\Psi_0^2(\boldsymbol{x}_i^{(n)})$ should grow with the system size (potentially in an exponential fashion). One has a signal only if there is some overlap in the configurations that have nonzero occupation in the two parallel simulations. In fact, simulating a chain of $40$ spins with $20,000$ walkers already displays this pathology (data not shown). It is possible however that using a good trial wave-function reduces considerably the number of walkers needed, thus opening the way for large scale simulations. However, in annealing simulations, at the end when quantum fluctuations have been completely removed, one expects that only few configurations are occupied, thus giving high chance of overlap. Therefore, this method is a promising tool in that context.




(a)




(b)

Figure 1.6: (a) Expectation value of the absolute value of the ground-state magnetization per spin of the quantum Ising chain with respect to $\Gamma$. The solid lines are exact diagonalization (ED) data taken on different estimators (see text). Blue and green filled symbols are data obtained with the simple PQMC with respectively pure and mixed estimators. (b) Average ground state potential energy per spin of the quantum Ising chain computed using simple PQMC. The labels are identical to the plot in (a). In both figures, the target number of walkers is $N_w = 20,000$ and the time step is $\Delta\tau = 0.01$. A chain of $N = 10$ spins is considered. Error bars are smaller than the symbol size.

# 2 Simulated quantum annealing of double-well and multi-well potentials

We analyze the performance of quantum annealing as a heuristic optimization method to find the absolute minimum of various continuous models, including landscapes with only two wells and also models with many competing minima and with disorder. The simulations performed using a projective quantum Monte Carlo (QMC) algorithm are compared with those based on the finite-temperature path-integral QMC technique and with classical annealing. We show that the projective QMC algorithm is more efficient than the finite-temperature QMC technique, and that both are inferior to classical annealing if this is performed with appropriate long-range moves. However, as the difficulty of the optimization problem increases, classical annealing looses efficiency, while the projective QMC algorithm keeps stable performance and is finally the most effective optimization tool. We discuss the implications of our results for the outstanding problem of testing the efficiency of adiabatic quantum computers using stochastic simulations performed on classical computers.

## 2.1 Introduction

Recent extraordinary developments in the technology of superconducting flux qubits give us well-grounded hope that adiabatic quantum computers capable to solve large-scale optimization problems via quantum annealing will be available in the near future [43, 44, 20]. However, the currently available quantum annealers did not demonstrate superiority with respect to state-of-the-art classical optimization algorithms [45, 46], and it is still under investigation whether their quantum features play a fundamental functional role in the optimization process [47, 48, 49, 50, 51]. In fact, it is not even clear if, at least under

certain circumstances, one should expect quantum annealing to be superior to classical methods [52], in particular to simulated (classical) annealing [53]. Some indications suggesting the supremacy of quantum annealing were originally provided by experiments performed on disordered magnetic materials [54]. Unfortunately, giving a definite answer to this issue using classical computers is not straightforward [1], since the computational times required to exactly simulate the real-time dynamics of the quantum annealing process (as defined in the framework of adiabatic quantum computing [2]) increase exponentially with the number of variables. Therefore, one has to resort to approximate simulation methods. The most relevant one consists in performing stochastic simulations based on quantum Monte Carlo (QMC) algorithms.

It is of outstanding importance to establish if and how computer simulations based on stochastic methods (hereafter referred to as simulated quantum annealing (SQA)) can be used to ascertain the superiority of quantum annealing versus classical algorithms. This would permit us to understand under which conditions quantum speed-up is attainable [45], and to identify the distinctive signatures of quantum effects to be sought for in a quantum device.

In this chapter, we analyze the performance of simulated quantum annealing (SQA) in finding the absolute minimum of simple double-well potentials and in more intricate multi-well potentials with disorder and with competing interaction terms. Compared to the Ising models, such continuous-space potentials allow us to tune more easily the difficulty of the optimization problem, an aspect which was indeed found to be of crucial importance for a fair assessment of the performance of quantum annealing [55]. Furthermore, effective double-well and multi-well potentials have recently been implemented on a D-Wave machine [18], thus demonstrating that such models could indeed have direct technological applications.

In most previous studies addressing analogous problems, the simulations were performed using the path integral Monte Carlo (PIMC) method (relevant exceptions are Refs. [11, 14, 56]). This is designed to simulate quantum many-body systems at finite temperatures, and is based on an effective classical model which evolves according to the stochastic dynamics defined by the Metropolis algorithm. Instead, here we employ a projective QMC technique, namely the diffusion Monte Carlo (DMC) algorithm, which was described in the previous chapter.

The layout of the chapter is as follows: in Section 2.2 we describe the implemen-

tation of SQA with the DMC algorithm, as well as the CA methods we employ to perform comparisons with SQA. In Section 2.3, we first consider the optimization of both symmetric and asymmetric double-well potentials with different types of SQA and CA methods. Then, we address more intricate models with many closely competing minima characterized by an increasing degree of difficulty, including: the multi-well washboard potential, the quasi-periodic (double-sinusoidal) potential, and a two-particle model with competing interaction terms. Our conclusions concerning the potential supremacy of quantum annealing and the possibility to analyze its efficiency with stochastic simulations are reported in Section 2.5.

## 2.2 Simulated quantum annealing and classical annealing methods

The DMC algorithm is one of the most powerful stochastic techniques to simulate the ground state of quantum many-body systems [57, 8]. It has proven to be extremely effective in numerous studies of divers systems, including electrons in solids, quantum fluids, nuclear matter, ultracold atoms, and also discrete lattice models.

In this chapter, we consider one-particle and two-particle continuous-space models in one spatial dimension. The Hamiltonian can be written in the generic form:

$$\hat{H} = -\frac{1}{2m}\sum_{i=1}^{N}\nabla_i^2 + V(\boldsymbol{x}), \tag{2.1}$$

where $m$ is the particles mass, $\boldsymbol{x} = (x_1, \ldots, x_N)$ denotes the particles configuration, with $x_i$ the position of the particle $i$ (with $i = 1, \ldots, N$), and $N$ is the particle number. We consider only the two cases $N = 1$ and $N = 2$. The total potential-energy operator

$$V(\boldsymbol{x}) = \sum_{i<j} v_{\text{int}}(|x_i - x_j|) + \sum_i v_{\text{ext}}(x_i), \tag{2.2}$$

is composed by the two-body inter-particle interaction $v_{\text{int}}(x)$ and by the external potential $v_{\text{ext}}(x)$.

Here, we employ the DMC algorithm exhaustively described in the subsection 1.3.1 of the previous chapter. The only difference being that the Hamiltonian

now is time-dependent. In this chapter, we are interested in using the DMC algorithm as a heuristic optimization method which searches for the optimal configuration $\boldsymbol{x}_{\min}$, where the potential attains its minimum value $V_{\min} = V(\boldsymbol{x}_{\min})$. This can be achieved by implementing a quantum annealing process, in which quantum fluctuations are gradually suppressed during the stochastic imaginary-time evolution. The suppression of quantum fluctuations can be enforced by reducing the diffusion coefficient $D = 1/2m$, which is equivalent to a particle mass increase. This reduces the quantum delocalization of the particle position, thus favoring random-walkers localization in the configuration $\boldsymbol{x}_{\min}$ corresponding to the classical absolute minimum.

$D = D(\tau)$ is now time-dependent in a step-wise manner (the imaginary-time can take only the discrete values $\tau = 0, \Delta\tau, 2\Delta\tau, \dots$) and, in each time-interval the Green's function $\tilde{G}(\boldsymbol{x}, \boldsymbol{x}', \Delta\tau)$ (in Equation 1.26) corresponding to a time-independent Hamiltonian is employed [14]. Equation (1.25) now defines an inhomogeneous Markov chain, since the transition matrix varies at each step, due to the (discrete) changes in $D(\tau)$. Rigorous sufficient conditions for the ergodicity and for the convergence of this quantum annealing method based on the inhomogeneous Markov chain have been derived in Ref. [31]. The corresponding conditions for CA were derived in Ref. [58].

Quantum annealing simulated using the DMC method is thus implemented using the following protocol:
First, we make the random-walkers population equilibrate by applying the standard DMC algorithm with a constant $D = D_{\text{ini}}$ for a sufficiently long equilibration time $\tau_{\text{eq}}$. The walkers then distribute according to

$$f(\boldsymbol{x}, \tau_{\text{eq}}) = \Psi_0^{\text{ini}}(\boldsymbol{x})\psi_T(\boldsymbol{x}), \tag{2.3}$$

where $\Psi_0^{\text{ini}}(\boldsymbol{x})$ is the ground-state wave function at $D = D_{\text{ini}}$.
Then, we run the DMC algorithm for a (long) annealing time $\tau_f$ while decreasing the effective diffusion coefficient after each time-step $\Delta\tau$ according to the step-wise linear law

$$D(\tau) = D_{\text{ini}} - \Delta D \frac{\tau}{\Delta\tau}, \tag{2.4}$$

where,

$$\Delta D = D_{\rm ini}\frac{\Delta\tau}{\tau_f}. \tag{2.5}$$

Here, the imaginary time $\tau$ is measured from the end of the equilibration time. At the end of the annealing process the diffusion coefficient vanishes $D(\tau_f) = 0$, while during the last DMC step it is $D(\tau_f - \Delta\tau) = \Delta D$. All quantum annealing simulations reported in this chapter start with $D_{\rm ini} = 0.5$ (this is equivalent to an initial mass $m = 1$).

For an infinitely slow quantum annealing process (corresponding to $\tau_f \to \infty$), the random-walkers population would follow the adiabatic ground-state wave function at $D(\tau)$ (multiplied times the trial function $\psi_T(\boldsymbol{x})$), which gradually shrinks in the minima of the potential landscape. Therefore, at the end of the quantum annealing process all random walkers would concentrate in the absolute minimum $\boldsymbol{x}_{\rm min}$ [31].

The key issue we investigate is how efficiently the absolute minimum is found for finite $\tau_f$. To quantify the efficiency of the optimization algorithm we measure the average of the potential energies computed in the configurations corresponding to the final random-walkers populations, formally written as:

$$\bar{V}(\tau_f) = \frac{\int {\rm d}\boldsymbol{x} V(\boldsymbol{x}) f(\boldsymbol{x},\tau_f)}{\int {\rm d}\boldsymbol{x} f(\boldsymbol{x},\tau_f)}. \tag{2.6}$$

In the standard DMC formalism this formula would correspond to the mixed estimator (refer to section 1.4) of the potential energy. In particular, we analyze the dependence of the residual energy

$$\varepsilon_{\rm res} = \bar{V}(\tau_f) - V_{\rm min}. \tag{2.7}$$

as a function of the total annealing time $\tau_f$. Notice that the total number of DMC steps in the annealing process is $\tau_f/\Delta\tau$ (we use fixed time-steps, in the range $2\Delta\tau D_{\rm ini} \in [0.01, 0.1]$), simply proportional to the total annealing time; this number determines the run time of the simulation on the classical computer.

In order to benchmark the performance of the DMC algorithm, we also perform CA simulations. In CA one uses the Metropolis algorithm to sample configurations

according to the Boltzmann canonical distribution

$$P(\boldsymbol{x}) = \frac{e^{-V(\boldsymbol{x})/T}}{Z}, \tag{2.8}$$

where,

$$Z = \int \mathrm{d}\boldsymbol{x} e^{-V(\boldsymbol{x})/T},$$

and $T$ is the temperature of a fictitious classical statistical system (we chose units such that the Boltzmann constant is $k_B = 1$). The temperature is gradually reduced during the simulation, thus removing thermal fluctuations. We adopt a linear annealing schedule of the temperature:

$$T(\tau) = T_{\mathrm{ini}}(1 - \frac{\tau}{\tau_f}), \tag{2.9}$$

with $\tau = 0, 1, \ldots, \tau_f$ an integer counting the Monte Carlo sweeps (a number of proposed updates equal to the number of variables), and $T_{\mathrm{ini}}$ the initial temperature.

The Markov chain is specified by the transition probability

$$W\left(\boldsymbol{x}, \boldsymbol{x}'\right) = A\left(\boldsymbol{x}, \boldsymbol{x}'\right) P\left(\boldsymbol{x}, \boldsymbol{x}'\right),$$

where $P\left(\boldsymbol{x}, \boldsymbol{x}'\right)$ is the probability to propose a move from the configuration $\boldsymbol{x}'$ to $\boldsymbol{x}$. For a symmetric proposal function, the acceptance probability is:

$$A\left(\boldsymbol{x}, \boldsymbol{x}'\right) = \min\left\{1, \exp\left[-(V(\boldsymbol{x}) - V(\boldsymbol{x}'))/T\right]\right\}. \tag{2.10}$$

We adopt two proposal functions. The first is the box distribution:

$$P\left(\boldsymbol{x}, \boldsymbol{x}'\right) = \frac{1}{2\sigma}\Theta(\sigma - |x_i - x'|)\ , \tag{2.11}$$

where $\Theta(x)$ is the Heaviside step function; the second is a Lorentzian distribution:

$$P\left(\boldsymbol{x}, \boldsymbol{x}'\right) = \frac{1}{\pi}\frac{\sigma}{(x_i - x_i')^2 + \sigma^2}. \tag{2.12}$$

The index $i$ labels the particle being (tentatively) displaced from $x_i'$ to $x_i$, and $\boldsymbol{x} = (x_1', \ldots, x_i, \ldots, x_N')$. In both cases the parameter $\sigma$ controls the range of the

proposed displacements.

However, the two distributions determine qualitatively different dynamics, the first one characterized by short-range moves with maximum range $\sigma$, the second one by long-range jumps due to the fat tail of the Lorentzian distribution. Following Ref. [59], we vary the range parameter during the annealing process according to the square root law

$$\sigma(T) = \sigma_{\rm ini}\sqrt{T/T_{\rm ini}},$$

where $\sigma_{\rm ini}$ is the initial range parameter. This schedule was found to generate reasonably constant acceptance rates close to the optimal value [59]. For the box updates, we adopt the initial range parameter $\sigma_{\rm ini} = 2$; for the Lorentzian updates, we use $\sigma_{\rm ini} = 2.9$ for all models considered in this chapter, apart for the two-particle model (see Section 2.3), for which we use $\sigma_{\rm ini} = 1.5$.

Analogously to the case of the DMC simulations, we perform the classical annealing using a large ensemble of random walkers. Before starting the annealing process, we let the population evolve according to the Metropolis algorithm at the constant temperature $T_{\rm ini} = 1$, so that the walkers population equilibrates at the Boltzmann thermal distribution. As a measure of the CA efficiency we consider the residual energy $\varepsilon_{\rm res}$, defined (see Eq. 2.7) as the average potential energy of the final random-walker population, minus $V_{\rm min}$. $\tau_f$ is here the number of Metropolis steps in the annealing process. Clearly, in the classical annealing case one could perform serial single-walker simulations. Both in DMC quantum annealing and in CA we determine the uncertainty on $\varepsilon_{\rm res}$ by repeating a few (typically $5$) simulations starting from different initial random-walker distributions and computing the standard deviation of the (small) population. The resulting error bar is typically smaller than the symbol sizes.

One of the key issues we address is whether SQA, which exploits quantum fluctuations to escape local minima, is more or less efficient than CA, which instead exploits thermal fluctuations. In particular, we analyze how rapidly $\varepsilon_{\rm res}$ decreases with the annealing time $\tau_f$. One should notice that the annealing times of the quantum and the classical annealing processes cannot be directly compared. Indeed, while in the former case $\tau_f$ is an imaginary-time, in the latter case it is just an integer counting the number of Monte Carlo sweeps during the annealing schedule. Furthermore, depending on the details of the

implementation (e.g., serial versus parallel simulations of the random walkers in DMC) the computational times can be different. However, in general, $\varepsilon_{\rm res}$ decays asymptotically as a power-law of $\tau_f$. Therefore, as in previous works [59, 23], we will compare the powers characterizing the asymptotic scalings of $\varepsilon_{\rm res}$ in the various annealing algorithms, thus obtaining a measure of their efficiency which is independent of the implementation details and of the scale chosen to measure $\tau_f$.

The potential sources of systematic errors in the DMC algorithm originate from the finite time-step $\Delta\tau$ and the finite walkers population $N_{\rm w}$. For all the models considered in this chapter, we carefully analyzed these effects, and we report data obtained with small enough values of $\Delta\tau$ (and symmetrized branching, see Eq. 1.28) and large enough values of $N_{\rm w} \in [10000, 20000]$ to be close to the asymptotic exact regime.

## 2.3 SQA of double-well potentials

We start by analyzing the performance of the DMC algorithm as a heuristic optimization method in the context of double-well potentials. We consider the two models introduced in Ref. [23]. The first is a symmetric double well:

$$V_{\rm sym} = V_0 \left(x^2 - a^2\right)^2 / a^4 + \delta x, \tag{2.13}$$

where $V_0 = 1$, $a = 1$ and $\delta = 0.1$ (see Fig. 2.1, top panel). Since $\delta a \ll V_0$, the difference in the two minima is $\Delta_V \simeq 2\delta a$, with the absolute minimum located at $x_{\rm min} \simeq -a$. The two wells have essentially the same widths.

We adopt the annealing protocol described in the Section 2.2, where the parameter $D$ is linearly reduced to zero (in a step-wise manner). We perform both DMC simulations with importance sampling using a Boltzmann-like trial wave function

$$\psi_T(x) = \exp\left(-\tilde{\beta} V(x)\right), \tag{2.14}$$

where $\tilde{\beta} = 0.8$ is a fictitious inverse temperature and also without importance sampling, setting $\psi_T(x) = 1$.

The results for the residual energy $\varepsilon_{\rm res}$ as a function of the annealing time $\tau_f$

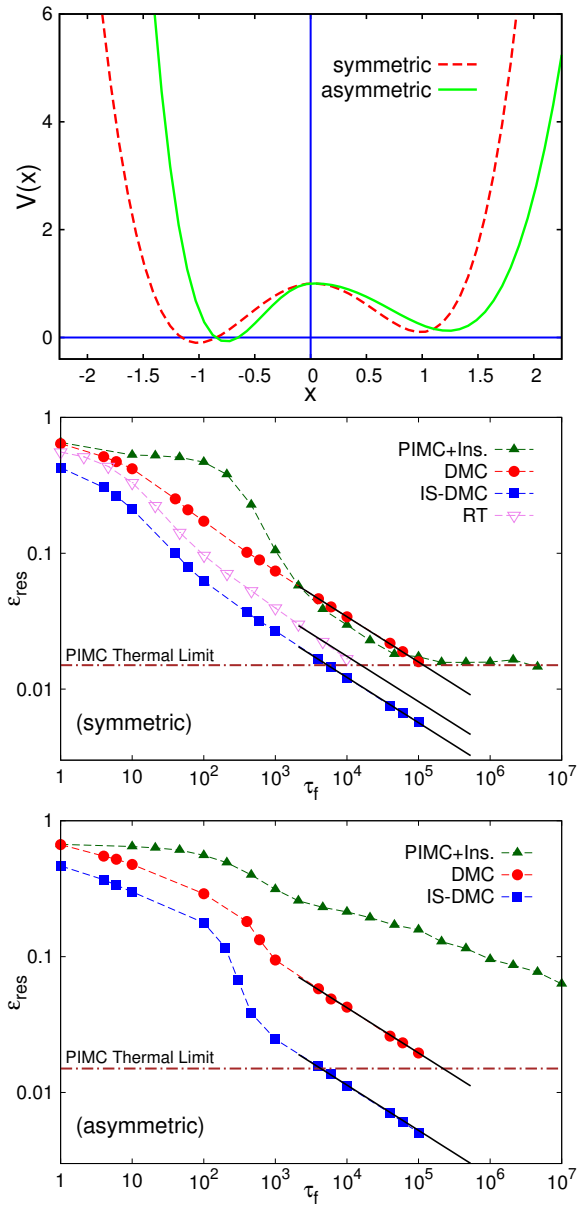



Figure 2.1: SQA of symmetric and asymmetric double-well potentials. Top panel: potential energy $V(x)$ versus particle coordinate $x$. Central panel: residual energy $\varepsilon_{\rm res}$ versus annealing time $\tau_f$ for the symmetric double well, obtained using the DMC algorithm, the DMC algorithm with importance sampling (IS-DMC), the PIMC algorithm with instanton move (PIMC+ins., from Ref. [59]), and via integration of the real-time Schrödinger equation (RT, from Ref. [23]). Bottom panel: as in the central panel (except for the RT data), for the asymmetric double-well. The horizontal brown dot-dashed lines indicate the lowest $\varepsilon_{\rm res}$ reachable in the PIMC simulation due to the finite temperature. The thick black solid segments indicate fits to the DMC asymptotic data according to the power-law scaling $\varepsilon_{\rm res} \sim \tau_f^{-1/3}$. The units of $\tau_f$ in PIMC and DMC simulations are different (see text). The thin dashed curves are guides to the eye. Here and in the other graphs the error bars are smaller than the symbol size.

are displayed in Fig. 2.1 (we recall that the number of DMC steps, and so the simulation run-time, is simply proportional to $\tau_f$). We observe that the use of importance sampling introduces a quantitative improvement, providing somewhat lower residual energies; however, the two approaches (with and without importance sampling) display the same asymptotic scaling

$$\varepsilon_{\rm res} \propto \tau_f^{-1/3}, \tag{2.15}$$

meaning that the efficiency of the optimization process is not affected in a qualitative manner. This power-law dependence with the characteristic power $-1/3$ was first found in Ref. [23] by exactly solving the imaginary-time Schrödinger equation for a single harmonic well; it appears to be a generic feature of DMC quantum annealing in the asymptotic regime.

Fig. 2.1 also displays the residual energies obtained in Ref. [59] using the PIMC method, employing a linear annealing protocol as in our DMC simulations. For the PIMC data, the annealing time $\tau_f$ represents the number of Monte Carlo sweeps. One observes that for large $\tau_f$ the PIMC data decay similarly to the DMC results; however, in the $\tau_f \to \infty$ limit they saturate at the energy corresponding to the temperature at which the PIMC simulations were performed. In principle, this finite temperature could be reduced arbitrarily close to zero (but at the expense of higher computational cost); therefore, we conclude that the DMC and the PIMC quantum annealing methods perform comparably well in this symmetric double-well problem.

A more interesting test problem is obtained by introducing an asymmetry in the widths of the two wells; specifically, we consider the asymmetric potential:

$$V_{\rm asym}(x) = \begin{cases} V_0 \left(x^2 - a_+^2\right)^2 / a_+^4 + \delta x & \text{if } x \geqslant 0 \\ V_0 \left(x^2 - a_-^2\right)^2 / a_-^4 + \delta x & \text{if } x < 0, \end{cases} \tag{2.16}$$

where the new constants are $a_- = 0.75$ and $a_+ = 1.25$. In this case, the well corresponding to the false minimum is wider than the well corresponding to the absolute minimum (which is located at $x_{\rm min} \simeq -a_- - \delta a_-^2/(8V_0)$, to linear order in $\delta$).

In the early state of the annealing process, where zero-point motion is large, the wave function weight is mostly located close to the false minimum, implying a larger probability to find the quantum particle in the wider well. Only in a

latter stage of the annealing process, where the annealing parameter $D$ is small (corresponding to a large particle mass), the wave function starts concentrating in the deepest (narrower) well. This is reflected in the dependence of $\varepsilon_{\rm res}$ versus $\tau_f$ obtained with the DMC algorithm, which displays different behaviors in the two stages, with the crossover taking place at $\tau_f \approx 10^3$. In the asymptotic regime $\tau_f \to \infty$, the residual energy decays again with the power-law $\varepsilon_{\rm res} \propto \tau_f^{-1/3}$, just like in the symmetric wells case, both with and without importance sampling.

It is interesting to observe that the PIMC data display a qualitatively different behavior: the residual energy decreases much slower than in the symmetric wells case, even well before the thermal limit is reached. This occurs in spite of the fact that the PIMC simulations were performed including the so-called instanton Monte Carlo update, which is designed to displace a significant portion of the path-integral across the energy barrier which separates the two wells, exploiting the knowledge of the potential landscape details. In principle, this kind of update should boost the performance of the PIMC simulations, strongly favoring equilibration. However, it is clear that in the framework of the annealing process this is not sufficient; in fact, the transfer to the deepest (but narrower) well is particularly slow. This dramatic change of efficiency going from symmetric to asymmetric double-wells appears to be a deficiency of the path-integral scheme, rather than a genuine feature of a perfect quantum annealer. Similarly, very recents results on random Ising chains also display pathological samplings of the PIMC [60]. There, different behaviors of the PIMC were obtained depending on the type of proposal move that was used to update the spin configurations.

Whereas, when quantum annealing is simulated via the DMC algorithm, the asymmetry of the two wells has essentially no effect on the optimization efficiency. In the DMC scheme, the random-walker distribution in the different wells easily equilibrates thanks to the branching process, making it perfectly suited to simulate the optimization of potential energy landscapes in situations where quantum tunneling across energy barriers plays a fundamental computational role.

Such a case was indeed recently implemented by researchers working with a D-Wave Two chip via an appropriate choice of the couplings between the quantum spins in two unit cells of the Chimera graph [18]. In this experiment, the effective double-well potential varies in time, with the false minimum appearing first, and the absolute minimum appearing at a latter time. While classical trajectories would remain trapped in the false minimum, quantum tunneling allows the

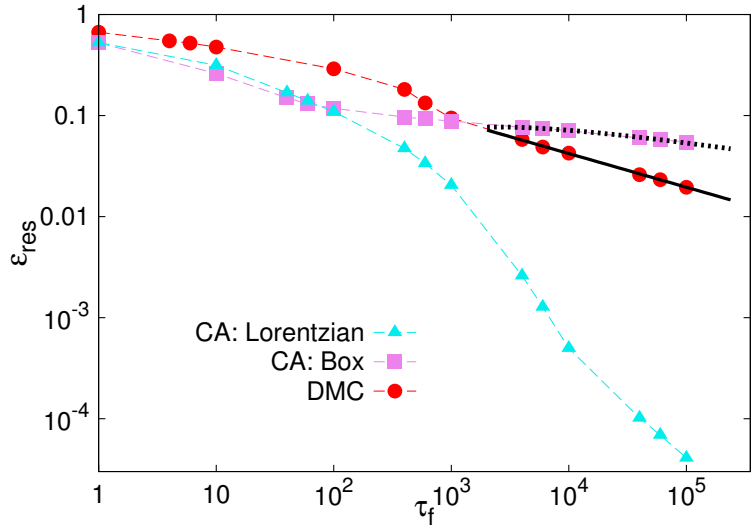



Figure 2.2: Classical annealing (CA) and SQA of the asymmetric double-well potential. Residual energy $\varepsilon_{\rm res}$ as a function of annealing time $\tau_f$ obtained using the Metropolis algorithm with Lorentzian proposed moves (cyan triangles) and box-type moves (violet squares), and using the DMC algorithm (red circles). The thick black solid segments is a fit to the DMC asymptotic data with the power-law $\varepsilon_{\rm res} \sim \tau_f^{-1/3}$. The dotted black curve is a fit to the CA box-moves data with the asymptotic law $\varepsilon_{\rm res}^{\rm HF}(\tau_f)$ (see text). The units of $\tau_f$ in CA and DMC simulations are different, see text.

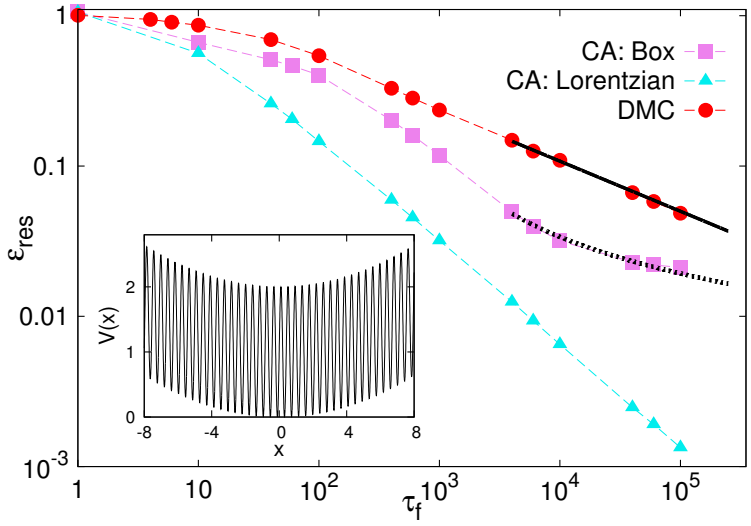



Figure 2.3: CA and SQA of the washboard potential (shown in the inset). The symbols are defined as in Fig. 2.2. The thick black solid segment is a fit to the DMC asymptotic data with the power-law $\varepsilon_{\rm res} \sim \tau_f^{-1/3}$. The dotted black curve is a fit to the CA data with the logarithmic law $\varepsilon_{\rm res} = c_1 \log^{-1}(c_2 \tau_f)$, where the fitting parameters are $c_1$ and $c_2$.

system to reach the absolute minimum. This setup is slightly different from the double-well model we address here: in our case the potential does not vary with time, but the system is initially attracted towards the well corresponding to the wrong minimum due to its larger width. It is possible that in the time-dependent potential case considered in Ref. [18] the PIMC and DMC algorithms would perform equally well.

The DMC simulations have a more direct connection with the quantum annealing as understood in the framework of adiabatic quantum computing, which assumes a real-time Schrödinger dynamics with a time-dependent Hamiltonian [2]. Indeed, inspired by the conjecture formulated in Ref. [23], Morita and Nishimori [32] showed that, from the imaginary-time version of the adiabatic theorem, it follows that the residual energy obtained from the imaginary-time Schrödinger dynamics has, for $\tau_f \to \infty$, the same asymptotic scaling form as the one obtained from the real-time Schrödinger dynamics. Since the DMC algorithm stochastically simulates the imaginary-time Schrödinger equation - with the difference that the annealing parameter is decreased in a step-wise manner (see Section 2.2) - it represents a more legitimate benchmark for the performance of the quantum annealing process as it would be implemented on an ideal (perfectly isolated)

quantum device operating at zero temperature.

The strong relation between imaginary-time and real-time dynamics has been, in fact, confirmed in nontrivial simulations of both clean and disordered Ising models driven to critical points [61, 56]. However, the conditions for the applicability of the adiabatic perturbation theory of Ref. [32], namely that the adiabatic ground-state contribution is the dominant one at all times, might be violated in certain, possibly relevant, cases. In fact, in clean and disordered Ising chains driven across the quantum critical point, the Kibble-Zurek scaling of the residual energy was shown to demonstrate completely different results for the real and imaginary time dynamics [24].

As an illustrative example, we show in Fig. 2.1 (central panel) the residual energies for the symmetric double-well case as obtained by performing quantum annealing via real-time Schrödinger dynamics. These data were obtained in Ref. [23] via exact (deterministic) numerical integration of the time-dependent Schrödinger equation. As predicted by the theory of Ref. [32], the real-time residual energies share the same asymptotic scaling behavior as the DMC imaginary-time data, the uniform shift being affected by the use of importance sampling and of the mixed energy estimator in the DMC simulations (see Chapter 1).

Therefore, one understands the importance of comparing the performance of the DMC algorithm with the one of CA-based methods in this and in more challenging optimization problems; this will help us understand in which situations quantum annealing has at least the potential to outperform classical algorithms. Also, it is useful to establish if and when the stochastic simulation of the imaginary-time Schrödinger dynamics on a classical computer via the DMC algorithm becomes unfeasible (e.g., due to the exponential growth on the required random-walker number [29]), because only in such a case one is actually forced to resort to quantum devices.

Fig. (2.2) shows the DMC and the CA data for the asymmetric double-well potential. CA is performed both with the short-range (box) and with the long-range (Lorentzian) proposed updates, as described in the Section 2.2. In the former case, the stochastic dynamics is well described by the Fokker-Planck equation [62], which - as shown by Huse and Fisher [63] - in a double-well problem leads to the following asymptotic decay of the residual energy (see

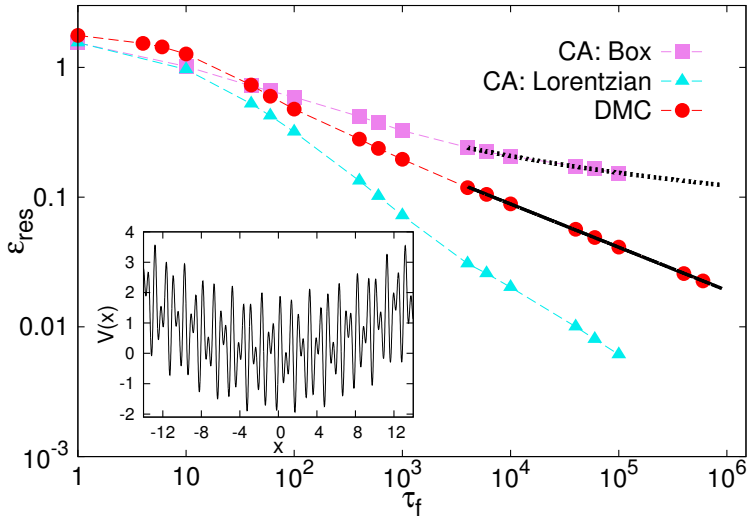



Figure 2.4: CA and SQA of the quasi-disordered potential (shown in the inset). The symbols and the curves are defined as in Fig. 2.2.

dotted curve in Fig. 2.2):

$$\varepsilon_{\text{res}}^{\text{HF}}(\tau_f) = c_1 \tau_f^{-\Delta_V/B} \left[\ln\left(c_2 \tau_f\right)\right]^{2\Delta_V/B}, \tag{2.17}$$

where $c_1$ and $c_2$ are fitting parameters, $\Delta_V = \delta(a_+ + a_-)$ is the splitting between the two minima and $B = V_0 - \Delta_V - V(x_{\min})$ is the energy barrier separating them. As discussed above, DMC quantum annealing displays the power-law asymptotic decay $\varepsilon_{\text{res}} \sim \tau_f^{-1/3}$, clearly outperforming CA with short-range updates. However, long-range (Lorentzian) updates strongly increase the efficiency of CA, making it more performant than the DMC algorithm.

## 2.4 SQA of multiwell potentials

It is now natural to wonder how CA and DMC quantum annealing perform in more challenging optimization problems. To address this question, we consider the multi-well problem defined by the following "washboard" potential (shown in the inset of Fig. 2.3):

$$V(x) = a_1 x^2 + a_2 \sin\left(a_0 x\right) + a_2, \tag{2.18}$$

where $a_0 = 15$, $a_1 = 0.01$ and $a_2 = 1$. This model was first studied in Refs. [64, 65], where it was suggested that CA should suffer from a pathological slowdown due to the presence of many well-separated minima. In a wide range of annealing times [23], the residual energy decay should be at best logarithmic: $\varepsilon_{\rm res}(\tau_f) \propto (\ln(\tau_f))^{-1}$.

The CA data with short-range (box) updates (shown in Fig. 2.3) are indeed consistent with this logarithmic upper bound, showing that, in general, with the CA dynamics it becomes problematic to equilibrate to the minimum energy configuration when many close solutions compete. Instead, DMC quantum annealing maintains its efficiency, displaying again the asymptotic decay $\varepsilon_{\rm res} \sim \tau_f^{-1/3}$ (we only display data obtained with the pure DMC algorithm, since importance sampling was again found not to affect the asymptotic efficiency). This $-1/3$ power appears to be the footprint of quantum annealing. Below, we will demonstrate the same behavior in even more intricate problems. This suggests that the identification of a $-1/3$ power-law decay in a quantum annealer could be interpreted as an evidence of quantum effects playing a fundamental computational role.

We also notice that, as in the double-well case, long-range updates boost the efficiency of CA at the point of outperforming the DMC algorithm. One might suspect that long-range updates do not provide the same boost in more complex problems with more variables. We will show latter on that this is indeed the case.

DMC quantum annealing has so far displayed a surprisingly stable efficiency. Clearly, the models we addressed previously do not contain one of the most relevant ingredients which make realistic optimizations problems difficult, namely a disordered distribution of the competing solutions. Disorder is indeed expected to hamper quantum annealing due to the Anderson localization phenomenon. This consists in the spatial localization of the Hamiltonian eigenstates, causing the absence of particle diffusion. While in three and more dimensions Anderson localization takes place only for sufficiently strong disorder [66], in one and two dimensions any amount of uncorrelated disorder induces localization.

A minimal model which contains (quasi) disorder is the following double-sine potential:

$$V(x) = K_0 x^2 + A\left[\sin(b_1 \pi x) + \sin(b_2 \pi x)\right], \tag{2.19}$$

where $K_0 = 0.01$ , $A = 1$, $b_1 = 2$, and $b_2 = 1 + 5^{1/2}$. Due to the irrational ratio

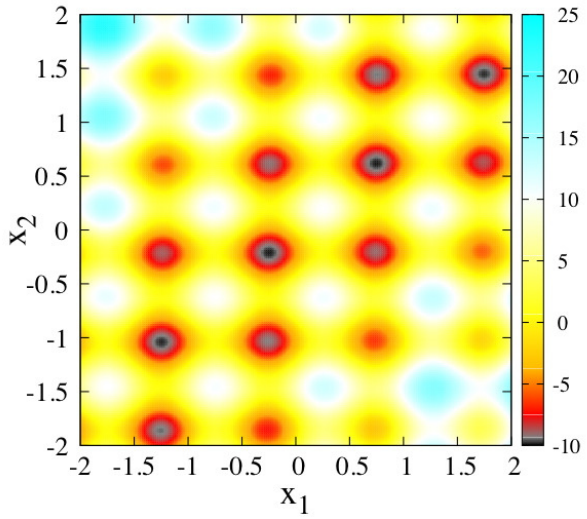



Figure 2.5: Potential energy $V(x_1, x_2)$ of the two particle model 2.20 in the case $A = 5$. The colorscale represents the potential intensity, $x_1$ and $x_2$ are the coordinates of the two particles. The absolute minimum is at $V(-0.25, -0.208) \cong -9.99759$.

of the wave-lengths of the two sinusoidal functions, this potential is aperiodic. However, it is deterministic and, therefore, not truly random. This kind of pseudo-randomness is conventionally referred to as quasi-disorder. In a tight binding scheme (which would be rigorously justified if one sinusoidal potential was much more intense than the other one) this incommensurate double-sine model could be approximated by the so-called Aubry-André Hamiltonian, provided one neglects the weak harmonic confinement. Differently from one-dimensional models with uncorrelated disorder, in the Aubry-André Hamiltonian, Anderson localization takes place at a finite disorder strength or, equivalently, when the particle mass exceeds a critical value [67]. We expect the weak harmonic term to play a minor role, at least well beyond the critical point, meaning that the Hamiltonian eigenstates would still be strongly localized in this regime. This strong spatial localization inhibits diffusion, preventing the particle from exploring the complete configuration space, possibly causing localization in local minima. It is therefore interesting to analyze whether in the last part of the quantum annealing process, where the annealing parameter is small and, correspondingly, the particle mass is large, Anderson localization deteriorates the efficiency of DMC quantum annealing.

In fact, the data displayed in Figure (2.4) show that the efficiency of DMC quantum annealing is not affected by the presence of quasi-disorder, demonstrating again the extreme stability of its performance (notice that here and in the following we only consider data obtained without importance sampling). As in the case of ordered minima (i.e., the washboard potential) CA with short-range updates displays a pathological slow-down of the annealing process at large annealing times, leading to a logarithmic decay of the residual energy. However, CA with long-range updates is still the most efficient optimization algorithm.

So far, the DMC algorithm has proven to be an effective and stable optimization method. However, the fact that it is outperformed by CA if appropriate (long-range) Monte Carlo updates are included is quite discouraging for the prospect of developing quantum enhanced optimization algorithms. As anticipated above, one might wonder whether it is always possible to obtain such a boost in the efficiency of CA via long-range updates. To address this question, we consider a model with two particles which move in an external potential and interact with each other. The potential energy is defined as:

$$\begin{aligned} V(x_1, x_2) =& K_0(x_1^2 + x_2^2)/2 + K_{\rm rel}(x_1 - x_2)^2/2 + \\ & A\left[\sin(b_1 \pi x_1) + \sin(b_2 \pi x_2)\right], \end{aligned} \tag{2.20}$$

where we set $K_0 = 0.01$, $K_{\rm rel} = 2$, $b_1 = 2$, and $b_2 = 1 + 2^{1/2}$. For the intensity parameter $A$, we will employ the three values $A = 1,\ 3,\ 5$.

The first term in Eq. (2.20) confines both particles in a global harmonic trap. The second term introduces an attractive harmonic interaction between the two particles. The last term is inspired by the incommensurate double-sine potential analyzed previously. However, here the two particles experience the two sinusoidal fields separately. The first sinusoidal field acts only on one particle, while the second one acts on the other particle. While the attractive interaction tends to localize the two particles in the same location, the sinusoidal terms have their minima in different points. The competition among the terms in Eq. (2.20) constitutes the minimal element of frustration.

If one interprets Equation (2.20) as the external potential of one particle moving in a two-dimensional system, one obtains the intricate landscape shown in Figure (2.5). One notices that there are several closely competing solutions, well separated by high energy barriers. By construction, there is no periodicity and the bottoms of the different valleys are at close but different levels. The height of

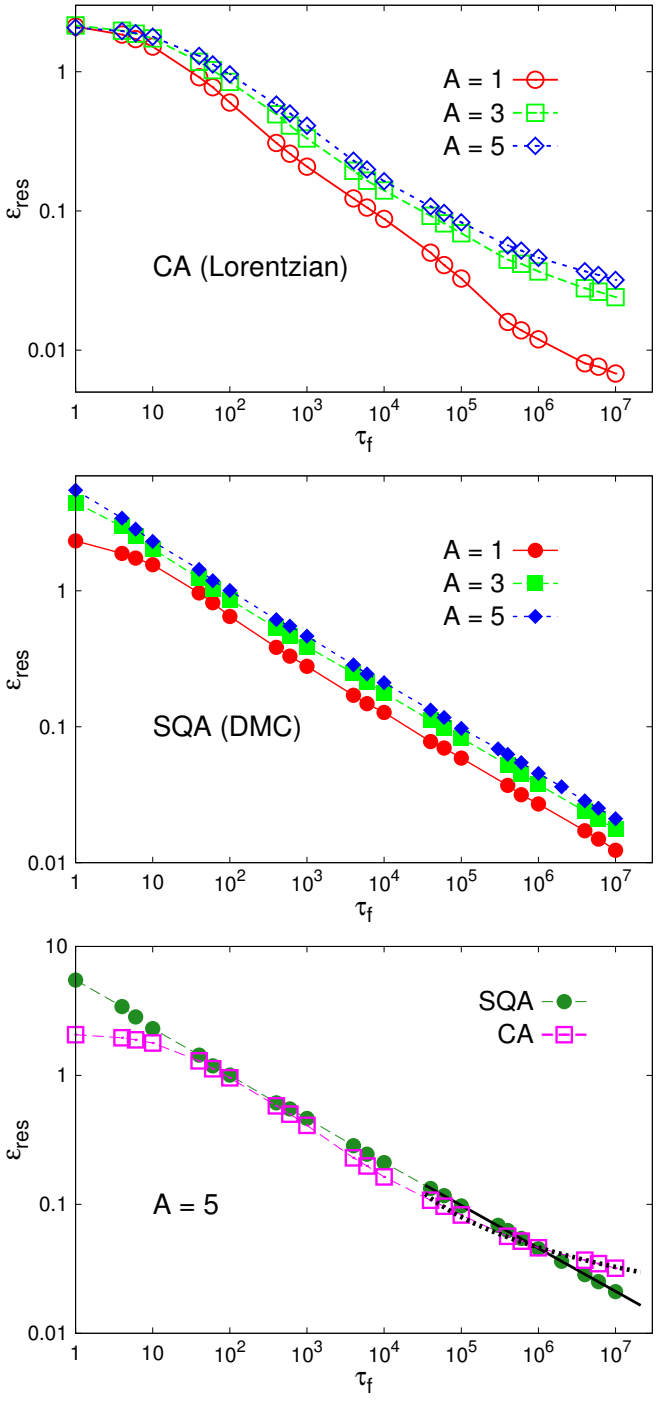



Figure 2.6: CA and SQA of the two particle model. Top panel: residual energy $\varepsilon_{\rm res}$ versus annealing time $\tau_f$ obtained using CA with Lorentzian updates, for different intensities $A$ of the sinusoidal part of the potential. Central panel: data analogous to those in the top panel, obtained using the DMC algorithm. Bottom panel: comparison between CA and SQA at $A = 5$. The thick black solid segment indicates the asymptotic scaling of the DMC data $\varepsilon_{\rm res} \sim \tau_f^{-1/3}$, while the black dotted curve the one of the CA data $\varepsilon_{\rm res} \sim \log^{-1}(c\tau_f)$.

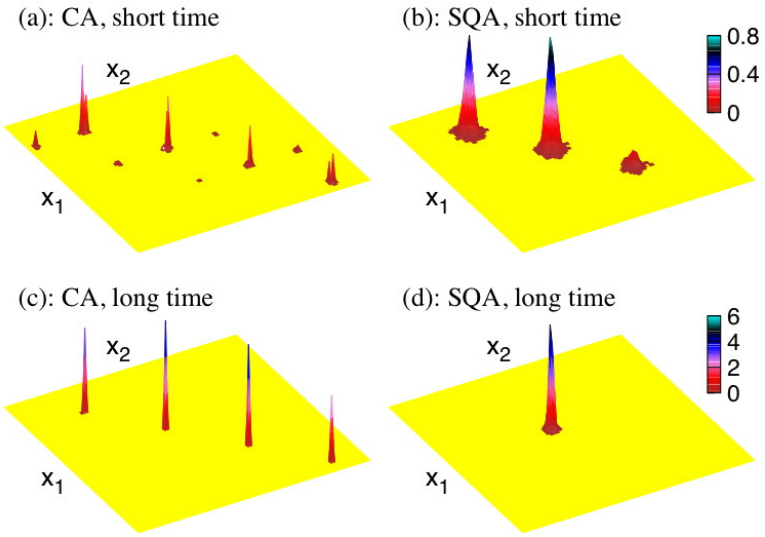



Figure 2.7: Final random-walker distribution in CA [panels (a) and (c)] and in SQA [panels (b) and (d)]. Panels (a) and (c) correspond to short annealing times $\tau_f = 10^3$, while panels (b) and (d) correspond to long annealing times $\tau_f = 10^5$. The colorscale corresponds to the random-walker density.

the barriers can be varied by changing the intensity parameter $A$, allowing us to tune the difficulty of the optimization problem.

The results of the optimization of this two-particle model are displayed in Fig. 2.6, both for CA (top panel) and SQA performed via the DMC algorithm (central panel). The CA simulations are performed with Lorentzian updates. We apply three kinds of updates: the first displaces only the first particle, the second the other particle, the third applies the same displacement to both particles. It is evident that when $A$ increases the asymptotic slope of the CA data diminishes, indicating a loss of efficiency of the optimization process.

Instead, the DMC data display the same asymptotic decay $\varepsilon_{\rm res} \sim \tau_f^{-1/3}$ for all values of $A$. As anticipated before, the independence of the asymptotic power-law decay of the residual energy appears to be the hallmark of quantum annealing. The comparison shown in the bottom panel of Fig. 2.6 (for $A = 5$) demonstrates that in the most challenging optimization problem DMC quantum annealing outperforms CA, suggesting that quantum annealing has indeed the potential to outperform classical algorithms in hard optimization problems.

It is particularly instructive to analyze how and why SQA outperforms CA. In

Fig. 2.7 we show the probability to reach a certain two-particle configuration $(x_1, x_2)$, both with CA and SQA, after a short and after a long annealing time. In the mixed estimator scheme (see section 1.4), this probability corresponds to the spatial distribution of the random-walker population at the end of the annealing process. We observe that, in CA, several false minima have a large probability to be selected. While the random-walker distribution in each individual well rapidly shrinks, even after a long annealing time quite a few competing solutions are still likely to be chosen. This indicates that CA has high chance to remain trapped close to false minima. The behavior of DMC quantum annealing is, in a sense, the opposite: already after a short annealing time only three wells are populated by random walkers. However, the distribution in each well is quite broad, indicating that SQA is slower in sinking to the very bottom of the well. After a long annealing time, only the well corresponding to the absolute minimum is populated, but the residual energy is still nonzero since the random walkers need further time to sink to the very bottom of the well, thus selecting the optimal configuration.

At this point, it is worth mentioning that other classical algorithms which potentially outperform CA for specific optimization problems, in particular for small systems, do exist (e.g., genetic algorithms). However, CA has proven to be one of the most powerful and versatile optimization methods [53], succeeding even in challenging continuous-variables problems with multiple minima (e.g., the optimization of the structure of Lennard-Jones clusters [68]) where gradient-based algorithms like the conjugate gradient method remain trapped in local minima. Therefore, CA represents a fair term of comparison for SQA.

## 2.5 Conclusions

We have analyzed the efficiency of SQA in finding the absolute minimum of different model potentials in continuous space, including symmetric and asymmetric double wells, and also more intricate models with many closely competing minima with both ordered and disordered spatial distribution of the wells. Contrarily to the finite-temperature path-integral Monte Carlo techniques adopted in several previous studies, the projective method employed in this work, namely the DMC algorithm, exhibits a stable performance which is not affected by details of the potential energy landscape like the asymmetry of the competing wells.

While the DMC quantum annealing is outperformed by CA in simple one-variable model potentials if one employs ad-hoc Monte Carlo updates which exploit the specific features of the potential landscapes, it is easy to construct more challenging optimization problems with more variables where such tricks in CA become ineffective and SQA turns out to be the most effective optimization method. This result is strongly encouraging for the prospect of developing quantum devices which exploits quantum fluctuations to enhance the efficiency of optimization methods. The stable performance of DMC quantum annealing, characterized in quite general continuous-space models by an asymptotic power-law decay of the residual energy, appears to be a hallmark of an ideal quantum annealer.

The DMC algorithm is designed to simulate the ground states of isolated quantum systems; it does not take into account thermal fluctuations nor the coupling to the environment. Including these effects in a projective QMC algorithm would provide us with an extremely useful tool to investigate the potential of realistic devices designed to perform adiabatic quantum computations. This is clearly an interesting direction for future research. Also, it would be important to identify the cases where the DMC simulations become infeasible due to, e.g., an exponential scaling of the required random-walker number [28]. Computational problems of this kind did not occur in any of the double-well and multi-well problems addressed in this work, probably due to the small particle number. However, in the context of standard (i.e., without annealing) quantum simulations of parahydrogen clusters, results suggesting an exponential scaling of the random-walker population with the system size have been reported [29, 69]. In a subsequent chapter we will discuss on a possibility to attenuate the scaling of the walkers population.

# 3 Understanding Quantum Tunneling using DMC Simulations

In simple ferromagnetic quantum Ising models characterized by an effective double-well energy landscape the characteristic tunneling time of path-integral Monte Carlo (PIMC) simulations has been shown to scale as the incoherent quantum-tunneling time, i.e., as $1/\Delta^2$, where $\Delta$ is the tunneling gap. Since incoherent quantum tunneling is employed by quantum annealers (QAs) to solve optimization problems, this result suggests there is no quantum advantage in using QAs with respect to quantum Monte Carlo (QMC) simulations. A counterexample is the recently introduced shamrock model, where topological obstructions cause an exponential slowdown of the PIMC tunneling dynamics with respect to incoherent quantum tunneling, leaving the door open for potential quantum speedup, even for stoquastic models. In this work, we investigate the tunneling time of projective QMC simulations based on the diffusion Monte Carlo (DMC) algorithm without guiding functions, showing that it scales as $1/\Delta$, i.e., even more favorably than the incoherent quantum-tunneling time, both in a simple ferromagnetic system and in the more challenging shamrock model. However a careful comparison between the DMC ground-state energies and the exact solution available for the transverse-field Ising chain points at an exponential scaling of the computational cost required to keep a fixed relative error as the system size increases.

## 3.1 Introduction

Adiabatic quantum computers, such as the quantum annealers (QAs) built using superconducting flux qubits [43, 44, 20] — or, potentially, with Rydberg atoms trapped in arrays of optical tweezers [70] — can be used to solve complex

combinatorial optimization problems. They implement a quantum annealing process [71, 4, 1], in which quantum mechanical tunneling through tall barriers is used to escape local minima, and quantum fluctuations are gradually removed by reducing to zero the transverse field of a quantum Ising model. While in problems with energy landscapes characterized by tall but thin barriers quantum tunneling definitely makes QAs more efficient than classical optimization methods such as simulated classical annealing [72, 73], certain examples are known where the opposite seems to be true [16, 17].

In a recent study [21], which aimed at shedding light on the relation between the dynamics of SQA and the one of QAs, it was found that the characteristic timescale of tunneling events occurring during path-integral Monte Carlo (PIMC) simulations increases with the system size as $1/\Delta^2$, where $\Delta$ is the energy gap between the ground state and the first exited state (see also Ref. [74]). This $1/\Delta^2$ scaling was found to hold in ferromagnetic quantum Ising models [21], which are characterized by an effective double-well energy landscape (the two symmetric minima are the ground states with opposite magnetizations), and also in one-dimensional and two-dimensional continuous-space double-well models relevant for quantum chemistry applications [22]. Remarkably, this is the same scaling of the time of *incoherent* quantum tunneling in symmetric double-well models [75].

Furthermore, according to the adiabatic theorem, also the annealing time required in a *coherent* adiabatic quantum computation to avoid diabatic transitions [76] to the first exited state must increase as the squared inverse of the smallest instantaneous gap. The similar scaling of the respective tunneling times — which was explained using an instanton theory [77] — suggests that PIMC simulations can efficiently simulate incoherent quantum tunneling. This latter phenomenon is supposed to be one of the empowering resources of QAs (although quantum superposition and entanglement might also be crucial ingredients). It allows them to explore different localized states far away in Hamming distance, like those typically emerging in the glassy phases characteristic of Ising glass models at small transverse field. On the one hand, this finding suggests that SQA can be used to predict the performance of QAs, providing us with a useful tool to guide the engineering of these devices. On the other hand, it might also imply that SQA has the same potential efficiency in solving complex optimization problems as QAs have, meaning that quantum speedup is unlikely to be achieved (apart for a prefactor), at least as long as the Hamiltonian under consideration is

stoquastic [1] [79, 80].

Latter on, Ref. [3] introduced the so-called "shamrock model", showing that due to topological obstructions [81] — which originate from frustrated couplings in this model — the PIMC tunneling time increases with the system size exponentially faster than the incoherent quantum tunneling time, giving one hope that QAs can outperform SQA, and so maybe also other heuristic optimization methods running on classical computers.

The performance of SQA in solving optimization problems crucially depends on the specific type of QMC algorithm as it was shown in the precedent chapter. The above mentioned study [3] of the QMC tunneling time for the shamrock model considered only finite-temperature PIMC algorithms. This naturally raises the following questions: can projective QMC methods efficiently simulate quantum mechanical tunneling? Would they be immune from the (exponential) pathological slowdown which affects the PIMC simulations in the shamrock model?

The main goal of this chapter is to address the above two questions. In order to do so, we implement a projective QMC method for quantum Ising models based on the DMC algorithm in which the stochastic dynamics is defined by the Trotter-decomposed imaginary-time evolution operator. Then, following Refs. [21, 22], we introduce a protocol to measure the characteristic time of tunneling events occurring in DMC simulations, and we analyze the scaling with the system size of the so-defined tunneling time, both in the ferromagnetic quantum Ising chain and in the shamrock model. Furthermore, in order to understand if the DMC algorithm allows one to *efficiently* simulate the behavior of QAs on classical computers, we analyze the computational cost of DMC ground-state simulations. In particular, we study the convergence of the systematic biases in calculations of the ground-state energy, using as a testbed the quantum Ising chain. It should be noted that we focus on the *simple* DMC algorithm, i.e., we do not consider the use of importance sampling techniques [82]. These are based on suitably constructed guiding wave functions which represents the worst case scenario in which no suitable guiding wave function that approximates the ground state can be defined. As we will show in the next chapter, importance sampling can indeed boost the algorithm efficiency and fasten the convergence of the systematic biases.

[1] A local Hamiltonian is called stoquastic [78] if all of its off-diagonal matrix elements defined in the standard basis of $n$ qubits are real and non-positive. This would imply that the corresponding Hamiltonian would not be affected by the sign problem in a Monte Carlo simulations. Examples of such Hamiltonians are the quantum Ising model and the ferromagnetic Heisenberg model.

The rest of the chapter is organized as follows: in Section 3.2, we describe the protocol used to measure the characteristic time of tunneling events occurring during the DMC simulations, and we provide the results for the ferromagnetic quantum Ising chain, making comparisons with exact diagonalization calculations of the gap, showing the $1/\Delta$ scaling of the tunneling time with the system size. In Section 3.3 the system-size scaling of the DMC tunneling time for the shamrock model is analyzed, showing also in this case the $1/\Delta$ scaling. Section 3.4 reports the analysis of the convergence of the systematic bias of the DMC algorithm due to the finite size of the random walker population. Our conclusions and the outlook are reported in Section 3.5.

The potential sources of systematic errors in the DMC algorithm originate from the finite time step $\Delta\tau$ and the finite number of random walkers $N_{\rm w}$. For what concerns the DMC tunneling times, we carefully analyzed these effects, and we report in Sections 3.2 and 3.3 only data obtained with small enough values of $\Delta\tau$ and large enough values of $N_{\rm w}$ to be in the asymptotic exact regime. For what concerns predictions of ground-state energies, a detailed analysis of the systematic bias due to the finite $N_{\rm w}$ is reported in Section 3.4. The systematic error in the ground-state energy due to the finite $\Delta\tau$ is less relevant and can be made smaller than statistical uncertainties with moderate computational effort, so its analysis is not reported in this work.

## 3.2 The Ferromagnetic quantum Ising Chain

In this section, we describe the protocol we use to measure the characteristic time of tunneling events occurring in a DMC simulation, and we present the results for the one-dimensional ferromagnetic transverse-field Ising model already mentioned in the chapter 1. In other to make this chapter more comprehensive, we will rewrite again its Hamiltonian (see Eq. 1.48), which is given as:

$$\hat{H} = -\sum_{i=1}^{N} J\sigma_i^z\sigma_{i+1}^z - \Gamma\sum_{i=1}^{N}\sigma_i^x, \tag{3.1}$$

where the coupling is $J > 0$ (set to $J = 1$) and $\Gamma$ is the intensity of the transverse field. Periodic boundary conditions are considered, i.e., $\sigma_{N+1}^a = \sigma_1^a$ where $a = x, y, z$. At zero temperature this model undergoes a quantum phase transition from a paramagnetic phase at $\Gamma > 1$ to a ferromagnetic phase at $\Gamma < 1$. In the $\Gamma \to 0$ limit quantum fluctuations are suppressed and one has two degenerate

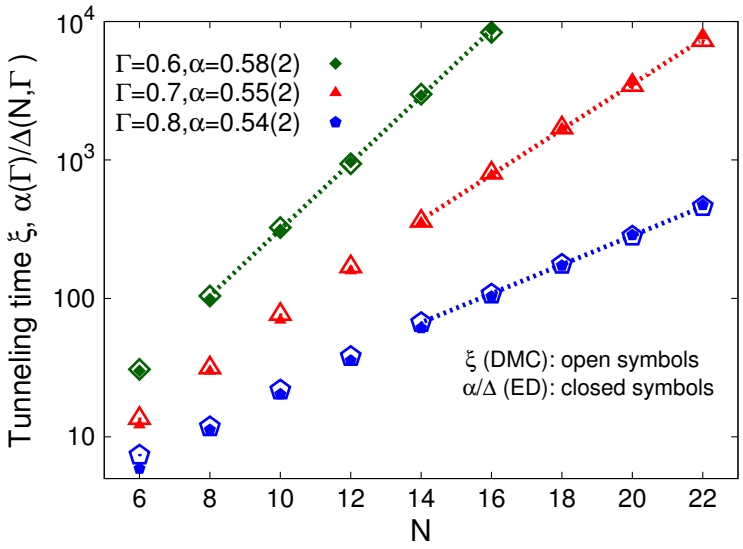



Figure 3.1: DMC tunneling time $\xi$ for the ferromagnetic Ising chain (open symbols) as a function of the number of spins $N$, for different values of the transverse field $\Gamma$ with $J = 1$. The closed symbols represent the inverse gap values $1/\Delta$ obtained with exact diagonalization and rescaled by a parameter $\alpha(\Gamma) = \mathrm{O}(1)$. The thin dashed curves represent exponential fits on the tunneling time $\xi$ in the large-$N$ regime. Here and in the other graphs of the chapter, the error bars are smaller than the symbol size if not visible.

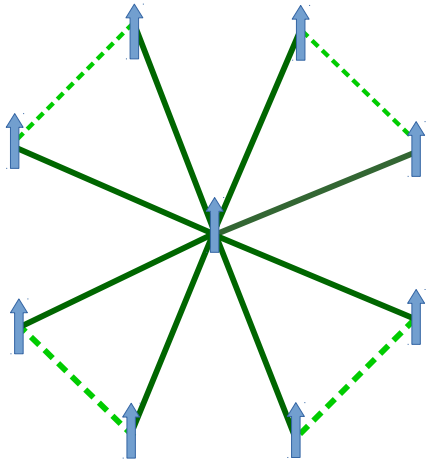


Figure 3.2: The shamrock, a model of $N$ frustrated spins in a transverse field. It is made up of $K = (N-1)/2$ leaves each having three spins. The solid dark-green lines depict ferromagnetic interactions (with interaction strength $J = 6$) between the central spin and all the other $N-1$ spins. The dashed light-green lines instead show the anti-ferromagnetic interactions (with interaction strength $J - \epsilon$) between the outer spins of the same leaf (see Eq. 3.2). The overall effect results in creating $2^K$ tunneling paths between the degenerate classical ground states in the incoherent quantum tunneling regime.

(classical) states with all spins up $|\uparrow\uparrow \ldots \uparrow\rangle$ or all spins down $|\downarrow\downarrow \ldots \downarrow\rangle$. In order to go from one state to the other, the system would have to overcome an energy barrier separating the two minima, with the magnetization playing the role of a one-dimensional reaction coordinate which parametrizes a symmetric double-well profile.

For small $\Gamma > 0$, in the thermodynamic limit there are still two degenerate ground states with opposite magnetizations, but in a finite chain the degeneracy is lifted by an exponentially small (in the system size) energy gap due to the quantum tunneling which couples the two states. This scenario is reminiscent of what happens in a QA towards the end of the annealing process when the transverse field is small and the system explores different well-separated local minima via incoherent quantum tunneling. For this reason, shedding light on how tunneling events take place in QMC simulations — even in the simple double-well scenario — is important to understand if QAs have the potential to outperform classical heuristic optimization algorithms, such as SQA.

We define the DMC quantum *tunneling time* $\xi$ by implementing the following

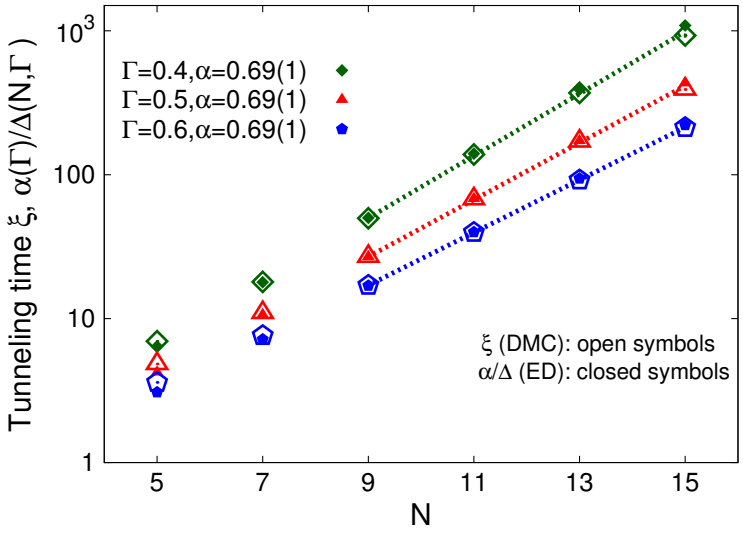



Figure 3.3: DMC tunneling time $\xi$ for the shamrock model DMC (open symbols) as a function of the number of spins $N$, for different values of the transverse field $\Gamma$. The other system parameters are $J = 6$ and $\epsilon = 0.2$. The filled symbols represent the inverse gap values $1/\Delta$ obtained with exact diagonalization and rescaled by a parameter $\alpha(\Gamma) = \mathrm{O}(1)$. The thin dashed curves are exponential fits to the tunneling time $\xi$ in the large-$N$ regime.

protocol: the simulation starts with all random walkers initialized in the basis state with all spins pointing up; we then measure the imaginary time $\tau$ (computed as time step $\Delta\tau$ times number of DMC iterations) required to first reach a certain percentage of walkers, somewhat arbitrarily taken to be $25\%$, with negative magnetization (majority of spins pointing down), meaning that they have crossed the energy barrier. This definition is analogous to the one employed in Refs. [21, 22, 3] in the case of PIMC simulations, where a certain percentage of imaginary-time slices, instead of walkers, is considered. The simulation is repeated approximately $250$ times for larger systems and small $\Gamma$ and approximately $2500$ for smaller systems and larger values of $\Gamma$. We then take the average value to define $\xi$ and its standard deviation to define the error bar.

The DMC tunneling times for the ferromagnetic Ising chain are shown in Figure (3.1), as a function of the number of spins $N$ and for different values of $\Gamma$. For large $N$ the data display an exponential growth, quite similar to the dependence of the inverse gap $1/\Delta$, which we obtain via exact diagonalization of the Hamiltonian matrix. In fact, by multiplying the inverse gap $1/\Delta$ by an appropriate numerical prefactor $\alpha$ we obtain precise matching between the two

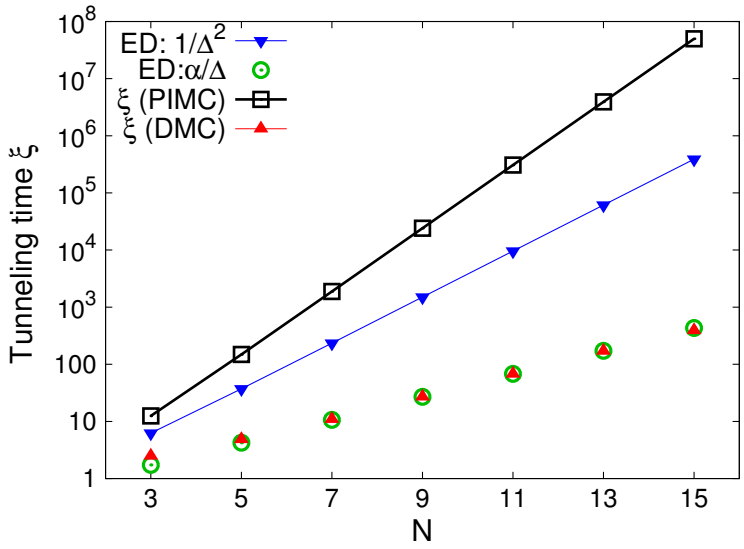



Figure 3.4: Comparison between the tunneling times of the finite-temperature PIMC algorithm and the DMC algorithm, for the shamrock model at $\Gamma = 0.5$. The PIMC data are obtained from the formula $\xi = 2^K/\Delta^2$, which was found in Ref [3]. The solid blue points represent the scaling $1/\Delta^2$, characteristic of incoherent quantum tunneling. The red triangles represent the DMC data. They are well described by the scaling law $\alpha/\Delta$ (green empty circles), where the gap $\Delta$ is obtained via exact diagonalization. The simulation parameters are $J = 6$ and $\epsilon = 0.2$.

datasets. The coefficient $\alpha$ turns out to be a number $\mathrm{O}(1)$. We also consider different definitions of DMC tunneling time, using percentages of walkers that have to cross the barrier between $10\%$ and $25\%$, obtaining again results which follow the $1/\Delta$ scaling but with a slightly different value of the prefactor $\alpha$.

The inverse-gap scaling displayed by the DMC tunneling times is similar to the result found in Ref. [21] using modified PIMC simulations performed using open boundary conditions in imaginary time. This is not surprising, since such modified PIMC method had been originally introduced as a computational tool to study ground-state properties [83, 84]. However, it is usually employed in combination with guiding wave functions that accurately describe the ground state, so that the convergence to the zero-temperature limit as a function of the total path length is quite rapid. How this algorithm converges to the ground state in the absence of the guiding wave function has not been analyzed in detail yet.

It is also worth stressing that in the PIMC formalism the tunneling time is defined by counting the number of Monte Carlo sweeps (a sweep corresponds to one

Monte Carlo step per spin) and, therefore, it does not bear the significance of imaginary time as in the DMC method employed in this thesis. In Ref. [21], also finite-temperature PIMC simulations (with periodic boundary conditions) have been performed, finding that the PIMC tunneling times scale as $1/\Delta^2$. This behavior was found in ferromagnetic Ising models, which are characterized by a one-dimensional reaction coordinate, and it was latter confirmed also in one-dimensional and two-dimensional continuous-space models [22], showing that it persists also when the reaction coordinate is multidimensional.

Considered together, the above findings suggest that QMC algorithms are either as efficient as (in the case of the finite-temperature PIMC algorithm) or quadratically faster than (in the case of the PIMC algorithm with open boundary conditions in imaginary time or of the DMC algorithm) QAs in tunneling through energy barriers and therefore, if one assumes that incoherent quantum tunneling is the major resource of QAs, also in solving optimization problems.

## 3.3 The shamrock model

The results for the ferromagnetic Ising chain presented in the previous section indicate that, in an effective double-well system, QMC simulations can efficiently simulate incoherent quantum tunneling and, therefore, they might potentially be as efficient as, or even faster than QAs in solving complex optimization problems. In order to understand if this finding is valid in a more general setup, the authors of Ref. [3] considered a model, named "shamrock", which contains the minimal elements of frustration. This model is described by the following Hamiltonian:

$$\hat{H} = -J \sum_{i=1}^{K} \sum_{j=2i}^{2i+1} \sigma_1^z \sigma_j^z + (J - \epsilon) \sum_{i=1}^{K} \sigma_{2i}^z \sigma_{2i+1}^z - \Gamma \sum_{i=1}^{N} \sigma_i^x. \tag{3.2}$$

The $N$ spins are grouped in $K$ rings, which form the leaves of the shamrock. See Fig. (3.2). Each ring is made of three spins and the $K$ rings all share one spin, which is placed in the center. The number of spins is related to the number of rings by the formula $N = 2K + 1$. In Eq. (3.2) $\epsilon \ll J$ is a small interaction energy.

The first term in Eq. (3.2) describes ferromagnetic interactions between the central spin and the outer $N - 1$ spins. The two outer spins of each ring are coupled to each other by an anti-ferromagnetic interaction, described by the second term in Eq. (3.2). The intensity of the transverse field in the last term in

Eq. (3.2) is $\Gamma$.

We investigate the DMC tunneling time using the protocol described in Section 3.2. The results are shown in Figure (3.3). They display the same $1/\Delta$ scaling already observed in the case of the ferromagnetic Ising chain, corresponding to a quadratic speedup with respect to incoherent quantum tunneling. The value of the prefactor $\alpha$ used to superimpose the inverse-gap data to the DMC tunneling time is, as in the ferromagnetic Ising chain, a number of $\mathrm{O}(1)$. This suggests that frustrated couplings do not play a fundamental role in the tunneling dynamics of DMC simulations.

In Figure (3.4) we also report the scaling of the tunneling times $\xi$ obtained in finite-temperature PIMC simulations in Ref. [3]. As opposed to the DMC data, which display the same $1/\Delta$ scaling in the ferromagnetic Ising chain and in the shamrock model, the PIMC results display, in the latter model, a faster growth of $\xi$ with the system size, very accurately described by the scaling law $\xi^{PIMC} \propto 2^K/\Delta^2$. Due to the $2^K$ term, this growth is exponentially faster than the scaling of the DMC tunneling time and of the timescale of incoherent quantum tunneling.

This pathological slowdown of PIMC simulations was indeed anticipated by the perturbation theory of Ref. [3]. This theory predicts that in frustrated models where the two competing ground states are connected by a number of homotopy-inequivalent paths which grows with system size, incoherent quantum tunneling can display a quantum speedup if many inter-path transitions are inhibited by topological obstructions (related to the obstructions discussed in Ref. [81]). The shamrock model was indeed introduced as an example of this scenario, with the PIMC simulations confirming the theoretical prediction also beyond the perturbative regime.

## 3.4 Analysis of the systematic bias in DMC simulations due to the finite random-walker population

As mentioned in Chapter 1, the ground-state energy obtained via DMC simulations is subject to two sources of possible systematic bias, originating from the finite time step $\Delta\tau$ and from the finite random-walker number $N_w$. The convergence to the $\Delta\tau \to 0$ limit is quadratic, and all results presented in this chapter have been

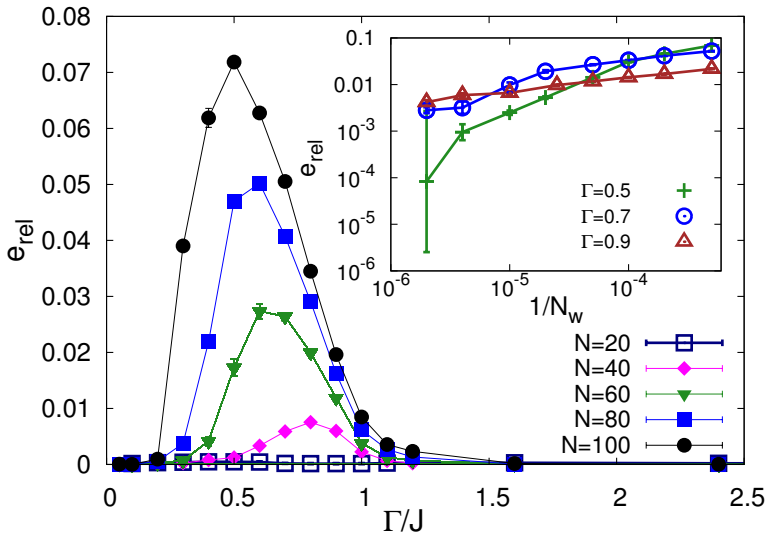



Figure 3.5: Main panel: Relative error $e_{\rm rel} = |E - E_{\rm JW}| \,/\, |E_{\rm JW}|$ of the DMC result $E$ with respect to the exact Jordan–Wigner theory $E_{\rm JW}$ as a function of the transverse field intensity $\Gamma$, for different system sizes $N$. The average number of random walkers is $N_w = 20000$. Inset: $e_{\rm rel}$ as a function of the inverse number of walkers $1/N_w$, for different transverse field intensity $\Gamma$. The size of the spin chain is N=60.

performed using sufficiently small $\Delta\tau$ to make its systematic effect negligible compared to the statistical uncertainty. In this section, we focus on the bias resulting from the finite value of $N_w$. All data reported here have been obtained with DMC simulations much longer than the equilibration time, meaning that there is no bias due to the initial random-walker configuration (see also the discussion on equilibration times reported in subsection 1.3.4).

We consider the ferromagnetic quantum Ising chain defined in Eq. (3.1). Its ground-state energy per site can be exactly determined via Jordan–Wigner transformation, obtaining in the thermodynamic limit

$$E_{\rm JW}/N = -\frac{2}{\pi}\,(1+\Gamma)\,\mathrm{E}\,(\theta)\,, \tag{3.3}$$

where $\mathrm{E}(x)$ is the complete elliptic integral of the second kind and [40]

$$\theta^2 = 4\Gamma/\left[(1+\Gamma)^2\right]. \tag{3.4}$$

As a reference for the outcomes of the DMC simulations, we use the ground-state

energy results obtained from the Jordan–Wigner transformation applied on the corresponding finite-size systems. In fact, for system sizes larger than $N = 30$, the relative discrepancy between the ground state energies of the finite systems and the one corresponding to the thermodynamic limit is below $0.02\%$, and is negligible on the scale considered in this Section.

In the main panel of Fig. (3.5), we plot the relative error

$$e_{\rm rel} = \frac{|E - E_{\rm JW}|}{|E_{\rm JW}|}, \tag{3.5}$$

of the DMC result $E$ with respect to the Jordan–Wigner theory as a function of the transverse field intensity, for different system sizes. These data correspond to a fixed random walker population $N_w = 20000$. One notices that in the paramagnetic phase $\Gamma > 1$, as well as in the $\Gamma \to 0$ limit, the systematic bias due to the finite $N_w$ is negligible. However, in the ferromagnetic phase $0 < \Gamma < 1$ a systematic bias is observable, and this bias increases with the system size $N$. The maximum relative error appears to drift to smaller $\Gamma$ values as $N$ increases, giving one the (wrong) impression that high accuracy is more difficult to achieve when quantum fluctuations vanish. However, by analyzing how the relative error vanishes in the $N_w \to \infty$ limit for different $\Gamma$ values at fixed system size (see inset of Fig. 3.5) one understands that in the asymptotic large-$N_w$ regime the largest relative error occurs for $\Gamma$ values close to (but smaller than) the quantum critical point $\Gamma = 1$. We attribute this effect to the large ferromagnetic fluctuations characteristic of the critical regime.

To better understand the effect of the finite random-walker population, we analyze in Fig. (3.6) the convergence to the exact Jordan–Wigner result in the $N_w \to \infty$ limit, considering different system sizes, at $\Gamma = 0.95$. The data are well described by power-law fitting functions of the type $e_{rel} = c/N_w^{\beta}$, where $c$ and $\beta$ are fitting parameters. The exponent $\beta$ decreases with the systems size, meaning that, as the system size increases, it takes a larger population of walkers to obtain accurate predictions. In order to quantify this dependence, in Figure (3.7) we show how the number of walkers required to have a fixed relative error increases with the system size. In the large-$N$ limit, the data are well described by an exponential fitting function, possibly indicating that the computational complexity of the simple DMC algorithm (i.e., without the use of the importance sampling technique) is exponential in the system size.

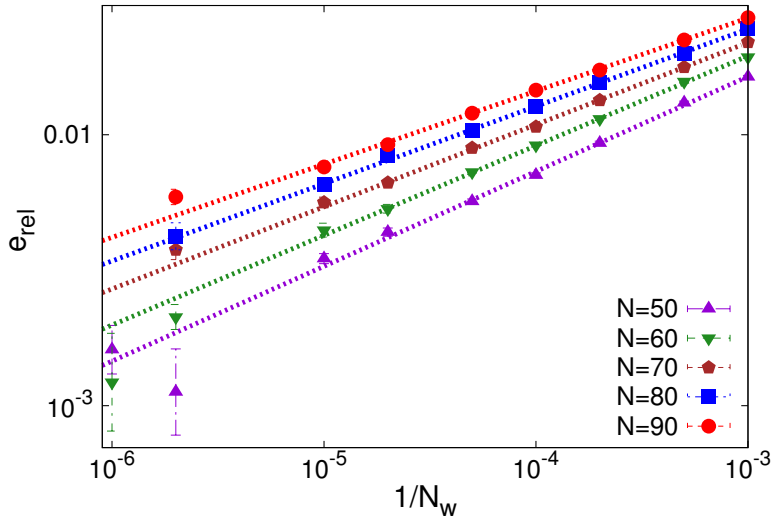



Figure 3.6: Relative error $e_{\rm rel}$ as a function of the inverse number of walkers $1/N_w$, for different system sizes. The transverse field intensity is $\Gamma = 0.95$. The dashed curves represent power-law fitting functions (see text).

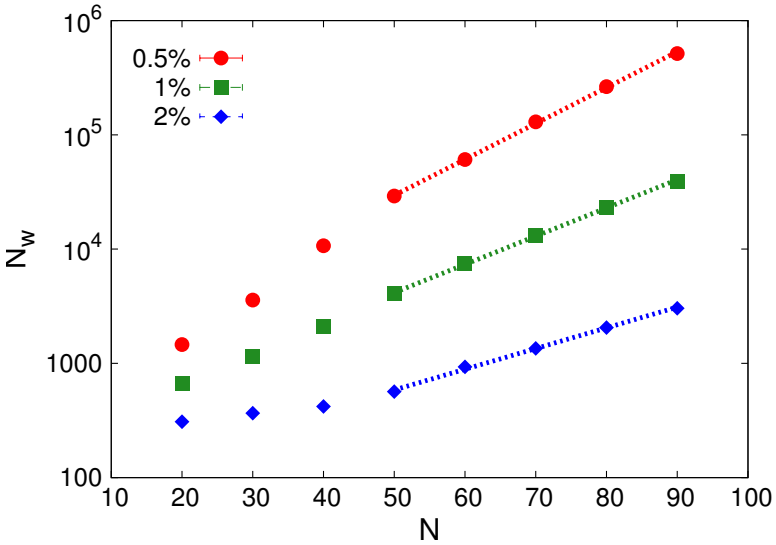



Figure 3.7: Random-walker number necessary to have $0.5\%$, $1\%$, and $2\%$ relative error as a function of the system size $N$. The transverse field intensity is $\Gamma = 0.95$.

## 3.5 Conclusions

We implemented a projective QMC method for quantum Ising models based on the DMC algorithm — in which the transition matrix is defined using a Trotter approximation of the Green's function — and we investigated the characteristic time of tunneling events in problems characterized by an effective double-well energy landscape. We found that the DMC tunneling time increases with the system size as the inverse of the gap, that is, more favorably than the incoherent tunneling time, which increases as the inverse gap squared. This inverse-gap scaling was found to hold both for a ferromagnetic quantum Ising chain and for the more challenging shamrock model. This is in contrast with previous studies based on finite-temperature PIMC simulations, where a pathological slowdown due to topological obstructions originating from frustrated interactions was found to cause, in the shamrock model, an increase of the PIMC tunneling time which is exponentially faster [3] than the inverse-gap squared scaling observed in the case of simple ferromagnetic models [21]. Our findings indicate that the DMC algorithm is not affected by the obstructions that slow down the PIMC tunneling dynamics, thus suggesting that this algorithm is a more efficient engine for SQA considered as a heuristic optimization method.

Motivated by the arguments of Ref. [3] — according to which a classical algorithm is to be considered an efficient simulation of QAs only if it reproduces both their tunneling dynamics and their equilibrium properties — we analyzed the computation time required by the DMC algorithm to accurately predict ground-state properties. The analysis of the systematic bias in the ground-state energy due to the finite random-walker population revealed an exponential increase of the population size and, therefore, of the computation time, required to keep a fixed relative error as the system size increases. This suggests that, in general, the computational effort required to simulate the behavior of QAs via simple DMC simulations running on classical computers scales exponentially, leaving the door open for potential quantum speedup.

The finding of this exponential scaling is consistent with the statement of Ref. [85] that the problem of estimating the ground state energy of a stoquastic Hamiltonian with a small additive error is at least NP-hard [2]. This statement is based

[2] In some cases, ground-state energies of stoquastic models might be obtained in polynomial time using continuous-time PIMC simulations. However, the path-integral length has to increase with the system size, likely leading to increasing equilibration and correlation times, and so perhaps to the exponential scaling of the computational cost in the worst case scenario. In fact, it

on the observation that any Hamiltonian diagonal in the computational basis is stoquastic, and that finding its ground state encompasses hard optimization problems such as $k$-SAT and MAX-CUT. This essentially rules out the possibility that a polynomially scaling algorithm applicable to generic stoquastic Hamiltonians can be found. Still, for certain ferromagnetic models, including the transverse-field Ising chain considered in this work, algorithms which — albeit being far from practical — have a provably polynomial scaling have recently been found [87]. However, since the DMC algorithm we employ in this chapter is not tailored to a specific (e.g., ferromagnetic) Ising model, it is natural to observe the exponential behavior corresponding to a generic model.

The use of importance sampling might also allow one to efficiently simulate the models described in Refs. [33, 88], for which simple (i.e., without importance sampling) projective QMC methods have been shown to fail due to the large discrepancy between the $L_1$-normalized ground-state wave function — which is the probability distribution sampled from in simple projective QMC simulations — and the $L_2$-normalized ground-state wave function, which is sampled from when performing a measurement on the ground state of the adiabatic process. We plan to investigate these issues in future works.

Building accurate trial wave functions for generic optimization problems is an important, but highly nontrivial, task. We argue that finding models where such importance sampling technique is not feasible (because no accurate and efficiently computable guiding wave function exists) could help us in identifying optimization problems where quantum advantage can be achieved. For the same purpose, it would be useful to identify which features of a Hamiltonian might cause a pathological slowdown of the DMC dynamics.

is known that close to the superfluid to Mott-insulator transition ground-state PIMC simulations are feasible only with the use of very accurate trial wave functions [86].

# 4 Projective quantum Monte Carlo simulations guided by unrestricted neural network states

We investigate the use of unrestricted Boltzmann machines (uRBMs), as guiding functions in projective quantum Monte Carlo (PQMC) simulations of quantum spin models. As a preliminary step, we investigate the accuracy of the uRBM as a variational Ansätze for the ground state of the ferromagnetic quantum Ising chain. We find that by optimizing just three variational parameters, independently on the system size, accurate ground-state energies are obtained, comparable to those previously obtained using restricted Boltzmann machines with few variational parameters per spin. Chiefly, we show that if one uses optimized uRBM as guiding functions for importance sampling the efficiency of the PQMC algorithms is greatly enhanced, drastically reducing the most relevant systematic bias, namely that due to the finite walker population. The scaling of the computational cost with the system size changes from the exponential scaling characteristic of PQMC simulations performed without importance sampling, to a polynomial scaling, apparently even at the quantum critical point. The important role of the protocol chosen to sample hidden-spins configurations, in particular at the critical point, is analyzed. We discuss the implications of these findings for what concerns the problem of simulating adiabatic quantum optimization using stochastic algorithms on classical computers.

## 4.1 Introduction

Quantum Monte Carlo (QMC) algorithms are generally believed to be capable of predicting equilibrium properties of quantum many-body systems at an affordable computational cost, even for relatively large system sizes, at least when the sign problem does not occur. However, it has recently been shown that the

computational cost to simulate the ground state of a quantum Ising model with a simple projective QMC (PQMC) algorithm that does not exploit importance sampling techniques scales exponentially with the system size, making large-scale simulations unfeasible [69]. This happens in spite of the fact that the Hamiltonian is sign-problem free.

This result stresses the importance of shedding light on the computational complexity of PQMC algorithms: if these computational techniques allowed one to simulate, with a polynomially scaling computational cost, both the ground-state properties of a model Hamiltonian, and also the tunneling dynamics of a quantum annealing device described by such Hamiltonian (as discussed in the previous chapter), then quantum speedup would be very unlikely to be achieved. We focus in this chapter on the first of the two aspects, specifically, on analyzing and improving the scaling of the computational cost to simulate ground-state properties of quantum Ising models.

It is well known that the efficiency of PQMC algorithms can be enhanced by implementing importance sampling techniques using as guiding functions accurate variational Ansätze [36]. However, building accurate variational wave-functions for generic many-body systems is a highly non trivial task. Recently, variational wave-functions that mimic the structure of neural networks have been shown to accurately describe ground-state properties of quantum spin and lattice models [25, 89, 90]. The authors of Ref. [25] considered neural network states that mimic restricted Boltzmann machines (RBM), i.e. such that no interaction among hidden spins is allowed. One very appealing feature of such restricted neural network states is that the role of the hidden spins can be accounted for analytically, without the need of Monte Carlo sampling over hidden variables.

Furthermore, such states provide very accurate ground-state energy predictions, which can be systematically improved by increasing the number of hidden spins per visible spin (latter on referred to as hidden-spin density). However, this high accuracy is obtained at the cost of optimizing a number of variational parameters that increases with the system size. This optimization task can be tackled using powerful optimization algorithms such as the stochastic reconfiguration method (see, e.g, Ref. [91]). Yet, having to optimize a large number of variational parameters is not desirable in the context of quantum annealing simulations, since one would be dealing with a variational optimization problem, potentially even more difficult than the original classical optimization problem.

In this chapter, we consider instead neural network states that mimic unrestricted Boltzmann machines (uRBMs), allowing intra-layer correlations among hidden spins, beyond the inter-layer hidden-visible correlations and the intra-layer visible-visible correlations (see Fig. 4.1). The structure of these states resembles the one of the shadow wave functions originally introduced to describe quantum fluids and solids [92, 93]. We test their representational power considering as a testbed the ferromagnetic quantum Ising chain. The main goal of this chapter is to show that the above-mentioned unrestricted neural network states can be used as a guide for importance sampling in PQMC simulations.

The rest of the chapter is organized as follows: in Section 4.2 we define the conventional Boltzmann-type variational wave functions and the unrestricted neural network states, and we then analyze how accurately they predict the ground-state energy of the quantum Ising chain via optimization of, respectively one and three, variational parameters. Section 4.3 deals with the continuous-time PQMC algorithm and with the implementation of importance sampling using both Boltzmann-type wave functions and, chiefly, unrestricted neural network states, showing how the systematic bias due to the finite random-walker population is affected, both at and away from the quantum critical point. The important effect of choosing different sampling protocols for the hidden spins is also analyzed. Our conclusions and the outlook are reported in Section 4.4.

## 4.2 Unrestricted neural network states for quantum Ising models

Once more, we consider as a test bed the one-dimensional ferromagnetic quantum Ising Hamiltonian defined in Eq. (1.48):

$$\hat{H} = \hat{H}_{\rm cl} + \hat{H}_{\rm kin}, \tag{4.1}$$

where $\hat{H}_{\rm cl} = -J \sum_{i=1}^{N} \sigma_i^z \sigma_{i+1}^z$ and $\hat{H}_{\rm kin} = -\Gamma \sum_{i=1}^{N} \sigma_i^x$.

Recall that the function $E_{\rm cl}(\boldsymbol{x}) = \langle \boldsymbol{x}|\hat{H}_{\rm cl}|\boldsymbol{x}\rangle$ (with $\boldsymbol{x} = (x_1, x_2, \ldots, x_N)$) corresponds to the Hamiltonian function of a classical Ising model, while the operator $\hat{H}_{\rm kin}$ introduces quantum (kinetic) fluctuations.

Our first goal is to develop trial wave functions that closely approximate the ground state wave function $\Psi_0(\boldsymbol{x}) = \langle \boldsymbol{x}|\Psi_0\rangle$ of the Hamiltonian (4.1). A simple

Ansatz can be defined as

$$\Psi_{\boldsymbol{\kappa}}(\boldsymbol{x}) = e^{-\beta E_{\mathrm{cl}}(\boldsymbol{x})} = e^{-K_1 \sum_{i=1}^{N} x_i x_{i+1}} \ . \tag{4.2}$$

$\boldsymbol{\kappa}$ is here a set of real variational parameters to be optimized. Their values are obtained by minimizing the average of the energy, as in standard variational quantum Monte Carlo approaches. In this case, only one parameter $K_1 = \beta$ is present, $\boldsymbol{\kappa} = \{K_1\}$. This choice is inspired by the classical Boltzmann distribution where $\beta$ would play the role of a fictitious inverse temperature. The above Ansatz will be referred to as Boltzmann-type wave function.

A more sophisticated Ansatz can be constructed by using a generative stochastic artificial neural network, namely an unrestricted Boltzmann Machine (uRBM), see Figure (4.1). Beyond the visible spin variables $\boldsymbol{x} = (x_1, x_2, \ldots, x_N)$, one introduces $N$ hidden spin variables $\boldsymbol{h} = (h_1, h_2, \ldots, h_N)$, taking values $h_i = \pm 1$ (with $i = 1, \ldots, N$). Periodic boundary conditions within the layers are also incorporated, i.e $x_{N+1} = x_1$ and $h_{N+1} = h_1$. The trial wave function is thus written in the following integral form:

$$\Psi_{\boldsymbol{\kappa}}(\boldsymbol{x}) = \sum_{\boldsymbol{h}} \phi_{\boldsymbol{\kappa}}(\boldsymbol{x}, \boldsymbol{h}) \ , \tag{4.3}$$

where,

$$\phi_{\boldsymbol{\kappa}}(\boldsymbol{x}, \boldsymbol{h}) = e^{-\sum_{i=1}^{N} (K_1 x_i x_{i+1} + K_2 h_i h_{i+1} + K_3 x_i h_i)} \ . \tag{4.4}$$

Notice that the architecture of this uRBM includes correlations between nearest-neighbor visible spins, between nearest-neighbor hidden spins, as well as between pairs of visible and hidden spins with the same index $i$. These three correlations are parametrized by the three constants $K_1$, $K_2$, and $K_3$, respectively. With this uRBM trial Ansatz, the set of variational parameters is $\boldsymbol{\kappa} = \{K_1, K_2, K_3\}$.

It is straightforward to generalize the uRBM Ansatz including more layers of hidden spins. Every additional hidden-spin layer adds two more variational parameters, and it effectively represents the application of an imaginary-time Suzuki-Trotter step $e^{-\Delta\tau \hat{H}}$ for a certain time step $\Delta\tau$. Thus, a deep neural network state with many hidden layers can represent a long imaginary-time dynamics, which projects out the ground state provided that the initial state is not orthogonal to it. In fact, the mapping between deep neural networks and the imaginary time projection has been exploited in Refs. [94, 95] to construct more complex

neural network states. In this work we consider only the single hidden-spin layer uRBM, since this Ansatz turns out to be adequate for the ferromagnetic quantum Ising chain. The multi hidden-spin layer Ansatz might be useful to address more complex models as, e.g, frustrated Ising spin glasses. Extensions along these lines are left as future work.

In a recent work [25], Carleo and Troyer considered a restricted Boltzmann machine (RBM), where direct correlations among hidden spins were not allowed. Their Ansatz included a larger number of hidden spins, as well as more connections between visible and hidden spins, leading to an extensive number of variational parameter proportional to $\alpha N$, where $\alpha = 1, 2, \ldots$. One advantage of the RBM, due to the absence of hidden-hidden correlations, is that the role of hidden spins can be analytically traced out.

The uRBM we employ, which is analogous to the shadow wave functions used to describe quantum fluid and solids, includes only three variational parameters, independently of the system size. However, their effect has to be addressed by performing sampling of hidden spins configurations, as described below. Note that one could map this specific uRBM in one-dimension to a matrix product state with small bond dimension [96, 97], allowing to integrate out the hidden-spin configurations. However, we are interested in a flexible QMC algorithms that straightforwardly extends to general setups, in particular to higher-dimensional geometries and to models with disorder and/or long-range interactions, since these are relevant for optimization problems. In these cases Monte Carlo sampling becomes essential [98].

It is worth pointing out that correlations beyond nearest-neighbor spins could also be included in the uRBM Ansatz, with straightforward modifications in the sampling algorithms described below. We mention here also that, as shown in Ref. [99], neural network states with intra-layer correlations can be mapped to deep neural networks with more hidden layers, but no intra-layer correlations.

In the case of an uRBM variational wave function, the average value of the energy

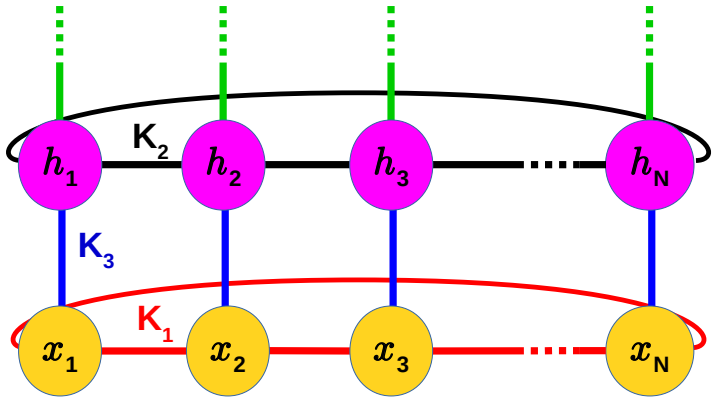



Figure 4.1: Structure of the unrestricted Boltzmann machine. The lower (yellow) nodes depict visible spins, the upper (magenta) nodes depict the hidden spins. The horizontal segments indicate intralayer visible-visible and hidden-hidden correlations. The vertical (blue) segments represent the interlayer correlations between the corresponding visible and hidden spins. The green lines allude to a possible extension to deep layers architectures.

$E = \langle \hat{H} \rangle$ is computed as follows

$$\begin{aligned} \langle \hat{H} \rangle &= \frac{\langle \Psi_{\boldsymbol{\kappa}} | \hat{H} | \Psi_{\boldsymbol{\kappa}} \rangle}{\langle \Psi_{\boldsymbol{\kappa}} | \Psi_{\boldsymbol{\kappa}} \rangle} = \frac{\sum_{\boldsymbol{x},\boldsymbol{x}'} \Psi_{\boldsymbol{\kappa}}(\boldsymbol{x}) H_{\boldsymbol{x},\boldsymbol{x}'} \Psi_{\boldsymbol{\kappa}}(\boldsymbol{x}')}{\sum_{\boldsymbol{x}} \Psi_{\boldsymbol{\kappa}}(\boldsymbol{x}) \Psi_{\boldsymbol{\kappa}}(\boldsymbol{x})} \\ &= \frac{\sum_{\boldsymbol{x},\boldsymbol{x}'} \left[\sum_{\boldsymbol{h}_{\mathbf{a}}} \phi_{\boldsymbol{\kappa}}(\boldsymbol{x}, \boldsymbol{h}_{\mathbf{a}})\right] H_{\boldsymbol{x},\boldsymbol{x}'} \left[\sum_{\boldsymbol{h}_{\mathbf{b}}} \phi_{\boldsymbol{\kappa}}(\boldsymbol{x}', \boldsymbol{h}_{\mathbf{b}})\right]}{\sum_{\boldsymbol{x}} \left[\sum_{\boldsymbol{h}_{\mathbf{a}}} \phi_{\boldsymbol{\kappa}}(\boldsymbol{x}, \boldsymbol{h}_{\mathbf{a}})\right] \left[\sum_{\boldsymbol{h}_{\mathbf{b}}} \phi_{\boldsymbol{\kappa}}(\boldsymbol{x}, \boldsymbol{h}_{\mathbf{b}})\right]} \\ &= \ll E_{\mathrm{loc}}(\boldsymbol{x}, \boldsymbol{h}_{\mathbf{b}}) \gg , \end{aligned} \tag{4.5}$$

where the local energy $E_{\mathrm{loc}}(\boldsymbol{x}, \boldsymbol{h})$ is defined as

$$E_{\mathrm{loc}}(\boldsymbol{x}, \boldsymbol{h}) = \frac{\sum_{\boldsymbol{x}'} H_{\boldsymbol{x},\boldsymbol{x}'} \phi_{\boldsymbol{\kappa}}(\boldsymbol{x}', \boldsymbol{h})}{\phi_{\boldsymbol{\kappa}}(\boldsymbol{x}, \boldsymbol{h})} , \tag{4.6}$$

with $H_{\boldsymbol{x},\boldsymbol{x}'} = \langle \boldsymbol{x} | \hat{H} | \boldsymbol{x}' \rangle$. $\boldsymbol{h}_{\mathbf{a}}$ and $\boldsymbol{h}_{\mathbf{b}}$ indicate two hidden spin configurations. Notice that the formula for the local energy can be symmetrized with respect to the two sets of hidden spins $\boldsymbol{h}_{\mathbf{a}}$ and $\boldsymbol{h}_{\mathbf{b}}$, providing results with slightly reduced statistical fluctuations. The double brackets $\ll \cdots \gg$ indicate the expectation value over the visible-spin configurations $\boldsymbol{x}$ and two sets of hidden spins configurations $\boldsymbol{h}_{\mathbf{a}}$

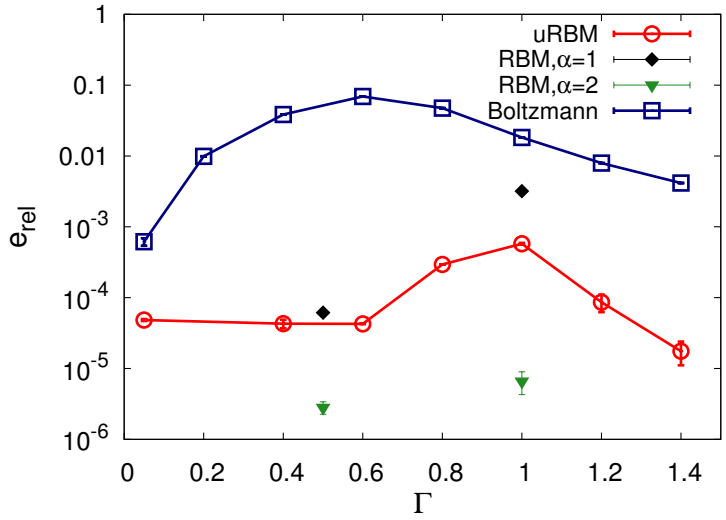



Figure 4.2: Relative error $e_{\rm rel}$ in the variational estimates of the ground-state energy as a function of the transverse field $\Gamma$, obtained using the simple Boltzmann wave function and for the unrestricted Boltzmann machine (uRBM) Ansatz. The system size is $N = 80$. For comparison, we also show the data corresponding to the restricted Boltzmann machine (RBM) from Ref [25], where $\alpha$ indicates the hidden-spin density. The thin lines are guides to the eyes.

and $\boldsymbol{h}_{\rm b}$, sampled from the following normalized probability distribution:

$$p(\boldsymbol{x}, \boldsymbol{h}_{\rm a}, \boldsymbol{h}_{\rm b}) = \frac{\phi_{\boldsymbol{\kappa}}(\boldsymbol{x}, \boldsymbol{h}_{\rm a})\phi_{\boldsymbol{\kappa}}(\boldsymbol{x}, \boldsymbol{h}_{\rm b})}{\sum_{\boldsymbol{x},\boldsymbol{h}_{\rm a},\boldsymbol{h}_{\rm b}} \phi_{\boldsymbol{\kappa}}(\boldsymbol{x}, \boldsymbol{h}_{\rm a})\phi_{\boldsymbol{\kappa}}(\boldsymbol{x}, \boldsymbol{h}_{\rm b})} \,. \tag{4.7}$$

As in standard Monte Carlo approaches, this expectation value is estimated as the average of $E_{\rm loc}(\boldsymbol{x}, \boldsymbol{h})$ over a (large) set of uncorrelated configurations, sampled according to $p(\boldsymbol{x}, \boldsymbol{h}_{\rm a}, \boldsymbol{h}_{\rm b})$. The statistical uncertainty can be reduced at will by increasing the number of sampled configurations. The optimal variational parameters $\boldsymbol{\kappa}_{\rm opt}$ that minimize the energy expectation value can be found using a stochastic optimization method. We adopt a relatively simple yet quite efficient one, namely the stochastic gradient descent algorithm (see, e.g., [27]). While more sophisticated algorithms exist as, e.g., the stochastic reconfiguration method [91], such methods are not necessary here since the Ansätze that we consider include a very small number of variational parameters, one or three. In fact, in these cases the optimal variational parameters can be obtained also by performing a scan on a fine grid. By doing so, we obtain essentially the same results provided by the stochastic gradient descent algorithm.

We assess the accuracy of the optimized variational wave functions by calculating the relative error (see also Eq. 1.50)

$$e_{\mathrm{rel}} = \frac{|E - E_{\mathrm{JW}}|}{|E_{\mathrm{JW}}|} \,, \tag{4.8}$$

in the obtained variational estimate $E$ of the ground state energy of the Hamiltonian in Eq. (4.1). $E_{\mathrm{JW}}$ is the exact finite size ground state energy of the quantum Ising chain defined in Equation (1.49).

Figure (4.2) displays the relative error $e_{\mathrm{rel}}$ in Eq. (4.8) corresponding to the variational wave functions introduced above, as a function of the transverse field $\Gamma$. The system size is $N = 80$, which is here representative of the thermodynamic limit. The Boltzmann-type Ansatz does not provide particularly accurate predictions. In the ferromagnetic phase $\Gamma < 1$, the relative error is up to $10\%$. The uRBM, instead, provides very accurate predictions. The relative error is always below $0.1\%$. The largest discrepancy occurs at the quantum critical point $\Gamma = 1$. Such high accuracy is remarkable, considering that the uRBM Ansatz involves only $3$ variational parameters. It is also worth mentioning that very similar accuracies are obtained also for different system sizes. Therefore, the uRBM Ansatz represents a promising guiding function for simulations of quantum annealing optimization of disordered models.

As a term of comparison, we show in Fig. (4.2) the results obtained in Ref. [25] using the RBM Ansatz. The relative errors corresponding to the RBM with hidden-unit density $\alpha = 1$ are larger than those corresponding to the uRBM, despite the fact that the RBM Ansatz involves a larger number of variational parameters. However, it is worth stressing that the RBM results can be systematically improved by increasing $\alpha$. For example, with $\alpha = 2$ the RBM relative errors are approximately an order of magnitude smaller than those corresponding to the uRBM Ansatz.

## 4.3 Importance sampling guided by unrestricted neural network states

In this section, we discuss how optimized variational wave functions can be utilized to boost the performance of PQMC simulations. The PQMC method that we implement is the continuous-time Green's function Monte Carlo (CTGFMC)

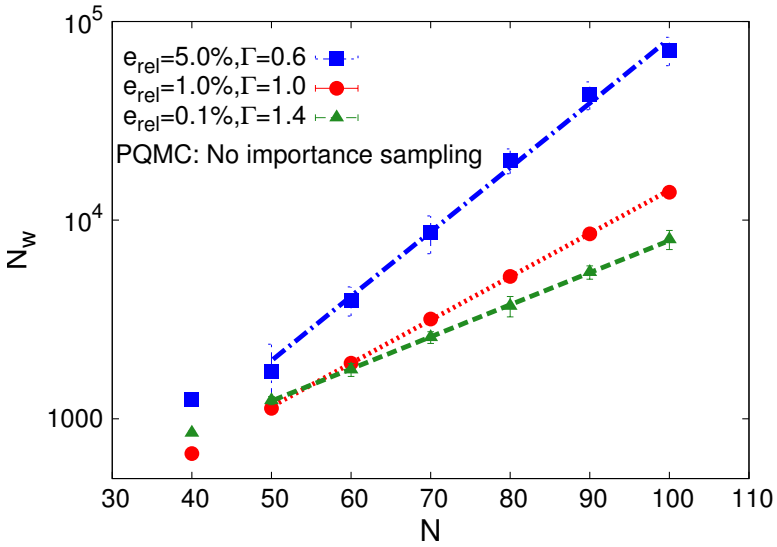



Figure 4.3: Number of random walkers $N_w$ required to determine, using the PQMC algorithm without importance sampling, the ground-state energy with a relative error $e_{\rm rel}$, as a function of the system size $N$. Different datasets correspond to different transverse field intensities $\Gamma$ and different relative errors. The lines represent exponential fitting functions.

method exhaustively described in chapter 1. The guiding wave-functions used for importance sampling in the CTGFMC simulations are the variational wave-functions $\Psi_{\boldsymbol{\kappa}}(\boldsymbol{x})$ introduced in Sec. 4.2 where, the variational parameters $\boldsymbol{\kappa}$ are set at their optimal values.

In order to employ the unrestricted neural-network states as guiding functions, the PQMC algorithm has to be modified. One has to implement a combined dynamics of the visible-spin configurations $\boldsymbol{x}$ and of the hidden-spin configurations $\boldsymbol{h}$. We will indicate the global configuration as $\boldsymbol{y} = (\boldsymbol{x}, \boldsymbol{h})$. The goal is to sample global configurations with the (normalized) probability distribution

$$p(\boldsymbol{y}) = \frac{\Psi_0(\boldsymbol{x})\phi_{\boldsymbol{\kappa}}(\boldsymbol{x},\boldsymbol{h})}{\sum_{\boldsymbol{x},\boldsymbol{h}} \Psi_0(\boldsymbol{x})\phi_{\boldsymbol{\kappa}}(\boldsymbol{x},\boldsymbol{h})} \,. \tag{4.9}$$

This allows one to compute the ground state energy as

$$E = \frac{1}{N_c} \lim_{N_c\to\infty} \sum_{i=1}^{N_c} E_{\rm loc}(\boldsymbol{x}_i, \boldsymbol{h}_i), \tag{4.10}$$

where $N_c$ is a number of uncorrelated configurations $\{\boldsymbol{y}_i\}$ sampled from $p(\boldsymbol{y})$.

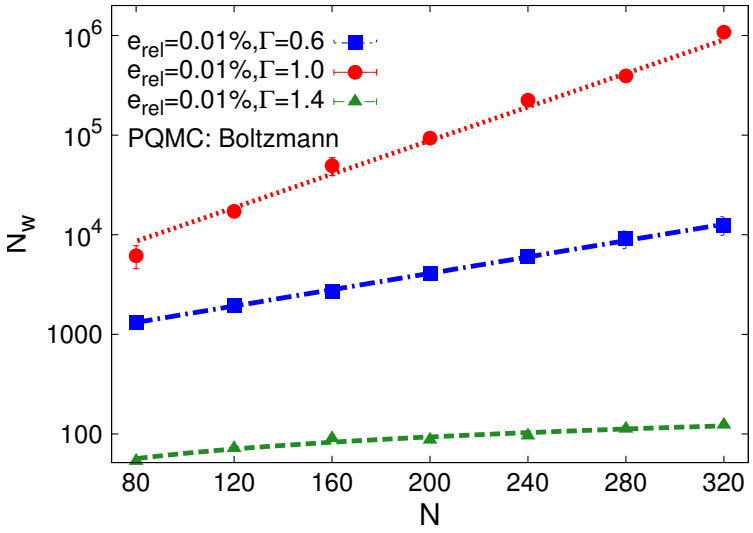



Figure 4.4: Number of random walkers $N_w$ required to determine, using the optimized Boltzmann-type wave function to guide importance sampling in the PQMC simulation, the ground-state energy with a relative error $e_{\mathrm{rel}}$, as a function of the system size $N$. Different datasets correspond to different transverse field intensities $\Gamma$. The (red) dotted and (blue) dot-dashed lines represent exponential fitting functions, while the (green) dashed line represents a power-law fit with power $b = 0.54(5)$.

The local energy $E_{\rm loc}(\boldsymbol{x},\boldsymbol{h})$ is defined as in Eq. (4.6).

A suitable algorithm was implemented in Ref. [100] in the case of the continuous-space Green's function Monte Carlo algorithm, where importance sampling was implemented using shadow wave functions. Here we modify the approach of Ref. [100] to address quantum spin models. The visible-spins configurations $\boldsymbol{x}$ are evolved according to the CTGFMC described above, keeping the hidden-spin configuration $\boldsymbol{h}'$ fixed. The modified imaginary-time Green's function is now

$$\tilde{G}(\boldsymbol{x},\boldsymbol{x}',\Delta\tau|\boldsymbol{h}') = G(\boldsymbol{x},\boldsymbol{x}',\Delta\tau)\frac{\phi_{\boldsymbol{\kappa}}(\boldsymbol{x},\boldsymbol{h}')}{\phi_{\boldsymbol{\kappa}}(\boldsymbol{x}',\boldsymbol{h}')}. \tag{4.11}$$

As discussed in subsection 1.3.3, this has to be rewritten as the product of a stochastic matrix, which defines how the visible-spin configurations updates are selected, and a weight term, which is taken into account with the branching process. The weight-update factor is

$$b_{\boldsymbol{y}'} = \sum_{\boldsymbol{x}} \tilde{G}(\boldsymbol{x},\boldsymbol{x}',\Delta\tau|\boldsymbol{h}').$$

The dynamics of the hidden-spins configurations is dictated by a (classical) Markov chain Monte Carlo algorithm.

Considering $\phi_{\boldsymbol{\kappa}}(\boldsymbol{x},\boldsymbol{h})$ as an unnormalized probability distribution allows one to write — for any fixed visible-spin configuration $\boldsymbol{x}$ — the Master equation:

$$\phi_{\boldsymbol{\kappa}}(\boldsymbol{x},\boldsymbol{h}) = \sum_{\boldsymbol{h}'} T(\boldsymbol{h},\boldsymbol{h}'|\boldsymbol{x})\phi_{\boldsymbol{\kappa}}(\boldsymbol{x},\boldsymbol{h}'), \tag{4.12}$$

where $T(\boldsymbol{h},\boldsymbol{h}'|\boldsymbol{x})$ is the transition matrix that defines the Markov process. Clearly, the following condition must be fulfilled

$$\sum_{\boldsymbol{h}} T(\boldsymbol{h},\boldsymbol{h}'|\boldsymbol{x}) = 1, \tag{4.13}$$

for any $\boldsymbol{x}$. Our choice is a single spin flip Metropolis algorithm, where the flip of a randomly selected spin is proposed, and accepted with the probability

$$A(\boldsymbol{h}'\rightarrow\boldsymbol{h}|\boldsymbol{x}) = \mathrm{Min}\left\{1, \frac{\phi_{\boldsymbol{\kappa}}(\boldsymbol{x},\boldsymbol{h})}{\phi_{\boldsymbol{\kappa}}(\boldsymbol{x},\boldsymbol{h}')}\right\}. \tag{4.14}$$

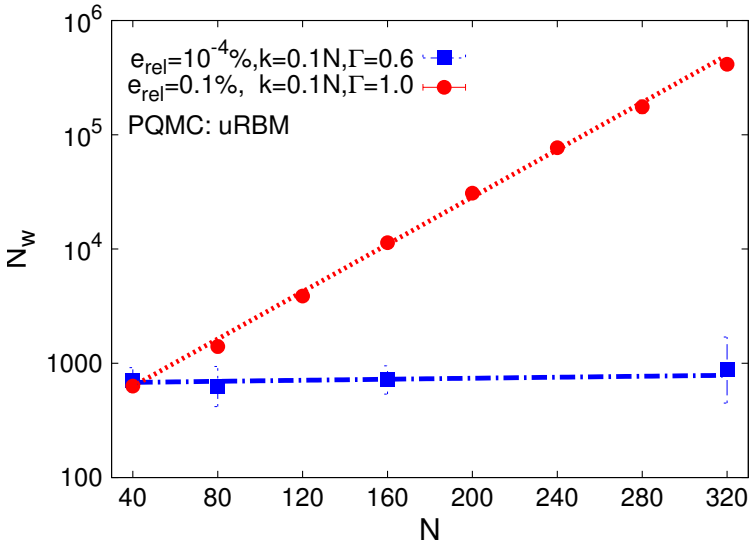



Figure 4.5: Number of random walkers $N_w$ required to determine, using the optimized uRBM Ansatz to guide importance sampling in the PQMC simulation, the ground-state energy with a relative error $e_{\rm rel}$, as a function of the system size $N$. The number of single-spin Metropolis updates of the hidden spins per CTGFMC hidden-spin update is $k = 0.1N$. The (red) dotted line represents and exponential fit, while the (blue) dot-dashed line represents a linear fit.

Here, $\boldsymbol{h}$ differs from $\boldsymbol{h}'$ only for the (randomly selected) flipped spin. One could perform a certain number, call it $k$, of Metropolis updates, without modifying the formalism. In fact, this turns out to be useful, as discussed below. The combined dynamics of the visible and the hidden spins is driven by the following equation:

$$f(\boldsymbol{y}, \tau + \Delta\tau) = \sum_{\boldsymbol{y}'} G(\boldsymbol{y}, \boldsymbol{y}', \Delta\tau) f(\boldsymbol{y}', \tau), \tag{4.15}$$

with

$$G(\boldsymbol{y}, \boldsymbol{y}', \Delta\tau) = T(\boldsymbol{h}, \boldsymbol{h}'|\boldsymbol{x}) \tilde{G}(\boldsymbol{x}, \boldsymbol{x}', \Delta\tau|\boldsymbol{h}'). \tag{4.16}$$

It can be shown [100] that the equilibrium probability distribution of this equation is the desired joint probability distribution $p(\boldsymbol{y})$ in Eq. (4.9). The stochastic process corresponding to this equation can be implemented with the following steps:

**i)** perform the visible-spin configuration update $\boldsymbol{x}' \to \boldsymbol{x}$, keeping $\boldsymbol{h}'$ fixed, according to the CTGFMC algorithm described above (including accumulation of

the weight factor);

**ii)** perform $k$ single-spin Metropolis updates of the hidden-spin configuration $\boldsymbol{h}'$, keeping $\boldsymbol{x}$ fixed;

**iii)** perform branching of the global configuration.

It is easily shown that the hidden-spin dynamics does not directly affect the weight factor since the normalization of the Green function of the combined dynamics is set by $b_{\boldsymbol{y}'}$.

Since the optimized uRBM describes the ground state wave function with high accuracy, one expects that its use as guiding function leads to a drastic reduction of the systematic errors due to the finite random walker population. However, one should take into account that there might be statistical correlations among subsequent hidden-spin configurations along the Markov chain. This might in turn affect the systematic error. Clearly, increasing the number of Metropolis steps $k$ per CTGFMC visible-spin configuration update allows one to suppress such correlations, possibly reducing the systematic error. This will indeed turn out to be important, in particular at the quantum critical point where statistical correlations along the Markov chain are more significant.

Following Ref. [69], we analyze the computational complexity of the PQMC algorithm by determining the number of walkers $N_w$ needed to determine the ground state energy of the Hamiltonian (4.1) with a prescribed accuracy. All data described below have been obtained with a time step $\Delta\tau = 0.1$, and all simulations have been run for a long enough total imaginary time to ensure equilibration.

First, we consider the simple PQMC algorithm i.e., performed without importance sampling. Fig. (4.3) displays the scaling with the system size $N$ of the number of walkers $N_w$ required to keep the relative error $e_{\mathrm{rel}}$, defined in Eq. (4.8), at the chosen threshold. This scaling is evidently exponential, below, above, and also at the quantum critical point. The most severe scaling comes from the ordered phase and could be attributed to the fact that the simple PQMC is formally equivalent to PQMC with a constant $\psi_T(\boldsymbol{x})$ for importance sampling. This turns out to be a very poor choice of the guiding function in the ordered regime given that it treats all configurations on an equal footing. Analogous results have been obtained in Ref. [69] using the diffusion Monte Carlo algorithm.

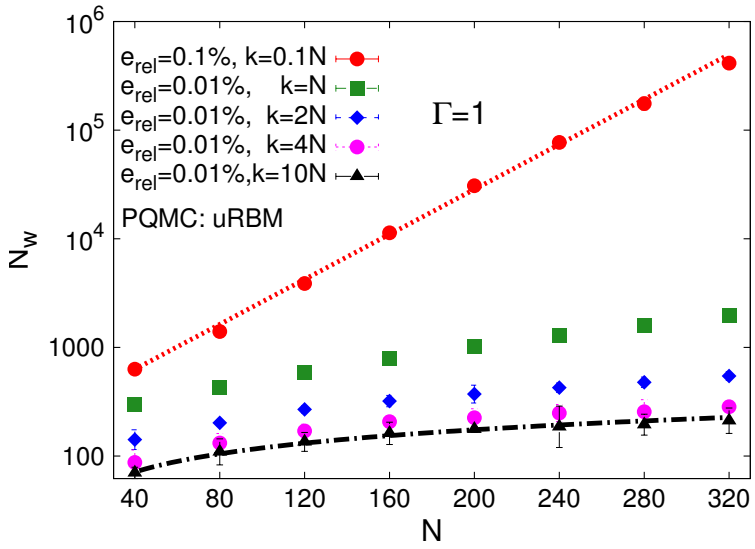



Figure 4.6: Number of random walkers $N_w$ required to determine, using the optimized uRBM to guide importance sampling in the PQMC simulation, the ground-state energy with a relative error $e_{\rm rel}$, as a function of the system size $N$. The transverse field intensity is set at the ferromagnetic quantum critical point $\Gamma = 1$. Different datasets correspond to different values of the the number of single-spin Metropolis updates $k$. The (red) dotted line represents an exponential fit, while the (black) dot-dashed line represents a power-law fit, with power $b = 0.55(1)$.

Introducing importance sampling using the optimized Boltzmann-type Ansatz as guiding function significantly reduces the systematic error due to the finite random walker population, allowing one to reach quite small relative errors. In particular, in the paramagnetic phase at $\Gamma = 1.4$, the scaling of $N_w$ versus $N$ is quite flat (see Fig. 4.4); it appears to be well described by the power-law $N_w \sim N^b$ with the small power $b = 0.54(5)$, rather than by an exponential. However, in the ferromagnetic phase at $\Gamma = 0.6$ and at the quantum critical point $\Gamma = 1$ the scaling is still clearly exponential. This means that the simple Boltzmann-type Ansatz is, in general, insufficient to ameliorate the exponentially scaling computational cost of the PQMC algorithm.

Figure (4.5) shows the scaling of $N_w$ obtained using the optimized uRBM Ansatz as the guiding function. The number of hidden-spin Metropolis steps per visible-spin update is set to a (small) fraction of the system size $N$, namely to $k = 0.1N$. At $\Gamma = 0.6$, the required walker population size $N_w$ turns out to be essentially independent on the system size $N$. It is worth noticing that the prescribed relative error is here as small as $e_{\mathrm{rel}} = 10^{-6}$, and that this high accuracy is achieved with a rather small walkers population $N_w \lesssim 1000$. However, at the quantum critical point, $N_w$ still displays an exponential scaling with system size. This effect can be traced back to the diverging statistical correlations among subsequent hidden-spin configurations along the Markov chain, due to quantum criticality. As anticipated above, these statistical correlations can be suppressed by increasing the number of hidden-spin updates $k$.

Figure (4.6) displays the scaling of $N_w$, at the quantum critical point, for different $k$ values. One observes that the scaling substantially improves already for moderately larger $k$ values, leading to a crossover from the exponential scaling obtained with $k = 0.1N$, to a scaling behavior that is well described by a square-root function $N_w \sim N^{0.55(1)}$ when $k = 10N$. While it is in principle possible that, for system sizes $N \gg 300$, $N_w$ still approaches an exponential scaling with a very small base, we argue that the data we provide indicate that the scaling of $N_w$ is not exponential, even at the ferromagnetic critical point, provided that $k$ is large enough.

It is important to point out that increasing $k$ implies a correspondingly increasing contribution to the global computational cost of the PQMC algorithm. However, since $k$ is here linear in the system size, this contribution does not modify, to leading order, the scaling of the global computational cost. Therefore, one

can conclude that the uRBM Ansatz is sufficient to change the scaling of the computational cost of the PQMC algorithm from exponential in the system size, to an amenable polynomial scaling. In the simulations presented here, single-spin flip Metropolis updates are employed for the hidden variables. It is possible that cluster spin updates would lead to an even faster convergence to the exact $N_w \rightarrow \infty$ limit, due to the more efficient sampling of the hidden-spin configurations. However, such cluster updates cannot always be implemented, in particular for frustrated disordered Hamiltonians relevant for optimization problems; therefore, we do not consider them here.

## 4.4 Conclusions

The accuracy of variational wave-functions that mimic unrestricted Boltzmann machines, which we refer to as unrestricted neural network states, has been analyzed using the one-dimensional ferromagnetic Ising model as a testbed. By optimizing just three variational parameters, ground-state energies with a relative error smaller than $10^{-3}$ have been obtained. The ferromagnetic quantum phase transition turns out to be the point where the relative error is the largest. This accuracy is comparable to the one previously obtained using restricted neural network states with few hidden variables per visible spin [25].

These restricted neural network states involve a number of variational parameters proportional to the system size, as opposed to the unrestricted neural network states considered here, where the (small) number of variational parameters is fixed. This feature of the unrestricted states makes them very suitable in the context of quantum annealing simulations for Ising-type models (which are sign-problem free). However, since one has to integrate over hidden-spins configurations via Monte Carlo sampling, as opposed to the case of the restricted neural network states [25] — for which the hidden-spin configurations can be integrated out — they represent a less promising approach to model ground-states of Hamiltonian where the negative sign-problem occurs. Indeed, in such case an accurate variational Ansatz might have to include also hidden-spins configurations with negative wave-function amplitude, making Monte Carlo integration via random sampling inapplicable.

We have found that unrestricted neural network states allow one to drastically reduce the systematic bias of the PQMC algorithm originating from the finite size

of the random-walker population. Specifically, the scaling of the population size required to keep a fixed relative error as the system size increases changes from the exponential scaling characteristic of simple PQMC simulations performed without guiding functions, to a polynomial scaling. This also implies a corresponding change in the scaling of the computational cost. This qualitative scaling change occurs above, below, and apparently also at the ferromagnetic quantum phase transition, at least for the system sizes $N \lesssim 300$ considered in this article. While we cannot strictly rule out a weak exponential scaling for $N \gg 300$, the numerical data we provide suggest that this is an unlikely scenario. Instead, a conventional variational Ansatz of the Boltzmann type was found to provide a significant improvement of the computational cost only above the critical point (in the paramagnetic phase), but to provide only a marginal improvement at and below the transition.

It is worth emphasizing that the use of unrestricted neural network states as guiding functions in PQMC simulations requires the sampling of both the visible and the hidden spins, using the combined algorithm described in Sec. 4.3 (more efficient variants might be possible). The role of the statistical correlations among hidden-spin configurations shows up in particular at the ferromagnetic quantum critical point. We found that these correlations can be eliminated by performing several single-spin updates, still without affecting, to leading order, the global computational complexity of the simulation.

In Ref. [87] it was proven that it is possible to devise polynomially-scaling numerical algorithms to determine the ground-state energy, with a small additive error, of various ferromagnetic spin models, including the ferromagnetic Ising chain considered here. However, practical implementations have not been provided. The numerical data we have reported in this manuscript indicate that the PQMC algorithm guided by an optimized unrestricted neural network state represents a practical algorithm with polynomial computational complexity for the ferromagnetic quantum Ising chain.

More in general, it was shown in Ref. [85] that the problem of estimating the ground-state energy of a generic sign-problem free Hamiltonian with a small additive error is at least NP-hard. Indeed, this task encompasses hard optimization problems such as $k-$SAT and MAX-CUT. This suggest that there might be relevant models where the unrestricted neural network states discussed here are not sufficient to make the computational cost of the PQMC simulations affordable.

Relevant candidates are Ising spin-glass models with frustrated couplings. Such systems might require more sophisticated guiding functions obtained, e.g., including more hidden-spin layers in the unrestricted neural network state, as discussed in Sec. 4.2. In future work we plan to search for models that make PQMC simulation problematic. We argue that this will help us in understanding if and for which models a systematic quantum speed-up in solving optimization problems using quantum annealing devices, instead of PQMC simulations performed on classical computer, could be achieved.

# 5 Simulating quantum annealing using PQMC: a preliminary exploration

In this chapter, we use the projective quantum Monte Carlo (PQMC) method described previously to simulate quantum annealing on the quantum Ising chain (QIC). We first implement the simple PQMC method, i.e, without importance sampling and show the results of the residual energies with respect to the total annealing time. In order to compare on an equal footing the dynamics of the quantum annealing simulations with the one obtained numerically by solving the imaginary-time Schroedinger equation (IT-SHE) [24], one needs the pure estimator of the average residual energies. We compute them using the techniques introduced in Chapter 1. We show that for finite size systems, the residual energy of the PQMC method scales with the total annealing time as what is obtained with the IT-SHE [24]. We study more in details the annealing simulations by looking at the distributions of the residual energies at the end of the annealing for different system sizes.

Next, for a fixed system size, we study two specific annealing simulations, one in which the classical ground state is recovered at the end of the annealing and the other one for which the simulations end up in the excited state of the classical Hamiltonian.

We move on to study the effect of the main source of systematic errors in the PQMC method which is the one due to the finite number of random-walkers. Previously, in Chapters 3 and 4, we studied how a finite number of walkers introduces a bias in equilibrium ground state simulations. In this chapter, we study how the number of walkers affects the results when the PQMC method is implemented to solve an optimization problem via a quantum annealing process in which the transverse field decreases in time thus gradually removing quantum

fluctuations.

Lastly, we present results of simulated quantum annealing (SQA) performed using the PQMC with importance sampling. We use the formalism described in the first chapter. Even though importance sampling is performed with a naive Boltzmann-like trial wave-function, we show that it indeed helps substantially to perform adequate simulations of the annealing dynamics. As it was in the case of continuous space-models, we observe that, though the use of a guiding wave-function reduces the residual energy, the scaling of the residual energy using a mixed estimator remains the same. Nevertheless, it seems to portray (if valid in the thermodynamic limit) a quadratic speed-up compared with the Kibble-Zurek scaling predicted for real-time dynamics of the Schroedinger equation [101], which path integral Monte Carlo was very recently shown to portray [60].

## 5.1 Introduction

The long term aim of our study is to solve hard optimization problems. In order to accomplish our goal, we develop a heuristic method to find the ground-state of the following type of Hamiltonian

$$\hat{H}_{\mathrm{cl}} = -\sum_{ij} J_{ij}\hat{\sigma_i^z}\hat{\sigma_j^z} - \sum_{i} h_i\hat{\sigma_i^z}, \tag{5.1}$$

where $J_{ij}$ denotes the interaction strength between spins at site $i$ and $j$ and $h_i$ is the strength of the longitudinal field at site $i$.

Many difficult optimization problems can be encoded using the form of the above Hamiltonian with the terms $J_{ij}$ and $h_i$ representing a typical instance of the problem to solve. The potential energy landscape of $\hat{H}_{\mathrm{cl}}$ is usually made-up of a very large number of hills and valleys and finding the deepest valley corresponds to solving the optimization problem.

Classical annealing (CA) [53] and quantum annealing (QA) [1] are the main heuristic methods used to address such optimization problems. In the latter case, it is achieved by introducing a fictitious dynamics in the configuration space which enables the system to explore different configurations in the system. The task is successfully completed if at the end of the search the configuration that minimizes the potential energy is found. The adiabatic theorem guarantees that it

is the case provided that the annealing time is much larger than the instantaneous gap squared of the Hamiltonian.

The dynamics is implemented through quantum fluctuations by introducing a fictitious kinetic energy term

$$\hat{H}_{\text{kin}} = -\Gamma(\tau) \sum_i \hat{\sigma_i^x}. \tag{5.2}$$

Thereafter, the system is initially prepared in the paramagnetic regime where $\Gamma(\tau) \gg J_{ij}, h_i$, then it is set to (hopefully) follow adiabatically the dynamics of the Hamiltonian

$$\hat{H}(\tau) = \hat{H}_{\text{cl}} + \hat{H}_{\text{kin}}(\tau). \tag{5.3}$$

The strength of the transverse field is progressively reduced to zero according to a given protocol. When all quantum effects are completely removed from the system, the ground state of the classical Hamiltonian $\hat{H}_{\text{cl}}$ is (in principle) found.

The QA method described above cannot be applied on classical computers due to the exponential amount of resources needed thus, one has to resort to the use of QMC methods. In the community, the above scheme is usually implemented using the path-integral Monte Carlo (PIMC) method. Using that, the $D$-dimensional Hamiltonian $\hat{H}(\tau)$ is mapped into an equivalent classical Hamiltonian in $D+1$ dimensions. Thereafter, a stochastic dynamics is performed on the resulting Hamiltonian using Metropolis Monte Carlo hence, it is not directly related to the Schroedinger dynamics.

In our approach, we aim at implementing an evolution more closely related to the imaginary-time Schroedinger dynamics using the PQMC method.

## 5.2 SQA used with simple PQMC for the quantum Ising model

In this section, we use as a test bed the quantum Ising model already studied in the previous chapters. The only difference being that the kinetic energy is now time dependent: $\Gamma \equiv \Gamma(\tau)$. $\tau$ here stands for an imaginary time. The derivations performed in Chapter 1 are still valid and we will use them both for the simple

PQMC method and for the PQMC with importance sampling (PQMC-IS).

Simulated quantum annealing is performed by preparing the system in the ground state of the QIC Hamiltonian where a very large value of the transverse field is chosen. All through this chapter it is set to $\Gamma_0 = 5$. Then the transverse field is gradually reduced following the typical protocol also employed in quantum annealers

$$\Gamma(\tau) = \Gamma_0(1 - \frac{\tau}{\tau_f}). \tag{5.4}$$

$\tau_f$ is the total imaginary-time during which the system is evolved and, is referred to as the annealing time. It is equal to the time step times the number of steps in which the annealing is performed. As in the continuous-time case, we measure the annealing time starting from configurations previously equilibrated at $\Gamma_0 = 5$. At the end of the annealing procedure, we compute the error $\varepsilon_{res}$ that the algorithm makes in finding the classical ground-state

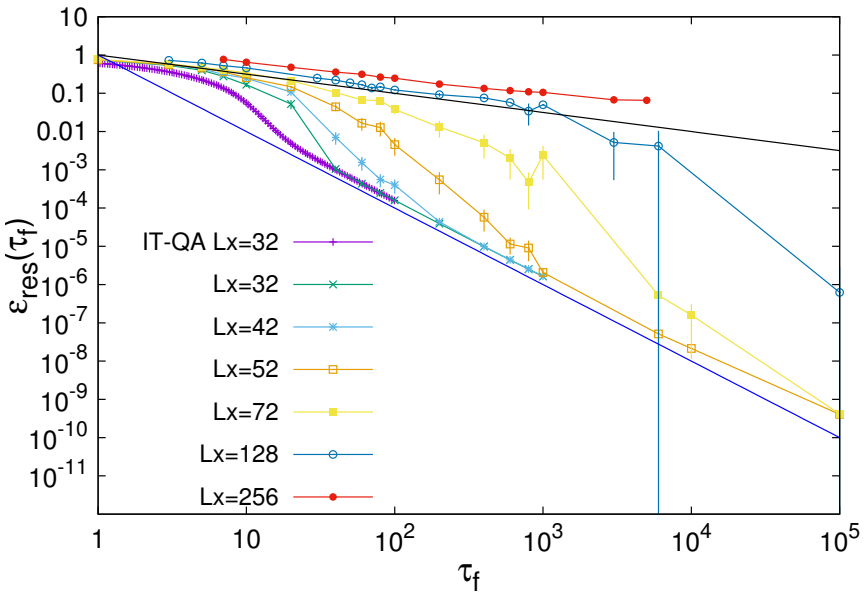
$$\varepsilon_{res}(\tau_f) = \frac{\langle \Psi_0(\tau_f)|\hat{H}_{\rm cl}|\psi_T(\tau_f)\rangle}{\langle \Psi_0(\tau_f)|\psi_T(\tau_f)\rangle} - E_0. \tag{5.5}$$

$\psi_T(\tau_f)$ here is a constant since we are using PQMC without importance sampling, however we find it useful to define it in such a way given that it will be used latter on when importance sampling is introduced. The average in Eq. 5.5 is estimated using typical averages given that the distribution of the solutions at the tail end of the annealing is not always normal as we will show latter on. The typical average of the residual energy is defined as

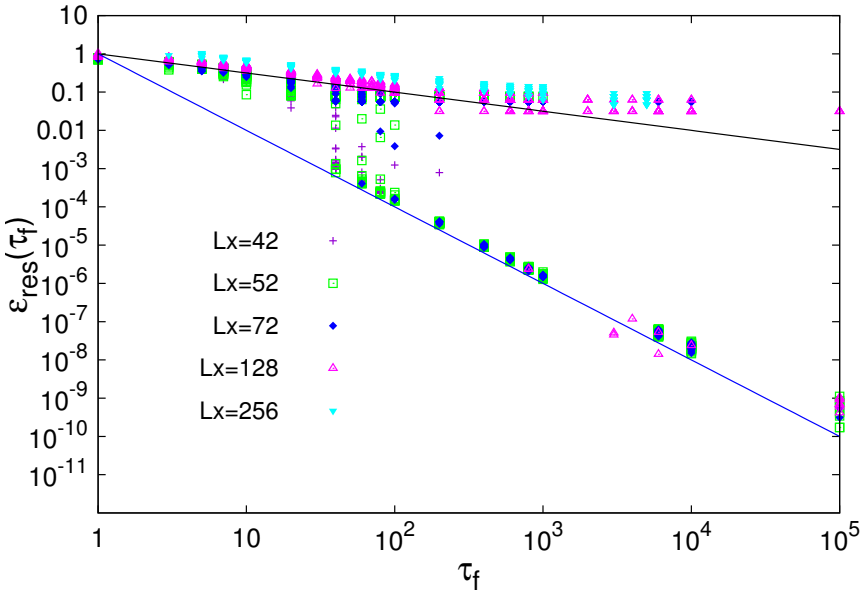
$$\Big[\varepsilon_{res}(\tau_f)\Big]_{typ} = e^{[\ln \varepsilon_{res}(\tau_f)^{estimator}]_{avg}}, \tag{5.6}$$

where $[.]_{avg}$ stands for the sum of elements in brackets divided by the total number of those elements. The label $estimator$ in Eq. (5.6) refers to the type of estimator that has been used.

Given that the PQMC algorithm gives a mixed estimator, we have to proceed to a re-weighting procedure at the tail end of the annealing in order to approximate pure estimators. We use at the end of the annealing schedule the novel method that we introduced in section 1.4. Except it is otherwise stated, in all the plots in this chapter, we use the "approximate" pure estimator computed with Eqs. 1.54 and 1.55. In this case of the QIC, the solution of the optimization problem




(a)




(b)

Figure 5.1: SQA of the quantum Ising model at different sizes $L_x$ of the chain. Top panel: Typical average of the residual energy $\varepsilon_{\rm res}$ versus annealing time $\tau_f$, obtained using the simple PQMC algorithm. Bottom panel: Residual energy of independent runs versus $\tau_f$. The simulations are equivalent to the ones that produce data of panel (a). The blue solid line indicate a function given by ${\tau_f}^{-2}$ whereas, the black line is the function ${\tau_f}^{-1}$. For all data, the walkers population is $N_w = 30,000$.

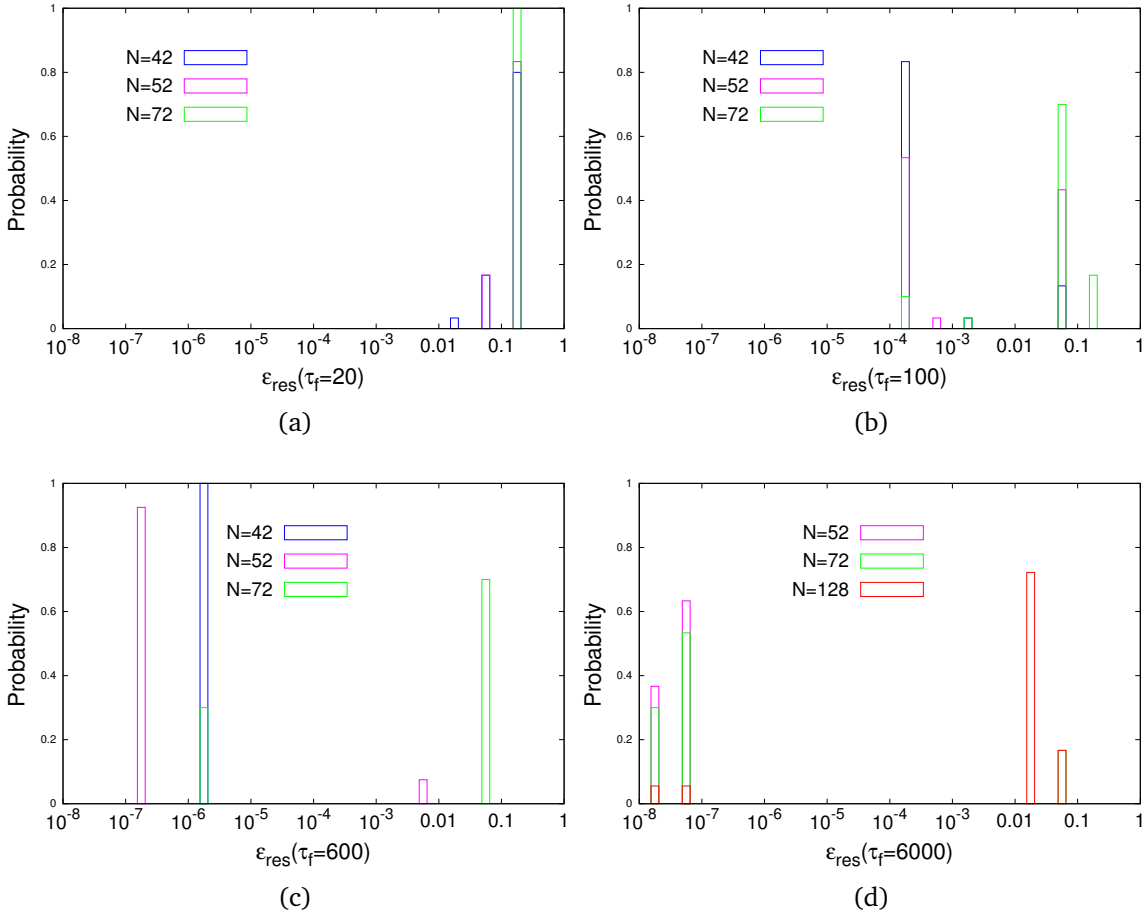



Figure 5.2: Probability distribution of the residual energy data taken from 5.1(b) at fixed annealing time $\tau_f$. Data are plotted for different system size $N$.

corresponds to finding one of the two equivalent classical states with all spins aligned up or down.

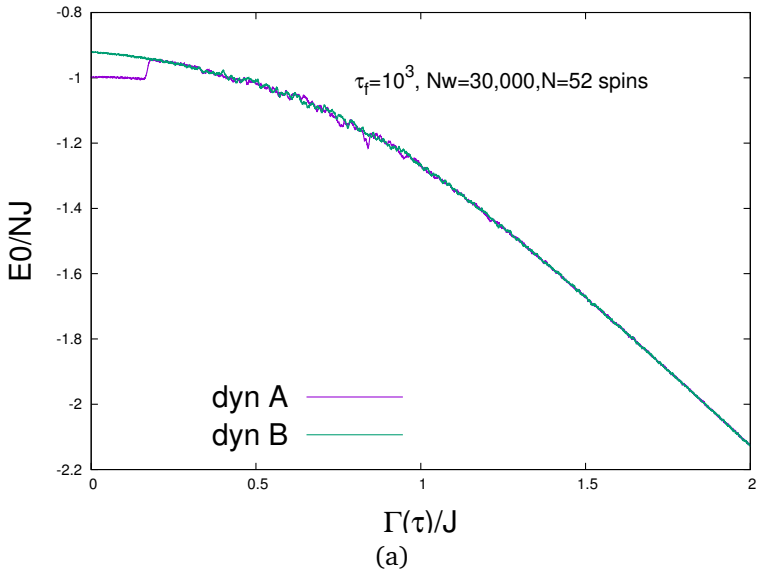
In Figure 5.1 we show results of the typical residual energy average with respect to the total annealing time $\tau_f$ at different system sizes. Averages are taken over many (of the order of $100$) different initial states. This is done by changing for each simulation the seed of the pseudo random number generator. In the panel (a), data labeled with IT-QA comes from [24] where the exact imaginary-time dynamics of the Schroedinger equation was performed. There, the authors used a novel approach based on a general imaginary-time dependent BCS state that has a Gaussian form [24].

The blue line in both plots corresponds to the power-law $\tau_f{}^{-2}$, it is chosen to represent the scaling of the IT-SHE calculated in [24]. System sizes of $32$ and $42$ spins seem to follow the dynamics of IT-QA at long annealing times whereas larger system sizes show pathological behaviors. In the panel (b), we plot individual instances of the residual energies for different runs of the annealing dynamics. We see that for large system sizes there are some instances for which $\varepsilon_{res}(\tau_f)$ falls on the blue line. Furthermore, the data at large annealing times the data give a seemingly bimodal distribution. The data that do not fall on that line describe a plateau which seems to indicate that the dynamics got stuck in highly excited classical states.

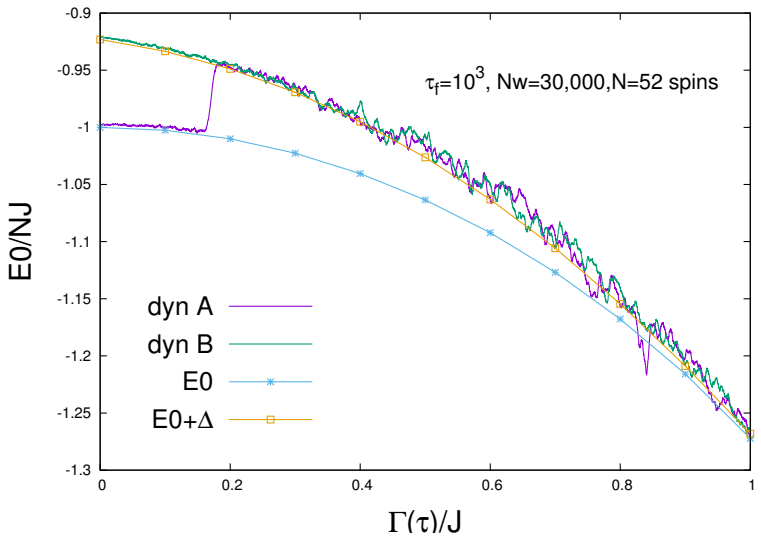
In order to understand deeper what is going on in the simulations, we draw an histogram of the residual energies that are obtained at the end of the annealing for each independent run of the simulations. Figure 5.2 shows the probability distribution of $\varepsilon_{res}(\tau_f)$ for different speed of the annealing at different system size $N$. A small $\tau_f$ is a sort of quantum quenching and as data in the panel (a) show, the algorithm is inefficient for those cases to find the classical ground state energy. When the annealing is performed long enough, the adiabatic regime (at least for small $N$) is reached and the ground state is found with higher accuracy. Nevertheless, the simple PQMC method seems to suffer from a metastability problem for large system sizes. In this plot, a clearer bimodal mode is observed in the data set, hence justifying the use of typical averages.

## 5.3 A case study of two distinct SQA runs

We then move on to study two different dynamics of the algorithm which are obtained from two independent runs at a fixed system size $N = 52$ spins. We will call them *dyn A* and *dyn B*. Figure 5.3 shows the dependence of the instantaneous ground state energy with respect to $\Gamma(\tau)$ as the annealing proceeds for the two dynamics previously mentioned. While the first one is able to project out the excited state at the tail end of the annealing, the second one stuck there. $\Delta$ on Figure 5.3(b) stands for the second gap which is the difference between the energies of the second excited state and the ground state obtained via Jordan-Wigner theory. Note that the first excited state will not be visible on this plot given that the first gap closes exponentially with the system size in the magnetic regime of the QIC. From the second panel of Figure 5.3, it seems that the PQMC simulations start going off the adiabatic regime after they have crossed the




(a)




(b)

Figure 5.3: (a) Dynamics of two distinct SQA runs performed using simple PQMC. The instantaneous ground state energy of the QIC is plotted versus the time dependent transverse field $\Gamma(\tau)$ (b) Zoom of the data in panel (a) to show details of the ferromagnetic regime. The blue line shows exact ground state energy data obtained using the Jordan–Wigner transformation. $\Delta$ stands for the gap between the second excited state and the ground state (see text).

quantum critical point. This is expected given that it is the point where the gap between the ground state and the first excited vanishes.

Figure 5.4 shows further details of the two dynamics. At the end of the annealing, the spin configurations are grouped into categories. The grouping rule is to put into the same category spin configurations that are equivalent under the $Z_2$ symmetry. In panels (a) and (b), we plot the classical energy with respect to the index labeling the different categories. *dyn A* has two order of magnitudes less categories than *dyn B*, indicating that most of its walkers did find the classical ground state. This is further verified in panels (c) and (d) where for *dyn A*, most of the walker fraction is in one spin configuration, which is the classical ground state. Whereas, *dyn B* has a much wider distribution of spin configurations.

## 5.4 Finite walkers population effect on the annealing simulations

In this section, we discuss on the effect of the finite number of random-walkers $N_w$ on SQA simulations. It was already shown in the previous chapters, using equilibrium simulations, that $N_w$ is the most severe systematic error of the PQMC method. Here, we are going to see that it is also very damaging for quantum annealing simulations.

In Figure (5.5) we plot the average residual energy with respect to the total annealing time for different values of the target number of walkers $N_w$. Simulations are done for a fixed (representative large) system size made up of $52$ spins. As it was previously shown in Fig. 5.1, panel (a) indicates averages of the $\varepsilon_{res}(\tau_f)$ taken for different runs of SQA and panel (b) depicts individuals values of $\varepsilon_{res}(\tau_f)$ at the end of the annealing. The figure clearly shows the dramatic effect that a finite number of walkers has on the simulations. With few walkers, say $N_w = 300$, the ground state energy is never reached even after runs as long as $\tau_f = 10^3$ thus, giving an erroneous scaling of the residual energies.

As the number of walkers increases, we observe a totally different behavior of the residual energy. When $N_w = 300,000$ walkers, the bimodal distribution previously seen for smaller $N_w$ vanishes in the long annealing time regime (see also Fig. 5.2). This indicates that the sampling crisis observed is actually caused by the insufficiency in the number of walkers. At annealing times of about $\tau_f = 100$,

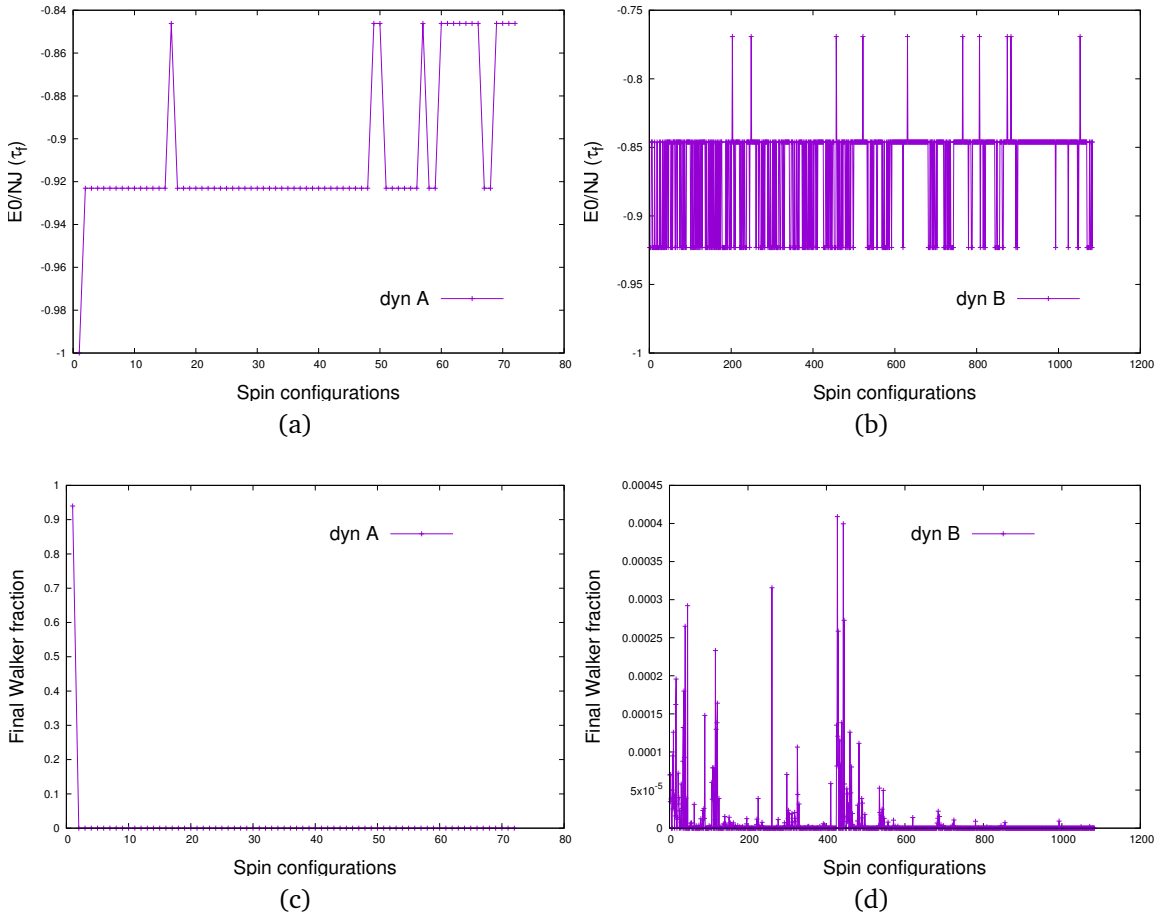



Figure 5.4: Panels (a) and (b) show the energy of distinct walkers at the end of the annealing for two different runs of the SQA simulations (see text). Panels (c) and (d) show the walker fraction of distinct walkers at the end of annealing for the two dynamics previously mentioned.

one already converges to the correct IT-QA scaling for the largest number of walkers. Given that the PQMC method is formally correct in the $N_w \to \infty$, we can conclude that a bimodal distribution is an artifact due to the incorrect implementation of the algorithm.

As discussed in Chapters 3 and 4, in equilibrium simulations, the scaling of the number of walkers required to have a fixed relative error is exponential in the system size. Therefore, one could suspect a similar behavior happening in the case of SQA. However we showed in Chapter 4 that a (good) guiding wavefunction may reduce the computational scaling of the algorithm, hence we introduce in the next section importance sampling.

## 5.5 SQA with a time-dependent trial wave function

In this section, we present the first attempt to implement SQA using a time-dependent trial wavefunction. The details of the formalism is taken from derivations performed in section 1.3. The PQMC method used is the continuous-time Green's function Monte Carlo (CTGFMC) algorithm presented in Chapter 1.

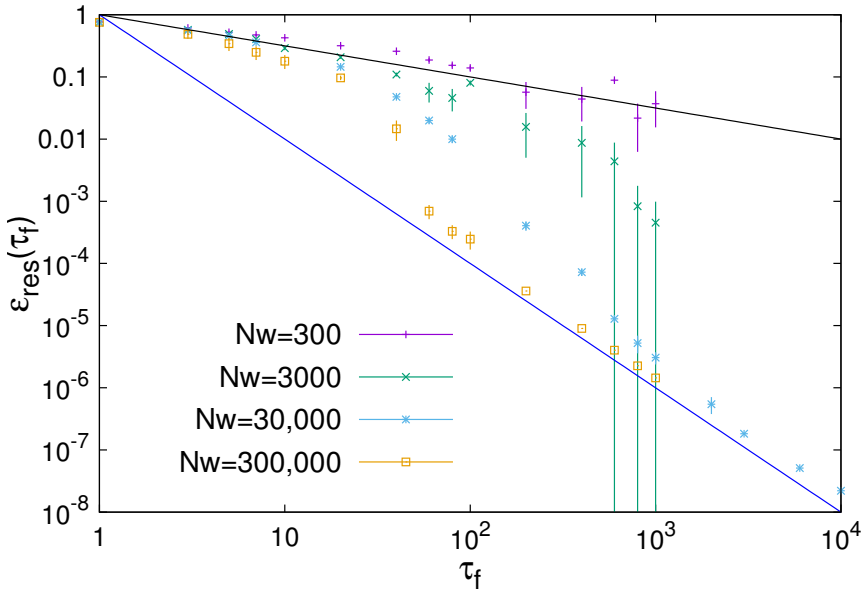
We chose a simple Boltzmann-like trial wave function for importance sampling, our choice is inspired by the work done in [14]. The trial wavefunction is given by

$$\psi_T(\boldsymbol{x}, \tau) = e^{-\frac{\widetilde{\beta}(\tau)}{2} E_{\mathrm{cl}}(\boldsymbol{x})}. \tag{5.7}$$

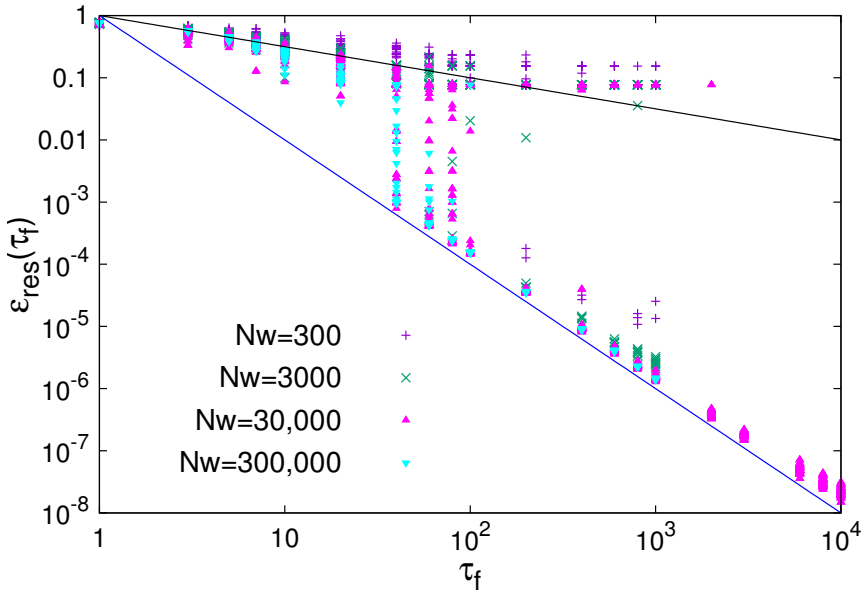
$E_{\mathrm{cl}}(\boldsymbol{x})$ is the classical energy of the QIC. $\widetilde{\beta}(\tau)$ is a fictitious time-dependent inverse-temperature given by:

$$\widetilde{\beta}(\tau) = \lambda \Big[1 - e^{-1/\Gamma(\tau)}\Big]^2. \tag{5.8}$$

$\lambda$ is a parameter that needs to be fixed. We do it by scanning different values of $\lambda$ in equilibrium simulations at the critical point of the quantum Ising chain. Simulations are performed using the CTGFMC method and the value of $\lambda$ that minimizes the standard deviation of the energy is chosen.




(a)




(b)

Figure 5.5: Similar plot as in Figure 5.1. A spin chain of length $N = 52$ is used at different values of the finite number of random-walker $N_w$.

The new local energy used in the branching is given by

$$E_{\rm loc}(\boldsymbol{x},\tau) = E_{\rm cl}(\boldsymbol{x}) + \tilde{V}(\boldsymbol{x},\tau) - \Gamma(\tau)\frac{\sum_{i=1}^{N}\psi_T(\tilde{\boldsymbol{x}}_i,\tau)}{\psi_T(\boldsymbol{x},\tau)}. \tag{5.9}$$

$\tilde{\boldsymbol{x}}_i$ is equivalent to the spin configuration $\boldsymbol{x}$ where the spin $i$ has been overturned. $\tilde{V}(\boldsymbol{x},\tau)$ is the fictitious time-dependent potential energy defined in Eq. (1.13).

The transition matrix is also modified, instead of it being given by Equation (1.47), it is now time-dependent:

$$t_{\boldsymbol{x},\boldsymbol{x}'}(\tau) = \begin{cases} \frac{\psi_T(\boldsymbol{x}_i,\tau)}{\sum_{i=1}^{N}\psi_T(\tilde{\boldsymbol{x}}_i,\tau)}, & \text{if } \boldsymbol{x}' \text{ is different from } \boldsymbol{x} \text{ of a single spin flip} \\ 0, & \text{otherwise.} \end{cases} \tag{5.10}$$

Fig. (5.6) shows the results on the SQA simulations performed on a chain of $52$ spins. Since we are using importance sampling, the procedure we introduce in the first chapter to approximate pure estimators cannot be used. We therefore plot values of the mixed estimators. We make comparisons with the "approximate" pure and mixed estimators obtained with simple PQMC simulations. For those cases, we plot the data in Fig. 5.5 for which the finite number of random-walkers is $N_w = 300,000$. It is interesting to note that our simulations with importance sampling use only $N_w = 5,000$, i.e, $60$ times less walkers than the case without importance sampling. Nevertheless, as it can be seen on the plot the data are well behaved.

Furthermore, we observe that using the mixed estimator with and without annealing gives the same scaling of the (typical) average residual energy at long imaginary times, i.e $\epsilon_{res} \backsim {\tau_f}^{-1}$. However, when an "approximate" pure estimator is used, the scaling is $\epsilon_{res} \backsim {\tau_f}^{-2}$, which is equivalent to the scaling of the IT-SHE.

Though the use of importance sampling renders the simulations more efficient, it still gives the same scaling as in the case when it is not used. However, if this scaling persists in the limit of infinite system sizes, it would display a quadratic speed-up with respect to the real-time dynamics of the Schroedinger equation (RT-SHE) as given by the Kibble-Zurek scaling, which is $\epsilon_{res} \backsim {\tau_f}^{-1/2}$. Note also that, this quadratic speed-up would apply also for finite temperature path-

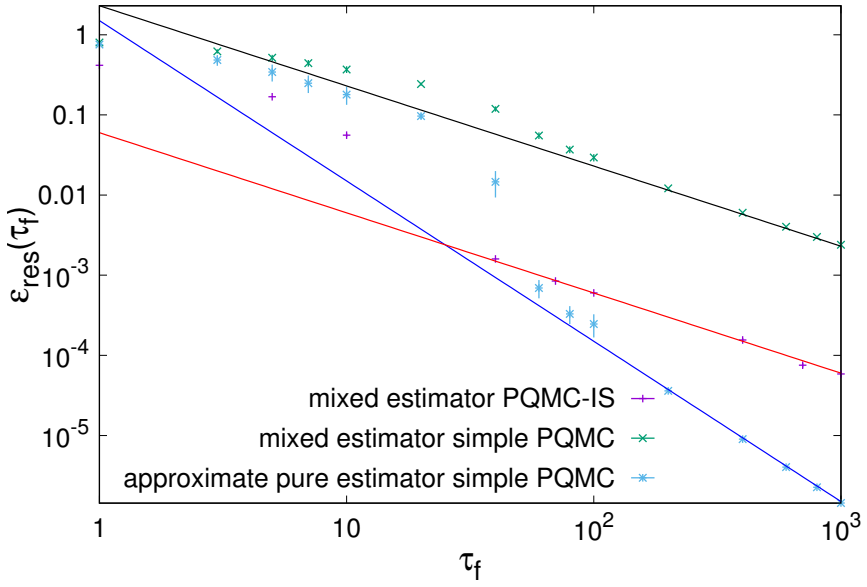



Figure 5.6: SQA of the quantum Ising model for both the simple PQMC and the PIMC with a time-dependent guiding wavefunction. Details of the plots are similar as in Fig. (5.1), however, the red line is a function that goes like $\tau_f{}^{-1}$. Different estimators are used to find the typical average of the residual energy as the data describe on the figure.

integral Monte Carlo that was very recently shown have the same scaling as the RT-SHE [60].

Notice that the Ceperley estimator could be used in this case, but we have not calculated it. It would indeed be interesting to check whether when using the Ceperley estimator, we are able to recover the scaling of the imaginary-time dynamics of the Schroedinger equation found in [24].

## 5.6 Conclusions

In this chapter, we have introduced an implementation of simulated quantum annealing using PQMC methods where importance sampling is performed using a time-dependent trial wavefunction. Comparisons with the simple PQMC were performed, and we showed that better estimates of the residual energies are found when using importance sampling even though the scaling $\epsilon_{res}$ remained the same when using a mixed estimator. Nevertheless, if this scaling is proven valid for infinite chains, it would imply a quadratic speed-up compared to both finite temperature path integral Monte Carlo simulations and the real-time dynamics of the Schroedinger equation. This result suggests that projective Monte Carlo methods may be a more efficient quantum-inspired optimization method compared with PIMC.

However, it would be interesting to see how PQMC simulations perform, when all parameters causing systematic errors are under control, in models that have elements of disorder. This is quite appealing given that in recent results on random Ising chains, finite temperature PIMC were shown to have a sampling crisis and were not able to recover the real-time dynamics of the Shrodinger equation.

In the simple PQMC case, we used the novel way of approximating pure estimators introduced in section 1.4. We found the scaling of the simple PQMC to be equivalent to the one of the imaginary-time dynamics of the Schroedinger equation. This suggests that the dynamics of PQMC simulations are indeed closer in spirit to the one of the imaginary-time dynamics of the Schroedinger equation.

# Summary and outlook

In this thesis we have developed algorithms based on the projective quantum Monte Carlo (PQMC) method, with a focus on applications to complex optimization problems. In particular, a new formalism to perform stochastic simulations of the modified imaginary-time Schroedinger equation taking into account a time-dependent trial wave function has been derived. Thus, extending previous formalisms that considered only time-independent trial wave-functions. A detailed description of the PQMC algorithm has been provided, shedding light on the origin of the possible sources of systematic errors which can affect the efficiency of the method. It introduces different types of propagators based on Trotter approximations or on a Taylor expansion, in order to stochastically simulate the imaginary-time Schroedinger equation aiming to reduce or eliminate some of those systematic errors. We have also proposed a novel way of estimating quantum mechanical average, in order to reduce or completely remove the bias that affects so-called mixed estimators.

In Chapter 2, using double-well models, we showed that simulated quantum annealing (SQA) performed with PQMC algorithm is a more effective solver compared to finite temperature path integral Monte Carlo (PIMC) algorithm even when it is used with Monte Carlo moves specifically tailored to the problem at hand. Though the PIMC method has been the QMC method of choice to perform comparisons with the outcomes of quantum annealers (QAs), it has been shown to display sampling problems in continuous models [59] and more recently in discrete based models [60]. We also shown that in models with more challenging potential energy landscapes, SQA driven by the PQMC algorithm has the potential to outperform simulated annealing (SA) even when the latter is performed with specifically tailored moves.

Inspired by recent studies [21, 22] in which the incoherent quantum tunneling time of a quantum annealer was compared to the tunneling time of PIMC simula-

tions, we studied the tunneling time of simple DMC simulations and compared it with the aforementioned methods. We showed that, differently from finite temperature PIMC algorithm, the DMC tunneling time kept the same inverse gap scaling, $1/\Delta$, both in the quantum Ising chain (QIC) and in a recently introduced model of frustrated spins dubbed Shamrock model [3]. The Shamrock model was introduced as an example where PIMC tunneling times scale worst than the incoherent tunneling time, indicating that quantum speed-up might be possible on QAs. Note that Google has recently announced to have plans of building a quantum annealer modeled after the Shamrock model therefore, our simulations will be a useful term of comparison with the ones performed on their chip.

In Chapter 4, using unrestricted Boltzmann machines (uRBMs) as new variational wave functions –with just $3$ parameters to optimize– we obtained ground state energies comparable to the recent results of the Restricted Boltzmann machines (RBMs) [25] on the QIC with the hidden-density unit $\alpha$ set to unity. $\alpha$ is the ratio between the number of hidden and visible units thus, corresponds to an extensive number of variational parameters. Then, we used optimized uRBMs as guiding wave functions for importance sampling in PQMC simulations with importance sampling to ameliorate the computational complexity of the simple DMC simulations shown in Chapter 3. We showed that the scaling of the computational cost goes from exponential in the system size to polynomial apparently even at the quantum critical point, provided that correlations among subsequent hidden spin variables are removed after an appropriate sampling protocol. It is worth emphasizing that the studies of the computational complexity of the QIC done in Chapter 3 are the the first to provide a clear numerical evidence of the exponential cost required to find the ground state energy of a stoquastic Hamiltonian at a fixed relative error.

Next, we addressed the problem of simulating the quantum annealing dynamics using PQMC algorithms with and without importance sampling on the quantum Ising chain. Using the mixed estimator, we found that the residual energy scales like the inverse of the total annealing time $\epsilon_{res} \backsim {\tau_f}^{-1}$, for finite size systems. Whereas, with the help of the new way of approximating pure estimators introduced in Chapter 1, we showed that for the simple PQMC method, the correct scaling of the imaginary-time dynamics of the QIC is recovered, namely $\epsilon_{res} \backsim {\tau_f}^{-2}$. This suggests that, if the same scaling remains for infinite system sizes, just using the mixed estimator, the PQMC method would render better results compared with the real-time dynamics of the Schrödinger equation for

which $\epsilon_{res} \backsim {\tau_f}^{-1/2}$. It is therefore primordial to check those assumptions in the future. Furthermore, the use of importance sampling was found to be crucial in those simulations given that in its absence, the bias due to the finite population of walkers was shown to render the simulations impracticable for large system sizes.

Therefore, a potential follow-up of this thesis would be to build efficient trial wave function which can be used in the context of annealing simulations to improve the performance of the PQMC method. Potential candidates are the uRMBs studied in this work. It would also be interesting to study how the efficiency in obtaining ground state properties depends on the number of deep layers in the uRBMs and, how that would affect tunneling times in PQMC simulations. We believe that this is a potential route to address more complex models such as frustrated spin glasses. Therefore, using such deep layer uRBMs in PQMC simulations could help in tackling difficult combinatorial problems in the quantum annealing framework.

A possible work in that direction would be to use SQA implemented via DMC to train deep Bolzmann neural networks. An attempt that has already been implemented using PIMC methods [102, 103].

# Bibliography


[1] Tadashi Kadowaki and Hidetoshi Nishimori. Quantum annealing in the transverse ising model. *Phys. Rev. E*, 58(5):5355, 1998.

[2] Edward Farhi, Jeffrey Goldstone, Sam Gutmann, Joshua Lapan, Andrew Lundgren, and Daniel Preda. A quantum adiabatic evolution algorithm applied to random instances of an np-complete problem. *Science*, 292(5516):472–475, 2001.

[3] Evgeny Andriyash and Mohammad H. Amin. Can quantum monte carlo simulate quantum annealing? *arXiv preprint arXiv:1703.09277*, 2017.

[4] Giuseppe E Santoro, Roman Martoňák, Erio Tosatti, and Roberto Car. Theory of quantum annealing of an ising spin glass. *Science*, 295(5564):2427–2430, 2002.

[5] Bettina Heim, Troels F Rønnow, Sergei V Isakov, and Matthias Troyer. Quantum versus classical annealing of ising spin glasses. *Science*, 348(6231):215–217, 2015.

[6] James B Anderson. A random-walk simulation of the schrödinger equation: H+ 3. *J. Chem. Phys.*, 63(4):1499–1503, 1975.

[7] Peter J Reynolds, David M Ceperley, Berni J Alder, and William A Lester Jr. Fixed-node quantum monte carlo for moleculesa) b). *J. Chem. Phys.*, 77(11):5593–5603, 1982.

[8] Malvin H Kalos and Paula A Whitlock. *Monte carlo methods*. John Wiley & Sons, 2008.

[9] WMC Foulkes, L Mitas, RJ Needs, and G Rajagopal. Quantum monte carlo simulations of solids. *Rev. Mod. Phys.*, 73(1):33, 2001.

[10] Jos Thijssen. *Computational physics*. Cambridge University Press, 2007.

[11] AB Finnila, MA Gomez, C Sebenik, C Stenson, and JD Doll. Quantum annealing: a new method for minimizing multidimensional functions. *Chem. Phys. Lett.*, 219(5):343–348, 1994.

[12] D. A. Battaglia, G. E. Santoro, and E. Tosatti. Optimization by quantum annealing: Lessons from hard satisfiability problems. *Phys. Rev. E*, 71(6):066707, 2005.

[13] Roman Martoňák, Giuseppe E Santoro, and Erio Tosatti. Quantum annealing by the path-integral monte carlo method: The two-dimensional random ising model. *Phys. Rev. B*, 66(9):094203, 2002.

[14] Lorenzo Stella and Giuseppe E Santoro. Quantum annealing of an ising spin-glass by green's function monte carlo. *Phys. Rev. E*, 75(3):036703, 2007.

[15] Roman Martoňák, Giuseppe E. Santoro, and Erio Tosatti. Quantum annealing of the traveling-salesman problem. *Phys. Rev. E*, 70:057701, Nov 2004.

[16] Demian A. Battaglia, Giuseppe E. Santoro, and Erio Tosatti. Optimization by quantum annealing: Lessons from hard satisfiability problems. *Phys. Rev. E*, 71:066707, Jun 2005.

[17] Cheng-Wei Liu, Anatoli Polkovnikov, and Anders W. Sandvik. Quantum versus classical annealing: Insights from scaling theory and results for spin glasses on 3-regular graphs. *Phys. Rev. Lett.*, 114:147203, Apr 2015.

[18] Sergio Boixo, Vadim N Smelyanskiy, Alireza Shabani, Sergei V Isakov, Mark Dykman, Vasil S Denchev, Mohammad Amin, Anatoly Smirnov, Masoud Mohseni, and Hartmut Neven. Computational role of collective tunneling in a quantum annealer. *arXiv preprint arXiv:1411.4036*, 2014.

[19] Vasil S. Denchev, Sergio Boixo, Sergei V. Isakov, Nan Ding, Ryan Babbush, Vadim Smelyanskiy, John Martinis, and Hartmut Neven. What is the computational value of finite-range tunneling? *Phys. Rev. X*, 6:031015, Aug 2016.

[20] Sergio Boixo, Troels F Rønnow, Sergei V Isakov, Zhihui Wang, David Wecker, Daniel A Lidar, John M Martinis, and Matthias Troyer. Evidence

for quantum annealing with more than one hundred qubits. *Nat. Phys.*, 10(3):218–224, 2014.

[21] Sergei V. Isakov, Guglielmo Mazzola, Vadim N. Smelyanskiy, Zhang Jiang, Sergio Boixo, Hartmut Neven, and Matthias Troyer. Understanding quantum tunneling through quantum monte carlo simulations. *Phys. Rev. Lett.*, 117:180402, Oct 2016.

[22] Guglielmo Mazzola, Vadim N. Smelyanskiy, and Matthias Troyer. Quantum monte carlo tunneling from quantum chemistry to quantum annealing. *Phys. Rev. B*, 96:134305, Oct 2017.

[23] Lorenzo Stella, Giuseppe E Santoro, and Erio Tosatti. Optimization by quantum annealing: lessons from simple cases. *Phys. Rev. B*, 72(1):014303, 2005.

[24] Tommaso Zanca and Giuseppe E. Santoro. Quantum annealing speedup over simulated annealing on random ising chains. *Phys. Rev. B*, 93:224431, Jun 2016.

[25] Giuseppe Carleo and Matthias Troyer. Solving the quantum many-body problem with artificial neural networks. *Science*, 355(6325):602–606, 2017.

[26] D. Blume, M. Lewerenz, P. Niyaz, and K. B. Whaley. Excited states by quantum monte carlo methods: Imaginary time evolution with projection operators. *Phys. Rev. E*, 55:3664–3675, Mar 1997.

[27] Federico Becca and Sandro Sorella. *Quantum Monte Carlo Approaches for Correlated Systems*. Cambridge University Press, 2017.

[28] Norbert Nemec. Diffusion monte carlo: Exponential scaling of computational cost for large systems. *Phys. Rev. B*, 81:035119, Jan 2010.

[29] Massimo Boninsegni and Saverio Moroni. Population size bias in diffusion monte carlo. *Phys. Rev. E*, 86(5):056712, 2012.

[30] Lode Pollet, Nikolay V. Prokof'ev, and Boris V. Svistunov. Stochastic lists: Sampling multivariable functions with population methods. *Phys. Rev. B*, 98:085102, Aug 2018.

[31] Satoshi Morita and Hidetoshi Nishimori. Convergence theorems for quantum annealing. *J. Phys. A: Math. Gen.*, 39(45):13903, 2006.

[32] Satoshi Morita and Hidetoshi Nishimori. Mathematical foundation of quantum annealing. *J. Math. Phys.*, 49(12):125210, 2008.

[33] Michael Jarret, Stephen P. Jordan, and Brad Lackey. Adiabatic optimization versus diffusion monte carlo methods. *Phys. Rev. A*, 94:042318, Oct 2016.

[34] J. Boronat. Monte carlo simulations at zero temperature: helium in one, two, and three dimensions. In Eckhard Krotscheck and Jesus Navarro, editors, *Microscopic Approaches to Quantum Liquids in Confined Geometries*, chapter 2, pages 21–90. World Scientific, Singapore, 2002.

[35] Harald A. Forbert and Siu A. Chin. Fourth-order diffusion monte carlo algorithms for solving quantum many-body problems. *Phys. Rev. B*, 63:144518, Mar 2001.

[36] WMC Foulkes, L Mitas, RJ Needs, and G Rajagopal. Quantum monte carlo simulations of solids. *Rev. Mod. Phys.*, 73(1):33, 2001.

[37] Sandro Sorella and Luca Capriotti. Green function monte carlo with stochastic reconfiguration: An effective remedy for the sign problem. *Phys. Rev. B*, 61(4):2599, 2000.

[38] 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, 2005.

[39] Sandro Sorella and Luca Capriotti. Green function monte carlo with stochastic reconfiguration: An effective remedy for the sign problem. *Phys. Rev. B*, 61:2599–2612, Jan 2000.

[40] Pierre Pfeuty. The one-dimensional ising model with a transverse field. *Ann. Phys. (N. Y.)*, 57(1):79–90, 1970.

[41] Nandini Trivedi and D. M. Ceperley. Ground-state correlations of quantum antiferromagnets: A green-function monte carlo study. *Phys. Rev. B*, 41:4552–4569, Mar 1990.

[42] J. Casulleras and J. Boronat. Unbiased estimators in quantum monte carlo methods: Application to liquid $^4$He. *Phys. Rev. B*, 52:3654–3661, Aug 1995.

[43] MW Johnson, MHS Amin, S Gildert, T Lanting, F Hamze, N Dickson, R Harris, AJ Berkley, J Johansson, P Bunyk, et al. Quantum annealing with manufactured spins. *Nature*, 473(7346):194–198, 2011.

[44] Sergio Boixo, Tameem Albash, Federico M Spedalieri, Nicholas Chancellor, and Daniel A Lidar. Experimental signature of programmable quantum annealing. *Nat. Commun.*, 4, 2013.

[45] Troels F Rønnow, Zhihui Wang, Joshua Job, Sergio Boixo, Sergei V Isakov, David Wecker, John M Martinis, Daniel A Lidar, and Matthias Troyer. Defining and detecting quantum speedup. *Science*, 345(6195):420–424, 2014.

[46] Sergei V Isakov, Ilia N Zintchenko, Troels F Rønnow, and Matthias Troyer. Optimised simulated annealing for ising spin glasses. *Comput. Phys. Commun.*, 192:265–271, 2015.

[47] John A Smolin and Graeme Smith. Classical signature of quantum annealing. *arXiv preprint arXiv:1305.4904*, 2013.

[48] Tameem Albash, Troels F Rønnow, Matthias Troyer, and Daniel A Lidar. Reexamining classical and quantum models for the d-wave one processor. *Eur. Phys. J. Spec. Top.*, 224(1):111–129, 2015.

[49] Lei Wang, Troels F Rønnow, Sergio Boixo, Sergei V Isakov, Zhihui Wang, David Wecker, Daniel A Lidar, John M Martinis, and Matthias Troyer. Comment on:" classical signature of quantum annealing". *arXiv preprint arXiv:1305.5837*, 2013.

[50] Seung Woo Shin, Graeme Smith, John A Smolin, and Umesh Vazirani. How" quantum" is the d-wave machine? *arXiv preprint arXiv:1401.7087*, 2014.

[51] T Lanting, AJ Przybysz, A Yu Smirnov, FM Spedalieri, MH Amin, AJ Berkley, R Harris, F Altomare, S Boixo, P Bunyk, et al. Entanglement in a quantum annealing processor. *Phys. Rev. X*, 4(2):021041, 2014.

[52] Arnab Das and Bikas K Chakrabarti. Colloquium: Quantum annealing and analog quantum computation. *Rev. Mod. Phys.*, 80(3):1061, 2008.

[53] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671–680, 1983.

[54] J Brooke, D Bitko, G Aeppli, et al. Quantum annealing of a disordered magnet. *Science*, 284(5415):779–781, 1999.

[55] Helmut G Katzgraber, Firas Hamze, and Ruben S Andrist. Glassy chimeras could be blind to quantum speedup: Designing better benchmarks for quantum annealing machines. *Phys. Rev. X*, 4(2):021008, 2014.

[56] Cheng-Wei Liu, Anatoli Polkovnikov, and Anders W Sandvik. Quantum versus classical annealing: Insights from scaling theory and results for spin glasses on 3-regular graphs. *Phys. Rev. Lett.*, 114(14):147203, 2015.

[57] James B Anderson. A random-walk simulation of the schrödinger equation: H+ 3. *J. Chem. Phys.*, 63(4):1499–1503, 1975.

[58] Stuart Geman and Donald Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intell.*, (6):721–741, 1984.

[59] Lorenzo Stella, Giuseppe E Santoro, and Erio Tosatti. Monte carlo studies of quantum and classical annealing on a double well. *Phys. Rev. B*, 73(14):144302, 2006.

[60] Glen B. Mbeng, Lorenzo Privitera, Luca Arceci, and Giuseppe E. Santoro. On the dynamics of simulated quantum annealing in random ising chains. *arXiv preprint arXiv:1809.02124*, 2018.

[61] Claudia De Grandi, Anatoli Polkovnikov, and AW Sandvik. Universal nonequilibrium quantum dynamics in imaginary time. *Phys. Rev. B*, 84(22):224303, 2011.

[62] Nicolaas Godfried Van Kampen. *Stochastic processes in physics and chemistry*, volume 1. Elsevier, 1992.

[63] David A Huse and Daniel S Fisher. Residual energies after slow cooling of disordered systems. *Phys. Rev. Lett.*, 57(17):2203, 1986.

[64] Yoshiyuki Kabashima and Shigeru Shinomoto. Asymptotic dependence of the residual energyon annealing time. *J. Phys. Soc. Jpn.*, 60(12):3993–3996, 1991.

[65] Shigeru Shinomoto and Yoshiyuki Kabashima. Finite time scaling of energy in simulated annealing. *J. Phys. A: Math. Gen.*, 24(3):L141, 1991.

[66] Elihu Abrahams, PW Anderson, DC Licciardello, and TV Ramakrishnan. Scaling theory of localization: Absence of quantum diffusion in two dimensions. *Phys. Rev. Lett.*, 42(10):673–676, 1979.

[67] Serge Aubry and Gilles André. Analyticity breaking and anderson localization in incommensurate lattices. *Ann. Israel Phys. Soc*, 3(133):18, 1980.

[68] L.T. Wille. Minimum-energy configurations of atomic clusters: new results obtained by simulated annealing. *Chem. Phys. Lett.*, 133(5):405 – 410, 1987.

[69] E. M. Inack, G. Giudici, T. Parolini, G. Santoro, and S. Pilati. Understanding quantum tunneling using diffusion monte carlo simulations. *Phys. Rev. A*, 97:032307, Mar 2018.

[70] Alexander W. Glaetzle, Rick M. W. van Bijnen, Peter Zoller, and Wolfgang Lechner. A coherent quantum annealer with rydberg atoms. *Nat. Commun.*, 8(10):15813, 2017.

[71] A. B. Finnila, M. A. Gomez, C. Sebenik, C. Stenson, and J. D. Doll. Quantum annealing: A new method for minimizing multidimensional functions. *Chem. Phys. Lett.*, 219:343, 1994.

[72] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. Quantum adiabatic evolution algorithms versus simulated annealing. *arXiv preprint quant-ph/0201031*, 2002.

[73] Elizabeth Crosson and Aram W Harrow. Simulated quantum annealing can be exponentially faster than classical simulated annealing. In *ANN. IEEE. SYMP. FOUND.*, pages 714–723. IEEE, 2016.

[74] Lucas T Brady and Wim van Dam. Quantum monte carlo simulations of tunneling in quantum adiabatic optimization. *Phys. Rev. A*, 93(3):032304, 2016.

[75] Ulrich Weiss, Hermann Grabert, Peter Hänggi, and Peter Riseborough. Incoherent tunneling in a double well. *Phys. Rev. B*, 35(18):9535, 1987.

[76] Edward Farhi, Jeffrey Goldstone, Sam Gutmann, and Michael Sipser. Quantum computation by adiabatic evolution. *arXiv preprint quant-ph/0001106*, 2000.

[77] Zhang Jiang, Vadim N. Smelyanskiy, Sergei V. Isakov, Sergio Boixo, Guglielmo Mazzola, Matthias Troyer, and Hartmut Neven. Scaling analysis and instantons for thermally assisted tunneling and quantum monte carlo simulations. *Phys. Rev. A*, 95:012322, Jan 2017.

[78] Sergey Bravyi, David P. DiVincenzo, Roberto I. Oliveira, and Barbara M. Terhal. The complexity of stoquastic local hamiltonian problems. *Quantum Inf. Comput.*, 8(5):0361–0385, 2008.

[79] Hidetoshi Nishimori and Kabuki Takada. Exponential enhancement of the efficiency of quantum annealing by non-stoquastic hamiltonians. *Frontiers in ICT*, 4:2, 2017.

[80] Layla Hormozi, Ethan W. Brown, Giuseppe Carleo, and Matthias Troyer. Nonstoquastic hamiltonians and quantum annealing of an ising spin glass. *Phys. Rev. B*, 95:184416, May 2017.

[81] Matthew B Hastings. Obstructions to classically simulating the quantum adiabatic algorithm. *Quantum Inf. & Comput.*, 13(11-12):1038–1076, 2013.

[82] Malvin H. Kalos, Dominique Levesque, and Loup Verlet. Helium at zero temperature with hard-sphere and other forces. *Phys. Rev. A*, 9:2178, 1974.

[83] A. Sarsa, K. E. Schmidt, and W. R. Magro. A path integral ground state method. *J. Chem. Phys.*, 113(4):1366–1371, 2000.

[84] Javier E. Cuervo, Pierre-Nicholas Roy, and Massimo Boninsegni. Path integral ground state with a fourth-order propagator: Application to condensed helium. *J. Chem. Phys.*, 122(11):114504, 2005.

[85] Sergey Bravyi. Monte carlo simulation of stoquastic hamiltonians. *Quantum Inf. Comput.*, 15(13-14):1122–1140, 2015.

[86] S Pilati and M Troyer. Bosonic superfluid-insulator transition in continuous space. *Phys. Rev. Lett.*, 108:155301, 2012.

[87] Sergey Bravyi and David Gosset. Polynomial-time classical simulation of quantum ferromagnets. *Phys. Rev. Lett.*, 119(10):100503, 2017.

[88] Jacob Bringewatt, William Dorland, Stephen P Jordan, and Alan Mink. Diffusion monte carlo versus adiabatic computation for local hamiltonians. *arXiv preprint arXiv:1709.03971*, 2017.

[89] Hiroki Saito. Solving the bose–hubbard model with machine learning. *J. Phys. Soc. Jpn.*, 86(9):093001, 2017.

[90] 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, 2018.

[91] 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, 2007.

[92] Silvio Vitiello, Karl Runge, and Malvin H Kalos. Variational calculations for solid and liquid he 4 with a" shadow" wave function. *Phys. Rev. Lett.*, 60(19):1970, 1988.

[93] L Reatto and GL Masserini. Shadow wave function for many-boson systems. *Phys. Rev. B*, 38(7):4516, 1988.

[94] Giuseppe Carleo, Yusuke Nomura, and Masatoshi Imada. Constructing exact representations of quantum many-body systems with deep neural networks. *arXiv preprint arXiv:1802.09558*, 2018.

[95] Nahuel Freitas, Giovanna Morigi, and Vedran Dunjko. Neural network operations and susuki-trotter evolution of neural network states. *arXiv preprint arXiv:1803.02118*, 2018.

[96] Elina Robeva and Anna Seigal. Duality of graphical models and tensor networks. *Inf. Inference*, 2017.

[97] Ivan Glasser, Nicola Pancotti, and J Ignacio Cirac. Supervised learning with generalized tensor networks. *arXiv preprint arXiv:1806.05964*, 2018.

[98] A. W. Sandvik and G. Vidal. Variational quantum monte carlo simulations with tensor-network states. *Phys. Rev. Lett.*, 99:220602, Nov 2007.

[99] Xun Gao and Lu-Ming Duan. Efficient representation of quantum many-body states with deep neural networks. *Nat. Commun.*, 8(1):662, 2017.

[100] Silvio A. Vitiello and Paula A. Whitlock. Green's-function monte carlo algorithm for the solution of the schrödinger equation with the shadow wave function. *Phys. Rev. B*, 44:7373–7377, Oct 1991.

[101] Anatoli Polkovnikov, Krishnendu Sengupta, Alessandro Silva, and Mukund Vengalattore. Colloquium: Nonequilibrium dynamics of closed interacting quantum systems. *Rev. Mod. Phys.*, 83:863–883, Aug 2011.

[102] Carlo Baldassi and Riccardo Zecchina. Efficiency of quantum vs. classical annealing in nonconvex learning problems. *Proceedings of the National Academy of Sciences*, 115(7):1457–1462, 2018.

[103] David Crawford, Anna Levit, Navid Ghadermarzy, Jaspreet S. Oberoi, and Pooya Ronagh. Reinforcement learning using quantum boltzmann machines. *arXiv preprint arXiv:1612.05695*, 2016.