UNIVERSITÀ DEGLI STUDI DI MILANO

Department of Physics "Aldo Pontremoli"

Ph.D. School in Physics, Astrophysics and Applied Physics
Cycle XXXVI

Ab-initio dynamics in strongly interacting many-body systems

Disciplinary Scientific Sector FIS/03

Ph.D. Thesis of:
Christian APOSTOLI

Director of the School: Prof. Aniello MENNELLA

Supervisor: Prof. Davide E. GALLI

Co-Supervisor: Dr. Gianluca BERTAINA

A.Y. 2022-2023

Commission of the final examination:

External Referee:
Prof. Francesco PEDERIVA

External Referee:
Prof. Sebastiano PILATI

Internal Member:
Prof. Davide E. GALLI

Final examination:

29 July 2024

Università degli Studi di Milano, Dipartimento di Fisica, Milano, Italy

To my parents Fabio and Cinzia, for their tangible love and unfailing support.

MIUR subjects:

FIS/03

Contents

List of Figures

Introduction

The understanding of many-body strongly interacting quantum systems is the focal point of a long-standing theoretical and experimental pursuit in modern Physics. The intricate quantum correlations arising in such systems result in highly nontrivial macroscopic properties. In condensed matter, strong quantum correlations are behind phenomena such as quantum phase transitions, collective excitations, topological quantum order, and high-temperature superconductivity. However, a full theoretical understanding of the mechanisms that lie behind these phenomena still defies our efforts, even when the microscopic description is characterized by deceptively simple interactions, like those at play in the paradigmatic Hubbard model.

Among these challenges, a special place is reserved to the description of the dynamics of strongly correlated systems, which is key to understanding out-of-equilibrium phenomena and the evolution of entanglement. The latter, besides its interest from the point of view of fundamental research, is also of great relevance to the rapidly developing field of quantum technologies, where a major challenge is the control and preservation of coherence and entanglement. A mature understanding of these phenomena is of paramount importance to the development of viable quantum computation and communication protocols that can rival or outperform their classical counterparts.

On the experimental side, the progress in the manipulation of ultracold atomic gases has made it possible to build quantum simulators in which interactions can be engineered to an unprecedented level of control. This allowed the observation of nonequilibrium phenomena like many-body localization [1, 2, 3] or dynamical quantum phase transitions [4, 5, 6].

On the theoretical and computational side, we need to tackle the exponential growth in the Hilbert space dimension as the number of degrees of freedom in a system increases. This makes it impossible to fully represent the wave function of a large many-body system on a classical computer, but effective alternative strategies have been devised. Many techniques exist to simulate ground states, and today the description of the ground-state properties of a large class of systems, mainly bosonic, can be calculated without approximations by means of quantum Monte Carlo (QMC) methods [7]. They allow the exact simulation of the imaginary-time dynamics, from which the dynamical correlation functions can be obtained by an inverse Laplace transform. However, this is an ill-posed inverse problem that has no unique solution, and needs to be tackled with analytic continuation techniques.

A possibly better way to obtain the dynamical properties would be to directly simulate the real-time dynamics. But this is a challenge, for which only particular solutions involving approximations are available. An approximate but very flexible approach,

that was originally devised to tackle ground states, is the variational Monte Carlo (VMC) method: it represents the quantum state by an Ansatz wave function that depends on a set of parameters, which are optimized in order to minimize the energy, in accordance with the variational principle of quantum mechanics. In recent years, G. Carleo extended this idea to the domain of dynamics simulations with the time-dependent variational Monte Carlo (t-VMC) method [8, 9], which provides a set of equations of motion for the parameters. This technique has been successfully applied especially to the class of variational Ansatzë known as the neural-network quantum states (NNQS), a wave function form inspired by the machine learning field. These Ansatzë have great descriptive ability, but usually use an enormous number of variational parameters, and their architecture is not directly grounded in the Physics of the system they try to describe.

In this Thesis, we extend the t-VMC methodologies to the class of shadow wave functions (SWF) [10, 11], that incorporate many-body correlations by introducing a set of auxiliary "shadow" variables. SWFs can be interpreted as particular neural networks in which the shadow variables represent a layer of neurons having a specific connectivity. What makes the SWFs attractive is their relation to path integration [12], which makes them an Ansatz with a solid ground in the Physics of the system. We introduce a specific SWF Ansatz for systems of spins on a lattice, the Baeriswyl-Shadow wave function (BSWF), whose structure takes a cue from the Physics of the transverse-field Ising model. Remarkably, the BSWF is very concise, using only three parameters.

We develop from scratch a C++ code that allows the t-VMC simulation of spin systems using various Ansatzë, including the BSWF and a particular type of NNQS, the restricted Boltzmann machine (RBM). We use it to investigate the strengths and the weaknesses of the BSWF Ansatz. We explore how it can be used in conjunction with other Ansatzë to improve their variational approximations. And we also introduce some variations of the BSWF and test their capabilities.

Finally, we leave the framework of t-VMC, but still stay in the context of dynamical properties. Using a radically different approach, we compute the dynamical structure factor of a two-dimensional dilute system of attractive fermions. Our motivation for investigating this system lies in its collective excitations: the spontaneous breaking of the $U(1)$ symmetry in its superfluid ground state results in two collective modes: the Nambu-Goldstone mode, corresponding to oscillations of the phase of the superfluid order parameter, and the Higgs mode, corresponding to oscillations of the amplitude of the order parameter. The amplitude mode is particularly elusive on the experimental side, but it has been suggested that its spin-spin dynamical structure factor may display a peak corresponding to the Higgs mode [13]. To investigate this possibility, we compute these dynamical correlations by leveraging generalized random phase approximation (GRPA) [14, 15, 16, 17, 13] and the unbiased auxiliary-field Monte Carlo method (AFQMC) [18, 19, 20], the latter coupled with a recent analytic continuation technique, the differential evolution for analytic continuation (DEAC) [21, 22].

Thesis overview

The present Thesis consists of five chapters. The first four chapters are devoted to the t-VMC method and its extension to SWF Ansatzë, and in particular the BSWF variational Ansatz that we introduce. The last one focuses the calculation of the dynamical properties of a 2D dilute gas of attractive fermions, relying on random phase approximation and unbiased QMC imaginary-time simulations.

Chapter 1: The time-dependent variational Monte Carlo method (t-VMC): We introduce G. Carleo's t-VMC method and present the derivation of its equations of motion, for both real and imaginary time (corresponding to the dynamics and the ground state optimization, respectively). We examine the geometrical meaning of the quantities involved in these equations, and the possibility of obtaining an *a posteriori* upper bound on the distance between the exact evolution and the variational one. We discuss some details of the integration of the t-VMC equations of motion, in particular the problem of regularizing the dynamics to avoid instabilities, and the possibility of using an adaptive time-step scheme in conjunction with an integration method for ordinary differential equations. We present two successful variational Ansatzë: the spin-Jastrow wave function and the restricted Boltzmann machine (RBM). Finally, we propose a method to compute dynamical structure factors using the t-VMC algorithm that we independently developed during this Thesis work, but is substantially equivalent to the one published by the Authors of Ref. [23].

Chapter 2: Shadow wave functions (SWF): We introduce the "shadow" form of variational wave functions and explain its connection with the path-integral formalism. We discuss how expectation values can be computed by sampling both the real and the "shadow" variables, introducing the reweighting technique that is needed in when the shadow-shadow or real-shadow correlations in the SWF are complex-valued. This is usually not the case for SWF approximations of the ground state, but it is, in general, for the real-time evolution. Then we present our extension of the t-VMC method to SWF Ansatzë.

Chapter 3: The Baeriswyl-Shadow wave function (BSWF): We exploit the connection between the SWFs and path integration to introduce the BSWF Ansatz, a shadow-type wave function for spin systems, that is very concise and is grounded in the Physics of the transverse-field Ising model. We also discuss the use of cluster moves in the Monte Carlo sampling of the real and shadow spin configurations for the BSWF.

Chapter 4: Results for the BSWF Ansatz: We present and discuss the results that we obtained for the BSWF Ansatz using the C++ code we developed to perform t-VMC simulations. We start from imaginary-time simulations and compare the BSWF approximation of the transverse-field Ising model ground state in 1D with the results obtained by RBM. We also compare the results of these Ansatzë for the variational dynamics of a quantum quench. We consider both a translation-invariant quantum quench and a non-translation-invariant one. Finally, we also present our results for two variations of the BSWF Ansatz, and for the imaginary-time propagation in the 2D case.

Chapter 5: Dynamical correlations of an attractive two-dimensional Fermi gas We present the model we study and the techniques used: the auxiliary-field quantum Monte Carlo (AFQMC), the differential evolution for analytic continuation (DEAC), and the generalized random phase approximation (GRPA). We compare the GRPA and AFQMC results for the dynamical structure factor in both the density and the spin channel, finding a peak in the spin channel that is compatible with the one observed in Ref. [13]. We complement the analysis with a GRPA study of the dynamical correlations of the order parameter, from which we conclude that this peak cannot be identified with the celebrated Higgs mode, as originally suggested in Ref. [13], and propose a different origin, rooted in kinematics.

CHAPTER 1

The time-dependent variational Monte Carlo method (t-VMC)

The time-dependent variational Monte Carlo method (t-VMC) is the main technique we employ and expand in this Thesis work. It addresses the exponentially-exploding complexity of the many-body Hilbert space by using a parametrized variational wave function. This way, the problem is reduced to finding the set of parameters that best represent the wave function. The t-VMC method prescribes a set of equations of motion for the parameters, so that the problem of computing the quantum dynamics is reduced to the integration of these equations for the evolution of the variational parameters.

In this Chapter we introduce the t-VMC method and discuss some of its features. In Sec. 1.1 we present the derivation of its equations of motion, discuss the geometrical meaning of the quantities involved, and explain how it is possible to compute an *a posteriori* error bound on the distance between the variational dynamics and the exact one. In Sec. 1.2 we discuss some details of the integration of the t-VMC equations of motion, in particular the regularization of the dynamics to avoid instabilities, and the possibility of using an adaptive time-step scheme. In Sec. 1.3 two successful variational Ansatzë for spin systems are introduced: the spin-Jastrow wave function and the restricted Boltzmann machine. Finally, in Sec. 1.4 we propose a technique to compute the dynamical structure factor of a system using the t-VMC method, that we developed independently but is substantially equivalent to a method published in 2023 by T. Mendes-Santos, M. Schmitt, and M. Heyl [23].

1.1 t-VMC basics

1.1.1 Equations of motion by minimizing the Hilbert-space distance

We represent the state of a quantum system by means of a variational wave function $\psi(\mathbf{x};\boldsymbol{\alpha})$, where $\mathbf{x}$ is the vector of the coordinates of the system (particle positions, or spin configurations, etc.) and $\boldsymbol{\alpha}$ is the vector of variational parameters, $\alpha_k \in \mathbb{C}$. It is assumed that $\psi(\mathbf{x};\boldsymbol{\alpha})$ is an analytic function of all the parameters. We denote by $|\psi(\boldsymbol{\alpha})\rangle$ the ket corresponding to wave function $\psi(\mathbf{x};\boldsymbol{\alpha})$ in the Hilbert space.

We want to use the variational wave function to approximate the evolution of the state under a certain Hamiltonian operator $\hat{H}$. If at a certain time t the state of the quantum system is represented by $|\psi(\boldsymbol{\alpha}(t))\rangle$, then the Schrödinger equation determines the *exact* time derivative of the state in that instant as

$$\left(\frac{d\,|\psi\rangle}{dt}\right)_{\text{ex}} = -\frac{i}{\hbar}\hat{H}\,|\psi(\boldsymbol{\alpha}(t))\rangle\ . \tag{1.1}$$

On the other hand, the time derivative of the state linked to the *variational* evolution of the parameters $\boldsymbol{\alpha}(t)$ is given by

$$\left(\frac{d\,|\psi\rangle}{dt}\right)_{\text{var}} = \sum_k \frac{\partial\,|\psi(\boldsymbol{\alpha}(t))\rangle}{\partial\alpha_k(t)}\dot{\alpha}_k(t)\,, \tag{1.2}$$

where $\dot{\alpha}_k(t)$ is the time derivative of parameter α_k at time t. Following Refs. [8, 24], our task is to determine the values of the time derivatives $\dot{\boldsymbol{\alpha}}(t)$ that minimize the normalized Hilbert-space distance

$$\Delta^2(t) = \frac{1}{\langle\psi(\boldsymbol{\alpha})|\psi(\boldsymbol{\alpha})\rangle}\left\|\left(\frac{d\,|\psi\rangle}{dt}\right)_{\text{var}} - \left(\frac{d\,|\psi\rangle}{dt}\right)_{\text{ex}}\right\|^2, \tag{1.3}$$

that is, we want to minimize the rate of departure between the variational state and the exact state. This way, the resulting variational evolution will be the best approximation of the exact one, at least locally in time. We point out that the condition of minimum $\Delta^2(t)$ is equivalent to the requirement of the McLachlan variational principle [25].

Plugging Eqs. (1.1) and (1.2) into Eq. (1.3), we obtain the expression

$$\Delta^2(t) = \frac{1}{\langle\psi(\boldsymbol{\alpha})|\psi(\boldsymbol{\alpha})\rangle}\left\|\sum_k \frac{\partial\,|\psi(\boldsymbol{\alpha})\rangle}{\partial\alpha_k}\dot{\alpha}_k + \frac{i}{\hbar}\hat{H}\,|\psi(\boldsymbol{\alpha})\rangle\right\|^2, \tag{1.4}$$

which we rewrite, by inserting an identity $\mathbb{1} = \sum_{\mathbf{x}} |\mathbf{x}\rangle\langle\mathbf{x}|$, as

$$\Delta^2(t) = \frac{1}{\langle\psi(\boldsymbol{\alpha})|\psi(\boldsymbol{\alpha})\rangle}\sum_{\mathbf{x}}\left|\sum_k \frac{\partial\,\langle\mathbf{x}|\psi(\boldsymbol{\alpha})\rangle}{\partial\alpha_k}\dot{\alpha}_k + \frac{i}{\hbar}\left\langle\mathbf{x}\middle|\hat{H}\middle|\psi(\boldsymbol{\alpha})\right\rangle\right|^2. \tag{1.5}$$

The above expression can be manipulated in order to factor out the squared modulus of the probability amplitude $|\langle\mathbf{x}|\psi(\boldsymbol{\alpha})\rangle|^2$:

$$\Delta^2(t) = \sum_{\mathbf{x}}\frac{|\langle\mathbf{x}|\psi(\boldsymbol{\alpha})\rangle|^2}{\langle\psi(\boldsymbol{\alpha})|\psi(\boldsymbol{\alpha})\rangle}\left|\sum_k \frac{1}{\langle\mathbf{x}|\psi(\boldsymbol{\alpha})\rangle}\frac{\partial\,\langle\mathbf{x}|\psi(\boldsymbol{\alpha})\rangle}{\partial\alpha_k}\dot{\alpha}_k + \frac{i}{\hbar}\frac{\left\langle\mathbf{x}\middle|\hat{H}\middle|\psi(\boldsymbol{\alpha})\right\rangle}{\langle\mathbf{x}|\psi(\boldsymbol{\alpha})\rangle}\right|^2. \tag{1.6}$$

The above expression suggests the definition of the *local energy*

$$\mathcal{E}(\mathbf{x}) = \frac{\left\langle\mathbf{x}\middle|\hat{H}\middle|\psi(\boldsymbol{\alpha})\right\rangle}{\langle\mathbf{x}|\psi(\boldsymbol{\alpha})\rangle}\,, \tag{1.7}$$

whose expectation value on the quantum state is the energy E, i.e., the expectation value of the Hamiltonian operator $\hat{H}$. We also define the *local operators*

$$\mathcal{O}_k(\mathbf{x};\boldsymbol{\alpha}) = \frac{1}{\langle\mathbf{x}|\psi(\boldsymbol{\alpha})\rangle}\frac{\partial\,\langle\mathbf{x}|\psi(\boldsymbol{\alpha})\rangle}{\partial\alpha_k} = \frac{\partial}{\partial\alpha_k}\ln\langle\mathbf{x}|\psi(\boldsymbol{\alpha})\rangle\,. \tag{1.8}$$

We also introduce the angle-brackets notation $\langle\cdot\rangle$ for the expectation values over the probability distribution given by the squared modulus of the normalized probability

amplitude $\langle \mathbf{x}|\psi(\boldsymbol{\alpha})\rangle$[1]. Using these definitions and notation, we rewrite Eq. (1.6) as

$$\Delta^2(t) = \left\langle \left| \sum_k \mathcal{O}_k(\mathbf{x};\boldsymbol{\alpha})\dot{\alpha}_k + \frac{i}{\hbar}\mathcal{E}(\mathbf{x};\boldsymbol{\alpha}) \right|^2 \right\rangle \tag{1.9}$$

From Eq. (1.9) we can see that $\Delta^2(t)$ depends on the time derivatives of the parameters $\dot{\alpha}_k$. We want to find the values of $\dot{\alpha}_k$ which minimize $\Delta^2(t)$. So we compute the derivatives of $\Delta^2(t)$ with respect to the real parts $\dot{\alpha}_k^{(R)}$ and the imaginary parts $\dot{\alpha}_k^{(I)}$ of the time derivatives of the parameters, yielding

$$\begin{aligned} \frac{\partial \Delta^2}{\partial \dot{\alpha}_k^{(R)}} &= \left\langle \sum_{k'} \mathcal{O}_k^*(\mathbf{x};\boldsymbol{\alpha})\mathcal{O}_{k'}(\mathbf{x};\boldsymbol{\alpha})\dot{\alpha}_{k'} + \frac{i}{\hbar}\mathcal{O}_k^*(\mathbf{x};\boldsymbol{\alpha})\mathcal{E}(\mathbf{x};\boldsymbol{\alpha}) \right\rangle + \text{c.c.} \\ &= 2\Re \left\langle \sum_{k'} \mathcal{O}_k^*(\mathbf{x};\boldsymbol{\alpha})\mathcal{O}_{k'}(\mathbf{x};\boldsymbol{\alpha})\dot{\alpha}_{k'} + \frac{i}{\hbar}\mathcal{O}_k^*(\mathbf{x};\boldsymbol{\alpha})\mathcal{E}(\mathbf{x};\boldsymbol{\alpha}) \right\rangle , \end{aligned} \tag{1.10}$$

$$\begin{aligned} \frac{\partial \Delta^2}{\partial \dot{\alpha}_k^{(I)}} &= -i \left\langle \sum_{k'} \mathcal{O}_k^*(\mathbf{x};\boldsymbol{\alpha})\mathcal{O}_{k'}(\mathbf{x};\boldsymbol{\alpha})\dot{\alpha}_{k'} + \frac{i}{\hbar}\mathcal{O}_k^*(\mathbf{x};\boldsymbol{\alpha})\mathcal{E}(\mathbf{x};\boldsymbol{\alpha}) \right\rangle + \text{c.c.} \\ &= 2\Im \left\langle \sum_{k'} \mathcal{O}_k^*(\mathbf{x};\boldsymbol{\alpha})\mathcal{O}_{k'}(\mathbf{x};\boldsymbol{\alpha})\dot{\alpha}_{k'} + \frac{i}{\hbar}\mathcal{O}_k^*(\mathbf{x};\boldsymbol{\alpha})\mathcal{E}(\mathbf{x};\boldsymbol{\alpha}) \right\rangle . \end{aligned} \tag{1.11}$$

Setting these derivatives equal to zero is equivalent to the equation

$$\left\langle \sum_{k'} \mathcal{O}_k^*(\mathbf{x};\boldsymbol{\alpha})\mathcal{O}_{k'}(\mathbf{x};\boldsymbol{\alpha})\dot{\alpha}_{k'} + \frac{i}{\hbar}\mathcal{O}_k^*(\mathbf{x};\boldsymbol{\alpha})\mathcal{E}(\mathbf{x};\boldsymbol{\alpha}) \right\rangle = 0 , \tag{1.12}$$

for every index k. We can reorganize Eq. (1.12) as

$$i\hbar \sum_{k'} \langle \mathcal{O}_k^* \mathcal{O}_{k'} \rangle \, \dot{\alpha}_{k'} = \langle \mathcal{O}_k^* \mathcal{E} \rangle \ . \tag{1.13}$$

The expectation values can be computed via Monte Carlo, and then the time derivatives of the parameters can be calculated by solving the linear system (1.13). So Eq. (1.13) defines the equations of motion for the parameters, whose solution yields an approximation of the exact quantum dynamics.

It is now useful to introduce an explicit parameter ϕ which has the role of a "complex phase", i.e., we consider a variational state that can be written as

$$|\psi(\phi,\boldsymbol{\alpha})\rangle = e^{\phi}|\tilde{\psi}(\boldsymbol{\alpha})\rangle , \tag{1.14}$$

where $\boldsymbol{\alpha}$ now denotes all the parameters *other than* ϕ, and $|\tilde{\psi}(\boldsymbol{\alpha})\rangle$ is an Ansatz that does not depend on ϕ. The complex phase ϕ controls the modulus and global phase of the quantum state, and the local operator corresponding to this parameter is the identity, i.e., $\mathcal{O}_\phi(\mathbf{x}) = \frac{\partial}{\partial\phi}\ln\psi \equiv 1$. The equations of motion (1.13) still hold, with the indices k and k' running over the full set of parameters $(\phi,\boldsymbol{\alpha})$. However, now that we have distinguished

[1] That is, the square modulus of the normalized wave function, since $\psi(\mathbf{x};\boldsymbol{\alpha}) \equiv \langle \mathbf{x}|\psi(\boldsymbol{\alpha})\rangle$. Please notice, however, that the approach that is here presented does not depend on the representation of the Hilbert space that one chooses to use. Any basis set could be used in place of the "position" eigenstates $|\mathbf{x}\rangle$.

ϕ from the other parameters, it is useful to rewrite Eq. (1.13) as two coupled equations in which the terms related to ϕ are treated separately:

$$i\hbar \sum_{k'} \langle \mathcal{O}_k^* \mathcal{O}_{k'} \rangle \, \dot{\alpha}_{k'} + i\hbar \, \langle \mathcal{O}_k^* \rangle \, \dot{\phi} = \langle \mathcal{O}_k^* \mathcal{E} \rangle \ , \tag{1.15}$$

$$i\hbar \sum_{k'} \langle \mathcal{O}_{k'} \rangle \, \dot{\alpha}_{k'} + i\hbar \dot{\phi} = \langle \mathcal{E} \rangle \ , \tag{1.16}$$

where the indices k and k' run only over the components of $\boldsymbol{\alpha}$. The first equation is the equivalent of Eq. (1.13) where the term in the summation corresponding to $\dot{\phi}$ has been isolated, and the second equation is the special case of Eq. (1.13) in which $\mathcal{O}_k$ is replaced by $\mathcal{O}_\phi \equiv 1$. Now, Eq. (1.16) can be solved for $\dot{\phi}$, and plugging the solution into Eq. (1.15) yields the final form of the equations of motion for the variational parameters:

$$\begin{aligned} & i\hbar \sum_{k'} \left(\langle \mathcal{O}_k^* \mathcal{O}_{k'} \rangle - \langle \mathcal{O}_k^* \rangle \, \langle \mathcal{O}_{k'} \rangle \right) \dot{\alpha}_{k'} = \langle \mathcal{O}_k^* \mathcal{E} \rangle - \langle \mathcal{O}_k^* \rangle \, \langle \mathcal{E} \rangle \ , \\ & \dot{\phi} = - \sum_{k'} \langle \mathcal{O}_{k'} \rangle \, \dot{\alpha}_{k'} - \frac{i}{\hbar} \, \langle \mathcal{E} \rangle \ . \end{aligned} \tag{1.17}$$

This is the final form of the t-VMC equations of motion. It should be noted that the first line of Eq. (1.17) is actually a system of equations running over the index k and only involving the time derivatives $\dot{\boldsymbol{\alpha}}$, which can thus be calculated without computing $\dot{\phi}$. This is reasonable, since the parameter ϕ only contains information about the norm and the global phase of the wave function, that are irrelevant to the physical state of the system. On the other hand, computing $\dot{\phi}$ requires knowledge of $\dot{\boldsymbol{\alpha}}$. Incidentally, it is interesting to notice that if we were to simulate the time evolution of an eigenstate of the Hamiltonian of energy E_n, then ϕ would be the only parameter changing in time, according to $\phi(t) = -\frac{i}{\hbar} E_n t$; in general, if ϕ were the only parameter, then its evolution, according to Eq. (1.17), would be $\phi(t) = -\frac{i}{\hbar} \, \langle \mathcal{E} \rangle \, t$.

Looking at Eq. (1.17), it is convenient to define the so-called *quantum geometric tensor*

$$S_{k,k'} = \langle \mathcal{O}_k^* \mathcal{O}_{k'} \rangle - \langle \mathcal{O}_k^* \rangle \, \langle \mathcal{O}_{k'} \rangle \ , \tag{1.18}$$

and the *force vector*

$$F_k = \langle \mathcal{O}_k^* \mathcal{E} \rangle - \langle \mathcal{O}_k^* \rangle \, \langle \mathcal{E} \rangle \ . \tag{1.19}$$

We observe that S is the covariance matrix of the local operators, so it is positive semi-definite. With this notation, the equations of motion for $\dot{\boldsymbol{\alpha}}$ can be synthetically written as

$$i\hbar \sum_{k'} S_{k,k'} \dot{\alpha}_{k'} = F_k \, . \tag{1.20}$$

Here it is useful to point out that, although in Eq. (1.17) and in the definition of F we could substitute the energy E for the expectation value $\langle \mathcal{E} \rangle$, we decide not to. Discarding, in $\langle \mathcal{E} \rangle$, the contribution of the imaginary part of $\mathcal{E}(\mathbf{x}; \boldsymbol{\alpha})$ apparently makes sense, since we analytically know that its expectation value is zero ($\langle \mathcal{E} \rangle = E$ must be real). However, in actual Monte Carlo calculations we observed that the stochastic noises of the estimators for $\langle \mathcal{O}_k^* \mathcal{E} \rangle$, $\langle \mathcal{O}_k^* \rangle$, and $\langle \mathcal{E} \rangle$ are correlated in such a way that we obtain a more accurate estimation of F_k if we do not discard the imaginary part of the average local energy, even if, admittedly, it only includes noise. In other words, the noise in the averaged imaginary part of $\mathcal{E}(\mathbf{x}; \boldsymbol{\alpha})$ is observed to partially cancel out other noisy contributions in the expression defining F. The same point is also made in Ref. [26].

1.1.2 Minimizing the distance between rays: the Fubini-Study metric and the significance of the complex phase

In the Hilbert space a physical state is represented by a ray, that is the set of all nonzero vectors that are proportional to some reference vector $|\psi\rangle$. The set of (pure) physical states is *not* in one-to-one correspondence with the set of vectors, but it is with the set of rays, which is called the *ray space* or *projective Hilbert space*. So, when evaluating how accurate a variational approximation is in describing some quantum state, we should inspect how close the ray of the variational state is to the ray of the exact state.

How can we measure it? We need some notion of distance between the rays of two vectors, say $|\psi\rangle$ and $|\varphi\rangle$. The Hilbert-space norm of the difference between the two vectors $\||\psi\rangle - |\varphi\rangle\|$ is not a good choice, because it depends on the norm of the vectors, and, even when they are normalized, it still depends on their relative phase. A solution is given by the Fubini-Study distance, defined as

$$\mathcal{D}(|\psi\rangle, |\varphi\rangle) = \arccos \sqrt{\frac{\langle\psi|\varphi\rangle \langle\varphi|\psi\rangle}{\langle\psi|\psi\rangle \langle\varphi|\varphi\rangle}}\,. \tag{1.21}$$

This quantity satisfies the axioms of a metric and is independent of both the norms and the phases of the chosen vectors, so it is a good choice for measuring the distance between the rays that $|\psi\rangle$ and $|\varphi\rangle$ represent.

According to these considerations, the procedure we followed to derive the t-VMC equations of motion in Subsec. 1.1.1 looks flawed, because it was based on minimizing the quantity $\Delta^2(t)$, defined in Eq. (1.3) as the squared norm of a difference between vectors. It looks like a better approach would be to minimize the Fubini-Study distance between the short-time evolutions of the variational state and of the exact state, and indeed this is the approach suggested in Ref. [27]. However, as we now show, the introduction of the complex phase parameter ϕ makes the two approaches equivalent, so that the final form (1.17) of the t-VMC equations rests on solid ground.

In the following, we will need an expression for the Fubini-Study distance in the limit in which two vectors are close to each other, say $|\varphi\rangle = |\psi\rangle + |\delta\psi\rangle$ with $|\delta\psi\rangle$ infinitesimal. We expand the square of the Fubini-Study distance (1.21) to second order in $|\delta\psi\rangle$ (the first-order term vanishes), yielding [2]

$$\begin{aligned}\mathcal{D}^2(|\psi\rangle, |\psi\rangle + |\delta\psi\rangle) &\simeq \frac{\langle\delta\psi|\delta\psi\rangle}{\langle\psi|\psi\rangle} - \frac{\langle\psi|\delta\psi\rangle \langle\delta\psi|\psi\rangle}{\langle\psi|\psi\rangle^2} \\ &= \frac{1}{\langle\psi|\psi\rangle} \left\| |\delta\psi\rangle - |\psi\rangle \frac{\langle\psi|\delta\psi\rangle}{\langle\psi|\psi\rangle} \right\|^2 .\end{aligned} \tag{1.22}$$

The quantity within the Hilbert-space norm, i.e., $|\delta\psi\rangle - |\psi\rangle \frac{\langle\psi|\delta\psi\rangle}{\langle\psi|\psi\rangle}$, is the component of $|\delta\psi\rangle$ which is orthogonal to $|\psi\rangle$. It can be defined as the state of the form $|\delta\psi\rangle - \lambda |\psi\rangle$ which has minimum norm. Thus, Eq. (1.22) can be rewritten as

$$\mathcal{D}^2(|\psi\rangle, |\psi\rangle + |\delta\psi\rangle) \simeq \min_{\lambda\in\mathbb{C}} \frac{1}{\langle\psi|\psi\rangle} \||\delta\psi\rangle - \lambda |\psi\rangle\|^2 . \tag{1.23}$$

[2] To obtain this expansion, one uses the relation $(1+\epsilon)^{-1} = 1 - \epsilon + \epsilon^2 + O(\epsilon^3)$ to show that

$$\frac{\langle\psi|\psi+\delta\psi\rangle \langle\psi+\delta\psi|\psi\rangle}{\langle\psi|\psi\rangle \langle\psi+\delta\psi|\psi+\delta\psi\rangle} = 1 - \frac{\langle\delta\psi|\delta\psi\rangle}{\langle\psi|\psi\rangle} + \frac{\langle\psi|\delta\psi\rangle \langle\delta\psi|\psi\rangle}{\langle\psi|\psi\rangle^2} + O\left(\delta\psi^3\right) .$$

Then the result is obtained using the relations $\sqrt{1+\varepsilon} = 1 + \frac{1}{2}\varepsilon + O(\varepsilon^2)$, and $[\arccos(1-\varepsilon)]^2 = 2\varepsilon + O(\varepsilon^2)$.

Let's now apply this formula to the distance between the variational and the exact evolution of a state. Starting from state $|\psi(\boldsymbol{\alpha})\rangle$ at time t, after a short time δt the variational state will be $|\psi(t+\delta t)\rangle_{\text{var}} = |\psi(\boldsymbol{\alpha})\rangle + \sum_k \frac{\partial|\psi(\boldsymbol{\alpha})\rangle}{\partial\alpha_k}\dot{\alpha}_k\delta t$, and the exact state will be $|\psi(t+\delta t)\rangle_{\text{ex}} = |\psi(\boldsymbol{\alpha})\rangle - \frac{i}{\hbar}\hat{H}\,|\psi(\boldsymbol{\alpha})\rangle\,\delta t$. So, the difference $|\delta\psi\rangle$ between the two short-time evolutions is given by

$$|\delta\psi(\boldsymbol{\alpha},\dot{\boldsymbol{\alpha}})\rangle = |\psi(t+\delta t)\rangle_{\text{var}} - |\psi(t+\delta t)\rangle_{\text{ex}} = \left[\sum_k \frac{\partial\,|\psi(\boldsymbol{\alpha})\rangle}{\partial\alpha_k}\dot{\alpha}_k + \frac{i}{\hbar}\hat{H}\,|\psi(\boldsymbol{\alpha})\rangle\right]\delta t \quad (1.24)$$

The corresponding squared Fubini-Study distance is obtained by plugging this expression for $|\delta\psi\rangle$ into Eq. (1.23). We divide it by δt^2 to get the squared rate of departure $\varepsilon^2(t)$ between the two evolutions, yielding

$$\begin{aligned}\varepsilon^2(t) &= \frac{1}{\delta t^2}\mathcal{D}^2(|\psi(t+\delta t)\rangle_{\text{var}}, |\psi(t+\delta t)\rangle_{\text{ex}}) \\ &\simeq \min_{\lambda\in\mathbb{C}} \frac{1}{\langle\psi(\boldsymbol{\alpha})|\psi(\boldsymbol{\alpha})\rangle}\left\|\frac{1}{\delta t}\,|\delta\psi(\boldsymbol{\alpha},\dot{\boldsymbol{\alpha}})\rangle - \lambda\,|\psi(\boldsymbol{\alpha})\rangle\right\|^2 \\ &= \min_{\lambda\in\mathbb{C}} \frac{1}{\langle\psi(\boldsymbol{\alpha})|\psi(\boldsymbol{\alpha})\rangle}\left\|\frac{\partial\,|\psi(\boldsymbol{\alpha})\rangle}{\partial\alpha_k}\dot{\alpha}_k + \frac{i}{\hbar}\hat{H}\,|\psi(\boldsymbol{\alpha})\rangle - \lambda\,|\psi(\boldsymbol{\alpha})\rangle\right\|^2 .\end{aligned} \quad (1.25)$$

The quantity $\varepsilon(t)$ is sometimes called the *error per unit time* or the *local-in-time error* [28].

In order to derive a set of equations of motion for the variational parameters $\boldsymbol{\alpha}$, it is now natural to minimize $\varepsilon^2(t)$ with respect to the time derivatives $\dot{\boldsymbol{\alpha}}$. On the other hand, in Subsec. 1.1.1 we found the t-VMC equations in a different way: we minimized the quantity $\Delta^2 = \frac{1}{\langle\psi|\psi\rangle}\left\|\frac{1}{\delta t}\,|\delta\psi\rangle\right\|^2$, and this is clearly *not* equivalent to the minimization of $\varepsilon^2(t)$, since the definition of the latter (1.25) includes an additional term $-\lambda\,|\psi(\boldsymbol{\alpha})\rangle$ in the norm and a minimization with respect to λ, not present in $\Delta^2(t)$. However, let's see how $\Delta^2(t)$ changes with the introduction of the complex phase ϕ.

With this parameter, the variational Ansatz becomes $e^{\phi}\,|\psi(\boldsymbol{\alpha})\rangle$, and its short-time variation changes to $e^{\phi}\sum_k \frac{\partial|\psi(\boldsymbol{\alpha})\rangle}{\partial\alpha_k}\dot{\alpha}_k\delta t + \dot{\phi}e^{\phi}\,|\psi(\boldsymbol{\alpha})\rangle\,\delta t$, while the variation of the exact evolution is simply $-\frac{i}{\hbar}e^{\phi}\hat{H}\,|\psi(\boldsymbol{\alpha})\rangle\,\delta t$. Thus, when the complex phase is included, the difference between the states becomes

$$\begin{aligned}|\delta\psi\rangle_{\phi} &= \left[e^{\phi}\sum_k \frac{\partial\,|\psi(\boldsymbol{\alpha})\rangle}{\partial\alpha_k}\dot{\alpha}_k + \dot{\phi}e^{\phi}\,|\psi(\boldsymbol{\alpha})\rangle + \frac{i}{\hbar}e^{\phi}\hat{H}\,|\psi(\boldsymbol{\alpha})\rangle\right]\delta t \\ &= e^{\phi}\,|\delta\psi(\boldsymbol{\alpha},\dot{\boldsymbol{\alpha}})\rangle + \dot{\phi}e^{\phi}\,|\psi(\boldsymbol{\alpha})\rangle\,\delta t\,,\end{aligned} \quad (1.26)$$

where $|\delta\psi(\boldsymbol{\alpha},\dot{\boldsymbol{\alpha}})\rangle$ is the vector difference in absence of the complex phase, given by Eq. (1.24). The new expression of $\Delta^2(t)$ is

$$\Delta^2(t) = \frac{1}{\langle\psi(\boldsymbol{\alpha})|\psi(\boldsymbol{\alpha})\rangle\,|e^{\phi}|^2}\left\|\frac{1}{\delta t}\,|\delta\psi\rangle_{\phi}\right\|^2 \quad (1.27)$$

$$= \frac{1}{\langle\psi(\boldsymbol{\alpha})|\psi(\boldsymbol{\alpha})\rangle}\left\|\sum_k \frac{\partial\,|\psi(\boldsymbol{\alpha})\rangle}{\partial\alpha_k}\dot{\alpha}_k + \dot{\phi}\,|\psi(\boldsymbol{\alpha})\rangle + \frac{i}{\hbar}\hat{H}\,|\psi(\boldsymbol{\alpha})\rangle\right\|^2 \quad (1.28)$$

$$= \frac{1}{\langle\psi(\boldsymbol{\alpha})|\psi(\boldsymbol{\alpha})\rangle}\left\|\frac{1}{\delta t}\,|\delta\psi(\boldsymbol{\alpha},\dot{\boldsymbol{\alpha}})\rangle + \dot{\phi}\,|\psi(\boldsymbol{\alpha})\rangle\right\|^2 , \quad (1.29)$$

that should be minimized with respect to *both* $\dot{\boldsymbol{\alpha}}$ *and* $\dot{\phi}$, in order to derive the t-VMC equations of motion. Now notice that the minimization with respect to $\dot{\phi}$ is equivalent to the minimization with respect to λ which appears in the squared Fubini-Study error per unit time $\varepsilon^2(t)$ (1.25), with $-\dot{\phi}$ assuming the exact same role as λ. Thus the two quantities $\Delta^2(t)$ and $\varepsilon^2(t)$ are simply related by

$$\varepsilon^2(t) = \min_{\dot{\phi}\in\mathbb{C}} \Delta^2(t) \tag{1.30}$$

We conclude that minimizing $\Delta^2(t)$ with respect to $\dot{\boldsymbol{\alpha}}$ *and* $\dot{\phi}$ is really equivalent to minimizing $\varepsilon^2(t)$ with respect to $\dot{\boldsymbol{\alpha}}$ alone. This reconciles the two approaches.

1.1.3 Meaning of the quantum geometric tensor

To illustrate the geometrical meaning of the quantum geometric tensor S (1.18), we compute the Fubini-Study distance between the variational state $|\psi(\boldsymbol{\alpha})\rangle$ and the state corresponding to a small variation of the parameters $|\psi(\boldsymbol{\alpha}+\delta\boldsymbol{\alpha})\rangle$. The difference between the states is given by

$$|\delta\psi\rangle = \sum_k \frac{\partial\,|\psi(\boldsymbol{\alpha})\rangle}{\partial\alpha_k}\delta\alpha_k \tag{1.31}$$

The Fubini-Study distance can be computed substituting this expression for $|\delta\psi\rangle$ into Eq. (1.22), yielding

$$\mathcal{D}^2(|\psi\rangle, |\psi\rangle + |\delta\psi\rangle) = \frac{1}{\langle\psi|\psi\rangle}\left\|\sum_k \frac{\partial\,|\psi\rangle}{\partial\alpha_k}\delta\alpha_k - |\psi\rangle\,\frac{1}{\langle\psi|\psi\rangle}\sum_k \langle\psi|\,\frac{\partial\,|\psi\rangle}{\partial\alpha_k}\delta\alpha_k\right\|^2 . \tag{1.32}$$

Inserting the identity $\mathbb{1} = \sum_{\mathbf{x}} |\mathbf{x}\rangle\langle\mathbf{x}|$ two times, it is possible to rewrite this distance as

$$\mathcal{D}^2(|\psi\rangle, |\psi\rangle + |\delta\psi\rangle) = \tag{1.33}$$

$$= \sum_{\mathbf{x}} \frac{|\langle\mathbf{x}|\psi\rangle|^2}{\langle\psi|\psi\rangle}\left|\sum_k \frac{1}{\langle\mathbf{x}|\psi\rangle}\frac{\partial\langle\mathbf{x}|\psi\rangle}{\partial\alpha_k}\delta\alpha_k - \sum_{\mathbf{y}}\frac{|\langle\mathbf{y}|\psi\rangle|^2}{\langle\psi|\psi\rangle}\sum_k \frac{1}{\langle\mathbf{y}|\psi\rangle}\frac{\partial\langle\mathbf{y}|\psi\rangle}{\partial\alpha_k}\delta\alpha_k\right|^2 \tag{1.34}$$

$$= \left\langle\left|\sum_k (\mathcal{O}_k - \langle\mathcal{O}_k\rangle)\delta\alpha_k\right|^2\right\rangle , \tag{1.35}$$

and a short calculation yields

$$\mathcal{D}^2(|\psi\rangle, |\psi\rangle + |\delta\psi\rangle) = \sum_{k,k'} \delta\alpha_k^* \left(\langle\mathcal{O}_k^*\mathcal{O}_{k'}\rangle - \langle\mathcal{O}_k^*\rangle\langle\mathcal{O}_{k'}\rangle\right)\delta\alpha_{k'} = \sum_{k,k'}\delta\alpha_k^* S_{k,k'}\delta\alpha_{k'} . \tag{1.36}$$

So the quantum geometric tensor is the metric matrix representing the Fubini-Study distance in the space of the parameters $\boldsymbol{\alpha}$.

1.1.4 Meaning of the force vector

The meaning of the force vector F (1.19) is related to the dependence of the energy E on the parameters. If we assume a small change $\delta\boldsymbol{\alpha}$ in the parameters, the first-order change in the energy is

$$\delta E = \sum_k \left[\frac{\partial E(\boldsymbol{\alpha})}{\partial\alpha_k^R}\,\delta\alpha_k^R + \frac{\partial E(\boldsymbol{\alpha})}{\partial\alpha_k^I}\,\delta\alpha_k^I\right] . \tag{1.37}$$

This expression can be easily shown to be equivalent to

$$\delta E = \sum_k \left[\frac{\partial E(\boldsymbol{\alpha}, \boldsymbol{\alpha}^*)}{\partial \alpha_k} \, \delta\alpha_k + \frac{\partial E(\boldsymbol{\alpha}, \boldsymbol{\alpha}^*)}{\partial \alpha_k^*} \, \delta\alpha_k^* \right] , \tag{1.38}$$

where the parameters $\boldsymbol{\alpha}$ and their complex conjugate $\boldsymbol{\alpha}^*$ are treated as independent variables. The energy depends on both, since $E = \langle \psi(\boldsymbol{\alpha}) | \hat{H} | \psi(\boldsymbol{\alpha}) \rangle / \| \psi(\boldsymbol{\alpha}) \|^2$, so E depends on $\boldsymbol{\alpha}$ through the ket $|\psi(\boldsymbol{\alpha})\rangle$, and on $\boldsymbol{\alpha}^*$ through the bra $\langle \psi(\boldsymbol{\alpha})|$. The corresponding derivatives with respect to α_k and α_k^* are related to the derivatives with respect to α_k^R and α_k^I by the relations

$$\frac{\partial E}{\partial \alpha_k^R} = \frac{\partial E}{\partial \alpha_k} + \frac{\partial E}{\partial \alpha_k^*} , \qquad \frac{\partial E}{\partial \alpha_k^I} = i \left[\frac{\partial E}{\partial \alpha_k} - \frac{\partial E}{\partial \alpha_k^*} \right] , \tag{1.39}$$

$$\frac{\partial E}{\partial \alpha_k} = \frac{1}{2} \left[\frac{\partial E}{\partial \alpha_k^R} - i \frac{\partial E}{\partial \alpha_k^I} \right] , \qquad \frac{\partial E}{\partial \alpha_k^*} = \frac{1}{2} \left[\frac{\partial E}{\partial \alpha_k^R} + i \frac{\partial E}{\partial \alpha_k^I} \right] . \tag{1.40}$$

Now we compute one of the derivatives:

$$\begin{aligned}
\frac{\partial E}{\partial \alpha_k^*} &= \frac{\partial}{\partial \alpha_k^*} \left[\frac{1}{\| \psi(\boldsymbol{\alpha}) \|^2} \left\langle \psi(\boldsymbol{\alpha}) \middle| \hat{H} \middle| \psi(\boldsymbol{\alpha}) \right\rangle \right] && (1.41) \\
&= \sum_{\mathbf{x}} \frac{1}{\| \psi(\boldsymbol{\alpha}) \|^2} \cdot \frac{\partial \langle \psi(\boldsymbol{\alpha}) | \mathbf{x} \rangle}{\partial \alpha_k^*} \cdot \left[\left\langle \mathbf{x} \middle| \hat{H} \middle| \psi(\boldsymbol{\alpha}) \right\rangle - \frac{\langle \mathbf{x} | \psi(\boldsymbol{\alpha}) \rangle}{\| \psi(\boldsymbol{\alpha}) \|^2} \left\langle \psi(\boldsymbol{\alpha}) \middle| \hat{H} \middle| \psi(\boldsymbol{\alpha}) \right\rangle \right] && (1.42) \\
&= \sum_{\mathbf{x}} \frac{| \langle \mathbf{x} | \psi(\boldsymbol{\alpha}) \rangle |^2}{\| \psi(\boldsymbol{\alpha}) \|^2} \cdot \mathcal{O}_k^*(\mathbf{x}; \boldsymbol{\alpha}) \cdot [\mathcal{E}(\mathbf{x}; \boldsymbol{\alpha}) - E(\boldsymbol{\alpha})] && (1.43) \\
&= \langle \mathcal{O}_k^* \mathcal{E} \rangle - \langle \mathcal{O}_k^* \rangle \, E && (1.44) \\
&= F_k \, . && (1.45)
\end{aligned}$$

where in the last line we recognized the definition of the force vector, Eq. (1.19). An analogous calculation yields $\frac{\partial E}{\partial \alpha_k} = F_k^*$. So we can rewrite the energy variation (1.38) as

$$\delta E = \sum_k \left(F_k^* \, \delta\alpha_k + F_k \, \delta\alpha_k^* \right) . \tag{1.46}$$

The force vector can thus be interpreted as the *energy gradient* with respect to the parameters.

1.1.5 The error per unit time

From Eq. (1.28) it is easy to show that $\Delta^2(t)$ can be written in terms of the local operators and the local energy as

$$\Delta^2(t) = \left\langle \left| \sum_k \mathcal{O}_k \dot{\alpha}_k + \dot{\phi} + \frac{i}{\hbar} \mathcal{E} \right|^2 \right\rangle , \tag{1.47}$$

which is a simple modification of Eq. (1.9) to take into account the complex phase. Minimizing $\Delta^2(t)$ with respect to $\dot{\phi}$ *alone* yields the second equation in (1.17), which expresses

$\dot{\phi}$ as a function of the time derivatives $\dot{\alpha}_k$ of the other parameters. Plugging that expression for $\dot{\phi}$ back into $\Delta^2(t)$ yields an expression for the Fubini-Study error per unit time

$$\varepsilon^2(t) = \left\langle \left| \sum_k \left(\mathcal{O}_k - \langle \mathcal{O}_k \rangle\right) \dot{\alpha}_k + \frac{i}{\hbar} \left(\mathcal{E} - E\right) \right|^2 \right\rangle . \tag{1.48}$$

Some manipulation [3] shows that this expression can be rewritten in terms of the quantum geometric tensor S (1.18) and the force vector F (1.19) as

$$\varepsilon^2(t) = \frac{1}{\hbar^2} \left(\left\langle \hat{H}^2 \right\rangle - E^2 \right) + \sum_{k,k'} \dot{\alpha}_k^* S_{k,k'} \dot{\alpha}_{k'} + \frac{i}{\hbar} \sum_k \dot{\alpha}_k^* F_k - \frac{i}{\hbar} \sum_k \dot{F}_k^* \dot{\alpha}_k \, . \tag{1.49}$$

Further minimization of $\varepsilon^2(t)$ with respect to $\dot{\boldsymbol{\alpha}}$ yields the full set of the t-VMC equations of motion (1.17). Assuming that $\dot{\boldsymbol{\alpha}}$ solves that system, we obtain the value of $\varepsilon^2(t)$ at the variational minimum:

$$\varepsilon^2_{\text{min}}(t) = \frac{1}{\hbar^2} \left(\left\langle \hat{H}^2 \right\rangle - E^2 \right) - \sum_{k,k'} \dot{\alpha}_k^* S_{k,k'} \dot{\alpha}_{k'} . \tag{1.50}$$

We observe that, since S is positive semi-definite, the quantity $\sum_{k,k'} \dot{\alpha}_k^* S_{k,k'} \dot{\alpha}_{k'}$ is greater than or equal to zero. Therefore, $\varepsilon^2_{\text{min}}(t)$ cannot be greater than $\frac{1}{\hbar^2} \left(\left\langle \hat{H}^2 \right\rangle - E^2 \right)$, i.e., it has a maximum possible value that is proportional to the variance of the energy observable. It can reach this maximum if and only if $\dot{\boldsymbol{\alpha}} \equiv 0$. This condition can be met for two reasons:

- the system is in a stationary state, which means that it is in an eigenstate of the energy; in this case the variance of the energy is zero and $\varepsilon^2_{\text{min}}(t) = 0$;
- the exact evolution is orthogonal to the variational manifold and cannot be represented by our algorithm; as a consequence, the variational state does not evolve.

1.1.6 An *a posteriori* error bound

An upper bound for the distance between the variationally evolved state and the exact state can be expressed in terms of the error per unit time (1.25). Here is a derivation adapted from Ref. [29]. We use the shorthand notation $\psi(t)$ for the variational state and $\Phi(t)$ for the exact state, with the assumption that the two evolutions start from the same initial point, i.e., $\psi(0) = \Phi(0)$. We start by calculating the distance between $\psi(t)$ and $\Phi(t)$

[3] Write the squared modulus $\left| \sum_k \left(\mathcal{O}_k - \langle \mathcal{O}_k \rangle\right) \dot{\alpha}_k + \frac{i}{\hbar} \left(\mathcal{E} - E\right) \right|^2$ as the product of the quantity $\sum_k \left(\mathcal{O}_k - \langle \mathcal{O}_k \rangle\right) \dot{\alpha}_k + \frac{i}{\hbar} \left(\mathcal{E} - E\right)$ and its complex conjugate. Then use the following, easily provable relations:

- $\left\langle |\mathcal{E} - E|^2 \right\rangle = \left\langle \hat{H}^2 \right\rangle - E^2$;
- $\left\langle \left(\mathcal{O}_k - \langle \mathcal{O}_k \rangle\right)^* \left(\mathcal{O}_{k'} - \langle \mathcal{O}_{k'} \rangle\right) \right\rangle = S_{k,k'}$;
- $\left\langle \left(\mathcal{O}_k - \langle \mathcal{O}_k \rangle\right)^* \left(\mathcal{E} - E\right) \right\rangle = F_k$.

after a short time δt:

$$\begin{aligned}\|\psi(t+\delta t)-\Phi(t+\delta t)\| &= \left\|\psi(t+\delta t)-e^{-\frac{i}{\hbar}\hat{H}\delta t}\Phi(t)\right\| \\ &\leq \left\|\psi(t+\delta t)-e^{-\frac{i}{\hbar}\hat{H}\delta t}\psi(t)\right\| + \left\|e^{-\frac{i}{\hbar}\hat{H}\delta t}\psi(t)-e^{-\frac{i}{\hbar}\hat{H}\delta t}\Phi(t)\right\| \\ &= \left\|\psi(t+\delta t)-e^{-\frac{i}{\hbar}\hat{H}\delta t}\psi(t)\right\| + \|\psi(t)-\Phi(t)\|\,,\end{aligned} \tag{1.51}$$

where we used: in the second line, the triangle inequality $\|\zeta-\xi\| \leq \|\zeta-\chi\| + \|\chi-\xi\|$, holding for any three vectors ζ, ξ, χ; in the third line, the unitarity of the evolution operator $e^{-\frac{i}{\hbar}\hat{H}\delta t}$.

Now we define the variation of the distance between the variational state and the exact state as

$$\delta\,\|\psi(t)-\Phi(t)\| = \|\psi(t+\delta t)-\Phi(t+\delta t)\| - \|\psi(t)-\Phi(t)\|\,, \tag{1.52}$$

and, using this definition, we rewrite the inequality (1.51) as

$$\delta\,\|\psi(t)-\Phi(t)\| \leq \left\|\psi(t+\delta t)-e^{-\frac{i}{\hbar}\hat{H}\delta t}\psi(t)\right\|\,. \tag{1.53}$$

Dividing this by δt, we get an inequality between time derivatives,

$$\begin{aligned}\frac{d}{dt}\|\psi(t)-\Phi(t)\| &\leq \left(\frac{d}{dt}\left\|\psi(t)-e^{-\frac{i}{\hbar}\hat{H}(t-\tilde{t})}\psi(\tilde{t})\right\|\right)_{\tilde{t}=t} \\ &= \left(\frac{d}{dt}\left\|\psi(t)-\psi_{\text{ex}}(t,\tilde{t})\right\|\right)_{\tilde{t}=t}\,,\end{aligned} \tag{1.54}$$

where we introduced the notation $\psi_{\text{ex}}(t,\tilde{t}) = e^{-\frac{i}{\hbar}\hat{H}(t-\tilde{t})}\psi(\tilde{t})$ for the exact evolution from time $\tilde{t}$ to time t, starting from the variational state $\psi(\tilde{t})$. Next, we need to compute the quantity

$$\begin{aligned}\left\|\psi(t)-\psi_{\text{ex}}(t,\tilde{t})\right\| \cdot \frac{d}{dt}\left\|\psi(t)-\psi_{\text{ex}}(t,\tilde{t})\right\| &= \frac{1}{2}\frac{d}{dt}\left\|\psi(t)-\psi_{\text{ex}}(t,\tilde{t})\right\|^2 \\ &= \Re\left\langle\psi(t)-\psi_{\text{ex}}(t,\tilde{t})\middle|\frac{d}{dt}\left[\psi(t)-\psi_{\text{ex}}(t,\tilde{t})\right]\right\rangle \\ &\leq \left\|\psi(t)-\psi_{\text{ex}}(t,\tilde{t})\right\| \cdot \left\|\frac{d\psi(t)}{dt}-\frac{d\psi_{\text{ex}}(t,\tilde{t})}{dt}\right\|\,.\end{aligned} \tag{1.55}$$

Dividing by $\left\|\psi(t)-\psi_{\text{ex}}(t,\tilde{t})\right\|$ and then evaluating in $\tilde{t}=t$, we obtain

$$\left(\frac{d}{dt}\left\|\psi(t)-\psi_{\text{ex}}(t,\tilde{t})\right\|\right)_{\tilde{t}=t} \leq \left\|\frac{d\psi(t)}{dt}-\left(\frac{d\psi_{\text{ex}}(t,\tilde{t})}{dt}\right)_{\tilde{t}=t}\right\| = \Delta(t)\,, \tag{1.56}$$

where we recognized the definition of $\Delta(t)$ from Eq. (1.3), written in a different notation. Finally, we combine the two inequalities (1.54) and (1.56), obtaining

$$\frac{d}{dt}\|\psi(t)-\Phi(t)\| \leq \Delta(t)\,, \tag{1.57}$$

and integration over time yields the *a posteriori* error bound

$$\|\psi(t)-\Phi(t)\| \leq \int_0^t \Delta(t')dt'\,. \tag{1.58}$$

In Subsec. 1.1.2 we showed that, at the variational minimum, the Fubini-Study error per unit time $\varepsilon(t)$ is identical to $\Delta(t)$, so $\Delta(t)$ can be replaced by $\varepsilon(t)$ in the error bound. This error bound is *a posteriori* since it is given in terms of the quantity $\Delta(t)$ yielded by the variational approximation $\psi(t)$, rather than the exact evolution $\Phi(t)$.

1.1.7 Norm conservation

One of the consequences of the introduction of the complex phase ϕ is that the variational dynamics induced by the equations of motion (1.17) conserves the norm. To prove it, let's carry out the calculation for the time derivative of the norm and show that, if Eq. (1.17) holds, it is zero:

$$\begin{aligned}\frac{d}{dt}\langle\psi|\psi\rangle &= \frac{d}{dt}\sum_{\mathbf{x}}\langle\psi|\mathbf{x}\rangle\langle\mathbf{x}|\psi\rangle \\ &= \sum_{\mathbf{x}}\langle\psi|\mathbf{x}\rangle\frac{d\langle\mathbf{x}|\psi\rangle}{dt} + \text{c.c.}\,.\end{aligned} \tag{1.59}$$

Now we evaluate the time derivative of the variational wave function in terms of the time derivatives of the parameters:

$$\begin{aligned}\frac{d\langle\mathbf{x}|\psi\rangle}{dt} &= \sum_k \frac{\partial\langle\mathbf{x}|\psi\rangle}{\partial\alpha_k}\dot{\alpha}_k + \frac{\partial\langle\mathbf{x}|\psi\rangle}{\partial\phi}\dot{\phi} \\ &= \langle\mathbf{x}|\psi\rangle\left[\sum_k \frac{1}{\langle\mathbf{x}|\psi\rangle}\frac{\partial\langle\mathbf{x}|\psi\rangle}{\partial\alpha_k}\dot{\alpha}_k + \frac{1}{\langle\mathbf{x}|\psi\rangle}\frac{\partial\langle\mathbf{x}|\psi\rangle}{\partial\phi}\dot{\phi}\right] \\ &= \langle\mathbf{x}|\psi\rangle\left[\sum_k \mathcal{O}_k(\mathbf{x})\dot{\alpha}_k + \mathcal{O}_\phi(\mathbf{x})\dot{\phi}\right] \\ &= \langle\mathbf{x}|\psi\rangle\left[\sum_k \mathcal{O}_k(\mathbf{x})\dot{\alpha}_k + \dot{\phi}\right],\end{aligned} \tag{1.60}$$

where we used the definition of the local operators (1.8) and the fact that $\mathcal{O}_\phi(\mathbf{x}) \equiv 1$. Plugging Eq. (1.60) into Eq. (1.59), we can continue the computation of $\frac{d}{dt}\langle\psi|\psi\rangle$, which we divide by $\langle\psi|\psi\rangle$ to normalize the probability distribution given by the squared modulus of the wave function:

$$\begin{aligned}\frac{1}{\langle\psi|\psi\rangle}\frac{d}{dt}\langle\psi|\psi\rangle &= \frac{1}{\langle\psi|\psi\rangle}\sum_{\mathbf{x}}\langle\psi|\mathbf{x}\rangle\langle\mathbf{x}|\psi\rangle\left[\sum_k \mathcal{O}_k(\mathbf{x})\dot{\alpha}_k + \dot{\phi}\right] + \text{c.c.} \\ &= \left[\sum_k \langle\mathcal{O}_k\rangle\dot{\alpha}_k + \dot{\phi}\right] + \text{c.c.}\,.\end{aligned} \tag{1.61}$$

Here we insert the expression for $\dot{\phi}$ from Eq. (1.17), and get

$$\begin{aligned}\frac{1}{\langle\psi|\psi\rangle}\frac{d}{dt}\langle\psi|\psi\rangle &= \left[\sum_k \langle\mathcal{O}_k\rangle\dot{\alpha}_k - \sum_k \langle\mathcal{O}_k\rangle\dot{\alpha}_k - \frac{i}{\hbar}E\right] + \text{c.c.} \\ &= -\frac{i}{\hbar}E + \text{c.c.} = 0\,,\end{aligned} \tag{1.62}$$

where the last expression vanishes because a purely imaginary number plus its complex conjugate is zero.

1.1.8 Energy conservation

A second consequence of the introduction of the complex phase ϕ is that the variational dynamics conserves the energy of the quantum state. To prove it, we compute the time derivative of the expectation value of the Hamiltonian on the variational state. In this calculation we use the conservation of the norm, already proven in Subsec.1.1.7, and assume, without loss of generality, that $\langle\psi|\psi\rangle = 1$.

$$\begin{aligned}\frac{dE}{dt} = \frac{d}{dt}\left\langle \hat{H}\right\rangle &= \frac{d}{dt}\sum_{\mathbf{x},\mathbf{x}'}\langle\psi|\mathbf{x}\rangle\left\langle \mathbf{x}\middle|\hat{H}\middle|\mathbf{x}'\right\rangle\langle\mathbf{x}'|\psi\rangle \\ &= \sum_{\mathbf{x},\mathbf{x}'}\langle\psi|\mathbf{x}\rangle\left\langle \mathbf{x}\middle|\hat{H}\middle|\mathbf{x}'\right\rangle\frac{d\langle\mathbf{x}'|\psi\rangle}{dt} + \text{c.c.}\,.\end{aligned} \tag{1.63}$$

Here we substitute the time derivative of the variational wave function (1.60), yielding

$$\begin{aligned}\frac{dE}{dt} &= \sum_{\mathbf{x},\mathbf{x}'}\langle\psi|\mathbf{x}\rangle\left\langle \mathbf{x}\middle|\hat{H}\middle|\mathbf{x}'\right\rangle\langle\mathbf{x}'|\psi\rangle\left[\sum_k \mathcal{O}_k(\mathbf{x}')\dot{\alpha}_k + \dot{\phi}\right] + \text{c.c.} \\ &= \sum_{\mathbf{x}'}\langle\mathbf{x}'|\psi\rangle\left[\sum_{\mathbf{x}}\langle\psi|\mathbf{x}\rangle\left\langle \mathbf{x}\middle|\hat{H}\middle|\mathbf{x}'\right\rangle\right]\left[\sum_k \mathcal{O}_k(\mathbf{x}')\dot{\alpha}_k + \dot{\phi}\right] + \text{c.c.} \\ &= \sum_{\mathbf{x}'}\langle\mathbf{x}'|\psi\rangle\left\langle \psi\middle|\hat{H}\middle|\mathbf{x}'\right\rangle\left[\sum_k \mathcal{O}_k(\mathbf{x}')\dot{\alpha}_k + \dot{\phi}\right] + \text{c.c.} \\ &= \sum_{\mathbf{x}'}\langle\mathbf{x}'|\psi\rangle\langle\psi|\mathbf{x}'\rangle\frac{\left\langle \psi\middle|\hat{H}\middle|\mathbf{x}'\right\rangle}{\langle\psi|\mathbf{x}'\rangle}\left[\sum_k \mathcal{O}_k(\mathbf{x}')\dot{\alpha}_k + \dot{\phi}\right] + \text{c.c.}\,.\end{aligned} \tag{1.64}$$

Here we recognize the expression $\frac{\langle\psi|\hat{H}|\mathbf{x}'\rangle}{\langle\psi|\mathbf{x}'\rangle}$ as the complex conjugate of the local energy $\mathcal{E}(\mathbf{x}')$, defined in Eq. (1.7). So we can write

$$\begin{aligned}\frac{dE}{dt} &= \sum_{\mathbf{x}'}|\langle\psi|\mathbf{x}'\rangle|^2\,\mathcal{E}^*(\mathbf{x}')\left[\sum_k \mathcal{O}_k(\mathbf{x}')\dot{\alpha}_k + \dot{\phi}\right] + \text{c.c.}\,. \\ \frac{dE}{dt} &= \sum_{\mathbf{x}'}|\langle\psi|\mathbf{x}'\rangle|^2\left[\sum_k \mathcal{E}^*(\mathbf{x}')\mathcal{O}_k(\mathbf{x}')\dot{\alpha}_k + \dot{\phi}\right] + \text{c.c.}\,. \\ &= \left[\sum_k \langle\mathcal{E}^*\mathcal{O}_k\rangle\,\dot{\alpha}_k + E\dot{\phi}\right] + \text{c.c.}\,.\end{aligned} \tag{1.65}$$

In this expression we substitute the equation of motion for $\dot{\phi}$ from Eq. (1.17), and get

$$\begin{aligned}\frac{dE}{dt} &= \left[\sum_k \left(\langle\mathcal{E}^*\mathcal{O}_k\rangle - \langle\mathcal{O}_k\rangle\,E\right)\dot{\alpha}_k - \frac{i}{\hbar}E^2\right] + \text{c.c.} \\ &= \left[\sum_k F_k^*\dot{\alpha}_k - \frac{i}{\hbar}E^2\right] + \text{c.c.}\,,\end{aligned} \tag{1.66}$$

where we used the definition (1.19) of the force vector F. Now we can substitute $F_k^* = -i\hbar \sum_{k'} S_{k,k'}^* \dot{\alpha}_{k'}^*$, i.e., the complex conjugate of the equation of motion (1.20), yielding

$$\begin{aligned} \frac{dE}{dt} &= \left[-i\hbar \sum_{k,k'} S_{k,k'}^* \dot{\alpha}_{k'}^* \dot{\alpha}_k - \frac{i}{\hbar} E^2 \right] + \text{c.c.} \\ &= \left[-i\hbar \sum_{k,k'} \dot{\alpha}_{k'}^* S_{k',k} \dot{\alpha}_k - \frac{i}{\hbar} E^2 \right] + \text{c.c.} \\ &= 0 \,, \end{aligned} \tag{1.67}$$

where we used the Hermiticity of matrix S, see Eq. (1.18), to replace the complex conjugate S^* with the transpose S^T. The last expression vanishes because the quantity in square brackets is purely imaginary (this follows again from the Hermitianity of S).

1.1.9 Imaginary-time propagation

From the t-VMC equations of motion, through a simple mathematical trick, we can elegantly obtain another set of equation representing a different kind of evolution, i.e., the imaginary-time propagation, that allows the variational calculation of the ground-state properties of a system. But what do we mean by imaginary-time propagation?

Consider the propagator for the real-time evolution, $e^{-\frac{i}{\hbar}\hat{H}t}$. The imaginary time propagator is the operator that we obtain from the ordinary real-time one by performing the substitution

$$t \longrightarrow -i\tau \,. \tag{1.68}$$

This operation, called a *Wick rotation*, is a particular kind of *analytic continuation*, since we are analytically extending "time domain" of the propagator to the imaginary axis. The imaginary-time propagator thus obtained reads

$$e^{-\frac{1}{\hbar}\hat{H}\tau} \,. \tag{1.69}$$

This is not unitary and, accordingly, the imaginary-time evolution it generates does not conserve scalar products and norms. From the definition (1.69) it is straightforward to show that, just like ordinary real-time propagation obeys the Schrödinger equation, imaginary-time propagation obeys

$$\hbar \frac{d}{dt} |\psi(t)\rangle = -\hat{H} |\psi(t)\rangle \,. \tag{1.70}$$

The same equation can also be obtained directly from a Wick rotation on the Schrödinger equation, that is, substituting $\frac{d}{dt} \to i\frac{d}{d\tau}$.

The most important property of this kind of evolution is related to the ground state $|E_0\rangle$ of $\hat{H}$, and it is expressed by the following projection formula:

$$|E_0\rangle = \lim_{\tau \to +\infty} \frac{e^{-\frac{1}{\hbar}\hat{H}\tau} |\psi_0\rangle}{\left\| e^{-\frac{1}{\hbar}\hat{H}\tau} |\psi_0\rangle \right\|} \,, \tag{1.71}$$

where $|\psi_0\rangle$ is *any* state having *non-vanishing overlap* with the ground state $|E_0\rangle$. Eq. (1.71) can be easily proven by expanding $|\psi_0\rangle$ as a linear combination of eigenvectors of $\hat{H}$:

$$|\psi_0\rangle = \sum_n |E_n\rangle \langle E_n|\psi_0\rangle \,, \tag{1.72}$$

where $\hat{H}|E_n\rangle = E_n|E_n\rangle$. If we apply the exponential of the Hamiltonian, we get:

$$\begin{aligned} e^{-\frac{1}{\hbar}\hat{H}\tau}|\psi_0\rangle &= \sum_n \langle E_n|\psi_0\rangle\, e^{-\frac{1}{\hbar}E_n\tau}|E_n\rangle \\ &= e^{-\frac{1}{\hbar}E_0\tau}\sum_n \langle E_n|\psi_0\rangle\, e^{-\frac{1}{\hbar}(E_n-E_0)\tau}|E_n\rangle\,. \end{aligned} \tag{1.73}$$

Every coefficient in the superposition is a decaying exponential. So, since $E_0 < E_1 \leq E_2 \leq \ldots$, for big enough τ the term proportional to $|E_0\rangle$ is the only one to survive in the superposition, while the coefficient $e^{-\frac{1}{\hbar}E_0\tau}$ that we factored out simplifies with the normalization appearing in Eq. (1.71). This proves the projection formula. We assumed a non-degenerate ground state, since this is the case we will always deal with in this work. If the ground state happens to be degenerate, the imaginary-time propagation performs a projection of $|\psi_0\rangle$ on the subspace of minimum energy.

The projection formula shows that if we are able to implement imaginary-time propagation, then we can obtain the ground-state properties of a system up to the expressiveness introduced via the chosen variational Ansatz. Say that we want to approximate an imaginary-time evolution with a variational wave function $|\psi(\boldsymbol{\alpha})\rangle$. We could start from the Wick-rotated Schrödinger equation (1.70) and derive a set of equations of motion for the parameters, just like we did in Subsec. (1.1.1) to derive the t-VMC equations from the ordinary Schrödinger equation. But there is an equivalent and much faster way to do it: just performing a Wick rotation on the t-VMC equations themselves! Performing the substitution $\frac{d}{dt} \to i\frac{d}{d\tau}$ in Eq. (1.17), we obtain

$$\begin{aligned} &\hbar\sum_{k'}\left(\langle\mathcal{O}_k^*\mathcal{O}_{k'}\rangle - \langle\mathcal{O}_k^*\rangle\langle\mathcal{O}_{k'}\rangle\right)\frac{d\alpha_{k'}}{d\tau} = -\langle\mathcal{O}_k^*\mathcal{E}\rangle + \langle\mathcal{O}_k^*\rangle E\,, \\ &\frac{d\phi}{d\tau} = -\sum_{k'}\langle\mathcal{O}_{k'}\rangle\frac{d\alpha_{k'}}{d\tau} - \frac{1}{\hbar}E\,, \end{aligned} \tag{1.74}$$

where the first line can be written in a more compact way in terms of the quantum geometric tensor S and the force vector F as

$$\hbar\sum_{k'} S_{k,k'}\frac{d\alpha_{k'}}{d\tau} = -F_k\,. \tag{1.75}$$

These are the equations for imaginary-time t-VMC, or τ-VMC.

Stochastic reconfiguration

It is noteworthy that the imaginary-time equations (1.74) historically predate the devising of the t-VMC algorithm, as they implement the optimization algorithm known as *stochastic reconfiguration* [30].

The idea behind stochastic reconfiguration is the following. Our goal is finding a variational approximation of the ground state, so we look for the set of variational parameters that minimize the expectation value of the Hamiltonian operator $\hat{H}$ on the variational state $|\psi(\boldsymbol{\alpha})\rangle$. Starting from some set of parameters $\boldsymbol{\alpha}$, we want to perform a sequence of updates of the form

$$\boldsymbol{\alpha'} = \boldsymbol{\alpha} + \delta\boldsymbol{\alpha}\,, \tag{1.76}$$

carefully choosing $\delta\boldsymbol{\alpha}$ so that the energy is reduced at each step. This way the energy will eventually reach a minimum, which we hope to be the global minimum. Our task is determining the best choice for $\delta\boldsymbol{\alpha}$. After a single update, the energy becomes

$$E(\boldsymbol{\alpha'}) = E(\boldsymbol{\alpha}) + \sum_k \left[\frac{\partial E(\boldsymbol{\alpha})}{\partial \alpha_k}\delta\alpha_k + \frac{\partial E(\boldsymbol{\alpha})}{\partial \alpha_k^*}\delta\alpha_k^*\right] + O(|\delta\boldsymbol{\alpha}|^2) \tag{1.77}$$

$$= E(\boldsymbol{\alpha}) + \sum_k \left(F_k^*\delta\alpha_k + F_k\delta\alpha_k^*\right) + O(|\delta\boldsymbol{\alpha}|^2)\,, \tag{1.78}$$

where we used Eq. (1.46) to introduce the force vector (1.19). So the variation of the energy, to first order in $\delta\boldsymbol{\alpha}$, is given by

$$\Delta E \equiv E(\boldsymbol{\alpha'}) - E(\boldsymbol{\alpha}) = \sum_k \left(F_k^*\delta\alpha_k + F_k\delta\alpha_k^*\right)\,. \tag{1.79}$$

Now we could try to determine the optimal choice for the vector $\delta\boldsymbol{\alpha}$ by imposing ΔE to be minimum. However, this condition alone is not sufficient to determine $\delta\boldsymbol{\alpha}$, since ΔE is actually unbound: it can be made arbitrarily negative by picking up a $\delta\boldsymbol{\alpha}$ in any direction in which the energy decreases, say the opposite of the energy gradient $\frac{\partial E}{\partial \alpha_k^*}$, and making the vector arbitrarily large. So, we need to add an additional constraint that limits the size of the update. The most obvious choice is imposing the condition that the quantity

$$\delta s^2 \equiv \sum_k |\delta\alpha_k|^2 \tag{1.80}$$

is equal to some small value along the whole optimization procedure. This constraint can be included in the minimization of ΔE by introducing of a Lagrange multiplier μ and minimizing the quadratic form

$$\Delta E + \mu\,\delta s^2 = \sum_k \left(F_k^*\delta\alpha_k + F_k\delta\alpha_k^*\right) + \mu\sum_k |\delta\alpha_k|^2\,. \tag{1.81}$$

Minimizing this with respect to the parameter increments yields the solution

$$\delta\alpha_k = -\frac{1}{\mu}F_k = -\frac{1}{\mu}\frac{\partial E(\boldsymbol{\alpha})}{\partial \alpha_k^*}\,, \tag{1.82}$$

This is the traditional *steepest descent* method, with $1/\mu$ playing the role of a "learning rate". Eq. (1.82) ensures $\Delta E \leq 0$, the equality holding only when $\frac{\partial E}{\partial \alpha_k} = 0$ for every k, i.e., when a local minimum has been reached.

Steepest descent is a sensible choice for an optimization method, but it may not always be the best approach. The reason is that a constraint on the quantity (1.80) treats all the parameters α_k on the same footing. However, for a variational wave function whose dependence on the parameters is highly non-linear, it may be that changes in some parameter affect the quantum state much more than changes in another one.

This observation prompts us to redefine δs^2 in such a way that it measures the change in the quantum state, and *not* only in the variational parameters, as Eq. (1.80) does. The discussion in Subsec. 1.1.2 makes it clear that a good measure of the variation of the quantum state can be obtained via the Fubini-Study distance $\mathcal{D}$ (1.21), so we define

$$\delta s^2 \equiv \mathcal{D}(|\psi(\boldsymbol{\alpha}+\delta\boldsymbol{\alpha})\rangle, |\psi(\boldsymbol{\alpha})\rangle) = \sum_{k,k'} \delta\alpha_k^* S_{k,k'}\delta\alpha_{k'}\,, \tag{1.83}$$

where we used Eq. (1.36) to introduce the quantum geometric tensor (1.18). Now, minimization of the quadratic form $\Delta E + \mu\, \delta s^2$ yields

$$\sum_{k'} S_{k,k'} \delta\alpha_{k'} = -\frac{1}{\mu} F_k \,, \tag{1.84}$$

that is just a finite-increment version of the τ-VMC equations (1.74). This is stochastic reconfiguration, a robust optimization algorithm that, in comparison with steepest descent [30], proved to be less prone to get stuck in local minima.

Nevertheless, there may be cases in which this approach is unstable because of the need to invert the S matrix, which could be singular. While there are techniques to regularize S and invert it (see Subsec. 1.2.1), one may also want to try different methods, like stochastic gradient descent to minimize the energy. In the case of approaches based on machine learning (ML) like neural-network quantum states (NNQS), one could also employ a reinforcement learning technique, like Adam, AdaGrad, or RMSProp.

1.1.10 Equations of motion from the time-dependent variational principle (TDVP)

As an alternative approach to the one delineated in Subsec.1.1.1, the same equations of motion for the t-VMC method can be derived from the time-dependent variational principle (TDVP) [8]. The TDVP [31] is formulated as the requirement of stationarity for the action

$$\mathcal{S} = \int_{t_1}^{t_2} \mathcal{L}[\psi^*(t), \psi(t)] dt \,, \tag{1.85}$$

where the Lagrangian is taken as

$$\mathcal{L}[\psi^*(t), \psi(t)] = \frac{i\hbar}{2} \left[\left\langle \psi(t) \middle| \dot{\psi}(t) \right\rangle - \left\langle \dot{\psi}(t) \middle| \psi(t) \right\rangle \right] - \left\langle \psi(t) \middle| \hat{H} \middle| \psi(t) \right\rangle \,. \tag{1.86}$$

If we plug a variational state $|\psi(t)\rangle \equiv |\psi(\boldsymbol{\alpha}(t))\rangle$ into the Lagrangian $\mathcal{L}$, the action $\mathcal{S}$ becomes a functional of the variational parameters $\alpha_k(t)$ and their time derivatives $\dot{\alpha}_k(t)$. Specifically, the Lagrangian becomes

$$\begin{aligned}
\mathcal{L}[\psi^*, \psi] &= \frac{i\hbar}{2} \left[\left\langle \psi \middle| \dot{\psi} \right\rangle - \text{c.c.} \right] - \left\langle \psi \middle| \hat{H} \middle| \psi \right\rangle && (1.87)\\
&= \frac{i\hbar}{2} \left[\langle \psi | \sum_k \frac{\partial |\psi\rangle}{\partial \alpha_k} \dot{\alpha}_k - \text{c.c.} \right] - E(\boldsymbol{\alpha}) && (1.88)\\
&= \frac{i\hbar}{2} \left[\sum_{\mathbf{x}} \langle \psi | \mathbf{x} \rangle \sum_k \frac{\partial \langle \mathbf{x} | \psi \rangle}{\partial \alpha_k} \dot{\alpha}_k - \text{c.c.} \right] - E(\boldsymbol{\alpha}) && (1.89)\\
&= \frac{i\hbar}{2} \left[\sum_{\mathbf{x}} |\langle \mathbf{x} | \psi \rangle|^2 \sum_k \frac{1}{\langle \mathbf{x} | \psi \rangle} \frac{\partial \langle \mathbf{x} | \psi \rangle}{\partial \alpha_k} \dot{\alpha}_k - \text{c.c.} \right] - E(\boldsymbol{\alpha}) && (1.90)\\
&= \frac{i\hbar}{2} \left[\sum_{\mathbf{x}} |\langle \mathbf{x} | \psi \rangle|^2 \sum_k \mathcal{O}_k(\mathbf{x}) \dot{\alpha}_k - \text{c.c.} \right] - E(\boldsymbol{\alpha}) && (1.91)\\
&= \frac{i\hbar}{2} \left[\sum_k \langle \mathcal{O}_k \rangle \, \dot{\alpha}_k - \text{c.c.} \right] - E(\boldsymbol{\alpha}) && (1.92)\\
&= \frac{i\hbar}{2} \left[\sum_k \langle \mathcal{O}_k \rangle \, \dot{\alpha}_k - \sum_k \langle \mathcal{O}_k^* \rangle \, \dot{\alpha}_k^* \right] - E(\boldsymbol{\alpha}) \,, && (1.93)
\end{aligned}$$

where for the sake of simplicity we assumed that $|\psi(\boldsymbol{\alpha})\rangle$ is normalized. If it is not, the expectation values $\langle\cdot\rangle$ are simply replaced by non-normalized $\langle\psi(\boldsymbol{\alpha})|\cdot|\psi(\boldsymbol{\alpha})\rangle$, and all the calculations proceed in the same way.

Now, the requirement of stationary action $\delta\mathcal{S} = 0$ with respect to variations of $\alpha_k(t)$ and $\alpha_k^*(t)$ leads to the well-known Euler-Lagrange equations, which in this case read

$$\frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{\alpha}_k^*}\right) - \frac{\partial\mathcal{L}}{\partial\alpha_k^*} = 0\,, \quad \frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{\alpha}_k}\right) - \frac{\partial\mathcal{L}}{\partial\alpha_k} = 0 \tag{1.94}$$

Let's compute the terms appearing in the first equation, starting from the total time derivative:

$$\frac{d}{dt}\left(\frac{\partial\mathcal{L}}{\partial\dot{\alpha}_k^*}\right) = -\frac{i\hbar}{2}\frac{d}{dt}\langle\mathcal{O}_k^*\rangle = -\frac{i\hbar}{2}\frac{d}{dt}\langle\psi|\mathcal{O}_k^*|\psi\rangle \tag{1.95}$$

$$= -\frac{i\hbar}{2}\sum_{\mathbf{x}}\mathcal{O}_k^*(\mathbf{x})\left[\langle\psi|\mathbf{x}\rangle\frac{d\langle\mathbf{x}|\psi\rangle}{dt} + \text{c.c.}\right] \tag{1.96}$$

$$= -\frac{i\hbar}{2}\sum_{\mathbf{x}}\mathcal{O}_k^*(\mathbf{x})\left[\langle\psi|\mathbf{x}\rangle\sum_{k'}\frac{\partial\langle\mathbf{x}|\psi\rangle}{\partial\alpha_{k'}}\dot{\alpha}_{k'} + \text{c.c.}\right] \tag{1.97}$$

$$= -\frac{i\hbar}{2}\sum_{\mathbf{x}}\mathcal{O}_k^*(\mathbf{x})\left[|\langle\mathbf{x}|\psi\rangle|^2\sum_{k'}\mathcal{O}_{k'}(\mathbf{x})\dot{\alpha}_{k'} + \text{c.c.}\right] \tag{1.98}$$

$$= -\frac{i\hbar}{2}\sum_{\mathbf{x}}|\langle\mathbf{x}|\psi\rangle|^2\,\mathcal{O}_k^*(\mathbf{x})\left[\sum_{k'}\mathcal{O}_{k'}(\mathbf{x})\dot{\alpha}_{k'} + \sum_{k'}\mathcal{O}_{k'}^*(\mathbf{x})\dot{\alpha}_{k'}^*\right] \tag{1.99}$$

$$= -\frac{i\hbar}{2}\left[\sum_{k'}\langle\mathcal{O}_k^*\mathcal{O}_{k'}\rangle\,\dot{\alpha}_{k'} + \sum_{k'}\langle\mathcal{O}_k^*\mathcal{O}_{k'}^*\rangle\,\dot{\alpha}_{k'}^*\right]. \tag{1.100}$$

The next term is

$$\frac{\partial\mathcal{L}}{\partial\alpha_k^*} = \frac{i\hbar}{2}\left[\sum_{k'}\frac{\partial\langle\mathcal{O}_{k'}\rangle}{\partial\alpha_k^*}\dot{\alpha}_k - \sum_{k'}\frac{\partial\langle\mathcal{O}_{k'}^*\rangle}{\partial\alpha_k^*}\dot{\alpha}_k^*\right] - \frac{\partial E}{\partial\alpha_k^*}\,, \tag{1.101}$$

for which we need the derivative of the energy:

$$\frac{\partial E}{\partial\alpha_k^*} = \frac{\partial}{\partial\alpha_k^*}\left\langle\psi\middle|\hat{H}\middle|\psi\right\rangle \tag{1.102}$$

$$= \frac{\partial}{\partial\alpha_k^*}\sum_{\mathbf{x}}\langle\psi|\mathbf{x}\rangle\left\langle\mathbf{x}\middle|\hat{H}\middle|\psi\right\rangle \tag{1.103}$$

$$= \sum_{\mathbf{x}}\frac{\partial\langle\psi|\mathbf{x}\rangle}{\partial\alpha_k^*}\left\langle\mathbf{x}\middle|\hat{H}\middle|\psi\right\rangle \tag{1.104}$$

$$= \sum_{\mathbf{x}}\langle\psi|\mathbf{x}\rangle\,\mathcal{O}_k^*(\mathbf{x})\left\langle\mathbf{x}\middle|\hat{H}\middle|\psi\right\rangle \tag{1.105}$$

$$= \sum_{\mathbf{x}}|\,\langle\mathbf{x}|\psi\rangle\,|^2\mathcal{O}_k^*(\mathbf{x})\mathcal{E}(\mathbf{x}) \tag{1.106}$$

$$= \langle\mathcal{O}_k^*\mathcal{E}\rangle\,. \tag{1.107}$$

An analogous calculations yields the other derivatives needed to compute $\frac{\partial \mathcal{L}}{\partial \alpha_k^*}$:

$$\frac{\partial \langle \mathcal{O}_{k'} \rangle}{\partial \alpha_k^*} = \langle \mathcal{O}_k^* \mathcal{O}_{k'} \rangle \ , \quad \frac{\partial \langle \mathcal{O}_{k'}^* \rangle}{\partial \alpha_k^*} = \langle \mathcal{O}_k^* \mathcal{O}_{k'}^* \rangle \ . \tag{1.108}$$

And plugging Eq. (1.107) and Eq. (1.108) into Eq. (1.101), we get

$$\frac{\partial \mathcal{L}}{\partial \alpha_k^*} = \frac{i\hbar}{2} \left[\sum_{k'} \langle \mathcal{O}_k^* \mathcal{O}_{k'} \rangle \, \dot{\alpha}_{k'} - \sum_{k'} \langle \mathcal{O}_k^* \mathcal{O}_{k'}^* \rangle \, \dot{\alpha}_{k'}^* \right] - \langle \mathcal{O}_k^* \mathcal{E} \rangle \ . \tag{1.109}$$

Now we have all the ingredients needed to write down the first one of the Euler-Lagrange equations in (1.94) in terms of the time derivatives of the variational parameters. Indeed, plugging Eq. (1.100) and Eq. (1.109) into the first equation in (1.94), we get

$$-i\hbar \sum_{k'} \langle \mathcal{O}_k^* \mathcal{O}_{k'} \rangle \, \dot{\alpha}_{k'} + \langle \mathcal{O}_k^* \mathcal{E} \rangle = 0, \tag{1.110}$$

which is equivalent to the t-VMC equations of motion (1.13). The second Euler-Lagrange equation turns out to simply be the complex conjugate of the first one.

1.2 Solving the t-VMC equations

A t-VMC or τ-VMC simulation requires the performing of the steps that are outlined by the pseudocode in Algorithm 1. There two important steps that need to be addressed. The first one is solving the linear system to obtain the time derivatives of the parameters $\dot{\boldsymbol{\alpha}}$: the stochastic noise involved in the Monte Carlo estimations and the redundancy that may be present in the quantum state parametrization can lead to instabilities. We discuss this difficulty and the available methods to tackle it in Subsec. 1.2.1. The second point that we discuss is the numerical solution of the equations of motion by means of an ordinary differential equation (ODE) integrator. This is addressed in Subsec. 1.2.2.

1.2.1 Solution of the linear system

We need to solve the linear system of equations (1.17) for a real-time simulation, or the similar system (1.74) for imaginary time. The formal solution of (1.17) is

$$\dot{\boldsymbol{\alpha}} = -iS^+ F \, , \tag{1.111}$$

where S^+ is the Moore-Penrose pseudoinverse of the quantum geometric tensor S [27]. The pseudoinverse is obtained by means of a singular value decomposition, that, in the case of a positive semi-definite matrix like S, is equivalent to an eigendecomposition:

$$S_{k,k'} = \sum_l V_{k,l} \sigma_l^2 V_{l,k'}^\dagger \, , \tag{1.112}$$

where σ_l^2 are the eigenvalues of S, and $V_{k,l}$ is the unitary matrix whose columns are the orthonormalized eigenvectors of S. Then the pseudoinverse is defined as

$$S^+_{k,k'} = \sum_l V_{k,l} \zeta_l^2 V_{l,k'}^\dagger \, , \quad \text{with } \zeta_l^2 = \begin{cases} 1/\sigma_l^2 & \text{if } \sigma_l^2 > 0 \\ 0 & \text{if } \sigma_l^2 = 0 \end{cases} . \tag{1.113}$$

Algorithm 1: Standard tVMC algorithm

Data: $e^{\phi(t=0)}\,|\psi(\boldsymbol{\alpha}(t=0))\rangle$.
Result: $\boldsymbol{\alpha}(t_n)$ and $\phi(t_n)$, with $t_n = n\,\delta_t$, where δ_t is the time step of the simulation.
while $t < t_n$ **do**
  Use the Metropolis algorithm to sample the configurations $\mathbf{x}$ from the probability distribution $\left|\,\langle \mathbf{x}|\psi(\boldsymbol{\alpha}(t))\rangle\,\right|^2$.
  Use the samples to compute the expectation values $\langle \mathcal{O}_k\rangle$, $\langle \mathcal{O}_k^* \mathcal{O}_{k'}\rangle$, $\langle \mathcal{E}\rangle$, $\langle \mathcal{O}_k^* \mathcal{E}\rangle$, and any observable of interest.
  if Real-time evolution **then**
    Find $\dot{\boldsymbol{\alpha}}$ and $\dot{\phi}$ according to Eqs. 1.17.
  else Imaginary-time evolution
    Find $\dot{\boldsymbol{\alpha}}$ and $\dot{\phi}$ according to Eqs. 1.74.
  end
  Compute $\boldsymbol{\alpha}(t+\delta t)$ and $\phi(t+\delta t)$ using an ODE integrator (see Subsec. 1.2.2).
  Set $t = t + \delta_t$.
end

The pseudoinverse is used in place of the inverse because the S matrix may have zero eigenvalues. This may happen because an Ansatz with a large number of variational parameters generally can give a redundant description of the quantum state. This means that, given a set of parameters $\boldsymbol{\alpha}$, change of these parameters along some directions may only affect gauge degrees of freedom of the quantum state [32]. The irrelevant directions correspond to the kernel of the S matrix, i.e., the eigenspace of S corresponding to the zero eigenvector. In order to obtain the minimum-norm solution $\dot{\boldsymbol{\alpha}}$, these components belonging to the kernel are set to zero through the definition of ζ_l^2 in Eq. (1.113).

Besides the existence of zero eigenvalues, redundancy in the parametrization also has another consequence: the nonzero part of the spectrum of S spans many orders of magnitude. As a consequence, the condition number of the linear system (the ratio between the largest and smallest eigenvalues of S) is high. This means that small perturbations in the F vector, due to the stochastic noise in the Monte Carlo averages, get amplified in the solution S^+F [32], since the small eigenvalues σ_l^2 of S lead to large eigenvalues ζ_l^2 in S^+. This generates instabilities that can in principle be eliminated by increasing the number of Monte Carlo samples. However, the computational cost of doing so is not always affordable, so the instabilities need to be addressed through a *regularization method*. We now consider three such methods: two of them address the problem by modifying the S matrix in such a way as to reduce the noise-amplifying effect of S^+; the third method we consider directly addresses the noise in the F vector, in order to suppress those components of $\dot{\boldsymbol{\alpha}}$ that are dominated by noise.

Diagonal-shift regularization

The dynamics can be regularized by adding a diagonal shift to the S matrix [27]:

$$\tilde{S} = S + \epsilon \mathbb{1}\,, \quad \text{with } \epsilon > 0\,. \tag{1.114}$$

Since S is positive semi-definite, this makes the eigenvalue spectrum of S bounded from below by ϵ, reducing the condition number.

This method biases the dynamics, but this is acceptable for imaginary-time evolution, where we are only interested in the final optimized state. Indeed, the equilibrium points $\bar{\boldsymbol{\alpha}}$ of the imaginary-time dynamics are identified by $F(\bar{\boldsymbol{\alpha}}) = 0$, and the transformation (1.114) does not change this equilibrium condition, meaning that the stationary points are unchanged too.

However, this regularization modifies all of the spectrum of S, so it is not ideal when the details of the evolution are important, like in a real-time simulation, in which every time step is important for the future dynamics. In this case other methods are preferable.

Threshold-regularized pseudoinverse

In order to reduce the noise-amplifying effects of the pseudoinverse S^+ without modifying all of its spectrum, we can only act on the more problematic eigenvalues ζ_l^2, i.e., on the largest ones, corresponding to the smallest eigenvalues σ_l^2 of S. To this end, calling σ_1^2 the largest eigenvalue of S, we choose a threshold $\lambda < 1$ and truncate all the eigenvalues σ_l^2 that are smaller than $\lambda\,\sigma_1^2$. So the regularized pseudoinverse is

$$\tilde{S}^+_{k,k'} = \sum_l V_{k,l}\tilde{\zeta}_l^2 V^\dagger_{l,k'}\,, \quad \text{with } \tilde{\zeta}_l^2 = \begin{cases} 1/\sigma_l^2 & \text{if } \sigma_l^2 > \lambda\,\sigma_1^2 \\ 0 & \text{otherwise} \end{cases}\,. \tag{1.115}$$

Of course there is trade-off between stability and accuracy: too high a λ will generate an inaccurate dynamics, while too small a λ does not solve the instability problem [32].

Signal-to-noise ratio (SNR) regularization

A more sophisticated approach consists in evaluating the size of the noise in the equations of motion, in order to suppress the components of the solution $\dot{\boldsymbol{\alpha}}$ that are too noisy.

Following Ref. [33], let's first focus on the noise in the S matrix. We may be concerned that, since its eigenvalues σ_l^2 are affected by Monte Carlo noise, the smallest ones could be more prone to this disturbance, worsening the problem of the low condition number. Let's consider this possibility. We define the quantities $\mathcal{Q}_k$, the local operator's fluctuations in the basis that diagonalizes S, as

$$\mathcal{Q}_k(\mathbf{x}) = \sum_{k'} V_{k,k'}\left(\mathcal{O}_{k'}(\mathbf{x}) - \langle\mathcal{O}_{k'}\rangle\right)\,. \tag{1.116}$$

Then it is readily shown, from the definition of S (1.18) and the diagonalization (1.112), that the eigenvalues of S are given by

$$\sigma_l^2 = \langle\mathcal{Q}_l^*\,\mathcal{Q}_l\rangle = \langle|\mathcal{Q}_l|^2\rangle\,. \tag{1.117}$$

Now we assume that the $\mathcal{Q}_k$ quantities behave like gaussian random variables. This implies the Gaussian property

$$\langle|\mathcal{Q}_l|^4\rangle = 3\sigma_l^4\,, \tag{1.118}$$

from which we can estimate the signal-to-noise ratio (SNR) for the eigenvalues of S, as

$$\mathrm{SNR}(\sigma_l^2) = \frac{\sigma_l^2}{\sqrt{\left(\langle|Q_l|^4\rangle - \sigma_l^4\right)/N_{\mathrm{MC}}}} = \frac{\sigma_l^2\sqrt{N_{\mathrm{MC}}}}{\sqrt{2\sigma_l^4}} = \sqrt{\frac{N_{\mathrm{MC}}}{2}}\,, \tag{1.119}$$

where N_{MC} is the number of Monte Carlo samples. The result is that the SNR of the eigenvalues of S is the same along the whole spectrum, so that the smallest ones are not

more noisy than the larger ones. For this reason, we conclude that the most problematic source of noise is the Monte Carlo evaluation of the F vector, which we now analyze. First of all, we switch to the representation of F in the basis in which S is diagonal, i.e., the vector

$$\rho_k = \sum_{k'} V^\dagger_{k,k'} F_{k'} \,. \tag{1.120}$$

The representation $\dot{\boldsymbol{\eta}}$ of the solution vector $\dot{\boldsymbol{\alpha}}$ in the same basis is defined by

$$\dot{\eta}_k = \sum_{k'} V^\dagger_{k,k'} \dot{\alpha}_{k'} \,. \tag{1.121}$$

We can then express the component-by-component solution of the t-VMC equations of motion as

$$\dot{\eta}_k = -i \frac{\rho_k}{\sigma_k^2} \,, \quad \forall k. \tag{1.122}$$

To evaluate the SNR of ρ_k, we divide our Monte Carlo samples in N_{blocks} blocks of equal size, and evaluate the components ρ_k in each block, obtaining $[\rho_k]_m$, with $m = 1, \ldots, N_{\text{blocks}}$. Then we compute the variance of these evaluations of ρ_k over the blocks $\mathrm{Var}([\rho_k])$, and estimate the SNR for the final value of ρ_k (involving the whole sampling) as

$$\mathrm{SNR}(\rho_k) = \frac{|\rho_k|}{\sqrt{\mathrm{Var}([\rho_k])/N_{\text{blocks}}}} \,. \tag{1.123}$$

We choose a threshold λ_{SNR} and regularize the solution (1.122) as

$$\dot{\eta}_k = \frac{-i\,(\sigma_k^2)^{-1} \rho_k}{1 + \left(\frac{\lambda_{\text{SNR}}}{\mathrm{SNR}(\rho_k)} \right)^6} \,, \quad \forall k. \tag{1.124}$$

This suppresses the components whose SNR is lower than λ_{SNR}, and the suppression factor is continuous as the value of $\mathrm{SNR}(\rho_k)$ changes, so that the components do not get switched on and off all of a sudden.

There may be situations in which the assumption that the $\mathcal{Q}_k$ quantities behave like gaussian random variables is not satisfied. This is the case of the Shadow Ansatzë that we introduce in Chap. 2 and Chap. 3. In such cases we measure the variance $\mathrm{Var}([\sigma_k^2])$ of the evaluations of the eigenvalues σ_k^2 over the blocks, just like we do to compute $\mathrm{Var}([\rho_k])$. Then, we need an approximate estimation of the variance of the quantity ρ_k/σ_k^2 that appears in the solution of the equations of motion, Eq. (1.122). To obtain such an estimation, we consider $[\rho_k]$ and $[\sigma_k^2]$ to be independent, which implies

$$\begin{aligned} \frac{1}{\mathrm{SNR}(\rho_k/\sigma_k^2)} = \frac{\sqrt{\mathrm{Var}(\rho_k/\sigma_k^2)}}{\langle \rho_k/\sigma_k^2 \rangle} &\simeq \frac{\sqrt{\mathrm{Var}([\rho_k])}}{\langle \rho_k \rangle} + \frac{\sqrt{\mathrm{Var}([\sigma_k^2])}}{\langle \sigma_k^2 \rangle} \\ &= \frac{1}{\mathrm{SNR}(\rho_k)} + \frac{1}{\mathrm{SNR}(\sigma_k^2)} \,. \end{aligned} \tag{1.125}$$

We estimate $\mathrm{SNR}(\rho_k/\sigma_k^2)$ in this way rather than directly measuring the variance of $[\rho_k]/[\sigma_k^2]$, because the division by $[\sigma_k^2]$ is dangerous: stochastic fluctuations may bring the value of $[\sigma_k^2]$ in a block close to zero, leading to abnormally large fluctuations in the ratio that invalidate the calculation. We regularize the t-VMC solution as

$$\dot{\eta}_k = \frac{-i\,(\sigma_k^2)^{-1} \rho_k}{1 + \left(\frac{\lambda_{\text{SNR}}}{\mathrm{SNR}(\rho_k/\sigma_k^2)} \right)^6} \simeq \frac{-i\,(\sigma_k^2)^{-1} \rho_k}{1 + \left(\frac{\lambda_{\text{SNR}}}{\mathrm{SNR}(\rho_k)} + \frac{\lambda_{\text{SNR}}}{\mathrm{SNR}(\sigma_k^2)} \right)^6} \,, \quad \forall k. \tag{1.126}$$

In general $[\rho_k]$ and $[\sigma_k^2]$ are not independent, and, by considering them to be so, we usually overestimate the SNR of their ratio. However we are only interested in a loose approximation of the SNR to be used in the regularization. The overestimation of the SNR results in a regularization that is stronger than necessary, but this can be fixed by lowering the threshold λ_{SNR}.

1.2.2 ODE integrators

Here we consider the general problem of integrating a system of coupled ordinary differential equations (ODE) of the first order:

$$\dot{\boldsymbol{\alpha}}(t) = f(\boldsymbol{\alpha}(t), t)\,, \tag{1.127}$$

where $\boldsymbol{\alpha}$ is the vector of the variational parameters. In this work we use two ODE integrators, the Euler method and the Heun method.

Euler method

The Euler method is the simplest ODE integrator. It is a first-order integrator, in the sense that it is based on a first-order Taylor expansion of the parameters at time $t + \delta t$:

$$\begin{aligned}\boldsymbol{\alpha}(t+\delta t) &= \boldsymbol{\alpha}(t) + \dot{\boldsymbol{\alpha}}\,\delta t + O(\delta t^2) \\ &= \boldsymbol{\alpha}(t) + f(\boldsymbol{\alpha}(t), t)\,\delta t + O(\delta t^2)\,.\end{aligned} \tag{1.128}$$

The Euler update rule is simply obtained by truncating the higher orders. Denoting by $\boldsymbol{\alpha}_t$ the parameters at time t, the Euler method prescribes that the parameters at the next time step are

$$\boldsymbol{\alpha}_{t+\delta t} = \boldsymbol{\alpha}_t + f(\boldsymbol{\alpha}_t, t)\,\delta t\,. \tag{1.129}$$

We truncated a term $O(\delta t^2)$, so the error made at each time step is proportional to the square of the time-step size δt.

This is the least accurate method, but it is also the fastest, and it is routinely used in optimization by imaginary-time propagation, where we are only interested in the final optimized set of parameters, and the details of the evolution are not important.

Heun method

When more accuracy is needed, higher-order methods are available, that are more appropriate when the details of the evolution are important, like in real-time propagation. One such method is the Heun method. It is a predictor-corrector method that first performs an auxiliary Euler-like step, then evaluates the function f in the new set of parameters and uses this information to correct the parameters themselves. The full update formula reads

$$\begin{aligned}\tilde{\boldsymbol{\alpha}}_{t+\delta t} &= \boldsymbol{\alpha}_t + f(\boldsymbol{\alpha}_t, t)\,\delta t\,, \\ \boldsymbol{\alpha}_{t+\delta t} &= \boldsymbol{\alpha}_t + \frac{1}{2}\left[f(\boldsymbol{\alpha}_t, t) + f(\tilde{\boldsymbol{\alpha}}_{t+\delta t}, t+\delta t)\right]\,\delta t\,.\end{aligned} \tag{1.130}$$

A Taylor expansion shows that the error made with this update is $O(\delta t^3)$, so the Heun method is a second-order integrator.

1.2.3 Adaptive time step

The size of the error made with a certain ODE integrator can be estimated by varying step sizes. This idea can be used to adaptively adjust the time step so that the integration accuracy is kept within a certain threshold along the computation. Here we follow the Supplemental Material of Ref. [33].

Consider an ODE of the first order

$$\dot{y}(t) = f(y(t)) \tag{1.131}$$

and an ODE integrator of order n. We know that the error made at each time step is proportional to δt^{n+1}. So, if $y(t+\delta t)$ denotes the exact solution of a one-step update, and $y_{t+\delta t}$ denote the approximation yielded by the ODE integrator, we have

$$y_{t+\delta t} = y(t+\delta t) + c\,\delta t^{n+1}\,, \tag{1.132}$$

with an unkown constant c. If, alternatively, we take two steps of size $\delta t/2$, we get

$$y'_{t+\delta t} = y(t+\delta t) + 2c\left(\frac{\delta t}{2}\right)^{n+1}\,. \tag{1.133}$$

Taking the difference of the two solutions, we obtain

$$\Delta y_{t+\delta t} = \left\| y_{t+\delta t} - y'_{t+\delta t} \right\| \tag{1.134}$$

$$= \left\| \frac{2^n - 1}{2^n}\, c\,\delta t^{n+1} \right\| \propto \delta t^{n+1}\,, \tag{1.135}$$

which can be used to estimate the magnitude of the integration error. Given a tolerance λ, representing a value of $\Delta y_{t+\delta t}$ that we consider acceptable, we can adjust the time step to be

$$\delta t' = \delta t\left(\frac{\lambda}{\Delta y_{t+\delta t}}\right)^{\frac{1}{n+1}}\,. \tag{1.136}$$

This way, when $\Delta y_{t+\delta t}$ is larger than the tolerance λ, the time step is reduced to lower the error; conversely, when $\Delta y_{t+\delta t}$ is smaller than λ, the time step is increased, since we can tolerate a larger error. So, $\Delta y_{t+\delta t}$ is kept close to the value λ.

In our case the variable y is the vector of the parameters $\boldsymbol{\alpha}$. We only need to define how to measure the difference $\Delta\boldsymbol{\alpha}_{t+\delta t}$, and it is natural to use the norm induced by the matrix S, which amounts to a Fubini-Study distance, as discussed in Subsec. 1.1.3. Following again Ref. [33], we normalize it by the number of parameters P, as

$$\Delta\boldsymbol{\alpha}_{t+\delta t} = \left\| \boldsymbol{\alpha}_{t+\delta t} - \boldsymbol{\alpha}'_{t+\delta t} \right\|\,, \qquad \text{with } \|x\| \equiv \frac{1}{P}\sqrt{\sum_{k,k'} x_k^* S_{k,k'} x_{k'}}\,. \tag{1.137}$$

Estimating this quantity requires the computation of both $\boldsymbol{\alpha}_{t+\delta t}$ and $\boldsymbol{\alpha}'_{t+\delta t}$, so we always need to perform an extra calculation of the value of the parameters with time step $\delta t/2$. This may look like a waste, since we are making two updates of the parameters, one less accurate (one time step δt) and the other one more accurate (two time steps $\delta t/2$), but the numerical dynamics uses the less accurate one. Equivalently, one could perform the extra calculation using time step $2\delta t$, thus reversing the roles of $\boldsymbol{\alpha}_{t+\delta t}$ and $\boldsymbol{\alpha}'_{t+\delta t}$, but leaving the adaptive update of δt given in Eq. (1.136) unchanged. In general, we choose the latter option.

1.3 Some variational Ansatzë for spin systems

In this section we briefly describe two successful variational Ansatzë for spin systems.

1.3.1 The spin-Jastrow Ansatz

Jastrow wave functions are a class of variational Ansatzë that only include pairwise correlations between the degrees of freedom of the system. In the case of a spin system, the so-called spin-Jastrow wave function [34] reads

$$\psi_{\text{Jas}}(\boldsymbol{\sigma}) = \exp\left(\sum_{i,j} \eta_{i,j}\sigma_i\sigma_j\right), \tag{1.138}$$

where $\boldsymbol{\sigma}$ is the spin configuration, and $\eta_{i,j}$ are the variational parameters. If we do not impose symmetry constraints on the wave function (like, e.g., translation invariance), the number of parameters for N spins is $N(N-1)/2$.

Only including pairwise correlations could look like a harsh approximation, but these wave functions have proven to be effective in describing the qualitative properties of quantum systems, and they are lightweight on the computational side.

1.3.2 The restricted Boltzmann machine Ansatz (RBM)

In recent years, machine learning and its applications have been on the rise in many fields, including fundamental research. In the context of the quantum many-body problem, it has been suggested to describe the quantum state by means of a neural network, introducing the variational class of the so-called *neural-network quantum states* (NNQS) [27]. This approach exploits the high expressiveness of neural-network architectures to overcome the problem of the exponential wall in the description of quantum many-body systems.

One such architecture is the restricted Boltzmann machine (RBM), used in Ref. [27] to demonstrate the effectiveness of NNQS wave functions. It is an artificial neural network made of one "visible" layer of N neurons that is fully connected to a "hidden" layer of M neurons. The visible neurons in our case represent the physical spins, that can assume the discrete values of up (+1) or down (-1). The hidden are a set of non-physical auxiliary spins, also assuming the values of +1 and -1. When we say that the layers are fully connected between each other, we mean that every visible spin is pairwise correlated in a Jastrow-like fashion with each hidden spin. A summation over all the possible configurations of the hidden spins is performed to obtain the probability amplitude. The wave function reads

$$\psi_{\text{RBM}}(\boldsymbol{\sigma}) = \sum_{\boldsymbol{h}} \exp\left(\sum_{j=1}^{N} a_j\sigma_j + \sum_{k=1}^{M} b_k h_k + \sum_{j=1}^{N}\sum_{k=1}^{M} W_{k,j}h_k\sigma_j\right), \tag{1.139}$$

where the σ_j are the visible variables and the h_k are the hidden variables. The variational parameters are the visible biases a_j, acting on single visible spins, the hidden biases b_k, acting on single hidden spins, and the weights $W_{j,k}$, connecting the two layers. The total number of parameters is $N + M + N \times M$.

Since this architecture, only allowing connection between the two layers, does not have any intra-layer connection, the sum over the hidden variables can be carried out

analytically, yielding

$$\psi_{\text{RBM}}(\boldsymbol{\sigma}) = \exp\left(\sum_{j=1}^{N} a_j \sigma_j\right) \times \prod_{k=1}^{M} 2\cosh\left(b_k + \sum_{j=1}^{N} W_{k,j}\sigma_j\right) . \tag{1.140}$$

This greatly simplify the task of evaluating the probability amplitude for a certain spin configuration $\boldsymbol{\alpha}$. The ratio $d = M/N$ is called the RBM *density*, and, as demonstrated in Ref. [27], the expressiveness of the RBM for a fixed N greatly increases with the density.

If the system is known to respect some symmetry, usually it is possible to greatly simplify its description by exploiting it to impose constraints on the variational Ansatz. A way to define a constrained RBM that is invariant under a certain group of symmetries is explained in the Supplementary Material of Ref. [27]. Say that the system is known to be invariant under the group of symmetry transformations $\{T_s\}_{s=1}^{S}$. We denote by $\sigma_j(s)$ the j^{th} spin in the configuration $T_s\boldsymbol{\sigma}$, i.e., the configuration obtained applying the transformation T_s to the configuration $\boldsymbol{\sigma}$. In order to define a symmetry-invariant RBM we organize the hidden variables in groups, each one having S variables (that is the cardinality of the symmetry group). We denote the hidden variables by $h_{f,s}$, where the index f runs over the groups, and the index s runs over the members of each group. The symmetry-invariant RBM reads

$$\begin{aligned}\psi_{\text{RBM, symm}}(\boldsymbol{\sigma}) = \sum_{\boldsymbol{h}} \exp\Bigg(&\sum_{j=1}^{N} a_j \sum_{s=1}^{S} \sigma_j(s) + \sum_{f=1}^{F} b_f \sum_{s=1}^{S} h_{f,s} \\ &+ \sum_{f=1}^{F}\sum_{s=1}^{S}\sum_{j=1}^{N} W_{f,j} h_{f,s} \sigma_j(s)\Bigg) ,\end{aligned} \tag{1.141}$$

where we observe that hidden variables belonging to the same group of index f have the same bias b_f, and the same weights $W_{f,j}$, but differentiate in the way that the weights are connected to the visible variables: the s^{th} hidden variable is connected to the transformed configuration $T_s\boldsymbol{\sigma}$.

The index s runs up to S, the cardinality of the symmetry group, and the index f runs up to a value F that can be chosen in such a way to achieve a certain density of hidden neurons. The number of hidden neurons is, by construction, $F \times S$, so the density is $d = F \times S/N$. Since S is fixed by the symmetry group, and F is an integer, the density is constrained to be a multiple of S/N.

As an example, consider a 1D chain of N spins in periodic boundary conditions, and assume that it is invariant under the group of translations. There are N translations, including the identity, so $S = N$. Therefore the density is $d = F \times N/N = F$ and it is constrained to be an integer.

1.4 A proposal for the computation of dynamical structure factors

In this Section, we propose a method to use the t-VMC algorithm to compute the dynamical structure factor of a system. We consider a continuous system of N particles, and denote by $\mathbf{x}_i$ the coordinates of the i^{th} particle, and by $\mathbf{X}$ the vector comprising the coordinates of all the N particles. The dynamical structure factor of a system governed

by the Hamiltonian $\hat{H}$ is defined as

$$S(\mathbf{q},\omega) = \int_0^{+\infty} dt\; e^{i\omega t}\,\langle \hat{\rho}_{\mathbf{q}}(t)\,\hat{\rho}_{-\mathbf{q}}(0)\rangle = \int_0^{+\infty} dt\; e^{i\omega t}\,\langle\psi_0|\, e^{\frac{i}{\hbar}\hat{H}t}\,\hat{\rho}_{\mathbf{q}}\, e^{-\frac{i}{\hbar}\hat{H}t}\hat{\rho}_{-\mathbf{q}}\,|\psi_0\rangle\ , \tag{1.142}$$

where the expectation value is computed over ground state $|\psi_0\rangle$ of $\hat{H}$, and $\hat{\rho}_{\mathbf{q}}$ is the Fourier transform of the density operator $\hat{n}(\mathbf{x})$, that can be written as a function of the coordinates of all the particles, as

$$\rho_{\mathbf{q}}(\mathbf{X}) = \sum_i e^{-i\mathbf{q}\cdot\mathbf{x}_i}\,. \tag{1.143}$$

1.4.1 Rewriting the dynamical structure factor

Denoting by E_0 the ground-state energy, and by $\omega_0 = E_0/\hbar$ the corresponding frequency, we may rewrite the dynamical structure factor as

$$S(\mathbf{q},\omega) = \int_0^{+\infty} dt\, e^{i\omega t}\, e^{i\omega_0 t}\,\langle\psi_0|\,\rho_{\mathbf{q}}\, e^{-\frac{i}{\hbar}\hat{H}t}\,\rho_{-\mathbf{q}}\,|\psi_0\rangle\ . \tag{1.144}$$

We now insert an identity $\mathbb{1} = \int d\mathbf{X}\,|\mathbf{X}\rangle\,\langle\mathbf{X}|$, obtaining

$$S(\mathbf{q},\omega) = \int_0^{+\infty} dt\; e^{i\omega t}\; e^{i\omega_0 t} \int d\mathbf{X}\,\langle\psi_0|\hat{\rho}_{\mathbf{q}}|\mathbf{X}\rangle\,\langle\mathbf{X}|\, e^{-\frac{i}{\hbar}\hat{H}t}\hat{\rho}_{-\mathbf{q}}\,|\psi_0\rangle \tag{1.145}$$

Exploiting the fact that $\hat{\rho}_{\mathbf{q}}$ is diagonal in the coordinate basis, see Eq. (1.143), we can rewrite it as

$$\begin{aligned} S(\mathbf{q},\omega) &= \int_0^{+\infty} dt\; e^{i\omega t}\; e^{i\omega_0 t} \int d\mathbf{X}\,\langle\psi_0|\mathbf{X}\rangle\,\rho_{\mathbf{q}}(\mathbf{X})\,\langle\mathbf{X}|\, e^{-\frac{i}{\hbar}\hat{H}t}\hat{\rho}_{-\mathbf{q}}\,|\psi_0\rangle = \\ &= \int_0^{+\infty} dt\; e^{i\omega t}\; e^{i\omega_0 t} \int d\mathbf{X}\,|\langle\mathbf{X}|\psi_0\rangle|^2\,\frac{\rho_{\mathbf{q}}(\mathbf{X})\,\langle\mathbf{X}|\, e^{-\frac{i}{\hbar}\hat{H}t}\hat{\rho}_{-\mathbf{q}}\,|\psi_0\rangle}{\langle\mathbf{X}|\psi_0\rangle}\,. \end{aligned} \tag{1.146}$$

Now we define the time-dependent function

$$\xi_{-\mathbf{q}}(\mathbf{X},t) = \langle\mathbf{X}|\, e^{-\frac{i}{\hbar}\hat{H}t}\hat{\rho}_{-\mathbf{q}}\,|\psi_0\rangle\ , \tag{1.147}$$

that, for $t=0$, reduces to

$$\xi_{-\mathbf{q}}(\mathbf{X},t=0) = \langle\mathbf{X}|\hat{\rho}_{-\mathbf{q}}|\psi_0\rangle = \rho_{-\mathbf{q}}(\mathbf{X})\;\psi_0(\mathbf{X})\,. \tag{1.148}$$

Plugging the definition (1.147) into Eq. (1.146), and denoting by $\psi_0(\mathbf{X}) \equiv \langle\mathbf{X}|\psi_0\rangle$ the ground-state wave function, we rewrite the dynamical structure factor as

$$S(\mathbf{q},\omega) = \int_0^{+\infty} dt\; e^{i\omega t}\; e^{i\omega_0 t} \int d\mathbf{X}\,|\psi_0(\mathbf{X})|^2\,\frac{\rho_{\mathbf{q}}(\mathbf{X})\;\xi_{-\mathbf{q}}(\mathbf{X},t)}{\psi_0(\mathbf{X})}\,. \tag{1.149}$$

1.4.2 Variational approximation

Expression (1.149) is an exact way to write $S(\mathbf{q},\omega)$. Now we consider how it can be calculated approximately using a variational Ansatz and t-VMC. We focus on the estimation of the expectation value

$$A(\mathbf{q},t) = \int d\mathbf{X}\,|\psi_0(\mathbf{X})|^2\,\frac{\rho_{\mathbf{q}}(\mathbf{X})\;\xi_{-\mathbf{q}}(\mathbf{X},t)}{\psi_0(\mathbf{X})}\,. \tag{1.150}$$

If we know this quantity, then we can compute $S(\mathbf{q},\omega)$ as the Fourier transform of $e^{i\omega_0 t}\, A(\mathbf{q},t)$, according to Eq. (1.149)

Our goal is approximating $\psi_0(\mathbf{X})$ and $\xi_{-\mathbf{q}}(\mathbf{X},t)$ using the t-VMC algorithm, so that we can compute $A(\mathbf{q},t)$ as an expectation value over the probability distribution $|\psi_0(\mathbf{X})|^2$, which we can sample using the Metropolis algorithm.

Given a variational ansatz $|\psi(\boldsymbol{\alpha})\rangle$, where $\boldsymbol{\alpha}$ is the vector of the variational parameters, we can propagate it in imaginary time using t-VMC, and obtain an approximation of the ground state:

$$\psi_0(\mathbf{X}) \simeq e^{\phi^{(0)}} \psi(\mathbf{X};\boldsymbol{\alpha}^{(0)})\,, \tag{1.151}$$

where we explicitly factored out the complex phase ϕ.

To estimate $\xi_{-\mathbf{q}}(\mathbf{X},t)$, we start from its value at time $t=0$ given by Eq. (1.148), where we can substitute the exact ground state $|\psi_0\rangle$ with the variational one $|\psi(\boldsymbol{\alpha}^{(0)})\rangle$:

$$\xi_{-\mathbf{q}}(\mathbf{X},t=0) = \rho_{-\mathbf{q}}(\mathbf{X})\ \psi_0(\mathbf{X}) \simeq \rho_{-\mathbf{q}}(\mathbf{X}) \cdot e^{\phi^{(0)}} \psi(\mathbf{X};\boldsymbol{\alpha}^{(0)})\,. \tag{1.152}$$

This suggests the following Ansatz for the function $\xi_{-\mathbf{q}}(\mathbf{X},t)$:

$$\xi_{-\mathbf{q}}(\mathbf{X};\phi(t),\boldsymbol{\alpha}(t)) \stackrel{?}{=} \rho_{-\mathbf{q}}(\mathbf{X}) \cdot e^{\phi(t)} \psi(\mathbf{X};\boldsymbol{\alpha}(t))\,, \tag{1.153}$$

with the initial conditions $\boldsymbol{\alpha}(0) = \boldsymbol{\alpha}^{(0)}$ and $\phi(0) = \phi^{(0)}$. However, this Ansatz has a major drawback: the factor $\rho_{-\mathbf{q}}(\mathbf{X})$ introduces nodal surfaces, i.e., hypersurfaces in the coordinate space where the $\rho_{-\mathbf{q}}(\mathbf{X}) = 0$ and thus the whole function (1.153) vanishes; with this choice of the Ansatz, the nodal surfaces cannot move during the evolution, because the factor $\rho_{-\mathbf{q}}(\mathbf{X})$ does not depend on any variational parameter, and it is thus fixed. A way to overcome this problem is the introduction of a set of *backflow* coordinates $\mathbf{X}'$, defined by [35, 36, 37]

$$\mathbf{x}'_i = \mathbf{x}_i + \sum_j \eta(|\mathbf{x}_i - \mathbf{x}_j|;\boldsymbol{\beta})(\mathbf{x}_i - \mathbf{x}_j)\,, \tag{1.154}$$

where $\eta(r;\boldsymbol{\beta})$ is a backflow correlation function that depends on the distance r between two particles, and on a set of parameters $\boldsymbol{\beta}$. So the backflow coordinates $\mathbf{X}'$ are a function of the original coordinates $\mathbf{X}$ and of the parameters $\boldsymbol{\beta}$, i.e., $\mathbf{X}' = \mathbf{X}'(\mathbf{X};\boldsymbol{\beta})$.

We thus define a new Ansatz for $\xi_{-\mathbf{q}}(\mathbf{X},t)$, in which the factor $\rho_{-\mathbf{q}}(\mathbf{X})$ is modified by plugging in it the backflow coordinates in place of the ordinary ones:

$$\xi_{-\mathbf{q}}(\mathbf{X};\phi(t),\boldsymbol{\alpha}(t),\boldsymbol{\beta}(t)) = \rho_{-\mathbf{q}}(\mathbf{X}'(\mathbf{X};\boldsymbol{\beta}(t))) \cdot e^{\phi(t)} \psi(\mathbf{X};\boldsymbol{\alpha}(t))\,, \tag{1.155}$$

that depend on two sets of variational parameters: $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, plus the complex phase ϕ. The initial parameters $\phi^{(0)}$, $\boldsymbol{\alpha}^{(0)}$ are chosen so that $e^{\phi^{(0)}}|\psi(\boldsymbol{\alpha}^{(0)})\rangle$ is an approximation of the ground state. The initial parameters $\boldsymbol{\beta}^{(0)}$ are chosen so that $\mathbf{X}'(\mathbf{X};\boldsymbol{\beta}^{(0)}) \equiv \mathbf{X}$, which means that the backflow correlation function vanishes: $\eta(r,\boldsymbol{\beta}^{(0)}) \equiv 0$. This ensures that at time $t=0$, the Ansatz satisfies Eq. (1.152).

1.4.3 Calculation of the dynamical structure factor via t-VMC

To actually compute the dynamical structure factor $S(\mathbf{q},\omega)$ for a chosen set of wave vectors $\{\mathbf{q}_k\}$, we propose the following steps.

1. Perform an imaginary-time propagation of the wave function $e^{\phi}\psi(\mathbf{X};\boldsymbol{\alpha})$, to obtain the parameters $\phi^{(0)}$ and $\boldsymbol{\alpha}^{(0)}$ that best approximate the ground state.

2. For each wave vector $\mathbf{q}_k$, perform the real-time evolution of the Ansatz function $\xi_{-\mathbf{q}_k}(\mathbf{X}; \phi, \boldsymbol{\alpha}, \boldsymbol{\beta})$ defined in Eq. (1.155). The starting values of the parameters are $\phi^{(0)}$, $\boldsymbol{\alpha}^{(0)}$, $\boldsymbol{\beta}^{(0)}$, where the backflow parameters are chosen in such a way that $\eta(r, \boldsymbol{\beta}^{(0)}) \equiv 0$. Notice that we need to perform as many real-time simulations as the number of wave vectors $\mathbf{q}_k$ that we want to investigate. We denote by $\phi_k(t)$, $\boldsymbol{\alpha}_k(t)$, $\boldsymbol{\beta}_k(t)$ the evolution of the parameters for the k^{th} wave vector.

3. Compute $A(\mathbf{q}_k, t)$ according to Eq. (1.150). This can be done by using the Metropolis algorithm to sample the probability distribution defined by the squared modulus of the variational ground state $\left| e^{\phi^{(0)}} \psi(\mathbf{X}; \boldsymbol{\alpha}^{(0)}) \right|^2$. Then $A(\mathbf{q}_k, t)$ can be computed as the average of the function

$$\frac{\rho_{\mathbf{q}_k}(\mathbf{X})\ \xi_{-\mathbf{q}_k}(\mathbf{X}; \phi_k(t), \boldsymbol{\alpha}_k(t), \boldsymbol{\beta}_k(t))}{e^{\phi^{(0)}} \psi(\mathbf{X}; \boldsymbol{\alpha}^{(0)})} \tag{1.156}$$

   over the Monte Carlo samples. Notice that we need to sample the same probability distribution for all the wave vectors $\mathbf{q}_k$, so only one sampling needs to be performed, over which we can calculate the averages of the functions (1.156) for each $\mathbf{q}_k$.

4. Compute $S(\mathbf{q}_k, \omega)$ as the Fourier transform over time of $e^{i\omega_0 t} A(\mathbf{q}_k, t)$, see Eq. (1.149).

This method addresses the computation of the dynamical structure factor directly from the real-time evolution, rather than using analytic continuation on imaginary-time correlations. In this respect, it is similar to the approach described in Ref. [38], that simulates the system's response to a weak probe pulse. Compared to that technique, the method we propose has the difficulty of having to simulate more than one real-time evolution, namely one for every investigated wave vector $\mathbf{q}_k$. However, it may also have an advantage: the probe pulse used in Ref. [38] is position-dependent, so, even if the unperturbed Hamiltonian is translation-invariant, the wave function Ansatz has to include one-body correlation terms to describe the dependence of the wave function on the positions of single particles, and not only on their mutual distances. This significantly increases the number of parameters needed for an accurate simulation. In our proposal, on the other hand, one may choose an Ansatz wave function $|\psi(\boldsymbol{\alpha})\rangle$ that only includes two-body correlations, and let the one-body dependence of $\xi_{-\mathbf{q}_k}$ be optimized solely through the backflow terms included in $\rho_{-\mathbf{q}_k}(\mathbf{X}')$. If this choice proved to be effective, it could significantly reduce the number of parameters needed.

The method we presented was derived by us independently during this Thesis' research work, but is substantially equivalent to the technique published in 2023 by T. Mendes-Santos, M. Schmitt, and M. Heyl [23], which they apply to spin systems. The main difference is that, in our proposal, we consider continuous systems of particles, that imply itinerant degrees of freedom. This prompted us to include the use of backflow correlations in the description of the one-body term, that are not needed for non-itinerant degrees of freedom like the ones described in Ref. [23].

CHAPTER 2

Shadow wave functions (SWF)

One of the main results of our research work is the extension of the t-VMC method to the class of variational Ansatzë called the *shadow wave functions* (SWF). Shadow wave functions (SWF) address the problem of capturing the quantum many-body correlations by introducing auxiliary summation variables in the definition of the variational Ansatz. They were first introduced to extend the standard Jastrow form of quantum systems in the continuum like solid and liquid ^{4}He [10, 11]. In this case each real particle is coupled to its "shadow" counterpart by a harmonic potential, and it was shown that this coupling can be thought as a variationally-optimized single step of path integration [12]. A similar proposal for discrete systems was introduced by D. Baeriswyl and collaborators [39, 40, 41].

In Sec. 2.1 we introduce this class of variational wave functions and explain the rationale behind this particular form. In Sec. 2.2 we expand on this by explaining the link between SWFs and the path-integral formalism, which provides a strong physical argument for the introduction of this kind of Ansatz. In Sec. 2.3 we describe how to compute the expectation values of operators on a SWF state. Since we are going to use SWFs with complex parameters, in the computation of these expectation values we face a sign problem, so we also introduce the reweighting method we use to tackle this difficulty. In Sec. 2.4 we present one of our main results: the extension of the t-VMC technique to SWF Ansatzë.

2.1 The "shadow" wave function form

Denoting by $\mathbf{x}$ the coordinates that describe the system, and by $\mathbf{y}$ the auxiliary or "shadow" variables, a SWF for a discrete system reads

$$\psi(\mathbf{x}) = \sum_{\mathbf{y}} \Xi(\mathbf{x}, \mathbf{y}) \,, \tag{2.1}$$

where $\Xi(\mathbf{x}, \mathbf{y})$ is a function describing correlations among the real and shadow variables. For continuous systems the summation over $\mathbf{y}$ is replaced by an integral $\int d\mathbf{y}$.

The rationale behind this approach can be better understood through an example. Consider an interacting spin system, whose "coordinates" $\mathbf{x}$ are substituted by the spin configuration $\boldsymbol{\sigma}$. A well-studied Ansatz for such a system is the so-called spin-Jastrow wave function that we briefly discussed in Subsec. 1.3.1, which reads

$$\psi_{\text{Jas}}(\boldsymbol{\sigma}) = \exp\left(\sum_{i,j} \eta_{i,j}\sigma_i\sigma_j\right) \,, \tag{2.2}$$

where $\eta_{i,j}$ are the variational parameters. This wave function is rather lightweight on the computational side, but its descriptive power is limited in that it only includes pairwise correlations between spins. Many-body correlations can be included by means of a shadow wave function of the form

$$\psi(\boldsymbol{\sigma}) = \sum_{\boldsymbol{\sigma'}} \Xi(\boldsymbol{\sigma}, \boldsymbol{\sigma'}) \,,$$

$$\text{with} \quad \Xi(\boldsymbol{\sigma}, \boldsymbol{\sigma'}) = \exp\left(\sum_{i,j} \eta_{i,j}\sigma_i\sigma_j + \sum_i \omega_i \sigma_i \sigma'_i + \sum_{i,j} \rho_{i,j}\sigma'_i\sigma'_j \right) \,, \quad (2.3)$$

which keeps the pairwise correlations between real spins, and adds pairwise correlations between the shadow spins $\boldsymbol{\sigma'}$ (parameters $\rho_{i,j}$) and between each real spin σ_i and its shadow counterpart σ'_i (parameters ω_i). In this way, $\Xi(\boldsymbol{\sigma}, \boldsymbol{\sigma'})$ only includes Jastrow-like pairwise correlations, but the sum over the shadow variables $\boldsymbol{\sigma'}$ indirectly introduces many-body correlations in a non-perturbative manner. It is relevant to observe that this kind of structure, used in SWF, closely resembles a neural-network, as pointed out in Ref. [42]. It is similar to a RBM, but with intra-layer connections allowed.

2.2 The relation between SWF and path integration

Let us briefly explain the path integral approach to the ground state description [43], in order to highlight its connection with the SWF variational form [12]. We consider a system described by the degrees of freedom, or "coordinates", $\mathbf{x}$ and the Hamiltonian $\hat{H}$. Again, here we consider a discrete system, but the formulas for a continuous system can be readily obtained by replacing the summations $\sum_{\mathbf{x}}$ with integrals $\int d\mathbf{x}$.

We now set the stage for the path integral description of the ground state. We assume that $\hat{H}$ can be written as the sum $\hat{H} = \hat{K} + \hat{V}$, where $\hat{V}$ is an operator that is diagonal in the coordinates $\mathbf{x}$, and $\hat{K}$ an operator that is not diagonal. So, the operator $\hat{V}$ is described by some "local" function of the coordinates $V(\mathbf{x})$.

If $|\psi_T\rangle$ is a state that has nonzero overlap with the ground state $|E_0\rangle$, then we can write the imaginary-time projection formula (1.71) as

$$|E_0\rangle \propto e^{-\tau\hat{H}}|\psi_T\rangle \,, \quad (2.4)$$

for large enough τ. Thus, the non-normalized ground state wave function can be written

$$\Psi_0(\mathbf{x}) = \langle \mathbf{x}|E_0\rangle = \langle \mathbf{x}|e^{-\tau\hat{H}}|\psi_T\rangle \quad (2.5)$$

$$= \sum_{\mathbf{y}} \langle \mathbf{x}|e^{-\tau\hat{H}}|\mathbf{y}\rangle\langle \mathbf{y}|\psi_T\rangle \,, \quad (2.6)$$

where the propagator function $\langle \mathbf{x}|e^{-\tau\hat{H}}|\mathbf{y}\rangle$ appears. The path integral approach revolves around a clever way of writing this propagator function. To find it, we manipulate the propagator as follows:

$$e^{-\tau\hat{H}} = e^{-\tau\left(\hat{K}+\hat{V}\right)} = \left[e^{-\frac{\tau}{M}\left(\hat{K}+\hat{V}\right)}\right]^M \,, \quad (2.7)$$

where we rewrote it as the composition of many propagators for small imaginary time τ/M. Now we "Trotterize" the small-imaginary-time propagators:

$$e^{-\frac{\tau}{M}\left(\hat{K}+\hat{V}\right)} \simeq e^{-\frac{\tau}{M}\hat{V}}\, e^{-\frac{\tau}{M}\hat{K}} \,, \quad (2.8)$$

a relation that becomes exact in the limit $M \to \infty$. We plug an identity $\sum_{\mathbf{y}} |\mathbf{y}\rangle\langle\mathbf{y}|$ in the Trotterized propagator, as

$$e^{-\frac{\tau}{M}\hat{V}}\, e^{-\frac{\tau}{M}\hat{K}} = \sum_{\mathbf{y}} |\mathbf{y}\rangle\, e^{-\frac{\tau}{M}V(\mathbf{y})}\, \langle\mathbf{y}| e^{-\frac{\tau}{M}\hat{K}}\,, \tag{2.9}$$

Plugging the decomposition (2.7) and the "Trotterization" (2.9) into Eq. (2.6), we obtain

$$\Psi_0(\mathbf{x}) = \sum_{\mathbf{y}^1,\mathbf{y}^2,\ldots,\mathbf{y}^M} e^{-\frac{\tau}{M}V(\mathbf{x})} \left\langle \mathbf{x}\middle| e^{-\frac{\tau}{M}\hat{K}} \middle|\mathbf{y}^1\right\rangle e^{-\frac{\tau}{M}V(\mathbf{y}^1)} \left\langle \mathbf{y}^1\middle| e^{-\frac{\tau}{M}\hat{K}} \middle|\mathbf{y}^2\right\rangle \ldots$$
$$e^{-\frac{\tau}{M}V(\mathbf{y}^{M-1})} \left\langle \mathbf{y}^{M-1}\middle| e^{-\frac{\tau}{M}\hat{K}} \middle|\mathbf{y}^{M}\right\rangle \langle \mathbf{y}^M|\psi_T\rangle\,. \tag{2.10}$$

So we rewrote the ground state wave function in terms of a summation over several layers $\mathbf{y}^0,\ldots,\mathbf{y}^M$ of auxiliary variables. Within each layer, the auxiliary variables "interact" via the potential $V(\mathbf{y})$, as shown by the terms $e^{-\frac{\tau}{M}V(\mathbf{y})}$ that are reminiscent of the Boltzmann weight of a classical interacting system. Two consecutive layers are related by the propagator of the non-diagonal term, $\langle \mathbf{y}^i|e^{-\frac{\tau}{M}\hat{K}}|\mathbf{y}^{i+1}\rangle$. Each layer constitutes a single *path integration step*. This representation of the ground state is exact in the limit of an infinite number of such steps; however, in many quantum Monte Carlo approaches, a finite number of steps is sufficient to make the systematic error smaller than the stochastic uncertainties.

An interesting particular case is represented by a system of particles for which $\mathbf{x}$ represent the positions, $\hat{V}$ is the interaction energy and $\hat{K}$ is the kinetic energy. In this case, the propagator of the kinetic term $\langle \mathbf{y}^i|e^{-\frac{\tau}{M}\hat{K}}|\mathbf{y}^{i+1}\rangle$ implements an attractive harmonic interaction between each particle of the i^{th} layer and one particular particle of the $(i+1)^{\text{th}}$ layer.

Now, to understand the relation between path integration and shadow wave functions, we consider a single step of path integration:

$$\Psi_0(\mathbf{x}) = \sum_{\mathbf{y}} e^{-\frac{\tau}{M}V(\mathbf{x})} \left\langle \mathbf{x}\middle| e^{-\frac{\tau}{M}\hat{K}} \middle|\mathbf{y}\right\rangle \langle \mathbf{y}|\psi_T\rangle\,. \tag{2.11}$$

Comparing with Eq. (2.1), we clearly see that this is a SWF in which $\psi_T(\mathbf{y})$ describes the correlations between the shadow variables, $e^{-\frac{\tau}{M}V(\mathbf{x})}$ describes the correlations between the real variables, and the propagator $\left\langle \mathbf{x}\middle| e^{-\frac{\tau}{M}\hat{K}} \middle|\mathbf{y}\right\rangle$ describes the real-shadow correlations. So, the SWF paradigm has a strong physical root. A single step of path integration may look reductive, but in the SWF context the three terms describing correlations, that we just described, are not fixed by the formalism, like they are in the path integral case. Rather, they are functions dependent on a set of variational parameters that are subsequently optimized to minimize the energy. So a SWF can be seen as a single step of path integration which is enhanced by variational optimization to get as close as possible to the true ground state.

2.3 Expectation values on shadow wave functions by Monte Carlo sampling

Now we discuss the problem of using Monte Carlo sampling to compute the expectation value of some operator $\hat{A}$ on the state described by a SWF $\psi(\mathbf{x})$ of the form (2.1). For simplicity, we assume that the state is normalized. We can do this without loss of generality,

since the Metropolis algorithm allows the sampling of a probability distribution without the need of knowing its normalization.

If it were possible to analytically carry out the summation (or the integral, for continuous systems) on the shadow variables, we could simply evaluate $|\psi(\mathbf{x})|^2$ for any configuration $\mathbf{x}$ and perform the usual Metropolis sampling of the real variables $\mathbf{x}$. However, in many interesting cases an analytical summation over the auxiliary variables is not possible, hence the need of sampling both the real and the shadow variables [10]. Let's see how this can be carried out by computing the expectation value of an operator $\hat{A}$ over a shadow state:

$$\left\langle \hat{A} \right\rangle = \left\langle \psi \middle| \hat{A} \middle| \psi \right\rangle = \sum_{\mathbf{x},\mathbf{x}'} \psi^*(\mathbf{x}) \left\langle \mathbf{x} \middle| \hat{A} \middle| \mathbf{x}' \right\rangle \psi(\mathbf{x}') \,. \tag{2.12}$$

Now we expand ψ according to its definition (2.1), yielding

$$\left\langle \hat{A} \right\rangle = \sum_{\mathbf{x},\mathbf{x}'} \sum_{\mathbf{y},\mathbf{y}'} \Xi^*(\mathbf{x},\mathbf{y}) \left\langle \mathbf{x} \middle| \hat{A} \middle| \mathbf{x}' \right\rangle \Xi(\mathbf{x}',\mathbf{y}') \tag{2.13}$$

$$= \sum_{\mathbf{x},\mathbf{x}',\mathbf{y},\mathbf{y}'} \Xi^*(\mathbf{x},\mathbf{y}) \Xi(\mathbf{x},\mathbf{y}') \left\langle \mathbf{x} \middle| \hat{A} \middle| \mathbf{x}' \right\rangle \frac{\Xi(\mathbf{x}',\mathbf{y}')}{\Xi(\mathbf{x},\mathbf{y}')} \tag{2.14}$$

$$= \sum_{\mathbf{x},\mathbf{y},\mathbf{y}'} \Xi^*(\mathbf{x},\mathbf{y}) \Xi(\mathbf{x},\mathbf{y}') A^{\mathrm{sh}}(\mathbf{x},\mathbf{y}') \,, \tag{2.15}$$

where we introduced the *shadow local estimator* $A^{\mathrm{sh}}(\mathbf{x},\mathbf{y})$, which is defined as

$$A^{\mathrm{sh}}(\mathbf{x},\mathbf{y}) = \sum_{\mathbf{x}'} \left\langle \mathbf{x} \middle| \hat{A} \middle| \mathbf{x}' \right\rangle \frac{\Xi(\mathbf{x}',\mathbf{y})}{\Xi(\mathbf{x},\mathbf{y})} \tag{2.16}$$

and is a function of both the real and the shadow coordinates. Moreover, we define

$$\mathcal{P}(\mathbf{x},\mathbf{y},\mathbf{y}') = \Xi^*(\mathbf{x},\mathbf{y}) \Xi(\mathbf{x},\mathbf{y}') \,. \tag{2.17}$$

Since Ξ is usually chosen to be a real nonnegative function, $\mathcal{P}$ can be interpreted as a classical probability distribution in most applications. However, since later we are going to deal with a complex Ξ, for us $\mathcal{P}$ is in general a complex quantity that satisfies

$$\sum_{\mathbf{x},\mathbf{y},\mathbf{y}'} \mathcal{P}(\mathbf{x},\mathbf{y},\mathbf{y}') = \sum_{\mathbf{x}} |\psi(\mathbf{x})|^2 = 1 \,, \tag{2.18}$$

$$\mathcal{P}^*(\mathbf{x},\mathbf{y},\mathbf{y}') = \mathcal{P}(\mathbf{x},\mathbf{y}',\mathbf{y}) \,. \tag{2.19}$$

Using these definitions, the expectation value (2.15) of $\hat{A}$ becomes

$$\left\langle \hat{A} \right\rangle = \sum_{\mathbf{x},\mathbf{y},\mathbf{y}'} \mathcal{P}(\mathbf{x},\mathbf{y},\mathbf{y}') A^{\mathrm{sh}}(\mathbf{x},\mathbf{y}') \,. \tag{2.20}$$

If Ξ is a real nonnegative function, then $\mathcal{P}$ is too and can be interpreted as a classical probability distribution. This is the case in many application of SWFs, e.g., when they are used to simulate ground states of bosonic systems [10, 11]. In this case, we can use the Metropolis algorithm to sample a set of real variables $\mathbf{x}$ and two sets of shadow variables $\mathbf{y}$ and $\mathbf{y}'$, according to the probability $\mathcal{P}(\mathbf{x},\mathbf{y},\mathbf{y}')$. Then the expectation value of the

operator $\hat{A}$ is just the average over the samples of the quantity $A^{\text{sh}}(\mathbf{x},\mathbf{y}')$; or, even better, $\frac{1}{2}\left[A^{\text{sh}}(\mathbf{x},\mathbf{y})+A^{\text{sh}}(\mathbf{x},\mathbf{y}')\right]$, exploiting the symmetry properties of a real $\mathcal{P}$, see Eq. (2.19).

However, in our applications Ξ is complex, and $\mathcal{P}$ is too. The main reason is that we want to use SWFs in real-time evolution, so we need to use complex parameters. Then $\mathcal{P}(\mathbf{x},\mathbf{y},\mathbf{y}')$ is not interpretable as a classical probability distribution, and we face a type of sign problem, in some respects analogous to the sign problem of fermionic SWFs [44]. We tackle it using a reweighting technique, for which some more manipulation of Eq. (2.20) is needed. We start by writing $\mathcal{P}$ in terms of its modulus ρ and phase θ:

$$\mathcal{P}(\mathbf{x},\mathbf{y},\mathbf{y}')=\rho(\mathbf{x},\mathbf{y},\mathbf{y}')e^{i\theta(\mathbf{x},\mathbf{y},\mathbf{y}')} \tag{2.21}$$

From the property (2.19) of $\mathcal{P}$, it follows that ρ and θ satisfy

$$\rho(\mathbf{x},\mathbf{y},\mathbf{y}')=\rho(\mathbf{x},\mathbf{y}',\mathbf{y})\,, \tag{2.22}$$

$$\theta(\mathbf{x},\mathbf{y},\mathbf{y}')=-\theta(\mathbf{x},\mathbf{y}',\mathbf{y})\,. \tag{2.23}$$

Substituting $\mathcal{P}=\rho e^{i\theta}$ into the expectation value of $\hat{A}$ in Eq. (2.20), and exploiting the (anti-)symmetry properties of ρ and θ under exchange of $\mathbf{y}$ and $\mathbf{y}'$, we write

$$\left\langle\hat{A}\right\rangle=\sum_{\mathbf{x},\mathbf{y},\mathbf{y}'}\rho(\mathbf{x},\mathbf{y},\mathbf{y}')e^{i\theta(\mathbf{x},\mathbf{y},\mathbf{y}')}A^{\text{sh}}(\mathbf{x},\mathbf{y}') \tag{2.24}$$

$$=\frac{1}{2}\left[\sum_{\mathbf{x},\mathbf{y},\mathbf{y}'}\rho(\mathbf{x},\mathbf{y},\mathbf{y}')e^{i\theta(\mathbf{x},\mathbf{y},\mathbf{y}')}A^{\text{sh}}(\mathbf{x},\mathbf{y}')+\sum_{\mathbf{x},\mathbf{y},\mathbf{y}'}\rho(\mathbf{x},\mathbf{y}',\mathbf{y})e^{i\theta(\mathbf{x},\mathbf{y}',\mathbf{y})}A^{\text{sh}}(\mathbf{x},\mathbf{y})\right] \tag{2.25}$$

$$=\sum_{\mathbf{x},\mathbf{y},\mathbf{y}'}\rho(\mathbf{x},\mathbf{y},\mathbf{y}')\frac{1}{2}\left[e^{i\theta(\mathbf{x},\mathbf{y},\mathbf{y}')}A^{\text{sh}}(\mathbf{x},\mathbf{y}')+e^{-i\theta(\mathbf{x},\mathbf{y},\mathbf{y}')}A^{\text{sh}}(\mathbf{x},\mathbf{y})\right] \tag{2.26}$$

$$=\sum_{\mathbf{x},\mathbf{y},\mathbf{y}'}\rho(\mathbf{x},\mathbf{y},\mathbf{y}')A^{\text{sh, rew}}(\mathbf{x},\mathbf{y},\mathbf{y}')\,, \tag{2.27}$$

where we introduced the *reweighted shadow local estimator* $A^{\text{sh, rew}}$ (the reason for this naming will be clear later), defined as

$$A^{\text{sh, rew}}(\mathbf{x},\mathbf{y},\mathbf{y}')=\frac{1}{2}\left[e^{i\theta(\mathbf{x},\mathbf{y},\mathbf{y}')}A^{\text{sh}}(\mathbf{x},\mathbf{y}')+e^{-i\theta(\mathbf{x},\mathbf{y},\mathbf{y}')}A^{\text{sh}}(\mathbf{x},\mathbf{y})\right]\,. \tag{2.28}$$

Since $\rho(\mathbf{x},\mathbf{y},\mathbf{y}')$ is real and nonnegative, Eq. (2.27) looks like an expectation value. But unfortunately it is not, because ρ is not normalized. Denoting by $\mathcal{N}=\sum_{\mathbf{x},\mathbf{y},\mathbf{y}'}\rho(\mathbf{x},\mathbf{y},\mathbf{y}')$ the normalization of ρ, we define the classical probability distribution

$$q(\mathbf{x},\mathbf{y},\mathbf{y}')=\frac{1}{\mathcal{N}}\rho(\mathbf{x},\mathbf{y},\mathbf{y}')\,. \tag{2.29}$$

In order to write Eq. (2.27) in terms of expectation values on the probability distribution q, we divide it by $1=\sum_{\mathbf{x},\mathbf{y},\mathbf{y}'}\mathcal{P}(\mathbf{x},\mathbf{y},\mathbf{y}')$, yielding

$$\left\langle\hat{A}\right\rangle=\frac{\sum_{\mathbf{x},\mathbf{y},\mathbf{y}'}\rho(\mathbf{x},\mathbf{y},\mathbf{y}')A^{\text{sh, rew}}(\mathbf{x},\mathbf{y},\mathbf{y}')}{\sum_{\mathbf{x},\mathbf{y},\mathbf{y}'}\mathcal{P}(\mathbf{x},\mathbf{y},\mathbf{y}')} \tag{2.30}$$

$$=\frac{\sum_{\mathbf{x},\mathbf{y},\mathbf{y}'}\rho(\mathbf{x},\mathbf{y},\mathbf{y}')A^{\text{sh, rew}}(\mathbf{x},\mathbf{y},\mathbf{y}')}{\sum_{\mathbf{x},\mathbf{y},\mathbf{y}'}\rho(\mathbf{x},\mathbf{y},\mathbf{y}')e^{i\theta(\mathbf{x},\mathbf{y},\mathbf{y}')}}\,, \tag{2.31}$$

and then we divide both numerator and denominator by the normalization $\mathcal{N}$ of ρ:

$$\left\langle \hat{A} \right\rangle = \frac{\frac{1}{\mathcal{N}} \sum_{\mathbf{x},\mathbf{y},\mathbf{y}'} \rho(\mathbf{x},\mathbf{y},\mathbf{y}') A^{\text{sh, rew}}(\mathbf{x},\mathbf{y},\mathbf{y}')}{\frac{1}{\mathcal{N}} \sum_{\mathbf{x},\mathbf{y},\mathbf{y}'} \rho(\mathbf{x},\mathbf{y},\mathbf{y}') e^{i\theta(\mathbf{x},\mathbf{y},\mathbf{y}')}} \tag{2.32}$$

$$= \frac{\sum_{\mathbf{x},\mathbf{y},\mathbf{y}'} q(\mathbf{x},\mathbf{y},\mathbf{y}') A^{\text{sh, rew}}(\mathbf{x},\mathbf{y},\mathbf{y}')}{\sum_{\mathbf{x},\mathbf{y},\mathbf{y}'} q(\mathbf{x},\mathbf{y},\mathbf{y}') e^{i\theta(\mathbf{x},\mathbf{y},\mathbf{y}')}} \tag{2.33}$$

$$= \frac{\left\langle A^{\text{sh, rew}} \right\rangle_q}{\left\langle e^{i\theta} \right\rangle_q} = \frac{\left\langle A^{\text{sh, rew}} \right\rangle_q}{\left\langle \cos\theta \right\rangle_q} , \tag{2.34}$$

where we denote by $\langle \cdot \rangle_q$ the expectation values over the probability distribution q, and we used the (anti-)symmetry properties of q and θ under exchange of $\mathbf{y}$ and $\mathbf{y}'$ to replace $\left\langle e^{i\theta} \right\rangle = \langle \cos\theta \rangle$.

So we can use the Metropolis algorithm to sample $\mathbf{x}$, $\mathbf{y}$, $\mathbf{y}'$ according to the probability $q(\mathbf{x},\mathbf{y},\mathbf{y}')$, and then compute $\left\langle A^{\text{sh, rew}} \right\rangle_q$ and $\langle \cos\theta \rangle_q$ as simple averages over the samples. In the end the expectation value of $A^{\text{sh, rew}}$ over q needs to be *reweighted*, dividing by $\langle \cos\theta \rangle_q$. This yields the expectation value of the operator $\hat{A}$ over the original state. However, since all the expectation values computed by Monte Carlo sampling are affected by statistical noise, this reweighting procedure can become problematic when the expectation value of $\cos\theta$ turns out to be close to zero, and thus Eq. (2.34) involves the division of a noisy number, $\left\langle A^{\text{sh, rew}} \right\rangle_q$, by a small and noisy number, $\langle \cos\theta \rangle_q$. This can lead to serious instabilities, so the value of $\langle \cos\theta \rangle_q$ in an actual calculation always needs to be monitored.

2.4 t-VMC with shadow wave functions

In this Subsection we extend the t-VMC algorithm to variational wave functions of the shadow form. To our knowledge, this extension is a novel contribution to the t-VMC techniques. We consider a discrete system and denote by $\mathbf{x}$ the coordinates used to describe it, which could be, for example, spin configurations or particle positions on a lattice, depending on the type of system considered. We variationally describe the state of the system by a SWF Ansatz of the form

$$\psi(\mathbf{x};\phi,\boldsymbol{\alpha}) = e^{\phi} \sum_{\mathbf{y}} \Xi(\mathbf{x},\mathbf{y};\boldsymbol{\alpha}) , \tag{2.35}$$

which basically is the definition of a shadow wave function (2.1) where we introduced a parametrization of function Ξ and added a complex phase ϕ.

The equations of t-VMC apply to any variational wave function that is an analytic function of a set of complex parameters, so, assuming that Ξ is analytic with respect to the parameters α_k, the t-VMC equations (1.17) hold also for a SWF. The problem lies in the computation of the ingredients that appear in the equations, which need to be adapted to the sampling of the shadow variables described in Sec. 2.3. All the definitions introduced there still apply, with a minor change: since we added the complex phase ϕ to the definition of the wave function, the nonclassical probability distribution defined in Eq. (2.17) becomes

$$\mathcal{P}(\mathbf{x},\mathbf{y},\mathbf{y}';\phi,\boldsymbol{\alpha}) = e^{\phi^* + \phi} \Xi^*(\mathbf{x},\mathbf{y};\boldsymbol{\alpha}) \Xi(\mathbf{x},\mathbf{y}';\boldsymbol{\alpha}) . \tag{2.36}$$

We need to compute the matrix $S_{k,k'}$ defined in Eq. (1.18) and the vector F_k defined in Eq. (1.19). The ingredients needed are the expectation values of the local operators $\langle \mathcal{O}_k \rangle$, the expectation value of the energy $E = \langle \mathcal{E} \rangle$, and the correlations $\langle \mathcal{O}_k^* \mathcal{O}_{k'} \rangle$ and $\langle \mathcal{O}_k^* \mathcal{E} \rangle$.

The easiest is the energy: we just need to apply the formulas derived in Sec. 2.3 to the Hamiltonian operator $\hat{H}$. So, in accordance with Eq. (2.16), we define the *shadow local estimator* for the energy as

$$\mathcal{E}^{\text{sh}}(\mathbf{x}, \mathbf{y}) = \sum_{\mathbf{x}'} \left\langle \mathbf{x} \middle| \hat{H} \middle| \mathbf{x}' \right\rangle \frac{\Xi(\mathbf{x}', \mathbf{y})}{\Xi(\mathbf{x}, \mathbf{y})} , \tag{2.37}$$

and its *reweighted* version, according to Eq. (2.28), is defined as

$$\mathcal{E}^{\text{sh, rew}}(\mathbf{x}, \mathbf{y}, \mathbf{y}') = \frac{1}{2} \left[e^{i\theta(\mathbf{x},\mathbf{y},\mathbf{y}')} \mathcal{E}^{\text{sh}}(\mathbf{x}, \mathbf{y}') + e^{-i\theta(\mathbf{x},\mathbf{y},\mathbf{y}')} \mathcal{E}^{\text{sh}}(\mathbf{x}, \mathbf{y}) \right] . \tag{2.38}$$

Then the expectation value of the energy is the reweighted average of $\mathcal{E}^{\text{sh, rew}}$, see Eq. (2.34).

The expectation values involving the local operators are trickier, since the evaluation of $\mathcal{O}_k(\mathbf{x})$ itself involves summation over the shadow variables, which in general has no analytical closed form. This is easily seen with a short calculation starting from the definition of local operator in Eq. (1.8):

$$\mathcal{O}_k(\mathbf{x}; \boldsymbol{\alpha}) = \frac{1}{\psi(\mathbf{x}; \phi, \boldsymbol{\alpha})} \frac{\partial}{\partial \alpha_k} \psi(\mathbf{x}; \phi, \boldsymbol{\alpha}) \tag{2.39}$$

$$= \frac{1}{\sum_{\mathbf{y}'} \Xi(\mathbf{x}, \mathbf{y}'; \boldsymbol{\alpha})} \sum_{\mathbf{y}} \frac{\partial}{\partial \alpha_k} \Xi(\mathbf{x}, \mathbf{y}; \boldsymbol{\alpha}) . \tag{2.40}$$

In order to write an estimator for $\langle \mathcal{O}_k \rangle$, it is convenient to define the *shadow local operators*

$$\mathcal{O}_k^{\text{sh}}(\mathbf{x}, \mathbf{y}; \boldsymbol{\alpha}) = \frac{\partial}{\partial \alpha_k} \ln \Xi(\mathbf{x}, \mathbf{y}; \boldsymbol{\alpha}) = \frac{1}{\Xi(\mathbf{x}, \mathbf{y}; \boldsymbol{\alpha})} \frac{\partial}{\partial \alpha_k} \Xi(\mathbf{x}, \mathbf{y}; \boldsymbol{\alpha}) , \tag{2.41}$$

which depend on both the real and the shadow coordinates. Then, from Eqs. (2.40) and (2.41) we can see that the usual local operators can be written in terms of the shadow ones as

$$\mathcal{O}_k(\mathbf{x}; \boldsymbol{\alpha}) = \frac{\sum_{\mathbf{y}} \Xi(\mathbf{x}, \mathbf{y}; \boldsymbol{\alpha}) \mathcal{O}_k^{\text{sh}}(\mathbf{x}, \mathbf{y}; \boldsymbol{\alpha})}{\sum_{\mathbf{y}'} \Xi(\mathbf{x}, \mathbf{y}'; \boldsymbol{\alpha})} . \tag{2.42}$$

We can now exploit this expression to write the expectation value of $\mathcal{O}_k$ in a form that

can be reconducted to a reweighted average:

$$\langle \mathcal{O}_k \rangle = \sum_{\mathbf{x}} |\psi(\mathbf{x};\phi,\boldsymbol{\alpha})|^2 \mathcal{O}_k(\mathbf{x};\boldsymbol{\alpha}) \tag{2.43}$$

$$= \sum_{\mathbf{x},\mathbf{y},\mathbf{v}} e^{\phi^*+\phi}\Xi^*(\mathbf{x},\mathbf{y};\boldsymbol{\alpha})\Xi(\mathbf{x},\mathbf{v};\boldsymbol{\alpha})\frac{\sum_{\mathbf{y}'}\Xi(\mathbf{x},\mathbf{y}';\boldsymbol{\alpha})\mathcal{O}_k^{\text{sh}}(\mathbf{x},\mathbf{y}';\boldsymbol{\alpha})}{\sum_{\mathbf{w}}\Xi(\mathbf{x},\mathbf{w};\boldsymbol{\alpha})} \tag{2.44}$$

$$= \sum_{\mathbf{x},\mathbf{y}} e^{\phi^*+\phi}\Xi^*(\mathbf{x},\mathbf{y};\boldsymbol{\alpha})\cancel{\sum_{\mathbf{v}}\Xi(\mathbf{x},\mathbf{v};\boldsymbol{\alpha})}\frac{\sum_{\mathbf{y}'}\Xi(\mathbf{x},\mathbf{y}';\boldsymbol{\alpha})\mathcal{O}_k^{\text{sh}}(\mathbf{x},\mathbf{y}';\boldsymbol{\alpha})}{\cancel{\sum_{\mathbf{w}}\Xi(\mathbf{x},\mathbf{w};\boldsymbol{\alpha})}} \tag{2.45}$$

$$= \sum_{\mathbf{x},\mathbf{y},\mathbf{y}'} e^{\phi^*+\phi}\Xi^*(\mathbf{x},\mathbf{y};\boldsymbol{\alpha})\Xi(\mathbf{x},\mathbf{y}';\boldsymbol{\alpha})\mathcal{O}_k^{\text{sh}}(\mathbf{x},\mathbf{y}';\boldsymbol{\alpha}) \tag{2.46}$$

$$= \sum_{\mathbf{x},\mathbf{y},\mathbf{y}'} \mathcal{P}(\mathbf{x},\mathbf{y},\mathbf{y}';\phi,\boldsymbol{\alpha})\mathcal{O}_k^{\text{sh}}(\mathbf{x},\mathbf{y}';\boldsymbol{\alpha}) \tag{2.47}$$

$$= \sum_{\mathbf{x},\mathbf{y},\mathbf{y}'} \rho(\mathbf{x},\mathbf{y},\mathbf{y}';\phi,\boldsymbol{\alpha})e^{i\theta(\mathbf{x},\mathbf{y},\mathbf{y}';\boldsymbol{\alpha})}\mathcal{O}_k^{\text{sh}}(\mathbf{x},\mathbf{y}';\boldsymbol{\alpha}) \tag{2.48}$$

$$= \frac{1}{2}\left[\sum_{\mathbf{x},\mathbf{y},\mathbf{y}'} \rho(\mathbf{x},\mathbf{y},\mathbf{y}';\phi,\boldsymbol{\alpha})e^{i\theta(\mathbf{x},\mathbf{y},\mathbf{y}';\boldsymbol{\alpha})}\mathcal{O}_k^{\text{sh}}(\mathbf{x},\mathbf{y}';\boldsymbol{\alpha}) + \sum_{\mathbf{x},\mathbf{y}',\mathbf{y}} \rho(\mathbf{x},\mathbf{y}',\mathbf{y};\phi,\boldsymbol{\alpha})e^{i\theta(\mathbf{x},\mathbf{y}',\mathbf{y};\boldsymbol{\alpha})}\mathcal{O}_k^{\text{sh}}(\mathbf{x},\mathbf{y};\boldsymbol{\alpha})\right] \tag{2.49}$$

$$= \sum_{\mathbf{x},\mathbf{y},\mathbf{y}'} \rho(\mathbf{x},\mathbf{y},\mathbf{y}';\phi,\boldsymbol{\alpha})\frac{1}{2}\left[e^{i\theta(\mathbf{x},\mathbf{y},\mathbf{y}';\boldsymbol{\alpha})}\mathcal{O}_k^{\text{sh}}(\mathbf{x},\mathbf{y}';\boldsymbol{\alpha}) + e^{-i\theta(\mathbf{x},\mathbf{y},\mathbf{y}';\boldsymbol{\alpha})}\mathcal{O}_k^{\text{sh}}(\mathbf{x},\mathbf{y};\boldsymbol{\alpha})\right] \tag{2.50}$$

In line (2.49) we sum the expression of line (2.48) to itself and divide by two, then we exploit the (anti-)symmetry properties of ρ and θ with respect to exchange of $\mathbf{y}$ and $\mathbf{y}'$ to obtain the last line. Comparing this last expression with Eqs. (2.26) and (2.27), that we encountered in the evaluation of the expectation value of a generic operator $\left\langle \hat{A} \right\rangle$, we are prompted define the *reweighted shadow local estimator* for local operator $\mathcal{O}_k$ as

$$\mathcal{O}_k^{\text{sh, rew}}(\mathbf{x},\mathbf{y},\mathbf{y}';\boldsymbol{\alpha}) = \frac{1}{2}\left[e^{i\theta(\mathbf{x},\mathbf{y},\mathbf{y}';\boldsymbol{\alpha})}\mathcal{O}_k^{\text{sh}}(\mathbf{x},\mathbf{y}';\boldsymbol{\alpha}) + e^{-i\theta(\mathbf{x},\mathbf{y},\mathbf{y}';\boldsymbol{\alpha})}\mathcal{O}_k^{\text{sh}}(\mathbf{x},\mathbf{y};\boldsymbol{\alpha})\right]. \tag{2.51}$$

Then the expectation value $\langle \mathcal{O}_k \rangle$ is the reweighted average of $\mathcal{O}_k^{\text{sh, rew}}$, see Eq. (2.34). We need to carry out a similar calculation for the expectation value of the product of two

local operators:

$$\langle \mathcal{O}_k^* \mathcal{O}_{k'} \rangle = \sum_{\mathbf{x}} |\psi(\mathbf{x};\phi,\boldsymbol{\alpha})|^2 \, \mathcal{O}_k^*(\mathbf{x};\boldsymbol{\alpha}) \mathcal{O}_{k'}(\mathbf{x};\boldsymbol{\alpha}) \tag{2.52}$$

$$= \sum_{\mathbf{x},\mathbf{v},\mathbf{v}'} e^{\phi^*+\phi} \Xi^*(\mathbf{x},\mathbf{v};\boldsymbol{\alpha}) \Xi(\mathbf{x},\mathbf{v}';\boldsymbol{\alpha}) \cdot \frac{\sum_{\mathbf{y}} \Xi^*(\mathbf{x},\mathbf{y};\boldsymbol{\alpha}) \mathcal{O}_k^{\mathrm{sh}*}(\mathbf{x},\mathbf{y};\boldsymbol{\alpha})}{\sum_{\mathbf{w}} \Xi^*(\mathbf{x},\mathbf{w};\boldsymbol{\alpha})} \frac{\sum_{\mathbf{y}'} \Xi(\mathbf{x},\mathbf{y}';\boldsymbol{\alpha}) \mathcal{O}_{k'}^{\mathrm{sh}}(\mathbf{x},\mathbf{y}';\boldsymbol{\alpha})}{\sum_{\mathbf{w}'} \Xi(\mathbf{x},\mathbf{w}';\boldsymbol{\alpha})} \tag{2.53}$$

$$= \sum_{\mathbf{x}} e^{\phi^*+\phi} \cancel{\sum_{\mathbf{v}} \Xi^*(\mathbf{x},\mathbf{v};\boldsymbol{\alpha})} \cancel{\sum_{\mathbf{v}'} \Xi(\mathbf{x},\mathbf{v}';\boldsymbol{\alpha})} \cdot \frac{\sum_{\mathbf{y}} \Xi^*(\mathbf{x},\mathbf{y};\boldsymbol{\alpha}) \mathcal{O}_k^{\mathrm{sh}*}(\mathbf{x},\mathbf{y};\boldsymbol{\alpha})}{\cancel{\sum_{\mathbf{w}} \Xi^*(\mathbf{x},\mathbf{w};\boldsymbol{\alpha})}} \frac{\sum_{\mathbf{y}'} \Xi(\mathbf{x},\mathbf{y}';\boldsymbol{\alpha}) \mathcal{O}_{k'}^{\mathrm{sh}}(\mathbf{x},\mathbf{y}';\boldsymbol{\alpha})}{\cancel{\sum_{\mathbf{w}'} \Xi(\mathbf{x},\mathbf{w}';\boldsymbol{\alpha})}} \tag{2.54}$$

$$= \sum_{\mathbf{x},\mathbf{y},\mathbf{y}'} e^{\phi^*+\phi} \Xi^*(\mathbf{x},\mathbf{y};\boldsymbol{\alpha}) \Xi(\mathbf{x},\mathbf{y}';\boldsymbol{\alpha}) \mathcal{O}_k^{\mathrm{sh}*}(\mathbf{x},\mathbf{y};\boldsymbol{\alpha}) \mathcal{O}_{k'}^{\mathrm{sh}}(\mathbf{x},\mathbf{y}';\boldsymbol{\alpha}) \tag{2.55}$$

$$= \sum_{\mathbf{x},\mathbf{y},\mathbf{y}'} \mathcal{P}(\mathbf{x},\mathbf{y},\mathbf{y}';\phi,\boldsymbol{\alpha}) \mathcal{O}_k^{\mathrm{sh}*}(\mathbf{x},\mathbf{y};\boldsymbol{\alpha}) \mathcal{O}_{k'}^{\mathrm{sh}}(\mathbf{x},\mathbf{y}';\boldsymbol{\alpha}) \tag{2.56}$$

$$= \sum_{\mathbf{x},\mathbf{y},\mathbf{y}'} \rho(\mathbf{x},\mathbf{y},\mathbf{y}';\phi,\boldsymbol{\alpha}) \frac{1}{2} \Big[e^{i\theta(\mathbf{x},\mathbf{y},\mathbf{y}';\boldsymbol{\alpha})} \mathcal{O}_k^{\mathrm{sh}*}(\mathbf{x},\mathbf{y};\boldsymbol{\alpha}) \mathcal{O}_{k'}^{\mathrm{sh}}(\mathbf{x},\mathbf{y}';\boldsymbol{\alpha}) + e^{-i\theta(\mathbf{x},\mathbf{y},\mathbf{y}';\boldsymbol{\alpha})} \mathcal{O}_k^{\mathrm{sh}*}(\mathbf{x},\mathbf{y}';\boldsymbol{\alpha}) \mathcal{O}_{k'}^{\mathrm{sh}}(\mathbf{x},\mathbf{y};\boldsymbol{\alpha}) \Big] , \tag{2.57}$$

and again from comparison with Eqs. (2.26) and (2.27) we define the reweighted shadow local estimator for $\langle \mathcal{O}_k^* \mathcal{O}_{k'} \rangle$ as

$$[\mathcal{O}_k^* \mathcal{O}_{k'}]^{\text{sh, rew}}(\mathbf{x},\mathbf{y},\mathbf{y}';\boldsymbol{\alpha}) = \frac{1}{2} \Big[e^{i\theta(\mathbf{x},\mathbf{y},\mathbf{y}';\boldsymbol{\alpha})} \mathcal{O}_k^{\mathrm{sh}*}(\mathbf{x},\mathbf{y};\boldsymbol{\alpha}) \mathcal{O}_{k'}^{\mathrm{sh}}(\mathbf{x},\mathbf{y}';\boldsymbol{\alpha}) + e^{-i\theta(\mathbf{x},\mathbf{y},\mathbf{y}';\boldsymbol{\alpha})} \mathcal{O}_k^{\mathrm{sh}*}(\mathbf{x},\mathbf{y}';\boldsymbol{\alpha}) \mathcal{O}_{k'}^{\mathrm{sh}}(\mathbf{x},\mathbf{y};\boldsymbol{\alpha}) \Big] . \tag{2.58}$$

An analogous calculation for $\langle \mathcal{O}_k^* \mathcal{E} \rangle$ leads to the definition of

$$[\mathcal{O}_k^* \mathcal{E}]^{\text{sh, rew}}(\mathbf{x},\mathbf{y},\mathbf{y}';\boldsymbol{\alpha}) = \frac{1}{2} \Big[e^{i\theta(\mathbf{x},\mathbf{y},\mathbf{y}';\boldsymbol{\alpha})} \mathcal{O}_k^{\mathrm{sh}*}(\mathbf{x},\mathbf{y};\boldsymbol{\alpha}) \mathcal{E}^{\mathrm{sh}}(\mathbf{x},\mathbf{y}';\boldsymbol{\alpha}) + e^{-i\theta(\mathbf{x},\mathbf{y},\mathbf{y}';\boldsymbol{\alpha})} \mathcal{O}_k^{\mathrm{sh}*}(\mathbf{x},\mathbf{y}';\boldsymbol{\alpha}) \mathcal{E}^{\mathrm{sh}}(\mathbf{x},\mathbf{y};\boldsymbol{\alpha}) \Big] . \tag{2.59}$$

Now we have all we need to compute matrix S and vector F of the t-VMC equations of motion (1.20): we perform a Metropolis sampling of a set $\mathbf{x}$ of real variables and two sets $\mathbf{y}$, $\mathbf{y}'$ of shadow variables according to the joint probability distribution $q(\mathbf{x},\mathbf{y},\mathbf{y}')$ defined in Eq. (2.29); then the expectation values we need to compute S and F can be obtained by computing the averages of the estimators (2.38), (2.51), (2.58), and (2.59) over the samples, and then dividing them by the average of $\cos\theta$ over the samples, according to the reweighting formula (2.34). The same sampling can be used to compute the expectation values of various observables over the state, by computing their own reweighted shadow local estimators, defined as in Eq. (2.28).

CHAPTER 3

The Baeriswyl-Shadow wave function (BSWF)

In this Chapter, we exploit the extension of the t-VMC method to shadow Ansatzë introduced in Chap. 2, to introduce a new Ansatz type, whose study and implementation is the main topic of this work. This novel Ansatz leverages the expressiveness of the SWFs and their link to path integration to formulate a very concise variational wave function, that is still effective because its form is grounded in the Physics of the system we want to simulate.

We first present, in Sec. 3.1, the model we will use in our simulations as a benchmark, the transverse-field Ising model (TFI). Then, in Sec. 3.2, we show how the connection between shadow wave functions and path integration allow us to introduce an Ansatz wave function for the Ising model that takes a cue from its physics. Finally, in Sec. 3.3, we discuss some details of the particular Monte Carlo sampling techniques we use in conjunction with this Ansatz.

3.1 Our benchmark: the transverse-field Ising model (TFI)

The transverse-field Ising model (TFI) describes a set of spins on a lattice $\mathcal{L}$ with Hamiltonian operator

$$\hat{H} = -J \sum_{\langle \mathbf{r}, \mathbf{r}' \rangle} \hat{\sigma}^z_{\mathbf{r}} \hat{\sigma}^z_{\mathbf{r}'} - h \sum_{\mathbf{r}} \hat{\sigma}^x_{\mathbf{r}} \,, \tag{3.1}$$

where the summation over $\langle \mathbf{r}, \mathbf{r}' \rangle$ runs over nearest-neighbor sites, $\hat{\sigma}^x_{\mathbf{r}}$, $\hat{\sigma}^z_{\mathbf{r}}$ are Pauli operators for the spin occupying site $\mathbf{r}$, J is a constant controlling the nearest-neighbor coupling along the z axis (we only consider the ferromagnetic case, $J > 0$), and h is the strength of the transverse external field along the x axis. For the moment, we do not specify the dimensionality of the lattice $\mathcal{L}$, which cold be one-, two-, or three-dimensional. We observe that one of the two couplings J and h can be factored out and used as the unit of measure for energy; so, the model actually depends on only one parameter, namely the ratio $g = |h/J|$.

In one dimension (1D), the TFI is exactly solved by means of the *Jordan-Wigner transformation* [45], so its ground state is exactly known. Moreover, the exact time-dependent evolution can be calculated through the time-dependent Bogoliubov-de Gennes equations [46]. So the 1D Ising chain represents an ideal benchmark against which we will test the capabilities of the Ansatz we are going to introduce.

In the thermodynamic limit, the 1D TFI model presents a quantum phase transition at $g = 1$ between a ferromagnetic phase ($g < 1$) and a paramagnetic phase ($g > 1$) [45]. In the ferromagnetic phase, the ground state presents finite long-range order, while in the paramagnetic phase the spin-spin correlations decay exponentially with the distance, and this decay is characterized by a correlation length. The condition $g = 1$ is very

interesting because it represents a quantum critical point: as $g = 1$ is approached from above, the correlation length diverges.

3.2 The Baeriswyl-Shadow wave function

In this section, we introduce a new shadow wave function taking a cue from the link between the SWF paradigm and path integration, explained in Sec. 2.2. To this end, we consider the functional form deriving from a single step of path integration for the transver Ising model. We rewrite the TFI Hamiltonian as

$$\begin{aligned} \hat{H} = \hat{K} + \hat{V}\,, \qquad \text{with } \hat{K} &= -h \sum_{\mathbf{r}} \hat{\sigma}^x_{\mathbf{r}}\,, \\ \hat{V} &= -J \sum_{\langle \mathbf{r},\mathbf{r}' \rangle} \hat{\sigma}^z_{\mathbf{r}} \hat{\sigma}^z_{\mathbf{r}'}\,. \end{aligned} \tag{3.2}$$

This way we isolated a term diagonal in the σ^z basis, the nearest-neighbor interaction $\hat{V}$, and a non-diagonal term, the interaction with the transverse field $\hat{K}$. So, in a path integral with step $\delta\tau$, we would represent the correlations within the i^{th} layer through the interaction weight

$$e^{-\frac{\delta\tau}{2} V(\boldsymbol{\sigma}^i)} = \exp\left(\frac{\delta\tau J}{2} \sum_{\langle \mathbf{r},\mathbf{r}' \rangle} \sigma^i_{\mathbf{r}} \sigma^i_{\mathbf{r}'} \right)\,, \tag{3.3}$$

where we denote by $\boldsymbol{\sigma}^i$ the spin configuration of the i^{th} layer. The correlations between two consecutive layers, on the other hand, are described by the propagator of the non-diagonal term

$$\langle \boldsymbol{\sigma}^i | e^{-\delta\tau \hat{K}} | \boldsymbol{\sigma}^{i+1} \rangle = \langle \boldsymbol{\sigma}^i | e^{\delta\tau\, h \sum_{\mathbf{r}} \hat{\sigma}^x_{\mathbf{r}}} | \boldsymbol{\sigma}^{i+1} \rangle\,. \tag{3.4}$$

To express this propagator explicitly as a function of the configurations $\boldsymbol{\sigma}^i$ and $\boldsymbol{\sigma}^{i+1}$, we perform some manipulations:

$$\langle \boldsymbol{\sigma}^i | e^{\delta\tau\, h \sum_{\mathbf{r}} \hat{\sigma}^x_{\mathbf{r}}} | \boldsymbol{\sigma}^{i+1} \rangle = \langle \boldsymbol{\sigma}^i | \prod_{\mathbf{r}} e^{\delta\tau\, h\, \hat{\sigma}^x_{\mathbf{r}}} | \boldsymbol{\sigma}^{i+1} \rangle = \prod_{\mathbf{r}} \langle \sigma^i_{\mathbf{r}} | e^{\delta\tau\, h\, \hat{\sigma}^x_{\mathbf{r}}} | \sigma^{i+1}_{\mathbf{r}} \rangle\,, \tag{3.5}$$

where $|\sigma^i_{\mathbf{r}}\rangle$ is the single-spin state on site $\mathbf{r}$ with orientation $\sigma^i_{\mathbf{r}}$. We readily see that this propagator correlates each spin in the i^{th} layer only with the spin that lies on the same lattice site $\mathbf{r}$ in the $(i+1)^{\text{th}}$ layer. To proceed further, we compute the exponential by expanding it in a Taylor series and using the identity $(\hat{\sigma}^x_{\mathbf{r}})^2 = \mathbb{1}$:

$$e^{\delta\tau\, h \hat{\sigma}^x_{\mathbf{r}}} = \sum_n \frac{1}{n!} (\delta\tau\, h\, \hat{\sigma}^x_{\mathbf{r}})^n \tag{3.6}$$

$$= \mathbb{1} \sum_{n \text{ even}} \frac{1}{n!} (\delta\tau\, h)^n + \hat{\sigma}^x_{\mathbf{r}} \sum_{n \text{ odd}} \frac{1}{n!} (\delta\tau\, h)^n \tag{3.7}$$

$$= \cosh(\delta\tau\, h)\, \mathbb{1} + \sinh(\delta\tau\, h)\, \hat{\sigma}^x_{\mathbf{r}}\,, \tag{3.8}$$

so, using the fact that $\hat{\sigma}^x_{\mathbf{r}}$ inverts the orientation of the spin along the z-axis, we can write

$$\begin{aligned}
\langle \sigma^i_{\mathbf{r}} | e^{\delta\tau\, h\, \hat{\sigma}^x_{\mathbf{r}}} | \sigma^{i+1}_{\mathbf{r}} \rangle &= \cosh(\delta\tau\, h)\, \langle \sigma^i_{\mathbf{r}} | \sigma^{i+1}_{\mathbf{r}} \rangle + \sinh(\delta\tau\, h)\, \langle \sigma^i_{\mathbf{r}} | \hat{\sigma}^x_{\mathbf{r}} | \sigma^{i+1}_{\mathbf{r}} \rangle && (3.9)\\
&= \cosh(\delta\tau\, h)\, \langle \sigma^i_{\mathbf{r}} | \sigma^{i+1}_{\mathbf{r}} \rangle + \sinh(\delta\tau\, h)\, \langle \sigma^i_{\mathbf{r}} | - \sigma^{i+1}_{\mathbf{r}} \rangle && (3.10)\\
&= \cosh(\delta\tau\, h)\, \delta_{\sigma^i_{\mathbf{r}}, \sigma^{i+1}_{\mathbf{r}}} + \sinh(\delta\tau\, h)\, \delta_{\sigma^i_{\mathbf{r}}, -\sigma^{i+1}_{\mathbf{r}}} && (3.11)\\
&= [\cosh(\delta\tau\, h) \sinh(\delta\tau\, h)]^{1/2} \\
&\quad \cdot \left[\tanh(\delta\tau\, h)^{-1/2} \delta_{\sigma^i_{\mathbf{r}}, \sigma^{i+1}_{\mathbf{r}}} + \tanh(\delta\tau\, h)^{1/2}\, \delta_{\sigma^i_{\mathbf{r}}, -\sigma^{i+1}_{\mathbf{r}}} \right] && (3.12)\\
&= [\cosh(\delta\tau\, h) \sinh(\delta\tau\, h)]^{1/2} \cdot \exp\left(\gamma\, \sigma^i_{\mathbf{r}} \sigma^{i+1}_{\mathbf{r}} \right) , && (3.13)
\end{aligned}$$

where in the last line we denote $\gamma = -\frac{1}{2} \ln [\tanh(\delta\tau\, h)]$, and the product $\sigma^i_{\mathbf{r}} \sigma^{i+1}_{\mathbf{r}}$ appears, of corresponding spins on consecutive layers. Dropping the irrelevant multiplicative constant $[\cosh(\delta\tau\, h) \sinh(\delta\tau\, h)]^{1/2}$ in Eq. (3.13), we can finally write the propagator of the non-diagonal term as

$$\langle \boldsymbol{\sigma}^i | e^{-\delta\tau\, \hat{K}} | \boldsymbol{\sigma}^{i+1} \rangle \propto \exp\left(\gamma \sum_{\mathbf{r}} \sigma^i_{\mathbf{r}} \sigma^{i+1}_{\mathbf{r}} \right) , \tag{3.14}$$

which describes a classical Ising-like coupling between corresponding spins on consecutive layers. The full one-step propagator thus reads

$$\begin{aligned}
G(\boldsymbol{\sigma}^i, \boldsymbol{\sigma}^{i+1}, \delta\tau) &= e^{-\frac{\delta\tau}{2} V(\boldsymbol{\sigma}^i)} \langle \boldsymbol{\sigma}^i | e^{-\delta\tau\, \hat{K}} | \boldsymbol{\sigma}^{i+1} \rangle e^{-\frac{\delta\tau}{2} V(\boldsymbol{\sigma}^{i+1})} && (3.15)\\
&= \exp\left(\frac{\delta\tau J}{2} \sum_{\langle \mathbf{r}, \mathbf{r}' \rangle} \sigma^i_{\mathbf{r}} \sigma^i_{\mathbf{r}'} + \gamma \sum_{\mathbf{r}} \sigma^i_{\mathbf{r}} \sigma^{i+1}_{\mathbf{r}} + \frac{\delta\tau J}{2} \sum_{\langle \mathbf{r}, \mathbf{r}' \rangle} \sigma^{i+1}_{\mathbf{r}} \sigma^{i+1}_{\mathbf{r}'} \right) . && (3.16)
\end{aligned}$$

If we were to perform a one-step path integration, the wave function would read

$$\psi(\boldsymbol{\sigma}) = \sum_{\boldsymbol{\sigma'}} G(\boldsymbol{\sigma}, \boldsymbol{\sigma'}, \delta\tau) . \tag{3.17}$$

Inspired by this functional form, we define a variational Ansatz of the shadow class by performing a one-step path integration in which the constant couplings $\frac{\delta\tau J}{2}$ and γ are replaced by variational parameters that can be optimized. Thus, denoting with $\mathbf{h}$ the shadow variables, we define the Baeriswyl-Shadow wave function (BSWF) as

$$\begin{aligned}
\psi_{\text{BSWF}}(\boldsymbol{\sigma}; \eta, \omega, \rho) &= \sum_{\boldsymbol{h}} \Xi_{\text{BSWF}}(\boldsymbol{\sigma}, \boldsymbol{h}) \\
&= \sum_{\boldsymbol{h}} \exp\left(\eta \sum_{\langle \mathbf{r}, \mathbf{r}' \rangle} \sigma_{\mathbf{r}} \sigma_{\mathbf{r}'} + \omega \sum_{\mathbf{r}} \sigma_{\mathbf{r}} h_{\mathbf{r}} + \rho \sum_{\langle \mathbf{r}, \mathbf{r}' \rangle} h_{\mathbf{r}} h_{\mathbf{r}'} \right) ,
\end{aligned} \tag{3.18}$$

where the variational parameters are denoted by η for the real-real coupling, ω for the real-shadow coupling, and ρ for the shadow-shadow coupling. In the choice of the name for the BSWF, we have a two-fold motivation: (1) it belongs to the class of shadow wave functions, and (2) it is reminiscent of a functional form for discrete systems that was first suggested, to our knowledge, by D. Baeriswyl and collaborators [39, 40, 41]. A similar Ansatz, with real parameters, was studied in a ground-state context by the Authors of

Refs. [47, 48]. Our study is different in that we extend this Ansatz to the time-dependent case. Among other things, this implies that we use complex variational parameters, which require the reweighting technique described in Sec. 2.3.

We remark that this wave function only depends on three parameters, but, as it is shown in Chapter 4, it can nevertheless give quantitatively accurate results for the ground state and for *short* real-time evolutions. It can achieve these results because it is grounded in the physics of the system, being inspired by the path-integral formalism.

3.2.1 The t-VMC and the BSWF

Since the BSWF is a shadow Ansatz, we can apply the t-VMC to it as we explained in Sec. 2.4. But, in order to do that, we need to compute the shadow local operators, the probability distribution we have to sample, and the quantity θ that enters in the reweighting formulas.

There are three shadow local operators corresponding to the three parameters η, ω, ρ, and defined following to Eq. (2.41):

$$\mathcal{O}^{\text{sh}}_{\eta} = \frac{\partial}{\partial \eta} \ln \Xi_{\text{BSWF}}(\boldsymbol{\sigma}, \boldsymbol{h}) = \sum_{\langle \mathbf{r}, \mathbf{r}' \rangle} \sigma_{\mathbf{r}} \sigma_{\mathbf{r}'} \,, \tag{3.19}$$

$$\mathcal{O}^{\text{sh}}_{\omega} = \frac{\partial}{\partial \omega} \ln \Xi_{\text{BSWF}}(\boldsymbol{\sigma}, \boldsymbol{h}) = \sum_{\mathbf{r}} \sigma_{\mathbf{r}} h_{\mathbf{r}} \,, \tag{3.20}$$

$$\mathcal{O}^{\text{sh}}_{\rho} = \frac{\partial}{\partial \rho} \ln \Xi_{\text{BSWF}}(\boldsymbol{\sigma}, \boldsymbol{h}) = \sum_{\langle \mathbf{r}, \mathbf{r}' \rangle} h_{\mathbf{r}} h_{\mathbf{r}'} \,. \tag{3.21}$$

The complex non-classical probability distribution is found from Eq. (2.17), or rather from its modified version that includes the complex phase, Eq. (2.36), and it reads

$$\begin{aligned} \mathcal{P}_{\text{BSWF}}(\boldsymbol{\sigma}, \boldsymbol{h}, \boldsymbol{h'}) &= e^{\phi^* + \phi} \Xi^*_{\text{BSWF}}(\boldsymbol{\sigma}, \boldsymbol{h}) \Xi_{\text{BSWF}}(\boldsymbol{\sigma}, \boldsymbol{h'}) \\ &= \exp \left\{ 2\Re(\phi) + 2\Re(\eta) \sum_{\langle \mathbf{r}, \mathbf{r}' \rangle} \sigma_{\mathbf{r}} \sigma_{\mathbf{r}'} + \sum_{\mathbf{r}} \sigma_{\mathbf{r}} \left(\omega^* \, h_{\mathbf{r}} + \omega h'_{\mathbf{r}} \right) \right. \\ &\qquad \left. + \sum_{\langle \mathbf{r}, \mathbf{r}' \rangle} \left(\rho^* h_{\mathbf{r}} h_{\mathbf{r}'} + \rho h'_{\mathbf{r}} h'_{\mathbf{r}'} \right) \right\} . \end{aligned} \tag{3.22}$$

The classical probability function (2.29) that we will actually sample is proportional to the modulus of $\mathcal{P}$:

$$\begin{aligned} q_{\text{BSWF}}(\boldsymbol{\sigma}, \boldsymbol{h}, \boldsymbol{h'}) &\propto |\mathcal{P}_{\text{BSWF}}(\boldsymbol{\sigma}, \boldsymbol{h}, \boldsymbol{h'})| \\ &= \exp \left\{ 2\Re(\phi) + 2\Re(\eta) \sum_{\langle \mathbf{r}, \mathbf{r}' \rangle} \sigma_{\mathbf{r}} \sigma_{\mathbf{r}'} + \Re(\omega) \sum_{\mathbf{r}} \sigma_{\mathbf{r}} \left(h_{\mathbf{r}} + h'_{\mathbf{r}} \right) \right. \\ &\qquad \left. + \Re(\rho) \sum_{\langle \mathbf{r}, \mathbf{r}' \rangle} \left(h_{\mathbf{r}} h_{\mathbf{r}'} + h'_{\mathbf{r}} h'_{\mathbf{r}'} \right) \right\} . \end{aligned} \tag{3.23}$$

When the parameters are complex, as it is the case during real-time evolution, we need to apply the reweighting technique exposed in Sec. 2.3, and for that we need the value

of the argument θ of $\mathcal{P}$, see Eq. (2.21), that reads

$$\begin{aligned}\theta_{\text{BSWF}}(\boldsymbol{\sigma}, \boldsymbol{h}, \boldsymbol{h'}) &= \arg\left\{\mathcal{P}_{\text{BSWF}}(\boldsymbol{\sigma}, \boldsymbol{h}, \boldsymbol{h'})\right\} \\ &= \Im(\omega) \sum_{\mathbf{r}} \sigma_{\mathbf{r}} \left(h'_{\mathbf{r}} - h_{\mathbf{r}}\right) + \Im(\rho) \sum_{\langle \mathbf{r}, \mathbf{r}' \rangle} \left(h'_{\mathbf{r}} h'_{\mathbf{r}'} - h_{\mathbf{r}} h_{\mathbf{r}'}\right) . \end{aligned} \tag{3.24}$$

These are all the ingredients we need to apply the t-VMC method to a BSWF. Now we focus on the probability q_{BSWF} defined by Eq. (3.23). It looks like the Boltzmann weight of a classical (d+1)-dimensional Ising model, where d is the dimension of the system we are simulating, and the two shadow configurations act like an extension of it in the $(d+1)^{\text{th}}$ dimension. The main difference is that the spin-spin couplings are not the same across the system, but assume one of the three values $2\Re(\eta)$, $\Re(\omega)$, $\Re(\rho)$. There exist algorithms that allow for the efficient Monte Carlo sampling of spin configurations according to such Boltzmann weights by performing *cluster* moves. In the next Section we examine some of them.

3.3 Sampling algorithms for cluster moves

3.3.1 The Wolff algorithm

The Wolff cluster algorithm [49] is a powerful Monte Carlo method used to sample spin configurations of a classical Ising model, by performing moves in which whole clusters of spins are flipped at once. In this exposition we follow Ref. [50]. We want to sample the Boltzmann probability distribution

$$P(\boldsymbol{\sigma}) \propto \exp\left(\eta \sum_{\langle \mathbf{r}, \mathbf{r}' \rangle} \sigma_{\mathbf{r}} \sigma_{\mathbf{r}'}\right) , \tag{3.25}$$

with $\eta > 0$ (ferromagnetic case). Usually $\eta = \beta J$, where J is the coupling and β is the inverse temperature. However, in our work, probability distributions of this kind arise from the functional form of some variational wave functions, for which the parameter η has a different meaning.

The cluster moves is performed as follows. First of all, a random spin is selected, uniformly over the lattice: it is the first spin, or the seed, of our cluster. Then, we consider all its neighboring spins, one at a time: for each neighboring spin, if it is oriented like the seed spin, we add it to the cluster with probability P_{add}; otherwise, if it is not oriented like the seed, we simply discard it. Next, for each added spin, we consider all of its neighbors that have not yet been added to the cluster and decide whether to add them according to the same criterion. These steps are iterated until all the similarly oriented spins have been added to the cluster or discarded. Notice that a spin may have more than one chance of being added to the cluster, precisely one chance for each neighboring spin that already belongs to the cluster. Finally, a flip of the whole cluster is proposed, and it is accepted according to the probability prescribed by the Metropolis algorithm.

What is the value of the acceptance probability? Let's call μ and ν the original configuration and the one with the flipped cluster, respectively. In each of the two configurations, some of the spins just outside the cluster point in the same way as the spins in the cluster. When the cluster is flipped, the "bonds" between these spins and the spins in the cluster are "broken". The bonds that are *not* broken in the cluster flip from μ to ν are broken in the reverse move from ν to μ, and vice versa. All other bonds in and out of the cluster stay the same, in both directions.

Now consider a particular move taking us from μ to ν, i.e., a move starting from a particular seed and then adding spins to the cluster in a particular sequence. Let's compare it with the reverse move from ν to μ, that starts from the same seed spin and makes the same sequence of choices to build the cluster. The probability of choosing the seed is the same in the two cases, and also the probability of adding the spins to the cluster in that very sequence is the same. The only difference lies in the probability of "breaking" bonds along the cluster border, i.e., the probability of discarding all the similarly oriented spins that are just out of this particular cluster. For each discarded spin, the probability of doing so is $1 - P_{\text{add}}$. So, if in the $\mu \to \nu$ move m such bonds are broken, the total probability is $(1 - P_{\text{add}})^m$. Accordingly, if in the reverse $\nu \to \mu$ move n bonds are broken, the probability of doing it is $(1 - P_{\text{add}})^n$. We can now write the ratio between the probabilities of proposing the two moves, that reads

$$\frac{T(\nu \to \mu)}{T(\mu \to \nu)} = (1 - P_{\text{add}})^{n-m} \,. \tag{3.26}$$

Considering that ratio of the probabilities (3.25) of the two configurations is

$$\frac{P(\nu)}{P(\mu)} = e^{2\eta(n-m)} \,, \tag{3.27}$$

we can compute the Metropolis acceptance probability for the $\mu \to \nu$ move, that is

$$A(\mu \to \nu) = \min\left\{1, \frac{P(\nu)\,T(\nu \to \mu)}{P(\mu)\,T(\mu \to \nu)}\right\} \tag{3.28}$$

$$= \min\left\{1, \left[\, e^{2\eta}\,(1 - P_{\text{add}})\,\right]^{n-m}\right\} \,. \tag{3.29}$$

Obviously, the algorithm will be more efficient if the cluster move is accepted as often as possible, and we observe that, carefully choosing the value of P_{add}, the acceptance probability can be made equal to one! This is achieved by making

$$P_{\text{add}} = 1 - e^{-2\eta} \,. \tag{3.30}$$

Finally, we can summarize the Wolff algorithm in the following steps:

1. Choose a seed spin at random from the lattice. This is the first spin of the cluster.

2. For each neighboring spin that is oriented like the seed, add it to the cluster with probability $P_{\text{add}} = 1 - e^{-2\eta}$.

3. For each spin that was added in the last step, consider all of its similarly oriented neighboring spins, and add each one according to the same probability $P_{\text{add}} = 1 - e^{-2\eta}$. Iterate this step until there are no more spins in the cluster whose neighbors have not been considered for addition. Notice that a spin may be considered for addition more than once: it gets a chance each time one of its neighboring spins is added to the cluster. Obviously, if a spin has already been added to the cluster, it will not be considered again for addition when one of its neighboring spins is added.

4. Flip the cluster (always!).

This algorithm satisfies detailed balance, by construction. It is also ergodic, since there is a finite probability of flipping just the seed spin, and any configuration can be reached by a finite sequence of single spin flips. We stress that this algorithm is especially efficient in the ferromagnetic phase. Deep into the paramagnetic phase, instead, the cluster will very often consist of just a single spin, so that the Wolff algorithm becomes equivalent to performing single spin flip moves via Metropolis (which in this limit have acceptance proability $\simeq 1$), but in a less efficient way because of the added computational cost of considering for addition (and, almost always, discarding) all the neighbors of the seed spin.

3.3.2 Cluster moves for spin glasses

We consider a classical spin system described by the Hamiltonian

$$H = -\sum_{\langle \mathbf{r},\mathbf{r}'\rangle} J_{\mathbf{r},\mathbf{r}'}\sigma_{\mathbf{r}}\sigma_{\mathbf{r}'}\,, \tag{3.31}$$

where the coupling $J_{\mathbf{r},\mathbf{r}'}$ is different and random for every pair of nearest-neighbors. This is a particular case of the class of *Edwards-Anderson spin glasses* [51]. Common choices for the couplings are $J_{\mathbf{r},\mathbf{r}'} = \pm 1$ randomly for each bond, or independent Gaussian variables with zero mean.

We are interested in cluster algorithms that allow to sample the Boltzmann probability distribution

$$P(\boldsymbol{\sigma}) \propto \exp\left(\sum_{\langle \mathbf{r},\mathbf{r}'\rangle} \eta_{\mathbf{r},\mathbf{r}'}\sigma_{\mathbf{r}}\sigma_{\mathbf{r}'}\right)\,, \tag{3.32}$$

where $\eta_{\mathbf{r},\mathbf{r}'} = \beta J_{\mathbf{r},\mathbf{r}'}$ and β is the inverse temperature. However, like we pointed out while discussing the Wolff algorithm, for us these forms of the probability distribution arise from some choices of the variational wave function, for which the parameters $\eta_{\mathbf{r},\mathbf{r}'} = \beta J_{\mathbf{r},\mathbf{r}'}$ have a different meaning.

A solution for sampling P was proposed by Dotsenko and collaborators [52]. It consists in building a cluster in a way that is very similar to the Wolff algorithm, but with a bond-dependent addition probability $P_{\text{add}}(\eta_{\mathbf{r},\mathbf{r}'})$. If the spin $\sigma_{\mathbf{r}}$ has already been added to the cluster, and the spin $\sigma_{\mathbf{r}'}$ is one of its nearest-neighbors that we are considering for addition, then we accept it with probability

$$P_{\text{add}}(\eta_{\mathbf{r},\mathbf{r}'}) = 1 - e^{-2|\eta_{\mathbf{r},\mathbf{r}'}|}\,. \tag{3.33}$$

With a reasoning analogous to the one made for the Wolff algorithm, it can be shown that the Metropolis acceptance probability for this type of cluster move is 1, so that we can always accept it.

For the study of actual spin glasses, this algorithm is not as efficient as other techniques that can be found in the literature (see, e.g., Ref. [53]). However, in this work we are not really studying spin glasses, but face the task of sampling probability distributions that can be expressed in the form (3.32), and found this algorithm to be helpful in our use cases.

3.3.3 Optimal cluster algorithms for general probability distributions

We consider the problem of devising a cluster algorithm of the Dotsenko type (see Subsec. (3.3.2)) to sample configurations from a generic probability distribution $P(\boldsymbol{\sigma})$. We write it in terms of its logarithm as $P(\boldsymbol{\sigma}) = e^{\ln P(\boldsymbol{\sigma})}$.

We know how to use cluster moves to sample probabilities of the form (3.32), so we rewrite $\ln P(\boldsymbol{\sigma})$ as

$$\ln P(\boldsymbol{\sigma}) = \sum_{\langle \mathbf{r},\mathbf{r}'\rangle} \eta_{\mathbf{r},\mathbf{r}'}\sigma_{\mathbf{r}}\sigma_{\mathbf{r}'} + \left(\ln P(\boldsymbol{\sigma}) - \sum_{\langle \mathbf{r},\mathbf{r}'\rangle} \eta_{\mathbf{r},\mathbf{r}'}\sigma_{\mathbf{r}}\sigma_{\mathbf{r}'} \right) \tag{3.34}$$

$$= \sum_{\langle \mathbf{r},\mathbf{r}'\rangle} \eta_{\mathbf{r},\mathbf{r}'}\sigma_{\mathbf{r}}\sigma_{\mathbf{r}'} + C(\boldsymbol{\sigma})\,, \tag{3.35}$$

where we defined $C(\boldsymbol{\sigma})$ as

$$C(\boldsymbol{\sigma}) = \ln P(\boldsymbol{\sigma}) - \sum_{\langle \mathbf{r},\mathbf{r}'\rangle} \eta_{\mathbf{r},\mathbf{r}'}\sigma_{\mathbf{r}}\sigma_{\mathbf{r}'}\,. \tag{3.36}$$

So the probability $P(\boldsymbol{\sigma})$ can be written as

$$P(\boldsymbol{\sigma}) = \exp\left[\sum_{\langle \mathbf{r},\mathbf{r}'\rangle} \eta_{\mathbf{r},\mathbf{r}'}\sigma_{\mathbf{r}}\sigma_{\mathbf{r}'} + C(\boldsymbol{\sigma}) \right]\,. \tag{3.37}$$

We build a cluster according to the Dotsenko method assuming a probability of the form

$$P_{\text{cluster}}(\boldsymbol{\sigma}) \propto \exp\left(\sum_{\langle \mathbf{r},\mathbf{r}'\rangle} \eta_{\mathbf{r},\mathbf{r}'}\sigma_{\mathbf{r}}\sigma_{\mathbf{r}'} \right)\,, \tag{3.38}$$

and we need to determine the Metropolis acceptance probability for the flipping of the cluster. Say that μ is the configuration before the cluster move and ν is the configuration after it. Since the Dotsenko algorithm satisfies the detailed balance condition, we know that the probabilities of proposing the $\mu \to \nu$ move and its reverse are related by

$$\frac{T(\nu \to \mu)}{T(\mu \to \nu)} = \frac{P_{\text{cluster}}(\mu)}{P_{\text{cluster}}(\nu)}\,. \tag{3.39}$$

Then the Metropolis acceptance probability is given by

$$A(\mu \to \nu) = \min\left\{ 1, \frac{P(\nu)\,T(\nu \to \mu)}{P(\mu)\,T(\mu \to \nu)} \right\} \tag{3.40}$$

$$= \min\left\{ 1, \frac{P_{\text{cluster}}(\mu)}{P(\mu)} \cdot \frac{P(\nu)}{P_{\text{cluster}}(\nu)} \right\}\,, \tag{3.41}$$

where

$$\frac{P_{\text{cluster}}(\mu)}{P(\mu)} \propto \frac{\exp\left(\sum_{\langle \mathbf{r},\mathbf{r}'\rangle} \eta_{\mathbf{r},\mathbf{r}'}\sigma^{(\mu)}_{\mathbf{r}}\sigma^{(\mu)}_{\mathbf{r}'} \right)}{\exp\left[\sum_{\langle \mathbf{r},\mathbf{r}'\rangle} \eta_{\mathbf{r},\mathbf{r}'}\sigma^{(\mu)}_{\mathbf{r}}\sigma^{(\mu)}_{\mathbf{r}'} + C(\mu) \right]} = e^{-C(\mu)}\,. \tag{3.42}$$

So the acceptance probability becomes

$$A(\mu \to \nu) = \min\left\{ 1, e^{C(\nu)-C(\mu)} \right\}\,. \tag{3.43}$$

Up to now the coefficients $\eta_{\mathbf{r},\mathbf{r}'}$ have been arbitrary values. We ideally want to choose them so that the acceptance probability stays as close to 1 as possible. This, roughly

speaking, means that we want the function $C(\boldsymbol{\sigma})$ to be as "small" as possible. To measure how "small" $C(\boldsymbol{\sigma})$ is, we use the l^2-norm:

$$|C|^2 = \sum_{\boldsymbol{\sigma}} C(\boldsymbol{\sigma})^2 \,, \tag{3.44}$$

and we want to minimize this norm with respect to the coefficients $\eta_{\mathbf{r},\mathbf{r}'}$. Let's start by computing the derivative

$$\frac{\partial}{\partial \eta_{\mathbf{r}_1,\mathbf{r}_2}} |C|^2 = 2 \sum_{\boldsymbol{\sigma}} \sigma_{\mathbf{r}_1}\sigma_{\mathbf{r}_2} \cdot \left[\sum_{\langle \mathbf{r}_3,\mathbf{r}_4 \rangle} \eta_{\mathbf{r}_3,\mathbf{r}_4} \sigma_{\mathbf{r}_3}\sigma_{\mathbf{r}_4} - \ln P(\boldsymbol{\sigma}) \right] . \tag{3.45}$$

Now, it can be easily shown that

$$\sum_{\boldsymbol{\sigma}} \sigma_{\mathbf{r}_1}\sigma_{\mathbf{r}_2} \cdot \sum_{\langle \mathbf{r}_3,\mathbf{r}_4 \rangle} \eta_{\mathbf{r}_3,\mathbf{r}_4} \sigma_{\mathbf{r}_3}\sigma_{\mathbf{r}_4} = 4\, \eta_{\mathbf{r}_1,\mathbf{r}_2} \,, \tag{3.46}$$

so that we can rewrite the derivative as

$$\frac{\partial}{\partial \eta_{\mathbf{r}_1,\mathbf{r}_2}} |C|^2 = 2 \left[4\, \eta_{\mathbf{r}_1,\mathbf{r}_2} - \sum_{\boldsymbol{\sigma}} \sigma_{\mathbf{r}_1}\sigma_{\mathbf{r}_2} \cdot \ln P(\boldsymbol{\sigma}) \right] . \tag{3.47}$$

Setting the derivative to zero finally yields the solution of our problem:

$$\eta_{\mathbf{r}_1,\mathbf{r}_2} = \frac{1}{4} \sum_{\boldsymbol{\sigma}} \sigma_{\mathbf{r}_1}\sigma_{\mathbf{r}_2} \cdot \ln P(\boldsymbol{\sigma}) \,. \tag{3.48}$$

We observe that this summation is the scalar product between the function $f_{\mathbf{r}_1,\mathbf{r}_2}(\boldsymbol{\sigma}) = \sigma_{\mathbf{r}_1}\sigma_{\mathbf{r}_2}$ and the function $\ln P(\boldsymbol{\sigma})$ divided by $|f_{\mathbf{r}_1,\mathbf{r}_2}|^2 = 4$. So, we are performing a projection: the resulting $\ln P_{\text{cluster}}(\boldsymbol{\sigma})$ with optimized coefficients, see Eq. (3.38), is the projection of the function $\ln P(\boldsymbol{\sigma})$ on the subspace spanned by the set of linearly independent functions $\{f_{\mathbf{r}_1,\mathbf{r}_2} | \mathbf{r}_1, \mathbf{r}_2 \in \mathcal{L}\}$.

We observe that, since the function $C(\boldsymbol{\sigma})$ enters the acceptance probability only through the difference $C(\nu) - C(\mu)$, it is actually defined up to a constant, so that we could instead try to minimize the l^2-norm of $C(\boldsymbol{\sigma}) + const$. However, we remark that the solution of the minimization problem does not change under such a translation of the function $C(\boldsymbol{\sigma})$. From this fact, one can easily show that *in our case* the minimization of the l^2-norm is equivalent to the minimization of the variance of the function $C(\boldsymbol{\sigma})$ with respect to a uniform probability distribution over the spin configurations $\boldsymbol{\sigma}$.

In a big system and for a generic $P(\boldsymbol{\sigma})$, the summation in Eq. (3.48) is not computationally tractable, but for specific forms of $\ln P(\boldsymbol{\sigma})$ it may be so. In our applications, the form of $\ln P(\boldsymbol{\sigma})$ depends on the choice of a variational form for the wave function. When this choice is such that the summation is tractable, we can find the optimal coefficients $\eta_{\mathbf{r},\mathbf{r}'}$ that define the spin-glass-like probability distribution $P_{\text{cluster}}(\boldsymbol{\sigma})$ of the form (3.38). We use this probability to build a cluster in the way described in Subsec. 3.3.2, and then accept or reject the proposed cluster flip according to the acceptance probability (3.43).

CHAPTER 4

Results for the BSWF Ansatz

In this Chapter we present our exploration of the BSWF Ansatz introduced in Chap. 3. We implemented from scratch a C++ code to perform t-VMC simulations in both real and imaginary time, using the BSWF Ansatz, as well as a variety of other Ansatzë that are found in the literature. In Sec. 4.1 we focus on imaginary-time optimization of the wave function for a one-dimensional (1D) system, and discuss some details of the strategies we adopt for the Monte Carlo sampling. We compare the BSWF results with those obtained by using the RBM wave functions, belonging to the class of neural-network quantum states. In Sec. 4.2 we compare these Ansatzë results for the real-time simulations of quantum quenches, considering both a translation-invariant quantum quench and a non-translation-invariant one. In Sec. 4.3 we present some novel Ansatzë that extend the BSWF to increase its capabilities, and examine some results they produce in both imaginary and real time. Finally, in Sec.4.4 we present some preliminary results for the BSWF in two-dimensional (2D) systems.

4.1 Imaginary time optimization

We use imaginary-time optimization to test the capability of the BSWF Ansatz to represent the ground state of the TFI model for various system sizes and values of the TFI parameter g (the ratio between the transverse field strength h and the spin-spin coupling J).

As an example, in Fig. 4.1 we show the optimization in imaginary time of a 1D ferromagnetic Ising chain of 32 spins with parameter $g = 0.5$. The three parameters are initialized to 1, and we observe that they smoothly converge to a final value. The energy converges exponentially, as expected from an imaginary-time propagation, and the final value has relative error of $\simeq 10^{-5}$ with respect to the exact energy.

We want to compare the descriptive ability of BSWF with other Ansatzë, but first we need to discuss some technicalities regarding the Monte Carlo sampling.

4.1.1 Monte Carlo sampling

We use the Metropolis algorithm to sample one real configuration $\boldsymbol{\sigma}$ and a pair of shadow configurations $\boldsymbol{h}$, $\boldsymbol{h'}$ according to the probability distribution $q_{\text{BSWF}}(\boldsymbol{\sigma}, \boldsymbol{h}, \boldsymbol{h'})$ defined in Eq. (3.23). Our Monte Carlo moves for a system of N spins consist in proposing the spin flip of N randomly-chosen real spins. They are picked one at a time and their flip is accepted or rejected according to the Metropolis prescription. The same procedure is applied to the two shadow configurations, proposing the spin flip of N randomly-chosen shadow variables for each one of the two shadow configurations. The resulting

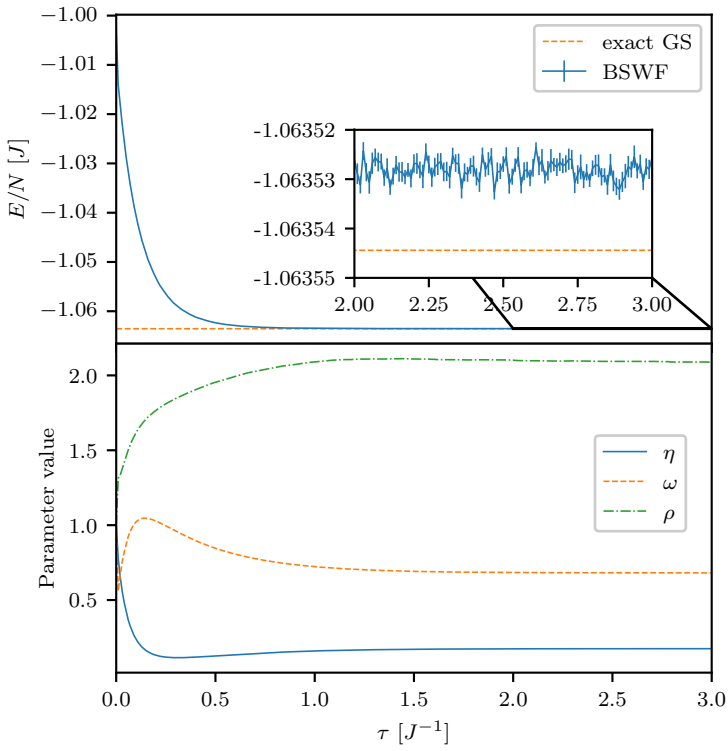

Figure 4.1: Imaginary-time evolution of a BSWF for a 32-spin Ising chain with $g = 0.5$ in PBC. The upper panel shows the convergence, in imaginary time τ, of the energy per spin (solid line), compared with the exact ground state energy (dashed horizontal line). For ease of comparison, an inset shows the details of the energy evolution once it has reached convergence, in the interval from $\tau = 2.0$ to 3.0. The lower panel shows the evolution of the parameters η (real-real coupling), ω (real-shadow coupling) and ρ (shadow-shadow coupling). The parameters are initialized to $\eta = \omega = \rho = 1$. This evolution is integrated via the Euler method with step $\delta\tau = 0.01$, and diagonal bias-10^{-4} regularization is used.

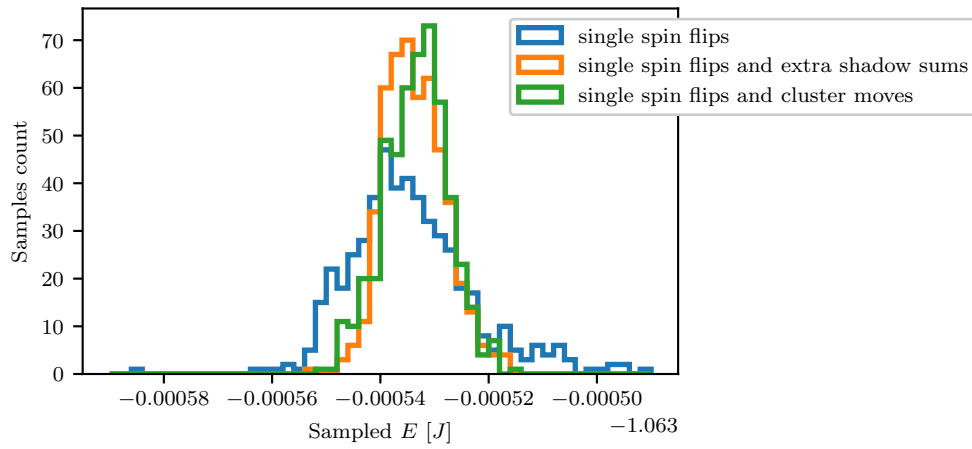

Figure 4.2: Histogram of the calculated BSWF ground-state energies E in a sequence of 500 independent Monte Carlo estimations for a 1D 16-spin TFI chain with $g = 0.5$. The results are shown for three different sampling strategies: "plain" single spin flips; single spin flips with $N_{\text{extra}} = 10$ extra shadow samplings for every sampled real configuration; single spin flips and cluster moves.

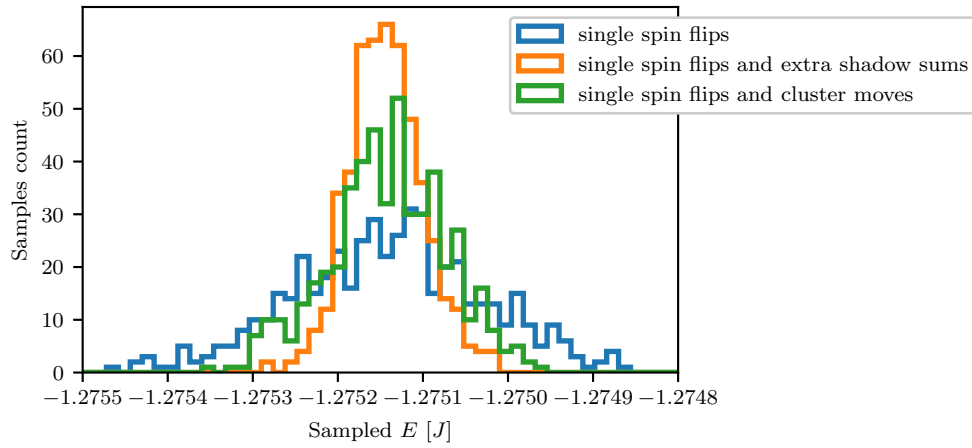

Figure 4.3: Histogram of the calculated BSWF ground-state energies E in a sequence of 500 independent Monte Carlo estimations for a 1D 16-spin TFI chain with $g = 1.0$. The results are shown for three different sampling strategies: "plain" single spin flips; single spin flips with $N_{\text{extra}} = 10$ extra shadow samplings for every sampled real configuration; single spin flips and cluster moves.

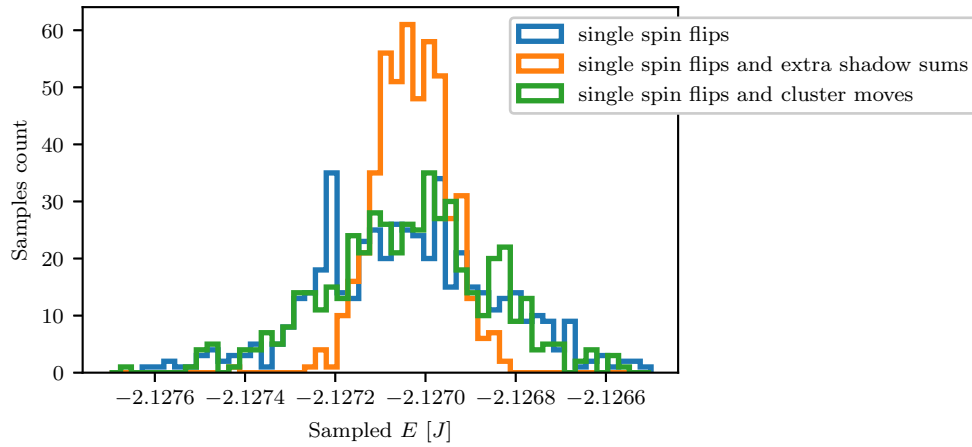

Figure 4.4: Histogram of the calculated BSWF ground-state energies E in a sequence of 500 independent Monte Carlo estimations for a 1D 16-spin TFI chain with $g = 2.0$. The results are shown for three different sampling strategies: "plain" single spin flips; single spin flips with $N_{\text{extra}} = 10$ extra shadow samplings for every sampled real configuration; single spin flips and cluster moves.

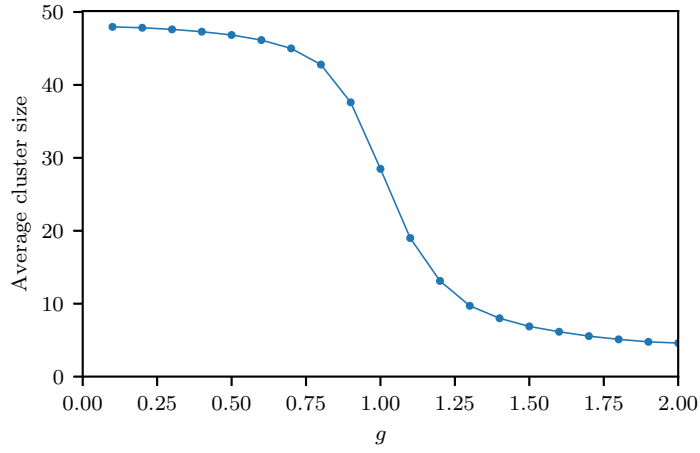

Figure 4.5: Average size of the clusters flipped in a cluster move, as a function of g, for the BSWF approximate ground state of a 1D 16-spin TFI chain. Since we have three layers (one real and two shadow) of 16 spins each, the maximum possible size for a cluster is $16 \times 3 = 48$. The markers represent the data, the lines only serve as visual guides.

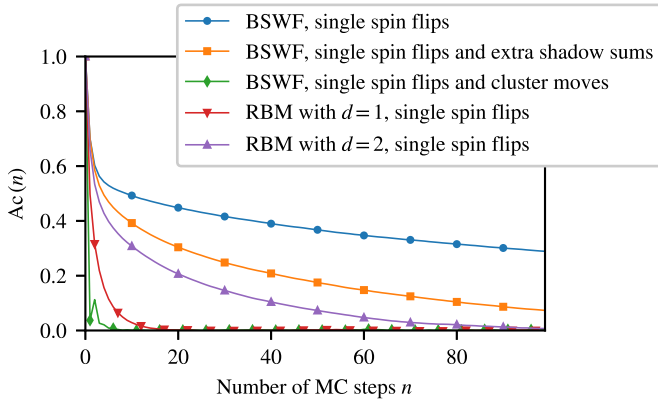

Figure 4.6: Autocorrelation $\mathrm{Ac}(n)$ between two real configurations that are sampled n Monte Carlo steps apart, for a 32-spin chain at $g = 1.0$. Each curve was calculated from a pool of 10^5 real configurations that were sampled for various Ansatzë and sampling methods.

$(\boldsymbol{\sigma}, \boldsymbol{h}, \boldsymbol{h'})$ configuration is then used to evaluate the Monte Carlo estimators that will be averaged in the end. This is what we call the "plain" single spin flip strategy.

With respect to the case of a non-shadow Ansatz, the additional necessity of sampling the shadow configurations makes the expectation values more noisy. A possible strategy to address this difficulty is sampling the shadow variables more often than the real variables, i.e., sampling N_{extra} shadow configurations every time a real configuration is sampled, and averaging the Monte Carlo estimators over the extra samples. Another strategy that we find useful consists in the use of cluster moves of the type described in Subsec. 3.3.2, that allow to flip entire clusters of spins in a single move.

In order to compare the two strategies, we perform 500 independent variational calculations of the ground state energy of a 1D 16-spin TFI chain for a certain value of the TFI parameter g, and plot the histogram of the calculated energies. Each calculation involves a pool of 1.25×10^6 sampled configurations $(\boldsymbol{\sigma}, \boldsymbol{h}, \boldsymbol{h'})$. We perform this calculations in three different ways: using only single spin flip moves; using single spin flips *and* cluster moves; and using single spin flips with extra samplings of the shadow variables. For this comparison, we sample $N_{\text{extra}} = 10$ shadow configurations for every real configuration sampled. Figs. 4.2, 4.3, and 4.4 show the results for $g = 0.5$, 1.0, and 2.0 respectively.

From Fig. 4.2, we observe that for $g = 0.5$, in the ferromagnetic phase, the cluster moves improve the precision of the calculated energy in a way comparable to the extra shadow sums. However, the extra shadow sums involve, in this case, a computational cost ~ 10 times greater with respect to plain single flip moves, while the cluster moves only involve a computational cost ~ 1.25 times greater, so they are much more convenient. In Fig. 4.3 we observe that at the critical point, $g = 1$, the cluster moves still improve the sampling, but not as much as performing extra shadow samplings. Still, the difference in computational cost make the cluster moves very attractive. Fig. 4.4 shows that for $g = 2$, in the paramagnetic phase, the cluster moves do not improve any more on the plain single spin flip strategy. Here the extra shadow sums are definitely helpful.

This behavior as g increases is expected, since in the ferromagnetic phase the spins exhibit long-range order and, as a consequence the cluster moves flip big sets of spins at once. Deep into the paramagnetic phase, on the other hand, the clusters become smaller

and smaller, eventually reducing to just one spin. In this limit, a cluster move is equivalent to a single spin flip. Fig. 4.5 shows this trend in more detail, plotting the average size of the flipped clusters as a function of g. The fast decrease in the cluster sizes around $g \simeq 1$ hints at the ferromagnetic-paramagnetic phase transition that occurs in the thermodynamic limit.

We also observe a clear increase of the variance near the phase transition. To further investigate the effect of the sampling strategy, Fig. 4.6 shows the autocorrelation $\mathrm{Ac}(n)$ between two real configurations that are sampled n Monte Carlo steps apart, for a 32-spin chain at $g = 1.0$. If $\sigma_i(m)$ is the value of the i^{th} spin that was sampled at the m^{th} Monte Carlo step, then we define the autocorrelation $\mathrm{Ac}(n)$ by

$$\mathrm{Ac}(n) = \frac{\left| \frac{1}{N_{MC}-n} \sum_{m=1}^{N_{MC}-n} \sigma_i(m)\sigma_i(m+n) - \left[\frac{1}{N_{MC}} \sum_{m=1}^{N_{MC}} \sigma_i(m) \right]^2 \right|}{\frac{1}{N_{MC}} \sum_{m=1}^{N_{MC}} \sigma_i(m)^2 - \left[\frac{1}{N_{MC}} \sum_{m=1}^{N_{MC}} \sigma_i(m) \right]^2}, \tag{4.1}$$

where N_{MC} is the total number of configurations that are sampled; since the system is translation-invariant, the choice of the spin index i is irrelevant. We sample $N_{MC} = 10^5$ configurations by various methods. The results in Fig. 4.6 show that the BSWF using only single spin flip moves has the highest autocorrelation between samples. The autocorrelation reduces when extra shadow sums are added, but the best result is obtained with cluster moves. In the same Figure, we also compare the autocorrelations obtained with a different Ansatz, an RBM; we consider two densities, $d = 1$ and $d = 2$. The RBM Ansatz does not use shadow variables, and, because of its functional form, does not allow the use of cluster moves, so we only use single spin flip moves. We observe that both the RBM with $d = 1$ and $d = 2$ have lower autocorrelations than the BSWF without cluster moves. This can be understood if we consider that the BSWF has to sample the shadow configurations as well as the real ones, and this makes the autocorrelation of the real configurations more persistent. On the other hand, when we use cluster moves with a BSWF, we move clusters of real and shadow spins at once, and we get lower autocorrelations than the RBMs. However, another observation is in order: we notice that the RBM with $d = 2$ has higher autocorrelation than the RBM with $d = 1$; since the $d = 2$ RBM is more descriptive, we conclude that a better description of the many-body correlations makes the autocorrelation in the Monte Carlo sampling more persistent. As we will see in Subsec. 4.1.3, the BSWF gives a better description of the ground state than the RBM with $d = 1$, but worse than the RBM with $d = 2$. So, while it is true that the efficiency of the sampling of the BSWF with cluster moves achieves the lowest autocorrelation in the Monte Carlo samples, we should also notice that the BSWF gives a less accurate description of the many-body correlations than the $d = 2$ RBM, probably due to the lack of explicit long-range terms in the BSWF functional form; this fact may also play a role in the low autocorrelation of the sampling of the BSWF with cluster moves.

When dealing with the BSWF Ansatz, in general we always use the cluster moves because they are very useful in the ferromagnetic phase, and even in the paramagnetic phase the slowdown they cause is acceptable. When more precision is needed, we usually resort to sampling more configuration to have a more significant statistics. Only rarely we resort to performing extra shadow sums, especially in the paramagnetic phase.

4.1.2 Choice of the initial parameters

Here we discuss how the outcome of the optimization process can depend on the choice of the initial parameters: there could exist bad starting points that generate instabilities

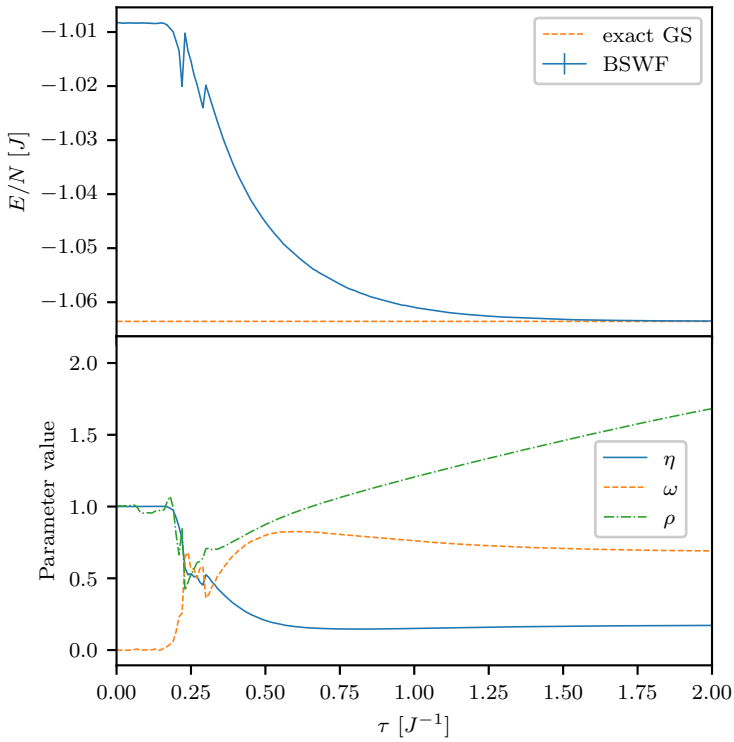

Figure 4.7: Imaginary-time evolution of a BSWF for a 32-spin Ising chain with $g = 0.5$ in PBC. The upper panel shows the evolution, in imaginary time τ, of the energy per spin (solid line), compared with the exact ground state energy (dashed horizontal line). The lower panel shows the evolution of the parameters η (real-real coupling), ω (real-shadow coupling) and ρ (shadow-shadow coupling). The parameters are initialized to $\eta = \rho = 1$ and $\omega = 0$. This evolution is integrated via the Euler method with step $\delta\tau = 0.01$, and diagonal bias-10^{-4} regularization is used.

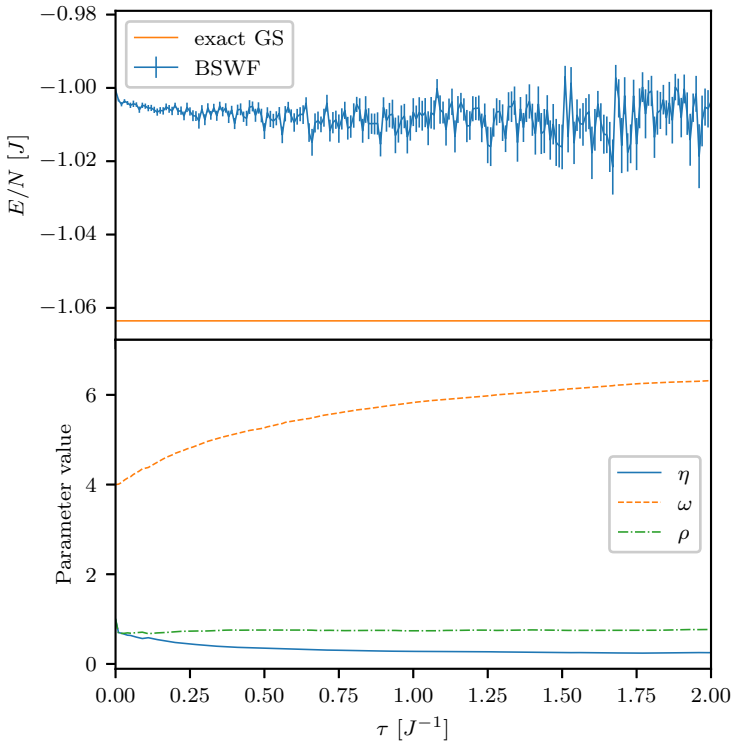

Figure 4.8: Imaginary-time evolution of a BSWF for a 32-spin Ising chain with $g = 0.5$ in PBC. The upper panel shows the evolution, in imaginary time τ, of the energy per spin (solid line), compared with the exact ground state energy (dashed horizontal line). The lower panel shows the evolution of the parameters η (real-real coupling), ω (real-shadow coupling) and ρ (shadow-shadow coupling). The parameters are initialized to $\eta = \rho = 1$, and $\omega = 4$. This evolution is integrated via the Euler method with step $\delta\tau = 0.01$, and diagonal bias-10^{-4} regularization is used.

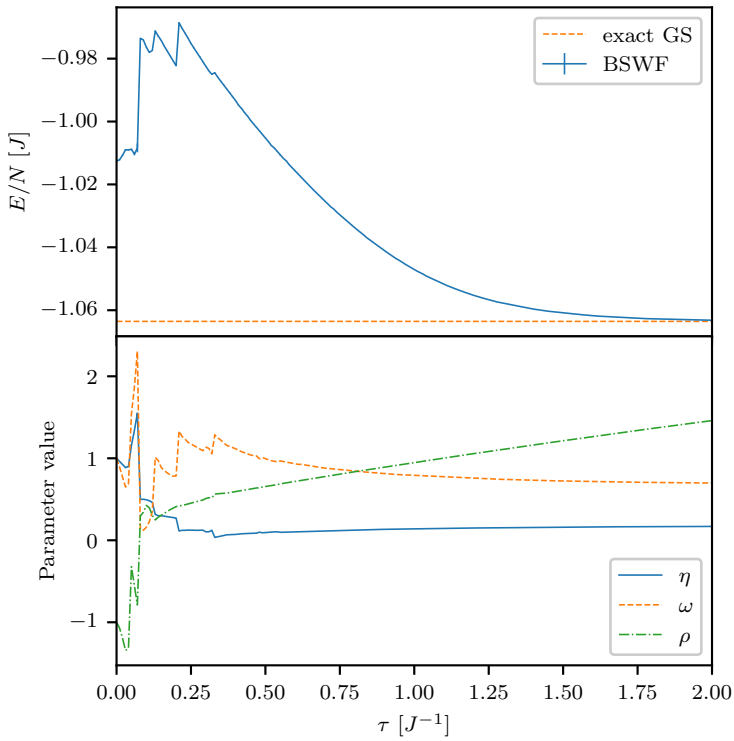

Figure 4.9: Imaginary-time evolution of a BSWF for a 32-spin Ising chain with $g = 0.5$ in PBC. The upper panel shows the evolution, in imaginary time τ, of the energy per spin (solid line), compared with the exact ground state energy (dashed horizontal line). The lower panel shows the evolution of the parameters η (real-real coupling), ω (real-shadow coupling) and ρ (shadow-shadow coupling). The parameters are initialized to $\eta = \omega = 1$ and $\rho = -1$. This evolution is integrated via the Euler method with step $\delta\tau = 0.01$, and diagonal bias-10^{-4} regularization is used.

or cause the state to get stuck in local minima. Let us discuss these pathological cases, so that we can identify strategies to avoid them.

A bad choice for the initial parameters consists in setting to zero the real-shadow coupling ω, as we purposefully did at the beginning of the optimization shown in Fig. 4.7, in which we try to optimize the Ising ground state with $g = 0.5$. In an initial state with $\omega = 0$ the shadow variables are not correlated with the real ones, so that the quantum state becomes independent of the shadow-shadow coupling ρ. As a consequence, ρ initially evolves erratically. In this particular case, eventually ω becomes significantly different from zero, and the wave function reaches a less problematic state, from which the optimization can continue smoothly; it is not shown, but it eventually reaches the same state as the succesful optimization of Fig. 4.1. Obviously, this lucky outcome cannot be controlled in general, and the state can become stuck in a bad region of parameter space. An example in which the optimization gets stuck is shown in Fig. 4.8. It's again $g = 0.5$, and the initial parameters are $\eta = 1$, $\omega = 4$, and $\rho = 1$. This time the initial real-shadow coupling ω is too large, so that the shadow variables are "locked" to their real counterparts, meaning that the probability of the shadow variables assuming different values from the real ones becomes negligible, and this compromises the Monte Carlo sampling. This causes the evolution to remain stuck to large ω values, and ω even increases over imaginary time. In this case the evolution cannot converge to the optimized value that can be reached with a more careful choice of the starting point. Such a choice was made for the successful optimization shown in Fig. 4.1, starting from $\eta = 1$, $\omega = 1$, and $\rho = 1$. This choice of ω ensures that the real degrees of freedom are correlated with the shadow ones but not "locked" together.

Another example of instability is shown in Fig. 4.9, where we are again trying to simulate the $g = 0.5$ ground state. The initial parameters are $\eta = 1$ for the real-real coupling, $\omega = 1$ for the real-shadow coupling, and $\rho = -1$ for the shadow-shadow coupling. The problem here is that the negative ρ causes the shadow variables to have an antiferromagnetic-like behavior, while the ground state we are looking for is ferromagnetic. This causes an unstable evolution, with violent jumps of the parameters. Eventually, a lucky (and unpredictable) jump ends up in a good region of parameter space, from which the optimization continues smoothly.

We conclude that, even if the initial parameters do not need to be fine-tuned, there exist problematic starting points to be avoided, and the initial parameters cannot be set randomly. As a rule of thumb, the real-shadow coupling ω should always be large enough to have significant correlations between the real and shadow variables, but not so large that they become "locked" together. The other two couplings, η (real-real) and ρ (shadow-shadow) should be chosen coherently with the expected phase of the state we want to describe.

In a more general setting, for example in higher dimensions, when the choice of the initial parameters is found to be problematic, it could be useful to first perform a preliminary optimization using different methods than imaginary-time evolution (or, equivalently, stochastic reconfiguration), like, e.g., stochastic gradient descent to minimize the energy. Then, one can conclude the optimization with a run of imaginary-time evolution, in order to benefit from its characteristic exponential convergence.

4.1.3 Imaginary time results and comparison with other Ansatzë

Fig. 4.10 shows the values of the optimized BSWF parameters for the ground state of a 32-spins TFI chain as a function of g, for g in the range $[0.1, 2.0]$. We find that the real-real coupling η has only small variations in the observed range of g. The shadow-shadow

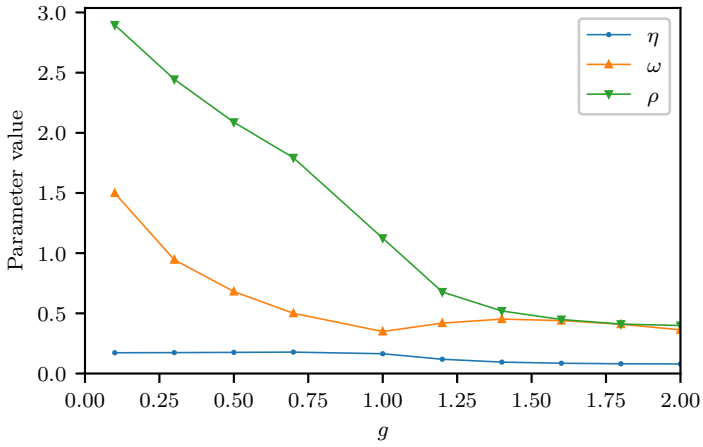

Figure 4.10: Variationally optimized BSWF parameters for the ground state of a 32-spins 1D TFI chain as a function of the TFI parameter g. η is the real-real coupling, ω the real-shadow coupling, and ρ the shadow-shadow coupling. The markers indicate the data, the lines only serve as visual guides.

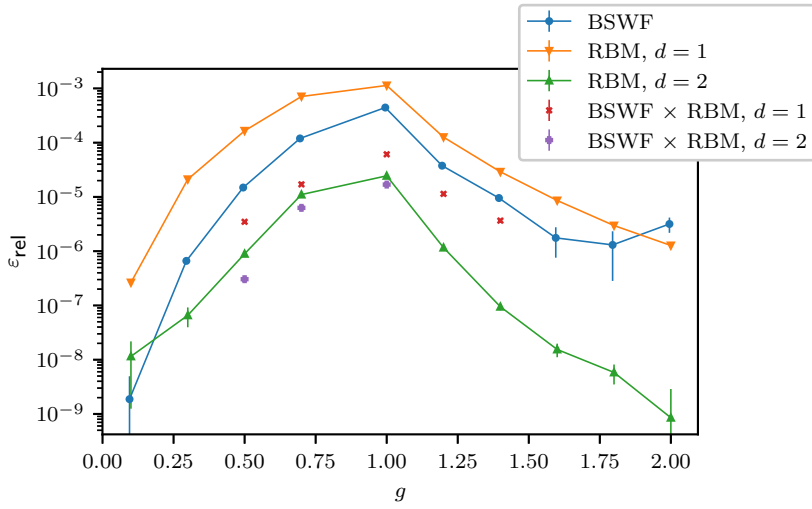

Figure 4.11: Relative error ε_{rel} in the variational energy of the 32-spin 1D TFI ground state as a function of g, for the BSWF, RBM, and BSWF×RBM wave functions. ε_{rel} is defined in Eq.(4.2). The RBM and BSWF×RBM are translation-invariant, and densities $d = 1$ and $d = 2$ are investigated. The error bars are often smaller than the plot markers.

coupling ρ is monotonic decreasing and it is the most variable of the three parameters. Interestingly, the real-shadow coupling ω is not monotonic: it has a minimum at $g \simeq 1.0$, which in the thermodynamic limit is the critical point. The variability of ρ and ω, opposed to the weak dependence of η on g, indicates that the algorithm is heavily relying on the shadow degrees of freedom to represent the many-body correlations in the system.

We measure the accuracy of the variational energy by the relative error ε_{rel}, which is defined in terms of the variational energy E_{var} and the exact ground-state energy E_{GS} as

$$\varepsilon_{\text{rel}} = \frac{E_{\text{var}} - E_{\text{GS}}}{|E_{\text{GS}}|}\,, \tag{4.2}$$

that is positive, since the variational energy is always an upper bound for the exact energy. Fig. 4.11 shows the relative error as a function of g in the range $[0.1, 2.0]$, comparing the BSWF error with the errors of the translation-invariant RBM's with densities 1 and 2. The hardest state to accurately represent for all the Ansatzë is the $g = 1.0$ ground state. The accuracy of BSWF is found to be, generally, between those of density-1 and density-2 RBM's. The BSWF is very accurate deep into the ferromagnetic phase, and for very low g its error is even lower than the density-2 RBM. On the other hand, it becomes less accurate for high g, where its error becomes comparable with or higher than the error of the density-1 RBM. We remark that, using only 3 parameters, the BSWF achieves an accuracy comparable to density-1 and density-2 translation-invariant RBM's, which, for this system of 32 spins, use 34 and 67 parameters, respectively.

4.1.4 The BSWF×RBM Ansatz

We want to investigate the possibility of using the concise BSWF Ansatz to "improve" other Ansatzë, using it as a multiplicative factor. This way, from a generic Ansatz we obtain a "BSWF-improved" that has only three more parameters, namely η, ω, and ρ. To this end, we introduce the BSWF×RBM Ansatz, defined as the product of a BSWF and an RBM:

$$\psi_{\text{BSWF}\times\text{RBM}}(\boldsymbol{\sigma}; \eta, \omega, \rho, \{a_j\}, \{b_k\}, \{W_{j,k}\}) = \\ \psi_{\text{BSWF}}(\boldsymbol{\sigma}; \eta, \omega, \rho) \cdot \psi_{\text{RBM}}(\boldsymbol{\sigma}; \{a_j\}, \{b_k\}, \{W_{j,k}\})\,. \tag{4.3}$$

This Ansatz depends on the parameters of both the BSWF and the RBM, so it has 3 extra parameters with respect to the RBM at a certain density. It includes both the BSWF and the RBM wave functions as special cases, since the parameters can be chosen so that $\psi_{\text{BSWF}} \equiv 1$, or $\psi_{\text{RBM}} \equiv 1$.

To optimize the BSWF×RBM wave function, η, ω, and ρ are initialized to the optimal BSWF values, obtained from the imaginary-time propagation of ψ_{BSWF} alone; the RBM biases a_j and b_k are initialized to zero and the weights $W_{j,k}$ are set to independent Gaussian-distributed values with zero mean. If the initial $W_{j,k}$ are too close to zero, the imaginary-time propagation easily gets stuck in a state that is very close to the pure ψ_{BSWF}; on the other hand, if they are too large, the initial state is too random and we lose the benefit of starting from the optimized BSWF parameters. So, some care is needed in the choice of the standard deviation of the Gaussian distribution from which we draw the initial weight $W_{j,k}$. For example, we found that a standard deviation of ~ 0.01 is too small, and we used ~ 0.1 for the optimizations that are shown here.

Along with the BSWF and RBM results, Fig. 4.11 also shows the relative error in the ground state energy for an optimized BSWF×RBM, where the RBM factor is translation-invariant. Optimizations were performed using a density-1 and a density-2 RBM factor.

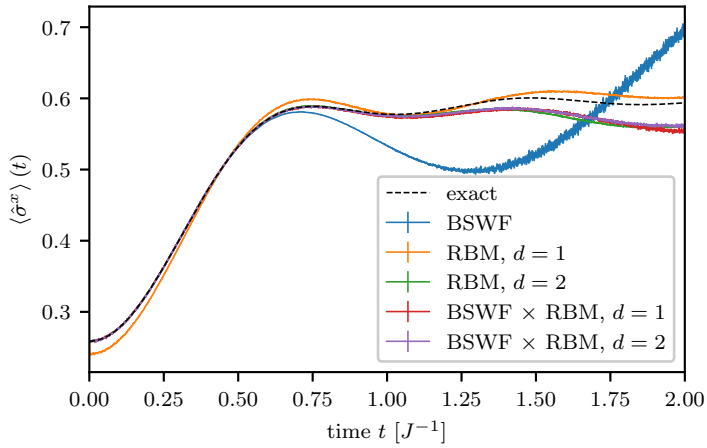

Figure 4.12: Transverse magnetization for a quantum quench of the 32-spin 1D TFI from $g = 0.5$ to $g = 1$. The dashed line shows the exact solution, obtained by numerical integration of the Bogoliubov-de Gennes equations. The solid lines display the variational approximations obtained by the following Ansatzë (see the legend): the BSWF, the translation-invariant RBM with density $d = 1$ and $d = 2$, and the translation-invariant BSWF×RBM with density $d = 1$ and $d = 2$.

We observe that the BSWF×RBM improves on both the BSWF and the RBM at the same density. The improvement is especially significant at density 1, where for $g = 1$ the relative error is reduced by about an order of magnitude; at density 2 the improvement is less significant, albeit still present. Since we are looking for Ansatzë that are as concise as possible, the density 1 results are particularly promising, since for 32 spins the BSWF×RBM at density 1 has 37 parameters, while at density 2 it has 70 parameters. These results show that the BSWF can be used to improve other Ansatzë increasing only by three the number of variational parameters, at the cost of sampling the shadow variables.

4.2 Real time

4.2.1 Translation-invariant quench

In a *quantum quench*, the system is prepared in the ground state of some Hamiltonian, and then it is evolved under a modified Hamiltonian. Typically one considers a Hamiltonian that depends on some parameters: the system is prepared in the ground state for a set of parameters, and evolved under a different set. This simulates the effect of a sudden change that drives the system out of equilibrium, generating nontrivial dynamics.

We consider the TFI Hamiltonian, which depends on the parameter $g = |h/J|$. We prepare a 32-spin 1D chain in the ground state for $g = 0.5$, and evolve it under $g = 1$. The $g = 0.5$ ground state is a ferromagnetic state protected by an energy gap Δ, while the critical $g = 1$ TFI has a gapless spectrum. Although this is true only in the thermodynamic limit, our 32-spin chain is big enough that going from $g = 0.5$ to $g = 1$ drastically reduces the energy gap by an order of magnitude, from $\Delta \simeq 2\,J$ to $\Delta \simeq 0.2\,J$, so we can expect a richer dynamics at $g = 1$. We integrate the equations of motion via the Heun algorithm, with time step $\delta t = 10^{-3}$.

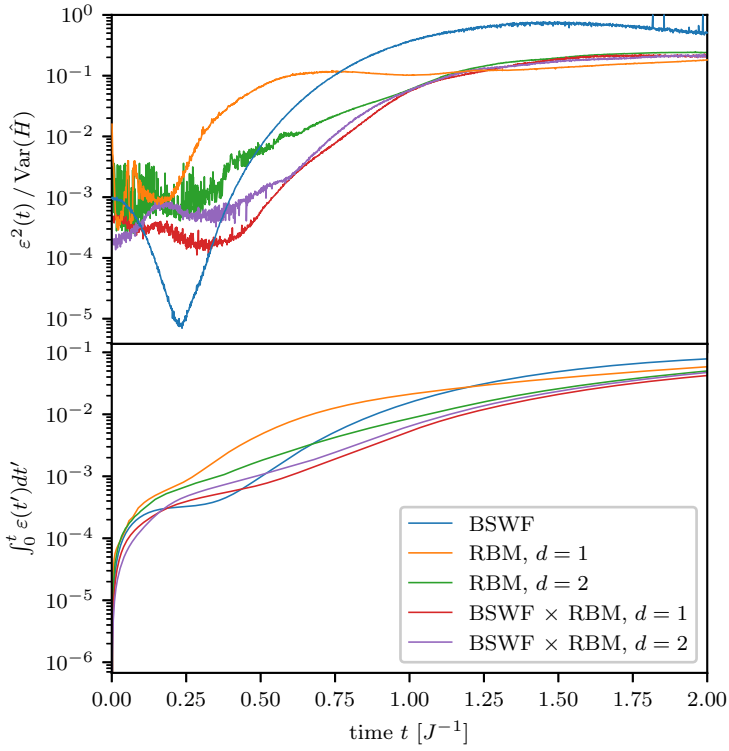

Figure 4.13: Squared error per unit time divided by the energy variance $\varepsilon^2(t)/\mathrm{Var}(\hat{H})$ (upper panel) and integral of $\varepsilon(t)$ over time (lower panel) for a quantum quench of the 32-spin 1D TFI from $g = 0.5$ to $g = 1$. The errors for the following Ansatzë are displayed (see the legend): the BSWF, the translation-invariant RBM with density $d = 1$ and $d = 2$, and the translation-invariant BSWF×RBM with density $d = 1$ and $d = 2$.

Fig. 4.12 shows the evolution of the transverse magnetization per spin

$$\langle \hat{\sigma}^x \rangle(t) = \frac{1}{N} \sum_{\mathbf{r}} \langle \hat{\sigma}^x_{\mathbf{r}} \rangle(t) \tag{4.4}$$

for various Ansatzë, compared with the exact evolution calculated by numerical integration of the Bogoliubov-de Gennes equations, resulting after Jordan-Wigner transformation of the spin operators [46].

First of all, we focus on the BSWF: for short times its transverse magnetization is very accurate, but around $t \simeq 0.5$ it becomes less and less consistent with the exact solution. Now we consider the translation-invariant RBM with density $d = 1$. From its transverse magnetization at $t = 0$, we see that the initial state of this Ansatz is visibly inaccurate compared to the other Ansatzë. However, for long times it is qualitatively more consistent than the BSWF with the exact solution. Again, we attribute this to the fact that the definition of the RBM contains long-range correlation terms. The translation-invariant RBM with density $d = 2$ starts from a much more accurate state, and it is very consistent with the exact solution for short times. It begins to be less consistent at around $t \simeq 1$, where the transverse magnetization starts to depart from the exact value, but it still qualitatively captures the oscillations of $\langle \hat{\sigma}^x \rangle(t)$. Finally, the BSWF×RBM Ansatzë for density 1 and 2 have very similar transverse magnetization to the $d = 2$ RBM, and the plots partially overlap in the figure making them not easily distinguishable.

Further details are provided by Fig. 4.13. In the upper panel, it shows the squared error per unit time normalized by the variance of the energy, $\varepsilon^2(t)/\mathrm{Var}(\hat{H})$, whose maximum possible value is 1; in the lower panel, it shows the integral over time of the error per unit time $\varepsilon(t)$, that gives an upper bound for the Fubini-Study distance between the exact evolution and the variational evolution (see Subsec. 1.1.6). The BSWF has very low $\varepsilon^2(t)$ for short times, and there is a time range in which it is even the lowest, by about two orders of magnitude, among the considered Ansatzë. However, around $t \simeq 0.5$ it begins to rapidly increase, and eventually it almost saturates to 1. The integral of $\varepsilon(t)$ is very low in the beginning, but eventually it becomes the largest. This failure of the BSWF for long times, that we already observed in the transverse magnetization, is likely due to the fact that the BSWF by definition lacks long-range correlation terms, making it unable to capture volume-law entanglement. Turning to the other Ansatzë, for long times their squared errors $\varepsilon^2(t)$ all approach a similar value, close to $2 \cdot 10^{-1}$. For intermediate times, the RBM with $d = 1$ has the largest error and largest integral of $\varepsilon(t)$, but eventually it gets surpassed in both respects by the BSWF; then come, in order from highest to lowest error, the RBM with $d = 2$, the BSWF×RBM with $d = 2$, and the BSWF×RBM with $d = 1$. The two BSWF×RBM, however, are very close to each other.

There are two facts that might look puzzling: first, the BSWF×RBM with $d = 1$ has lower $\varepsilon^2(t)$ and lower $\int_0^t \varepsilon(t')\, dt'$ than the BSWF×RBM with $d = 2$, except for very short times; second, there is a time range in which the pure BSWF has lower error than both the BSWF×RBM wave functions, even if by construction it is less expressive. However, it should be pointed out that different Ansatzë go through different states $\psi_{\text{var}}(t)$ during their evolution, and $\varepsilon^2(t)$ measures the error with respect to the *local-in-time* exact evolution, that is, the Schrödinger evolution that at time t goes through state $\psi_{\text{var}}(t)$. So, even if these quantities are useful to get an idea of the quality of the variational evolution, we cannot assume that the best wave function will always have the lowest $\varepsilon^2(t)$. A similar consideration can be made for the integral $\int_0^t \varepsilon(t')\, dt'$. It provides a useful error bound, as explained in Subsec. 1.1.6, but, being an upper bound, we cannot assume that it will always be lowest for the evolution that is actually the closest one to the exact trajectory.

Taking into account these caveats, the reduction in both $\varepsilon^2(t)$ and $\int_0^t \varepsilon(t')\,dt'$ still suggests that the BSWF×RBM ansatz improves over the RBM results, especially for density 1: comparing the $d = 1$ BSWF×RBM (that has 37 parameters) with the $d = 2$ RBM (that has 70 parameters), we observe that the accuracy of in the evolution of the transverse magnetization is comparable, and both the error per unit time and its time integral are lower for the BSWF×RBM Ansatz. We remark again that this improvement with respect to the $d = 1$ RBM was achieved increasing the number of parameters only by three, at the cost of sampling the shadow variables along with the real variables.

4.2.2 Localized quench

In these quenches we consider the following Hamiltonian:

$$\hat{H} = -J \sum_{\langle \mathbf{r},\mathbf{r}' \rangle} \hat{\sigma}^z_{\mathbf{r}} \hat{\sigma}^z_{\mathbf{r}'} - h \sum_{\mathbf{r}} \hat{\sigma}^x_{\mathbf{r}} - \delta h\, \hat{\sigma}^x_{\mathbf{r}=\mathbf{0}}\,, \tag{4.5}$$

where δh introduces a localized inhomogeneity in the magnetic field, that is only present on the $\mathbf{r} = \mathbf{0}$ site. We prepare a 16-spins 1D TFI chain in the ground state with $\delta h = 0$ for some value of $g = |h/J|$. Then the inhomogeneity is switched on, setting δh to a nonzero value, and the state is evolved in real time. This system is not translation-invariant, so we use the non-translation-invariant BSWF, that in one dimension is defined by

$$\psi(\boldsymbol{\sigma}; \{\eta_j\}, \{\omega_j\}, \{\rho_j\}) = \sum_{\boldsymbol{h}} \exp\left(\sum_j \eta_j \sigma_j \sigma_{j+1} + \sum_j \omega_j \sigma_j h_j + \sum_j \rho_j h_j h_{j+1} \right)\,, \tag{4.6}$$

where j is the integer index identifying a site in the 1D lattice. This wave function has $3N$ parameters for N spins. In our case $N = 16$ and we have 48 parameters.

We compare the BSWF results with those obtained by a non-translation-invariant RBM with density $d = 1$. Since we do not impose any constraint to impose translation symmetry, this RBM for 16 spins has 288 parameters. So, it is much more expressive (and computationally expensive) than the translation-invariant RBM at the same density, that only has 18 parameters. This holds also for the initial state. The exact ground state with $\delta h = 0$ is translation-invariant, but since no constraint is imposed, the non-translation-invariant RBM achieves a state that is only *approximately* translation-invariant. However, it exploits its large parameter space to obtain a variational approximation that is much closer to the true ground state than the one obtained by the translation-invariant RBM at the same density.

During the dynamics, we monitor the fluctuation in the transverse magnetization of the i^{th} spin, defined as

$$\delta\langle \hat{\sigma}^x_i \rangle(t) = \langle \hat{\sigma}^x_i \rangle(t) - \langle \hat{\sigma}^x_i \rangle(0)\,. \tag{4.7}$$

The additional transverse field δh is applied to the site of index $i = 0$.

The color maps in Fig. 4.14 show the absolute value of this fluctuation for a quench with $g = 0.5$ and $\delta h = 1\,J$, comparing the BSWF, the RBM, and the exact results obtained via integration of the Bogoliubov-de Gennes equations. The evolutions are very similar, so we look at the detail of the transverse magnetization of the spins on sites $i = 0$, 1, and 2 in Fig. 4.15. We observe that both the BSWF and the RBM are very accurate for short times, then, around $t \simeq 1.25$, the BSWF becomes less quantitatively consistent while still capturing the qualitative behavior. The error per unit time and its integral over time, plotted in Fig. 4.16, show that at the beginning of the simulation, the error of the BSWF

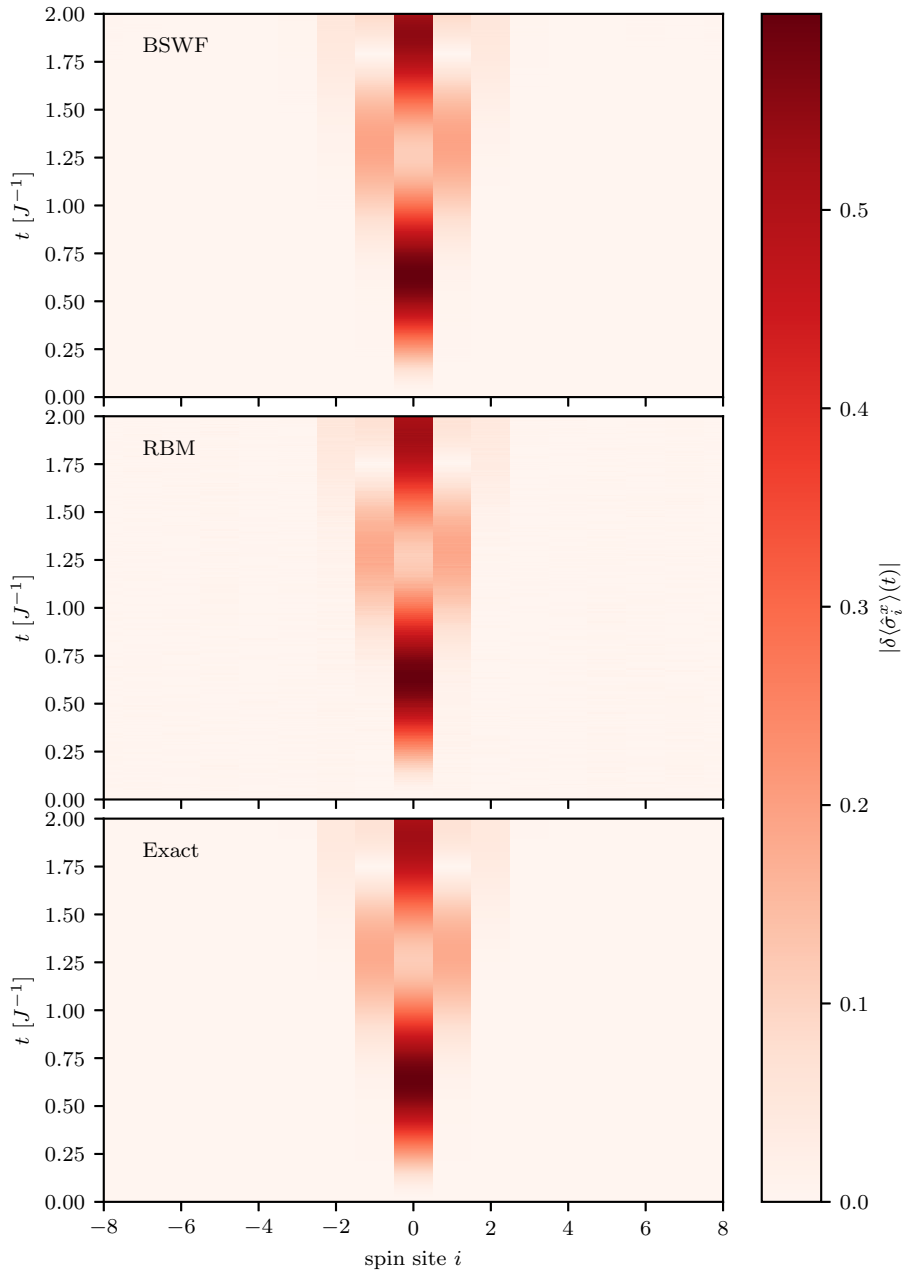

Figure 4.14: Absolute value of the fluctuations in the transverse magnetization of the spins of a 16-spin chain after a localized quench. The system is prepared in the TFI ground state with $g = 0.5$, and then an additional transverse field $\delta h = 1\,J$ on site $i = 0$ is switched on, and the system is evolved in real time. The results are shown for the BSWF Ansatz (upper panel), the RBM Ansatz (middle panel), and the exact evolution (lower panel).

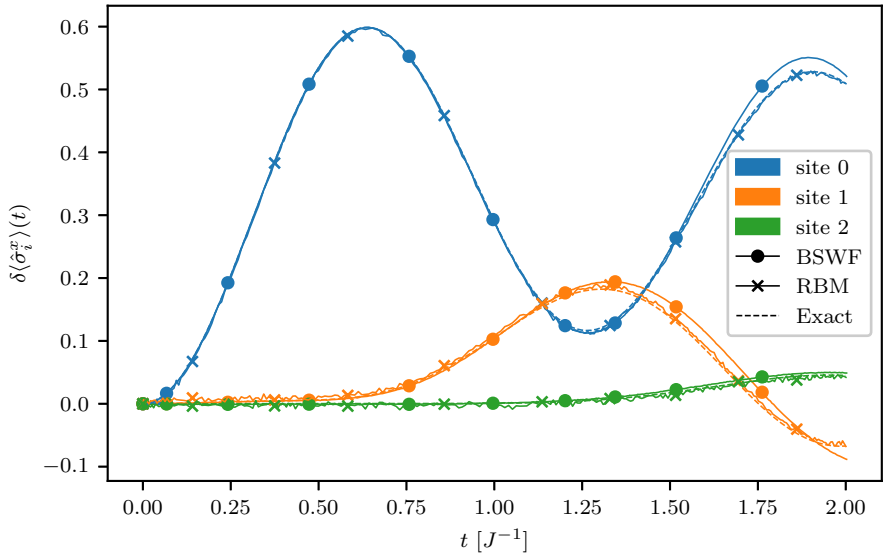

Figure 4.15: Fluctuations in the transverse magnetization of the spins of a 16-spin chain after a localized quench. The system is prepared in the TFI ground state with $g = 0.5$, and then an additional transverse field $\delta h = 1\,J$ on site $i = 0$ is switched on, and the system is evolved in real time. The figure shows the evolution of the spins on site $i = 0$, 1, and 2. The results are shown for the exact evolution, the BSWF Ansatz, and the RBM Ansatz.

is comparable with the RBM's one, but then it increases more rapidly. For these calculations, we employed the adaptive time-step scheme described in Subsec. 1.2.3 applied to the Heun ODE integrator, and the value of the time step during the simulations is shown in Fig. 4.17. The tolerance for the adaptive scheme was set to $\lambda = 10^{-6}$, and we observe that both the calculations benefited from the possibility of enlarging the time step in certain time ranges. This is especially true for the RBM, whose time-step change spans about an order of magnitude during the evolution. The reason probably lies in the greater expressiveness of the RBM, that has more parameters.

The Figs. 4.18, 4.19, and 4.20 show the same information for a quench with $g = 1$ and $\delta h = 1\,J$. The comparison of the accuracies of the BSWF and the RBM is similar to the $g = 0.5$ case, but the departure of the BSWF from the exact values happens somewhat earlier. From the color maps of the transverse magnetization in the two quenches, we observe that the BSWF correctly captures the speed of propagation of the perturbation caused by the switching on of δh, which is faster for $g = 1$ than for $g = 0.5$, as it is expected for the critical regime compared to the ferromagnetic phase. The adaptive time step variability, shown in Fig. 4.21, is similar to the $g = 0.5$ case for the RBM, but it is less variable for the BSWF.

We remark that the BSWF produces quantitatively accurate results for short times, and for long times they are still qualitatively consistent with the exact results. This is achieved with only 48 parameters, compared to the 288 parameters of the non-translation-invariant RBM with density 1.

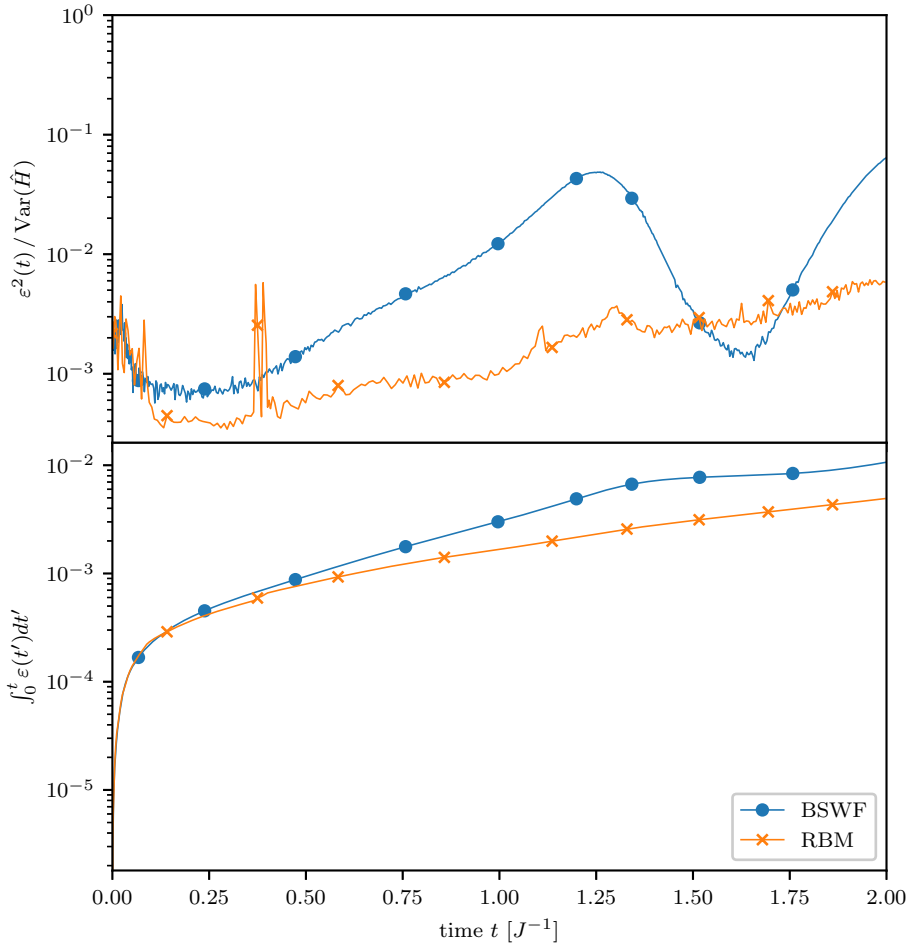

Figure 4.16: Squared error per unit time divided by the energy variance $\varepsilon^2(t)/\mathrm{Var}(\hat{H})$ (upper panel) and integral of $\varepsilon(t)$ over time (lower panel) for a 16-spin TFI chain under a localized quench. The system is prepared in the TFI ground state with $g = 1.0$, and then an additional transverse field $\delta h = 1\,J$ on site $i = 0$ is switched on, and the system is evolved in real time. The figure shows the evolution of the spins on site $i = 0$, 1, and 2. The results are shown for the exact evolution, the BSWF Ansatz, and the RBM Ansatz.

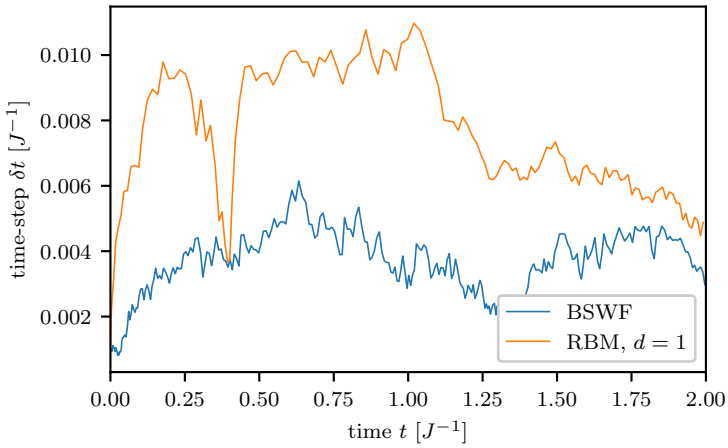

Figure 4.17: Value of the adaptive time step during the simulation of a 16-spin TFI chain under a localized quench. The system is prepared in the TFI ground state with $g = 0.5$, and then an additional transverse field $\delta h = 1\,J$ on site $i = 0$ is switched on, and the system is evolved in real time. The results are shown the BSWF Ansatz and the RBM Ansatz. The tolerance used to adaptively change the time step of the Heun integrator is $\lambda = 10^{-6}$.

4.3 Other BSWF-inspired Ansatzë

4.3.1 The multilayer BSWF

The BSWF is the variationally-optimized version of a single path-integral step. We may similarly introduce an Ansatz consisting of more than one variationally-optimized path-integral steps. This introduces additional layers of shadow variables, and for N_L layers the wave function reads

$$\begin{aligned}\psi_{\text{BSWF},N_L}(\boldsymbol{\sigma};\eta,\{\omega^{(k)}\},\{\rho^{(k)}\}) = \\ \sum_{\boldsymbol{h}^{(1)},\ldots,\boldsymbol{h}^{(N_L)}} \exp\left(\eta \sum_{\langle \mathbf{r},\mathbf{r}'\rangle}\sigma_{\mathbf{r}}\sigma_{\mathbf{r}'} + \omega^{(1)}\sum_{\mathbf{r}}\sigma_{\mathbf{r}}h^{(1)}_{\mathbf{r}} + \rho^{(1)}\sum_{\langle \mathbf{r},\mathbf{r}'\rangle} h^{(1)}_{\mathbf{r}}h^{(1)}_{\mathbf{r}'}\right. \\ \left. + \sum_{l=2}^{N_L}\omega^{(l)}\sum_{\mathbf{r}} h^{(l-1)}_{\mathbf{r}}h^{(l)}_{\mathbf{r}} + \sum_{l=2}^{N_L}\rho^{(l)}\sum_{\langle \mathbf{r},\mathbf{r}'\rangle} h^{(l)}_{\mathbf{r}}h^{(l)}_{\mathbf{r}'}\right),\end{aligned} \tag{4.8}$$

where $h^{(l)}_{\mathbf{r}}$ denotes the shadow variable on site $\mathbf{r}$ in the l^{th} shadow layer. For N spins, this wave function has $2N_L + 1$ parameters. We call this Ansatz the *multilayer BSWF*, or the N_L*-layers BSWF*. Notice that with $N_L = 1$ we recover the usual BSWF. Also, it is noteworthy that the cluster moves described in Subsec. (3.3.2) can be readily applied to the Monte Carlo sampling for this wave function.

Imaginary time

Fig. 4.22 shows an optimization with $N_L = 2$ for a 32-spins TFI chain with $g = 0.5$. The initial state is prepared in such a way that it is close to the optimized BSWF with $N_L = 1$

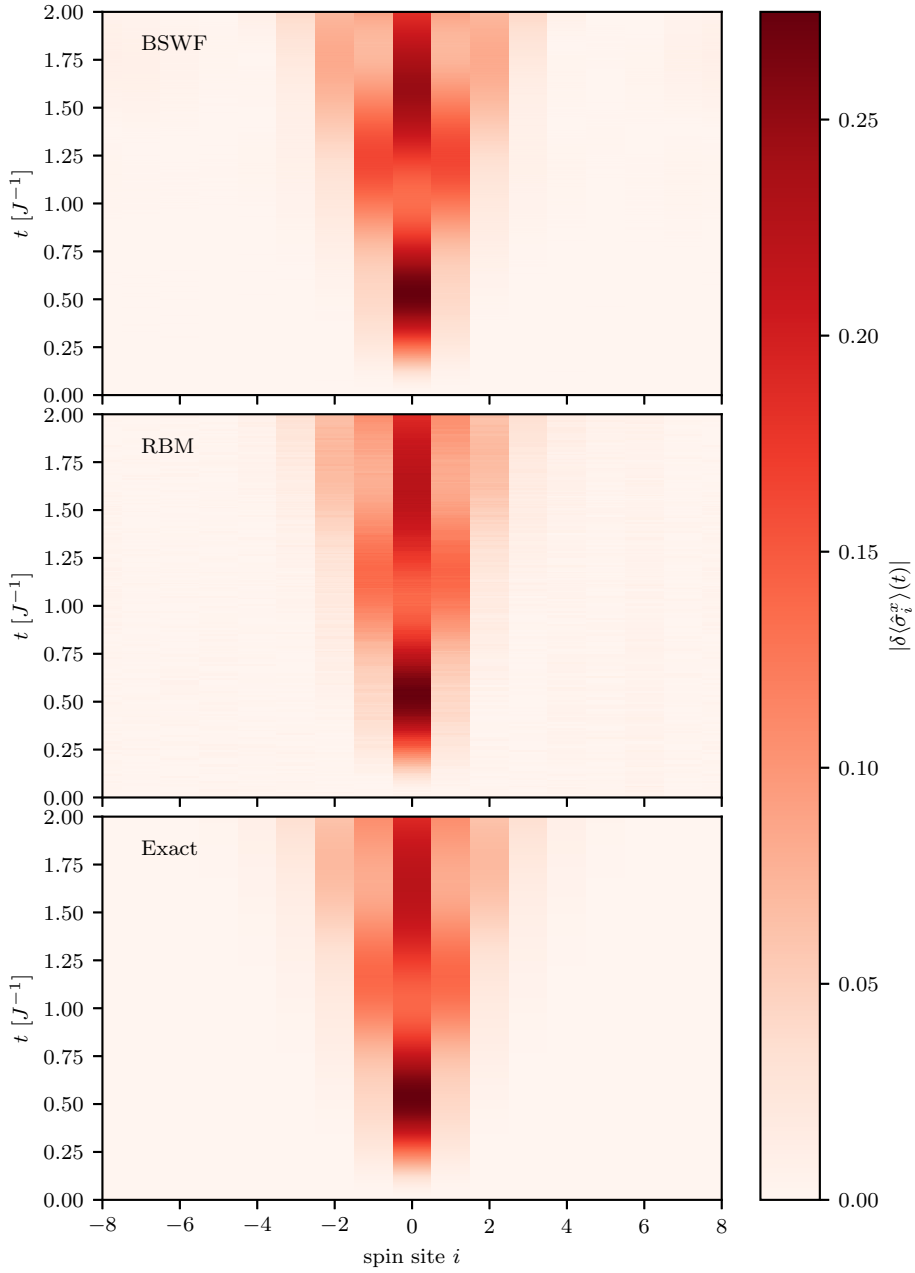

Figure 4.18: Absolute value of the fluctuations in the transverse magnetization of the spins of a 16-spin chain after a localized quench. The system is prepared in the TFI ground state with $g = 1$, and then an additional transverse field $\delta h = 1\,J$ on site $i = 0$ is switched on, and the system is evolved in real time. The results are shown for the BSWF Ansatz (upper panel), the RBM Ansatz (middle panel), and the exact evolution (lower panel).

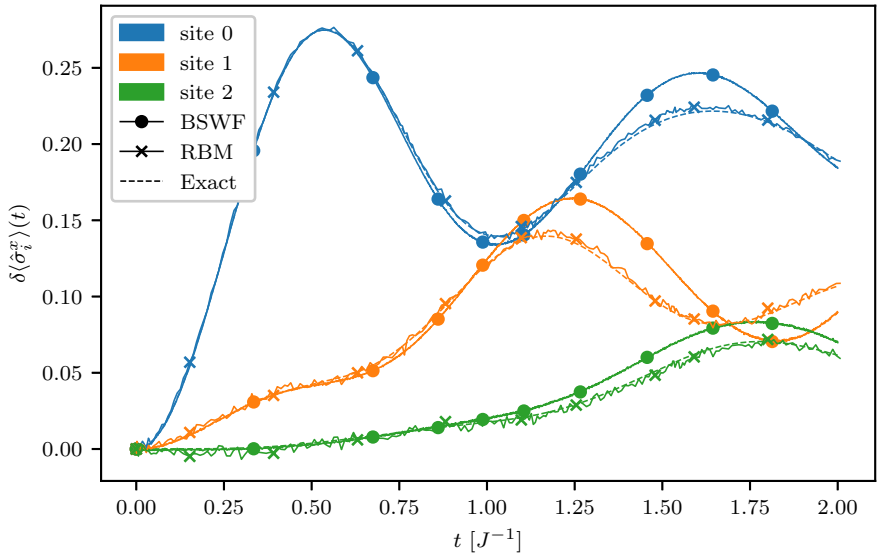

Figure 4.19: Fluctuations in the transverse magnetization of the spins of a 16-spin chain after a localized quench. The system is prepared in the TFI ground state with $g = 1$, and then an additional transverse field $\delta h = 1\,J$ on site $i = 0$ is switched on, and the system is evolved in real time. The figure shows the evolution of the spins on site $i = 0$, 1, and 2. The results are shown for the exact evolution, the BSWF Ansatz, and the RBM Ansatz.

(i.e. the standard BSWF): denoting by η_1, ω_1, ρ_1 the optimized parameters of a standard BSWF, we set the initial parameters of the $N_L = 2$ BSFW as follows. The parameters $\rho^{(2)}$, $\omega^{(2)}$, $\rho^{(1)}$, that involve only the two shadow layers, are initialized as $\rho^{(2)} = \rho_1$, $\omega^{(2)} = \omega_1$, $\rho^{(1)} = \eta_1$. In this way, the two shadow layers reproduce the same "path-integral" step as the optimized standard BSWF. Starting from that wave function, an additional step is performed, controlled by the parameters $\omega^{(1)}$ and η. We choose $\omega^{(1)} = 2$, which is large enough to make the real variables very correlated to the shadows of the first layer, but not so large that they become locked together; and we choose $\eta = 0$, so that there are no other correlations among the real variables. The "path-integral" step generated by this choice of $\omega^{(1)}$ and η perform only a small modification to the step performed by the two shadow layers alone, so that our starting point is close to the state represented by the optimized standard BSWF. This makes the ensuing optimization more stable than it would be with a random starting point.

The upper panel of Fig. 4.22 shows the convergence of the energy per spin, compared with the exact ground state energy and the energy of the optimized standard BSWF. We observe that this Ansatz improves over the standard BSWF; picking up a range of imaginary time in the region in which the state has converged, and averaging over the estimated energies, we find that the relative error (4.2) for the 2-layers BSWF is $\simeq 5 \cdot 10^{-7}$, whereas the relative error for the standard BSWF is $\simeq 1.5 \cdot 10^{-5}$. The lower panel of Fig. 4.22 shows the convergence of the parameters.

We also tried to similarly perform the optimization of a 3-layers BSWF for $g = 0.5$. We start from the state in which the three shadow layers perform the same two "path-integral" steps as the optimized 2-layers BSWF state, and the additional step has $\omega^{(1)} = 2$

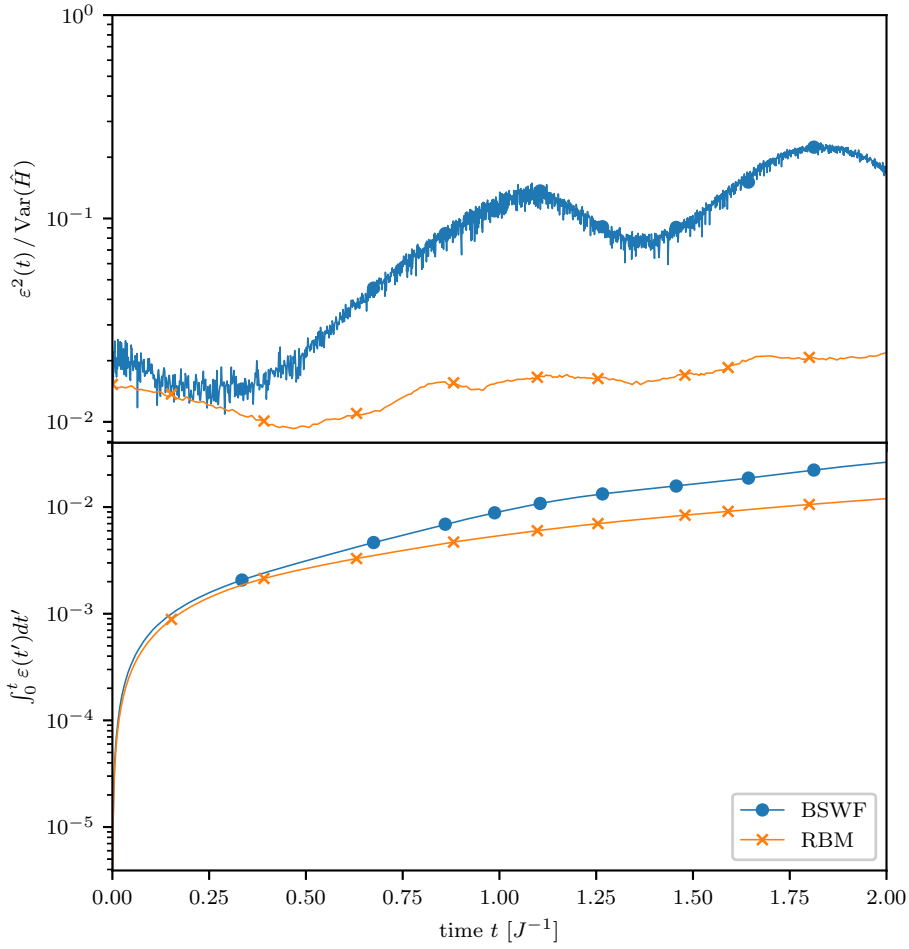

Figure 4.20: Squared error per unit time divided by the energy variance $\varepsilon^2(t)/\mathrm{Var}(\hat{H})$ (upper panel) and integral of $\varepsilon(t)$ over time (lower panel) for a 16-spin TFI chain under a localized quench. The system is prepared in the TFI ground state with $g = 1$, and then an additional transverse field $\delta h = 1\,J$ on site $i = 0$ is switched on, and the system is evolved in real time. The figure shows the evolution of the spins on site $i = 0$, 1, and 2. The results are shown for the exact evolution, the BSWF Ansatz, and the RBM Ansatz.

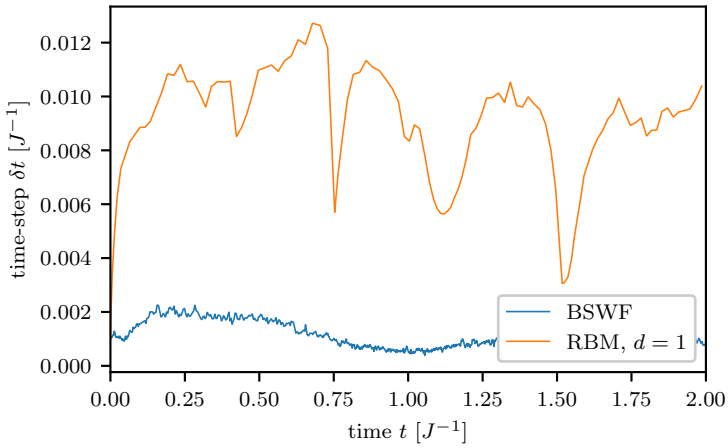

Figure 4.21: Value of the adaptive time step during the simulation of a 16-spin TFI chain under a localized quench. The system is prepared in the TFI ground state with $g = 1$, and then an additional transverse field $\delta h = 1\,J$ on site $i = 0$ is switched on, and the system is evolved in real time. The results are shown the BSWF Ansatz and the RBM Ansatz. The tolerance used to adaptively change the time step of the Heun integrator is $\lambda = 10^{-6}$.

and $\eta = 0$. However, the optimized variational energy resulted compatible with that of the 2-layers case. We conclude that, in order to properly optimize the 3-layers BSWF, we need a larger number of Monte Carlo samples. Indeed, the number of samples not only determines the precision with which we can compute the variational energy, but also the accuracy of the simulated imaginary-time propagation.

Real time

We perform a quantum quench from $g = 0.5$ to $g = 1.0$ with the 2-layers BSWF for the 32-spins TFI chain, and monitor the transverse magnetization (4.4). Fig. 4.23 shows the results, compared to the exact ones and the results obtained by the standard BSWF. We observe that around $t = 0.5$ both the t-VMC calculations start to depart from the exact evolution, but the 2-layers BSWF stays closer to the exact curve than the standard BSWF. This picture is confirmed by the error per unit $\varepsilon(t)$ time and its integral $\int_0^t \varepsilon(t')dt'$, plotted in Fig. 4.24: for the most part of the evolution, the 2-layers BSWF has smaller error than the BSWF.

On the other hand, we observe that the noise in the 2-layers BSWF rapidly increases after $t \simeq 0.75$. As the noise increases, also our computations of the S matrix and the F vector become more noisy, thus making the dynamics less reliable. To understand the reason of this noise explosion, we investigate the value of $\langle \cos\theta \rangle(t)$, where θ is the argument of the shadow non-classical probability distribution (2.21). The value of the average cosine is important since it is used in the reweighting technique to compute the expectation values of any operator, as in Eq. (2.34). Since $\langle \cos\theta \rangle$ appears at the denominator in the reweighting formula, when it becomes close to zero, the stochastic noise in the expectation values is greatly amplified. Fig. 4.25 compares the values of $\langle \cos\theta \rangle(t)$ of the standard BSWF and the 2-layers BSWF. We observe that, starting from $\langle \cos\theta \rangle(0) = 1$, for the 2-layers BSWF it decreases more rapidly and it gets closer to zero. This explains

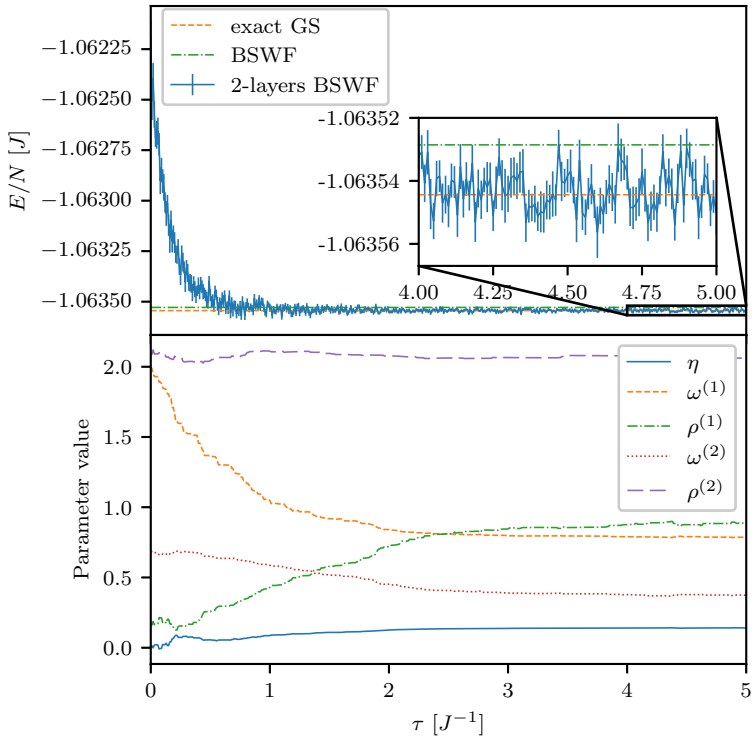

Figure 4.22: Imaginary-time evolution of a 2-layers BSWF for a 32-spin Ising chain with $g = 0.5$ in PBC. The upper panel shows the convergence, in imaginary time τ, of the energy per spin (solid line), compared with the exact ground state energy (dashed horizontal line) and the variational energy of the optimized standard BSWF (dash-dotted line). For ease of comparison, an inset shows the details of the energy evolution once it has reached convergence, in the interval from $\tau = 4.0$ to 5.0. The lower panel shows the evolution of the parameters. The parameters are initialized so that the initial state is a slightly perturbed version of the optimized standard BSWF state, as explained in the text. This evolution is integrated via the Euler method with step $\delta\tau = 0.01$, and the SNR regularization with threshold $\lambda_{\text{SNR}} = 1$ is used.

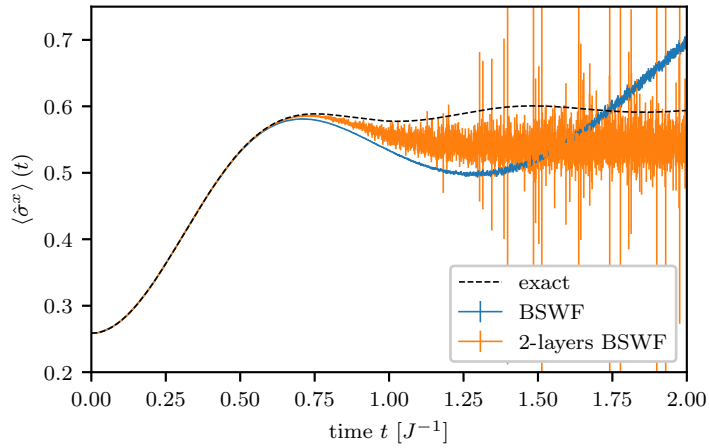

Figure 4.23: Transverse magnetization for a quantum quench of the 32-spin 1D TFI from $g = 0.5$ to $g = 1$. The dashed line shows the exact solution, obtained by numerical integration of the Bogoliubov-de Gennes equations. The solid lines display the variational approximations obtained by the BSWF and the 2-layers BSWF.

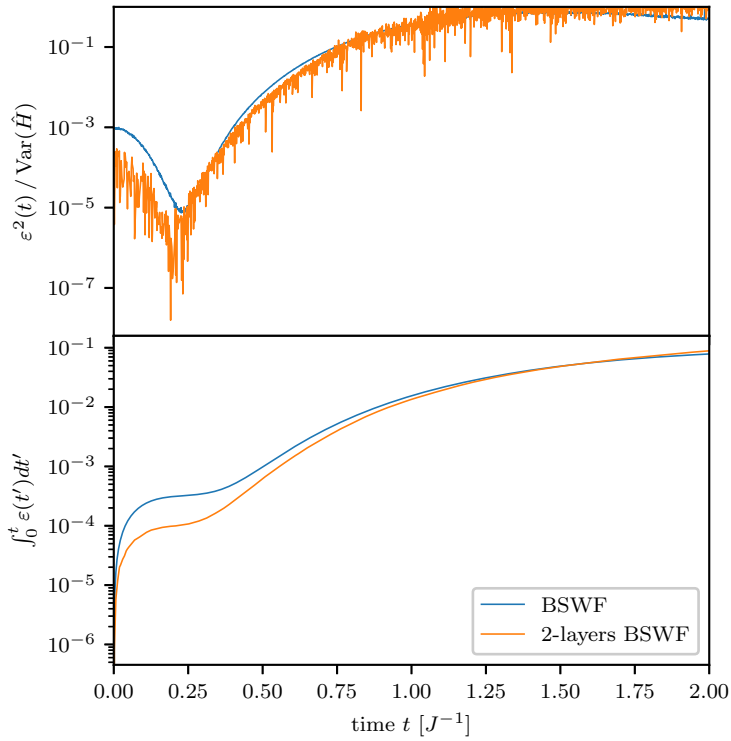

Figure 4.24: Squared error per unit time divided by the energy variance $\varepsilon^2(t)/\mathrm{Var}(\hat{H})$ (upper panel) and integral of $\varepsilon(t)$ over time (lower panel) for a quantum quench of the 32-spin 1D TFI from $g = 0.5$ to $g = 1$. The errors for the variational approximations obtained by the BSWF and the 2-layers BSWF are displayed.

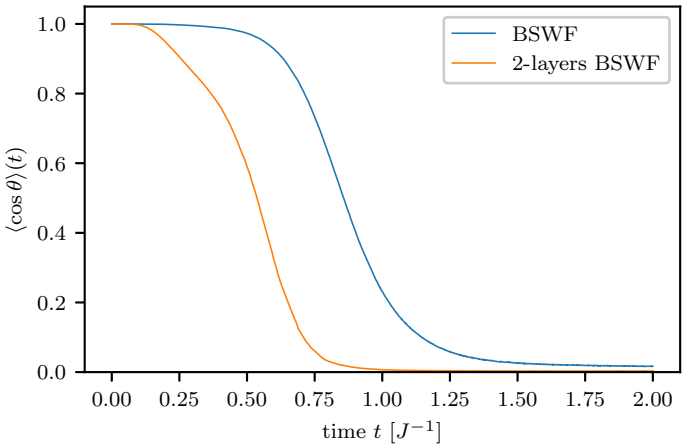

Figure 4.25: Reweighting normalization for a quantum quench of the 32-spin 1D TFI from $g = 0.5$ to $g = 1$. The results for the BSWF and the 2-layers BSWF are displayed.

the explosion of the noise, and suggests that the reweighting technique we use is not optimal after a certain time. Further investigations into different sampling techniques could reveal strategies that are more effective for long times.

4.3.2 The long-range BSWF (LRBSWF)

One of the major shortcomings of the BSWF is that it lacks long-range correlation terms. For this reason, we introduce the long-range BSWF (LRBSWF), defined as

$$\psi_{\text{LRBSWF}}(\boldsymbol{\sigma}; \eta, \{\omega^{(k)}\}, \{\rho^{(k)}\}) = \sum_{\boldsymbol{h}} \exp\left(\eta \sum_{\langle \mathbf{r},\mathbf{r}' \rangle} \sigma_{\mathbf{r}}\sigma_{\mathbf{r}'} + \omega \sum_{\mathbf{r}} \sigma_{\mathbf{r}} h_{\mathbf{r}} + \sum_{\langle \mathbf{r},\mathbf{r}' \rangle} \rho(d(\mathbf{r},\mathbf{r}')) h_{\mathbf{r}} h_{\mathbf{r}'} \right), \tag{4.9}$$

where the shadow-shadow coupling $\rho_{d(\mathbf{r},\mathbf{r}')}$ depends on the distance $d(\mathbf{r},\mathbf{r}')$ between the sites $\mathbf{r}$ and $\mathbf{r}'$. If the systems has periodic boundary conditions (PBC), then d has to be defined accordingly. In a 1D chain of N spins in PBC, where the site $\mathbf{r}$ is identified by an integer index i, the distance is defined as

$$d(i,j) = \begin{cases} |i-j| & \text{if } |i-j| \leq N/2\,, \\ N - |i-j| & \text{otherwise.} \end{cases} \tag{4.10}$$

We work in the translation-invariant case, in which this Ansatz has $\lfloor N/2 \rfloor + 2$ parameters.

Imaginary time

Fig. 4.27 shows an optimization for a 32-spins TFI chain with $g = 0.5$. The initial state is prepared in the same state as the optimized BSWF: denoting by η_1, ω_1, ρ_1 the optimized parameters of a standard BSWF, we set the initial parameters of the LRBSFW as $\eta = \eta_1$, $\omega = \omega_1$, $\rho(d = 1) = \rho_1$, and all the other shadow-shadow couplings $\rho(d(i,j))$ are set to

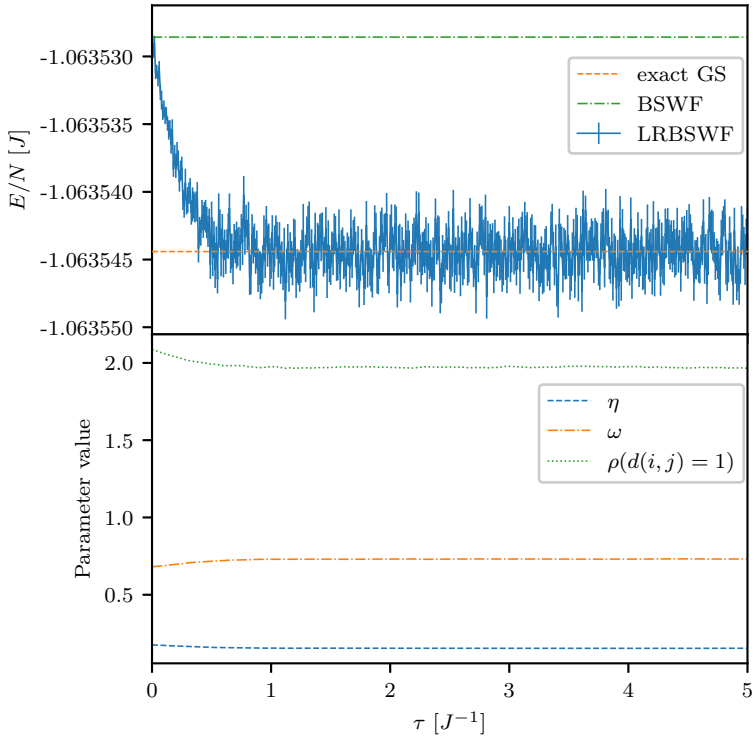

Figure 4.26

Figure 4.27: Imaginary-time evolution of a LRBSWF for a 32-spin Ising chain with $g = 0.5$ in PBC. The upper panel shows the convergence, in imaginary time τ, of the energy per spin (solid line), compared with the exact ground state energy (dashed horizontal line) and the variational energy of the optimized standard BSWF (dash-dotted line). The lower panel shows the evolution of the parameters, highlighting the real-real coupling η (dashed line), the real-shadow coupling ω (dash-dotted line) and the nearest-neighbor shadow-shadow coupling $\rho(d(i,j) = 1)$ (dotted line); the evolution of the other shadow-shadow couplings is not displayed. The parameters are initialized so that the initial state is the optimized standard BSWF state, as explained in the text. This evolution is integrated via the Euler method with step $\delta\tau = 0.01$, and the diagonal bias-10^{-4} regularization is used.

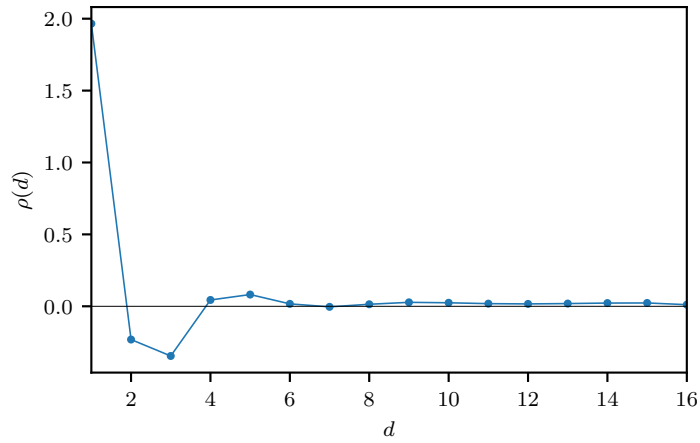

Figure 4.28: The optimized shadow-shadow couplings $\rho(d)$ of the LRBSWF as a function of the distance d between two shadow variables in lattice spacings. The system is a 32-spin Ising chain with $g = 0.5$ in PBC. The markers represent the data, the lines only serve as visual guides.

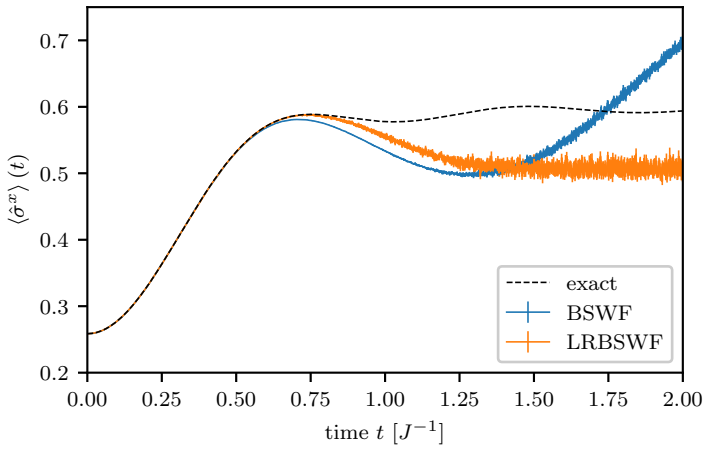

Figure 4.29: Transverse magnetization for a quantum quench of the 32-spin 1D TFI from $g = 0.5$ to $g = 1$. The dashed line shows the exact solution, obtained by numerical integration of the Bogoliubov-de Gennes equations. The solid lines display the variational approximations obtained by the BSWF and the LRBSWF.

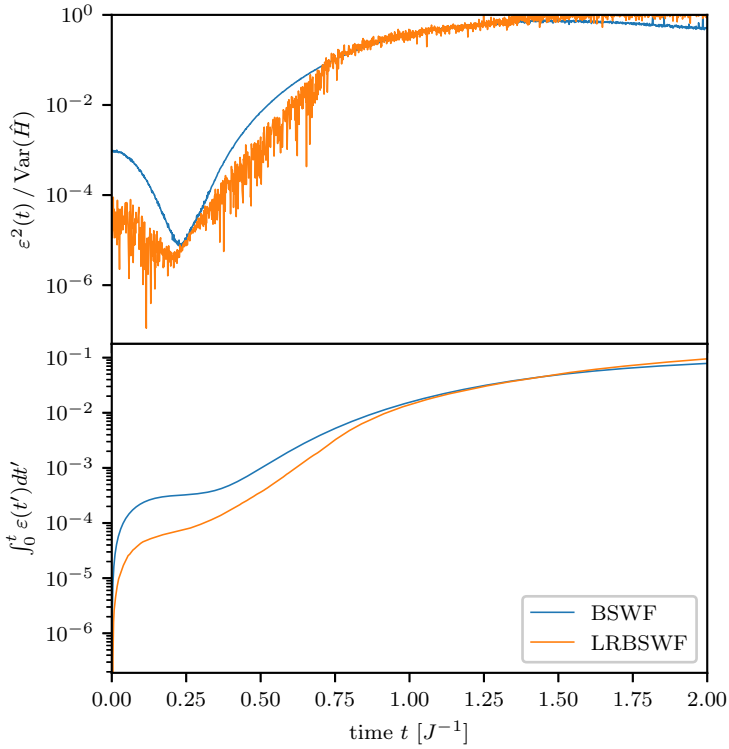

Figure 4.30: Squared error per unit time divided by the energy variance $\varepsilon^2(t)/\mathrm{Var}(\hat{H})$ (upper panel) and integral of $\varepsilon(t)$ over time (lower panel) for a quantum quench of the 32-spin 1D TFI from $g = 0.5$ to $g = 1$. The errors for the variational approximations obtained by the BSWF and the LRBSWF are displayed.

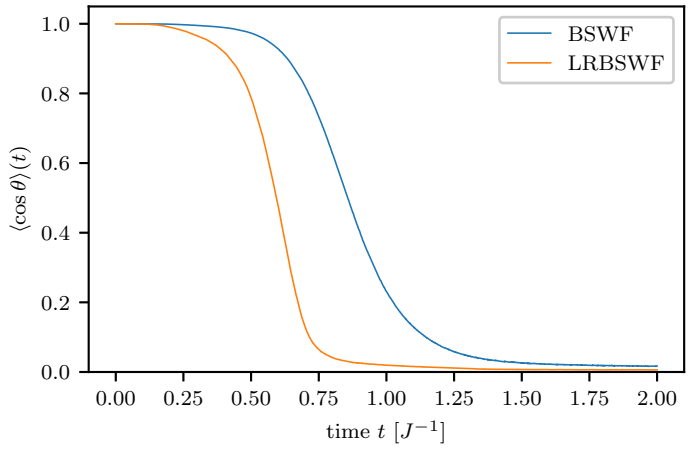

Figure 4.31: Reweighting normalization for a quantum quench of the 32-spin 1D TFI from $g = 0.5$ to $g = 1$. The results for the BSWF and the LRBSWF are displayed.

zero. This makes the ensuing optimization more stable than it would be with a random starting point.

The upper panel of Fig. 4.27 shows the convergence of the energy per spin, compared with the exact ground state energy and the energy of the optimized standard BSWF. Picking up a range of imaginary time in the region in which the state has converged, and averaging over the estimated energies, we find that the relative error (4.2) for the LRBSWF is $\simeq 5 \cdot 10^{-8}$. For comparison, the relative error for the standard BSWF is $\simeq 1.5 \cdot 10^{-5}$, and for the 2-layers BSWF is $\simeq 5 \cdot 10^{-7}$. The lower panel of Fig. 4.22 shows the convergence of the parameters η, ω, and $\rho(d(i,j) = 1)$, and we observe that they stay very close to the BSWF's optimized values. The final optimized value of all the shadow-shadow couplings ρ is displayed in Fig. 4.28.

Real time

We perform a quantum quench from $g = 0.5$ to $g = 1.0$ with the LRBSWF for the 32-spins TFI chain, and monitor the transverse magnetization (4.4). Fig. 4.29 shows the results, compared to the exact ones and the results obtained by the standard BSWF. We observe that around $t = 0.5$ both the t-VMC calculations start to depart from the exact evolution, but the LRBSWF stays closer to the exact curve than the standard BSWF, and it also performs better than the 2-layers BSWF (cfr. Fig. 4.23). Fig. 4.24 plots the error per unit $\varepsilon(t)$ time and its integral $\int_0^t \varepsilon(t')dt'$, confirming that the LRBSWF has lower error than the BSWF for the most part of the evolution.

We observe that, like for the 2-layers BSWF, the LRBSWF too suffers from a rapid noise increase after $t \simeq 0.75$, but in a less severe way. Fig. 4.31 compares the values of $\langle \cos\theta \rangle(t)$ of the standard BSWF and the LRBSWF, showing that the origin of this noise increase lays in the reweighting procedure. The value of $\langle \cos\theta \rangle(t)$ for the LRBSWF decreases more rapidly and it gets closer to zero, with respect to the BSWF. However, the problem is less severe than the 2-layers BSWF case (cfr. Fig. 4.25). As we noted while discussing the 2-layers BSWF, further investigation into different sampling techniques could help mitigate this problem.

4.4 Two-dimensional BSWF

We performed a calculation for the two-dimensional (2D) TFI model on a square grid in PBC. The BSWF is defined as in Eq. (3.18) and it still has three parameters. This model in the thermodynamic limit exhibits a quantum phase transition at $g \simeq 3.04438$ [54].

We optimize the 2D BSWF wave function for a 4×4 lattice at $g = 1$, in the ferromagnetic phase. We stick to this small size because this allows us to compute the exact ground state energy by exact diagonalization. The parameters are initialized to $\eta = 0.1$ and $\omega = \rho = 1$. Fig. 4.32 shows the convergence of the energy compared to the exact ground-state energy (upper panel), and the convergence of the variational parameters (lower panel). Picking up a range of imaginary time in the region in which the state has converged, and averaging over the estimated energies, we find that the relative error (4.2) for the BSWF is $\simeq 10^{-6}$. For comparison, the relative error obtained by a non-translation-invariant RBM at density 1, having 288 parameters, is $\simeq 3 \cdot 10^{-7}$.

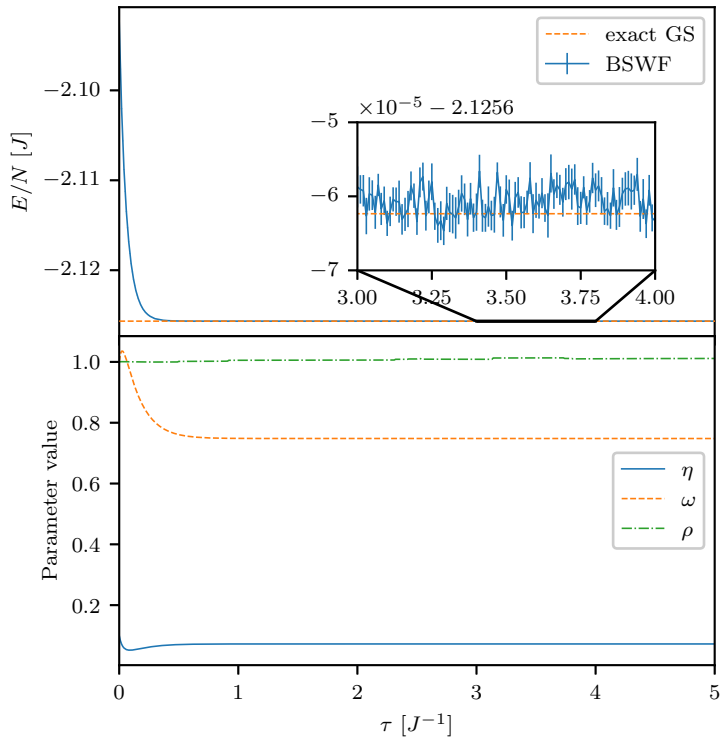

Figure 4.32: Imaginary-time evolution of a BSWF for a 2D 4×4 TFI model with $g = 1$ in PBC. The upper panel shows the convergence, in imaginary time τ, of the energy per spin (solid line), compared with the exact ground state energy (dashed horizontal line). For ease of comparison, an inset shows the details of the energy evolution once it has reached convergence, in the interval from $\tau = 3.0$ to 4.0. The lower panel shows the evolution of the parameters η (real-real coupling), ω (real-shadow coupling) and ρ (shadow-shadow coupling). The parameters are initialized to $\eta = 0.1$ and $\omega = \rho = 1$. This evolution is integrated via the Euler method with step $\delta\tau = 0.01$, and SNR regularization with threshold $\lambda_{\text{SNR}} = 1$ is used.

CHAPTER 5

Dynamical correlations of an attractive two-dimensional Fermi gas

In this Chapter, we approach the problem of computing the dynamical properties of a system from a different perspective. We focus on a system of particular physical interest, a two-dimensional (2D) dilute gas of attractive fermions, and compute its dynamical correlation functions using two methods: the generalized random phase approximation (GRPA) and the exact auxiliary-field quantum Monte Carlo method (AFQMC), the latter coupled with an analytic continuation technique.

As explained in Sec. 5.1, the main motivation for this research lies in the possibility, suggested in Ref. [13], of detecting the elusive Higgs mode of Fermi superfluids in the spin dynamical correlations. In Sec. 5.2 we introduce the Hamiltonian of the model, and present the regularization technique we use to deal with the singularities arising from contact interactions. In Sec. 5.3 we explain the quantum Monte Carlo technique we use to perform exact imaginary-time propagation, the auxiliary-field quantum Monte Carlo (AFQMC). In Sec. 5.4 we present the differential evolution for analytic continuation (DEAC) method, that we use to obtain the dynamical structure factors from the imaginary-time correlations. In Sec. 5.5 we briefly discuss the standard Bardeen-Cooper-Schrieffer (BCS) theory of superconductivity, in preparation for the introduction of the generalized random phase approximation (GRPA) in the subsequent Sec. 5.6. Finally, in Sec. 5.7 we present our GRPA and QMC results for the dynamical structure factors of the system, and discuss them in comparison with each other.

5.1 Motivation

The task of calculating the dynamical correlation functions of a strongly correlated quantum system is a famously difficult challenge, which is of great interest from both the theoretical and the experimental standpoints, since it gives access to critical information about the excitations of the system and its response to external probes. One important application is the detection and study of the collective excitations that arise from spontaneous broken symmetry in many-body systems. A particularly interesting case is represented by Fermi superfluids, in which the spontaneous breaking of the $U(1)$ symmetry gives rise to two collective modes: the first one is the Nambu-Goldstone mode, a massless excitation that corresponds to oscillations in the phase of the superfluid order parameter; the second one, which we want to focus on, is the celebrated Higgs mode, corresponding to the oscillations in the amplitude of the order parameter. The latter has low-momentum dispersion of the form $\omega(q) = 2\Delta + \alpha q^2$, where q is the momentum, Δ is the superfluid gap and α is a coefficient [55, 56].

The Higgs or amplitude mode is particularly elusive to observation, and its measure has been at the center of a long-standing experimental effort, both in condensed matter physics [57, 58] and in atomic physics. In ultracold Bose gases, experimental results were obtained by using lattice shaking or cavity-enhanced Bragg spectroscopy [59, 60, 61]. In the realm of ultracold fermions, on the other hand, devising an effective experimental probe for the Higgs mode has proven to be particularly hard, because of the very slow-decaying spectral tail that this mode has in Fermi systems [62, 63], that translates into strong damping due to the proliferation of quasiparticle excitations into which the amplitude mode can decay. Only very recently experimental observations were reported, using direct coherent excitation of the mode with radiofrequency pulses, Bragg spectroscopy following an interaction quench, or interaction modulation via periodic tuning of the magnetic field [64, 65, 66, 67].

Theoretical and computational studies of the Higgs mode in Fermi systems have employed various methodologies, including, e.g., the time-dependent Bogoliubov-de Gennes equations [68, 69, 70, 71], time-dependent density functional theory [72], the functional-integral method [62, 73, 74], and exact diagonalization (for few-body systems) [75]. A recent study by Zhao et al. [13], based on generalized random phase approximation (GRPA) [14, 15, 16, 17], suggested that the dynamical correlations of the spin degrees of freedom could be a good probe of the amplitude mode. This claim, if proven correct, presents the exciting possibility of experimentally probing the collective mode by measuring the spin dynamical structure factor through spin-sensitive Bragg spectroscopy [76, 77], thus extending the insightful investigations that can be performed in the density fluctuations sector [66, 78, 79, 80]. Zhao et al. [13] also argued that the two-dimensional (2D) geometry, as contrasted to the three-dimensional one mostly studied so far, presents a particularly visible Higgs signal.

Inspired by this suggestion, in this chapter we study a 2D dilute system of attractive fermions. We systematically run generalized random phase approximation (GRPA) calculations [14, 15, 16, 17, 13] for the dynamical structure factor in the channels of the particle density $S(\mathbf{q},\omega)$, the spin density $S^s(\mathbf{q},\omega)$, the order parameter's phase $S^p(\mathbf{q},\omega)$, and the order parameter's amplitude $S^a(\mathbf{q},\omega)$. Additionally, we check the density and spin spectra against those yielded by an exact QMC approach, to explore the effect of many-body correlations on these channels. Specifically, we leverage auxiliary-field quantum Monte Carlo (AFQMC) [18, 19, 20], an unbiased technique that yields exact results within statistical uncertainties, coupled with an analytic continuation algorithm, differential evolution for analytic continuation (DEAC) [21, 22].

5.2 The model studied and the regularization

We study a two-dimensional (2D) unpolarized system of N attractive spin-$1/2$ fermions of equal mass m. The two spin species, denoted by $\uparrow$ and $\downarrow$, represent, in the context of cold gases, the two lowest hyperfine states of ^{6}Li or ^{40}K atoms. We focus on the dilute regime, where the detail of the inter-particle forces can be neglected and they can be modeled by an attractive zero-range interaction $v_{\uparrow\downarrow}(\mathbf{r_1},\mathbf{r_2}) = -g\delta(\mathbf{r_1}-\mathbf{r_2})$. We write the Hamiltonian $\hat{H}$ as

$$\hat{H} = \hat{K} + \hat{V}\,, \tag{5.1}$$

with kinetic energy

$$\hat{K} = \int d\mathbf{r} \sum_\sigma \hat{\psi}^\dagger_\sigma(\mathbf{r}) \left(-\frac{\hbar^2\nabla^2}{2m}\right) \hat{\psi}_\sigma(\mathbf{r}) \tag{5.2}$$

and interaction energy

$$\hat{V} = -g \int d\mathbf{r}\, \hat{\psi}^\dagger_\uparrow(\mathbf{r})\, \hat{\psi}^\dagger_\downarrow(\mathbf{r})\, \hat{\psi}_\downarrow(\mathbf{r})\, \hat{\psi}_\uparrow(\mathbf{r})\,. \tag{5.3}$$

We want to study the zero-temperature density and spin density dynamical structure factors for various regimes of interaction strength, measured by the dimensionless parameter $\log(k_F a)$, where k_F is the Fermi momentum and a the s-wave scattering length. Since there is more than one way to define the scattering length in 2D, we adopt the convention that the dimer binding energy ϵ_b of the contact model is related to a by $|\epsilon_b| = \hbar^2/ma^2$ [81].

The zero-temperature density dynamical structure factor is defined as

$$S(\mathbf{q},\omega) = \int_0^{+\infty} \frac{dt}{2\pi N}\, e^{i\omega t}\, \left\langle e^{i\hat{H}\frac{t}{\hbar}}\, \hat{n}_{\mathbf{q}}\, e^{-i\hat{H}\frac{t}{\hbar}}\, \hat{n}_{-\mathbf{q}} \right\rangle\,, \tag{5.4}$$

where $\hat{n}_{\mathbf{q}}$ is the Fourier transform of the particle density operator $\hat{n}(\mathbf{r}) = \hat{n}_\uparrow(\mathbf{r}) + \hat{n}_\downarrow(\mathbf{r}) = \hat{\psi}^\dagger_\uparrow(\mathbf{r})\hat{\psi}_\uparrow(\mathbf{r}) + \hat{\psi}^\dagger_\downarrow(\mathbf{r})\hat{\psi}_\downarrow(\mathbf{r})$. Similarly, we define the spin dynamical structure factor $S^s(\mathbf{q},\omega)$ by replacing the particle density $\hat{n}(\mathbf{r})$ with the spin density $\hat{n}^s(\mathbf{r}) = \frac{1}{2}\left(\hat{n}_\uparrow(\mathbf{r}) - \hat{n}_\downarrow(\mathbf{r})\right)$.

The contact interaction leads to singularities that we address through lattice regularization, as follows. We introduce a high-momentum cutoff $|k_x|, |k_y| \leq \pi/b$ in the kinetic energy term:

$$\hat{K} \simeq \sum_\sigma \int_{[-\pi/b,\pi/b)^2} d\mathbf{k}\, \varepsilon(\mathbf{k})\, \hat{\psi}^\dagger_\sigma(\mathbf{k})\hat{\psi}_\sigma(\mathbf{k})\,, \tag{5.5}$$

where $\hat{\psi}_\sigma(\mathbf{k}) = \frac{1}{\sqrt{M}} \sum_{\mathbf{r}} e^{-i\mathbf{k}\cdot\mathbf{r}} \hat{\psi}_\sigma(\mathbf{r})$ is the Fourier transform of the field operator $\hat{\psi}_\sigma(\mathbf{r})$, and the dispersion relation is $\varepsilon(\mathbf{k}) = \frac{\hbar^2 k^2}{2m}$. This is equivalent to integrating over the first Brillouin zone of a square lattice of spacing b. So, by imposing this high-momentum cutoff, we have effectively discretized the position space to the lattice $(b\mathbb{Z})^2$. We perform a further regularization step by introducing a supercell of side L, multiple of b. Thus we consider only positions belonging to the finite square lattice $\mathcal{L} = (b\mathbb{Z})^2 \cap \left[-\frac{L}{2}, \frac{L}{2}\right)$, that has $M = L/b \times L/b$ sites, and impose periodic boundary conditions (PBC) on it. This restricts the integration in (5.5) to a summation over the set of momenta $\mathbf{k} = \frac{2\pi}{L}\mathbf{n}$, with $\mathbf{n} \in \mathbb{Z}^2 \cap \left[-\frac{L}{2b}, \frac{L}{2b}\right)^2$, so that the range of $\mathbf{k}$ respects the cutoff. So we rewrite the kinetic energy as

$$\hat{K} \simeq \sum_{\mathbf{k},\sigma} \varepsilon(\mathbf{k})\, \hat{\psi}^\dagger_\sigma(\mathbf{k})\hat{\psi}_\sigma(\mathbf{k})\,, \tag{5.6}$$

where we can express the dispersion in "Hubbard-like" notation as $\varepsilon(\mathbf{k}) = t|b\mathbf{k}|^2$ with hopping amplitude given by $t = \frac{\hbar^2}{2mb^2}$.

Now that the position space has been discretized, the interaction term $\hat{V}$ is regularized by introducing a contact interaction expressed by means of the discrete Kronecker delta as $v_{\uparrow\downarrow}(\mathbf{r}_1, \mathbf{r}_2) = -\frac{U}{b^2}\delta_{\mathbf{r}_1,\mathbf{r}_2}$, leading to

$$\hat{V} = -U \sum_{\mathbf{r}} \hat{\psi}^\dagger_\uparrow(\mathbf{r})\hat{\psi}_\uparrow(\mathbf{r})\hat{\psi}^\dagger_\downarrow(\mathbf{r})\hat{\psi}_\downarrow(\mathbf{r}) = -U \sum_{\mathbf{r}} \hat{n}_\uparrow(\mathbf{r})\hat{n}_\downarrow(\mathbf{r})\,, \tag{5.7}$$

where $\hat{n}_\sigma(\mathbf{r}) = \hat{\psi}^\dagger_\sigma(\mathbf{r})\hat{\psi}_\sigma(\mathbf{r})$ is the operator counting the number of particles of spin σ on site $\mathbf{r}$. The strength $U > 0$ is tuned in such a way that the lattice two-body problem has

the same zero-energy s-wave scattering length a as the problem in the continuum. The value of U corresponding to a given scattering length a is obtained from the relation [82]

$$\frac{U}{t} = \frac{4\pi}{\log(k_F a) + \log(\mathcal{C}\sqrt{n})}\,, \tag{5.8}$$

where $n = N/M$ is the particle density in the lattice, $k_F = \frac{2\pi n}{b}$ is the Fermi momentum, and $\mathcal{C} = 0.80261$ is a constant that depends on the form of the free-particle dispersion $\varepsilon(\mathbf{k})$ [83].

We have thus mapped the original system into a lattice model described by the "Hubbard-like" Hamiltonian

$$\hat{H}_{\mathcal{L}} = \sum_{\mathbf{k},\sigma} \varepsilon(\mathbf{k})\, \hat{\psi}^{\dagger}_{\sigma}(\mathbf{k})\hat{\psi}_{\sigma}(\mathbf{k}) - U \sum_{\mathbf{r}} \hat{n}_{\uparrow}(\mathbf{r})\hat{n}_{\downarrow}(\mathbf{r})\,. \tag{5.9}$$

5.3 Auxiliary-field quantum Monte Carlo (AFQMC)

The advantage of mapping our system into a lattice model as described in Sec. 5.2 is the possibility, for an unpolarized system of attractive fermions described by the Hamiltonian (5.9), of computing exactly the intermediate scattering function in imaginary time,

$$F(\mathbf{q},\tau) = \frac{1}{N}\left\langle e^{\tau \hat{H}_{\mathcal{L}}}\, \hat{n}_{\mathbf{q}}\, e^{-\tau \hat{H}_{\mathcal{L}}}\, \hat{n}_{-\mathbf{q}} \right\rangle\,, \quad \tau \geq 0\,, \tag{5.10}$$

both in the density and in the spin channels (substituting $\hat{n}$ with $\hat{n}^s$). $F(\mathbf{q},\tau)$ can then be used to retrieve the dynamical structure factor, as we discuss in Sec. 5.4. We compute $F(\mathbf{q},\tau)$ by leveraging unbiased auxiliary-field quantum Monte Carlo techniques (AFQMC) to sample at the same time both the ground state wave function of the system $|\Psi_0\rangle$ and the propagator in imaginary time $\exp(-\tau \hat{H}_{\mathcal{L}})$.

The AFQMC methodologies are extensively described in Refs. [18, 19, 20]. Here we only present a concise summary. The starting point of the AFQMC technique is the projection formula

$$|\Psi_0\rangle \propto \lim_{\tau\to+\infty} e^{-\tau(\hat{H}_{\mathcal{L}} - E_0)}|\phi_0\rangle\,, \tag{5.11}$$

where E_0 is a preliminary estimate of the fundamental energy and $|\phi_0\rangle$ is an approximation of the ground state, chosen to be a Slater determinant that is not orthogonal to the ground state $|\Psi_0\rangle$. The simplest choice for $|\phi_0\rangle$ is the ground state wave function of the non-interacting ($U = 0$) model.

The AFQMC method exploits a Trotter decomposition

$$e^{-\tau(\hat{H}_{\mathcal{L}} - E_0)} = \left(e^{-\delta\tau(\hat{H}_{\mathcal{L}} - E_0)}\right)^S, \quad \delta\tau = \frac{\tau}{S} \tag{5.12}$$

that writes the projector as a composition of propagators for small imaginary time $\delta\tau$, to which we can then a apply a Hubbard-Stratonovich transformation:

$$e^{-\delta\tau(\hat{H}_{\mathcal{L}} - E_0)} \simeq \int d\mathbf{x}\, p(\mathbf{x})\, \hat{B}(\mathbf{x}) \quad \text{for } \delta\tau \to 0\,, \tag{5.13}$$

where $\mathbf{x}$ is an auxiliary Ising field defined on the lattice, $x(\mathbf{r}) = \pm 1$; $p(\mathbf{x})$ is just the uniform probability distribution over the possible configurations of the auxiliary field:

$p(\mathbf{x}) = \frac{1}{2^M}$; and the integration runs over all the possible configurations of $\mathbf{x}$. The operator $\hat{B}(\mathbf{x})$ is defined as

$$\hat{B}(\mathbf{x}) = e^{\delta\tau E_0}\, e^{-\delta\tau \hat{K}/2} \prod_{\mathbf{r}} \hat{b}_{\mathbf{r}}(x(\mathbf{r}))\, e^{-\delta\tau \hat{K}/2} \tag{5.14}$$

where $\hat{K}$ is the kinetic energy (5.6) and

$$\hat{b}_{\mathbf{r}}(x) = e^{\delta\tau U(\hat{n}(\mathbf{r})-1)/2}\, e^{\gamma x(\hat{n}(\mathbf{r})-1)} \tag{5.15}$$

with $\cosh(\gamma) = \exp\left(\frac{\delta\tau|U|}{2}\right)$. From Eqs. (5.14) and (5.15) one can see that the Hubbard-Stratonovich transformation has mapped the interacting system onto an ensemble of systems of non-interacting fermions that move in random external potentials defined by the auxiliary field $\mathbf{x}$. This mapping can be exploited to compute the expectation value of an operator $\hat{O}$

$$\langle\, \hat{O}\, \rangle = \frac{\langle \Psi_0 | \hat{O} | \Psi_0 \rangle}{\langle \Psi_0 | \Psi_0 \rangle}\,. \tag{5.16}$$

Using the projection formula (5.11), we write

$$\langle\, \hat{O}\, \rangle = \frac{\langle \phi_0 | e^{-\tau \hat{H}_{\mathcal{L}}} \hat{O} e^{-\tau \hat{H}_{\mathcal{L}}} | \phi_0 \rangle}{\langle \phi_0 | e^{-\tau \hat{H}_{\mathcal{L}}} e^{-\tau \hat{H}_{\mathcal{L}}} | \phi_0 \rangle}\,. \tag{5.17}$$

Then we plug in the Trotterization of the imaginary-time propagator (5.12) and perform a Hubbard-Stratonovich transformation (5.13) on every $\delta\tau$ propagator, yielding

$$\langle\, \hat{O}\, \rangle \simeq \frac{\int d\mathbf{x}_1 \dots d\mathbf{x}_{2S}\, \pi(\mathbf{x}_0, \dots, \mathbf{x}_{2S}) \mathcal{O}(\mathbf{x}_0, \dots, \mathbf{x}_{2S})}{\int d\mathbf{x}_1 \dots d\mathbf{x}_{2S}\, \pi(\mathbf{x}_0, \dots, \mathbf{x}_{2S})}\,, \tag{5.18}$$

where

$$\begin{aligned} \pi(\mathbf{x}_1, \dots, \mathbf{x}_{2S}) &= \prod_{i=1}^{2S} p(\mathbf{x}_i) \langle \phi_L \,|\phi_R\rangle\,, \\ \mathcal{O}(\mathbf{x}_1, \dots, \mathbf{x}_{2S}) &= \frac{\langle \phi_L \,|\hat{O}\,|\, \phi_R\rangle}{\langle \phi_L \,|\phi_R\rangle}\,, \end{aligned} \tag{5.19}$$

and $|\phi_L\rangle$ and $|\phi_R\rangle$ are defined as

$$\begin{aligned} \langle \phi_L \,| &= \langle \phi_0 \,|\, \hat{B}(\mathbf{x}_{2S}) \dots \hat{B}(\mathbf{x}_{S+1})\,, \\ |\phi_R\rangle &= \hat{B}(\mathbf{x}_S) \dots \hat{B}(\mathbf{x}_1) |\phi_0\rangle\,. \end{aligned} \tag{5.20}$$

The fact that the operator $\hat{B}(\mathbf{x})$ defined in (5.14) is the exponential of a one-body operator has crucial consequence: $\hat{B}(\mathbf{x})$ maps Slater determinants into Slater determinants. So each successive application of $\hat{B}(\mathbf{x}_i)$ in Eq. (5.20), where the fields $\mathbf{x}_i$ are sampled randomly, generate a new random Slater determinants. This maps the imaginary-time evolution (5.11) into a random walk in the manifold of N particles' Slater determinants.

During an AFQMC calculation, we sample random configurations of the auxiliary field $\mathbf{x}$ and use them to perform random walks in the manifold of Slater determinants. The starting point $|\phi_0\rangle$ can be, in the simplest case, a Slater determinant of plane waves

describing a Fermi sphere. Then, the successive Slater determinants are obtained by successive applications of $\hat{B}(\mathbf{x})$ with randomly sampled fields $\mathbf{x}$. In the end, an average over random walks allows the calculation of expectation values over the ground state $|\Psi_0\rangle$, see Eq. (5.18). Moreover, since each application of $\hat{B}(\mathbf{x})$ performs a small step $\delta\tau$ of imaginary-time propagation (after averaging over $\mathbf{x}$), we can compute quantities that evolve along the propagation, like the intermediate scattering function (5.10). And this computation is unbiased and exact, in the sense that it is always possible to choose the time step $\delta\tau$ and the total projection time $S\,\delta\tau$ in such a way as to make the systematic error smaller than the statistical uncertainties.

Remarkably, for a system of attractive unpolarized fermions, this technique is not affected by the infamous sign problem, and a calculation can be carried out in a computational time that scales polynomially in the size of the system, precisely as $O(N^2M)$. Moreover, several techniques can be used to improve the efficiency of the procedure, e.g. importance sampling.

5.4 Analytic continuation and the DEAC algorithm

The dynamical structure factor $S(\mathbf{q},\omega)$ defined in Eq. (5.4) is related to the intermediate scattering function $F(\mathbf{q},\tau)$ defined in Eq. (5.10) by a Laplace transform:

$$F(\mathbf{q},\tau) = \int_0^{+\infty} d\omega\, e^{-\tau\omega} S(\mathbf{q},\omega)\,. \tag{5.21}$$

So, after calculating $F(\mathbf{q},\tau)$ through AFQMC, we need to invert the Laplace transform to retrieve $S(\mathbf{q},\omega)$. This inversion is a challenging problem. First of all, the estimation of $F(\mathbf{q},\tau)$ that we obtain from AFQMC is noisy, although unbiased. Moreover, the exponentially-decaying Laplace kernel $e^{-\tau\omega}$ quickly dampens to insignificance the high-ω and high-τ contributions. It follows that this inverse problem is ill-posed, in the sense that there is a large ensemble of spectral functions $S(\mathbf{q},\omega)$ whose Laplace transforms are all compatible with the measured $F(\mathbf{q},\tau)$, and a small variation in the data will greatly affect the ensemble of compatible spectra.

The problem that we face is called *analytic continuation*, since, if we were able to extend the domain of $F(\mathbf{q},\tau)$ to complex values of the variable τ (in particular to the imaginary axis, thus performing a so-called "Wick rotation"), then we could obtain $S(\mathbf{q},\omega)$ as the straightforward Fourier transform of $F(\mathbf{q},i\tau)$, as can be readily seen from the definitions of these functions.

Many methods have been proposed and used to tackle this difficult problem. We use the differential evolution for analytic continuation (DEAC) algorithm [21], and more specifically the code that was released by A. Del Maestro's group [22]. The DEAC is an evolutionary algorithm inspired by the genetic inversion via falsification of theories (GIFT) [84, 85], and it has the peculiarity of adaptively optimizing its hyperparameters. The DEAC belongs to the class of heuristic evolutionary algorithms called *differential evolution*, in which a population of candidate solutions is mantained and updated according to a simple rule. Then the new candidates are rated against a fitness criterion and added to the population or discarded accordingly, so that the average fitness of the population improves at every iteration.

More specifically, a candidate solution is represented by a spectrum $\tilde{S}_m(\mathbf{q},\omega)$ discretized over a fixed set of frequencies. Its fitness is evaluated by computing its Laplace transform $\tilde{F}_m(\mathbf{q},\tau_i)$ for the set of times $\{\tau_i\}_{i=0}^{N_\tau}$ on which the exact $F(\mathbf{q},\tau)$ has been mea-

sured. Then the following chi-squared distance is computed:

$$\chi^2 = \sum_{i=0}^{N_\tau} \frac{1}{N_\tau} \frac{(\tilde{F}_m(\tau_i) - F(\tau_i))^2}{\sigma^2(\tau_i)}\,, \tag{5.22}$$

where for simplicity we are omitting the dependence on $\mathbf{q}$, and $\sigma^2(\tau_i)$ is the square of the statistical uncertainty on the QMC-computed $F(\tau_i)$. The smaller χ^2 is, the higher the fitness of $S_m(\mathbf{q},\omega)$ is. If we have some prior knowledge about the solution, we can make the algorithm more efficient by including this knowledge in the definition of the fitness. In our case, we use the sum rule for the first moment of the spectrum, defined by

$$m_1(\mathbf{q}) \equiv \int_0^{+\infty} \omega\, S(\mathbf{q},\omega)\, d\omega\,. \tag{5.23}$$

The sum rule reads [86, 87]:

$$m_1(\mathbf{q}) = \frac{\hbar^2 q^2}{2m}\,. \tag{5.24}$$

Our prior knowledge of this exact relation can be included in the fitness by adding to the chi-squared distance (5.22) a term proportional to $(\tilde{m}_1 - m_1)^2$, where $\tilde{m}_1$ is the first moment of the candidate solution, and m_1 is the exact value. For the spin dynamical structure factor, a similar sum rule holds: denoting by $m_1^s(\mathbf{q})$ the first moment of the spin spectrum, we simply have $m_1^s(\mathbf{q}) = m_1(\mathbf{q})/4$.

At each iteration of the algorithm, also called a *generation*, a new population of candidate solutions is generated, according to the following rule. For each spectrum $\tilde{S}_m$ in the population, we randomly pick up three other spectra $\tilde{S}_j$, $\tilde{S}_k$, $\tilde{S}_l$ such that the four spectra are all different. Then a potential new member $\tilde{S}_n$ is constructed by iteration over the discretized frequencies ω_i. For each ω_i, a random number $\mathcal{U}(0,1)$ uniformly distributed between 0 and 1 is sampled and compared to a threshold P^c, the *crossover probability*. Then the value of $\tilde{S}_n(\omega_i)$ is chosen according to

$$\tilde{S}_n(\omega_i) = \begin{cases} \tilde{S}_j(\omega_i) + \gamma[\tilde{S}_k(\omega_i) - \tilde{S}_l(\omega_i)] & \text{for } \mathcal{U}(0,1) \leq P^c \\ \tilde{S}_m(\omega_i) & \text{otherwise,} \end{cases} \tag{5.25}$$

where γ is the *differential weight*. Then the fitness of the candidate $\tilde{S}_n$ is checked against the fitness of the old member $\tilde{S}_m$. If it is greater, then $\tilde{S}_n$ replaces $\tilde{S}_m$, otherwise $\tilde{S}_n$ is discarded.

The hyperparameters P^c and γ are self-adaptively updated by embedding them within the genome of the candidate solutions, such that every spectrum $\tilde{S}_m$ has its own P_m^c and γ_m. The crossover probability is updated with probability 0.1, according to

$$P_n^c = \begin{cases} \mathcal{U}(0,1) & \text{for } \mathcal{U}(0,1) \leq 0.1 \\ P_m^c & \text{otherwise,} \end{cases} \tag{5.26}$$

and the differential weight is also updated with probability 0.1:

$$\gamma_n = \begin{cases} \mathcal{U}(0,2) & \text{for } \mathcal{U}(0,1) \leq 0.1 \\ \gamma_m & \text{otherwise.} \end{cases} \tag{5.27}$$

This avoids the need of carefully choosing the hyperparameters, since they are optimized together with the population. This way, beneficial changes to P^c and γ are preserved, i.e., changes that increase the probability of producing spectra of high fitness.

Several iterations are performed, until an individual of the population reaches a fitness equal or better than a certain tolerance level. Then the simulation is terminated and that individual is returned as the optimal spectrum. Several runs are performed and the results are averaged to produce the final spectrum.

5.5 BCS theory

In this Section we concisely present the standard Bardeen-Cooper-Schrieffer theory (BCS) for the lattice model we introduced, since the GRPA approach, which we want to compare with our AFQMC results, is built from the starting point of BCS theory.

We begin by using the Fourier transform $\hat{\psi}_\sigma(\mathbf{r}) = \frac{1}{\sqrt{M}} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}} \hat{\psi}_\sigma(\mathbf{k})$ to write the interaction term (5.7) as

$$\hat{V} = -\frac{U}{M^2} \sum_{\mathbf{r}} \sum_{\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4} e^{i(-\mathbf{k}_1+\mathbf{k}_2-\mathbf{k}_3+\mathbf{k}_4)\cdot\mathbf{r}}\, \hat{\psi}^\dagger_\uparrow(\mathbf{k}_1)\hat{\psi}_\uparrow(\mathbf{k}_2)\hat{\psi}^\dagger_\downarrow(\mathbf{k}_3)\hat{\psi}_\downarrow(\mathbf{k}_4)\,. \tag{5.28}$$

From the relation $\frac{1}{M}\sum_{\mathbf{r}} e^{i\mathbf{k}\cdot\mathbf{r}} = \delta_{\mathbf{k},0}$ we find the momentum conservation condition $\mathbf{k}_1 + \mathbf{k}_3 = \mathbf{k}_2 + \mathbf{k}_4$. It is then convenient to rewrite $\hat{V}$ with the notation $\mathbf{k}_1 \equiv \mathbf{k}$, $\mathbf{k}_3 \equiv -\mathbf{k} + \mathbf{q}$, $\mathbf{k}_2 \equiv \mathbf{p}$, $\mathbf{k}_4 \equiv -\mathbf{p} + \mathbf{q}$, yielding the momentum representation of $\hat{V}$:

$$\begin{aligned} \hat{V} &= -\frac{U}{M} \sum_{\mathbf{k},\mathbf{p},\mathbf{q}} \hat{\psi}^\dagger_\uparrow(\mathbf{k})\hat{\psi}_\uparrow(\mathbf{p})\hat{\psi}^\dagger_\downarrow(-\mathbf{k}+\mathbf{q})\hat{\psi}_\downarrow(-\mathbf{p}+\mathbf{q}) \\ &= -\frac{U}{M} \sum_{\mathbf{k},\mathbf{p},\mathbf{q}} \hat{\psi}^\dagger_\uparrow(\mathbf{k})\hat{\psi}^\dagger_\downarrow(-\mathbf{k}+\mathbf{q})\hat{\psi}_\downarrow(-\mathbf{p}+\mathbf{q})\hat{\psi}_\uparrow(\mathbf{p})\,. \end{aligned} \tag{5.29}$$

The term $-\frac{U}{M}\hat{\psi}^\dagger_\uparrow(\mathbf{k})\hat{\psi}^\dagger_\downarrow(-\mathbf{k}+\mathbf{q})\hat{\psi}_\downarrow(-\mathbf{p}+\mathbf{q})\hat{\psi}_\uparrow(\mathbf{p})$ describes the interaction between a pair of particles in a singlet state and with center-of-mass momentum $\mathbf{q}$. We now make our first approximation, i.e., we assume that only pairs of zero center-of-mass momentum interact:

$$\hat{V} \simeq -\frac{U}{M} \sum_{\mathbf{k},\mathbf{p}} \hat{\psi}^\dagger_\uparrow(\mathbf{k})\hat{\psi}^\dagger_\downarrow(-\mathbf{k})\hat{\psi}_\downarrow(-\mathbf{p})\hat{\psi}_\uparrow(\mathbf{p})\,. \tag{5.30}$$

Then we break up the pair destruction operator $\hat{\psi}_\downarrow(-\mathbf{p})\hat{\psi}_\uparrow(\mathbf{p})$ as a fluctuation over a mean-field value

$$\hat{\psi}_\downarrow(-\mathbf{p})\hat{\psi}_\uparrow(\mathbf{p}) = \left\langle \hat{\psi}_\downarrow(-\mathbf{p})\hat{\psi}_\uparrow(\mathbf{p}) \right\rangle + \delta\left(\hat{\psi}_\downarrow(-\mathbf{p})\hat{\psi}_\uparrow(\mathbf{p})\right)\,, \tag{5.31}$$

and we perform our second approximation: we neglect the terms in $\hat{V}$ which are quadratic in the fluctuations, obtaining

$$\hat{V}_{BCS} = const - \frac{U}{M} \sum_{\mathbf{k},\mathbf{p}} \left(\left\langle \hat{\psi}^\dagger_\uparrow(\mathbf{k})\hat{\psi}^\dagger_\downarrow(-\mathbf{k}) \right\rangle \hat{\psi}_\downarrow(-\mathbf{p})\hat{\psi}_\uparrow(\mathbf{p}) + \hat{\psi}(\mathbf{k})^\dagger_\uparrow \hat{\psi}^\dagger_\downarrow(-\mathbf{k}) \left\langle \hat{\psi}_\downarrow(-\mathbf{p})\hat{\psi}_\uparrow(\mathbf{p}) \right\rangle \right)\,. \tag{5.32}$$

We now define the BCS *superfluid gap* as

$$\Delta_{BCS} = \frac{U}{M}\sum_{\mathbf{p}}\left\langle\hat{\psi}_{\downarrow}(-\mathbf{p})\hat{\psi}_{\uparrow}(\mathbf{p})\right\rangle = \frac{U}{M}\sum_{\mathbf{r}}\left\langle\hat{\psi}_{\downarrow}(\mathbf{r})\hat{\psi}_{\uparrow}(\mathbf{r})\right\rangle , \tag{5.33}$$

and thus we get the final BCS lattice Hamiltonian:

$$\hat{H}_{BCS} = const + \sum_{\mathbf{k},\sigma}\left(\varepsilon(\mathbf{k}) - \mu_{BCS}\right)\hat{\psi}^{\dagger}_{\sigma}(\mathbf{k})\hat{\psi}_{\sigma}(\mathbf{k}) \\ - \sum_{\mathbf{k}}\left(\Delta^{*}_{BCS}\hat{\psi}_{\downarrow}(-\mathbf{k})\hat{\psi}_{\uparrow}(\mathbf{k}) + \Delta_{BCS}\hat{\psi}^{\dagger}_{\uparrow}(\mathbf{k})\hat{\psi}^{\dagger}_{\downarrow}(-\mathbf{k})\right) , \tag{5.34}$$

where we have introduced the chemical potential μ_{BCS}, since we are going to deal with states that have no definite number of particles. This mean-field Hamiltonian is quadratic in the creation and annihilation operators, so it can be easily diagonalized. We assume for simplicity that Δ_{BCS} is real, and rewrite the Hamiltonian as

$$\hat{H}_{BCS} = const + \sum_{\mathbf{k}}\begin{pmatrix} \hat{\psi}^{\dagger}_{\uparrow}(\mathbf{k}) & \hat{\psi}_{\downarrow}(-\mathbf{k}) \end{pmatrix}\begin{pmatrix} \xi(\mathbf{k}) & -\Delta_{BCS} \\ -\Delta_{BCS} & -\xi(\mathbf{k}) \end{pmatrix}\begin{pmatrix} \hat{\psi}_{\uparrow}(\mathbf{k}) \\ \hat{\psi}^{\dagger}_{\downarrow}(-\mathbf{k}) \end{pmatrix} , \tag{5.35}$$

where we introduced the shorthand $\xi(\mathbf{k}) = \varepsilon(\mathbf{k}) - \mu_{BCS}$. We diagonalize the matrix, which yields

$$\begin{pmatrix} \xi(\mathbf{k}) & -\Delta_{BCS} \\ -\Delta_{BCS} & -\xi(\mathbf{k}) \end{pmatrix} = \\ \begin{pmatrix} u(\mathbf{k}) & v(\mathbf{k}) \\ -v(\mathbf{k}) & u(\mathbf{k}) \end{pmatrix}\begin{pmatrix} E(\mathbf{k}) & 0 \\ 0 & -E(\mathbf{k}) \end{pmatrix}\begin{pmatrix} u(\mathbf{k}) & -v(\mathbf{k}) \\ v(\mathbf{k}) & u(\mathbf{k}) \end{pmatrix} , \tag{5.36}$$

where we define the quasiparticle dispersion

$$E(\mathbf{k}) = \sqrt{\xi(\mathbf{k})^2 + \Delta^2_{BCS}} = \sqrt{\left(\varepsilon(\mathbf{k}) - \mu_{BCS}\right)^2 + \Delta^2_{BCS}} \tag{5.37}$$

and the quantities

$$u^2(\mathbf{k}) = \frac{1}{2}\left(1 + \frac{\xi(\mathbf{k})}{\sqrt{\xi(\mathbf{k})^2 + \Delta^2_{BCS}}}\right) , \quad v^2(\mathbf{k}) = \frac{1}{2}\left(1 - \frac{\xi(\mathbf{k})}{\sqrt{\xi(\mathbf{k})^2 + \Delta^2_{BCS}}}\right) . \tag{5.38}$$

The final step is the definition of the Bogoliubov operators:

$$\begin{aligned} \hat{\gamma}^{\dagger}_{\uparrow}(\mathbf{k}) &= u(\mathbf{k})\hat{\psi}^{\dagger}_{\uparrow}(\mathbf{k}) - v(\mathbf{k})\hat{\psi}_{\downarrow}(-\mathbf{k}) , \\ \hat{\gamma}_{\downarrow}(-\mathbf{k}) &= v(\mathbf{k})\hat{\psi}^{\dagger}_{\uparrow}(\mathbf{k}) + u(\mathbf{k})\hat{\psi}_{\downarrow}(-\mathbf{k}) . \end{aligned} \tag{5.39}$$

Using the diagonalization (5.36) and the definitions (5.39), the Hamiltonian (5.35) can be rewritten as

$$\hat{H}_{BCS} = const + \sum_{\mathbf{k}} E(\mathbf{k})\left(\hat{\gamma}^{\dagger}_{\uparrow}(\mathbf{k})\hat{\gamma}_{\uparrow}(\mathbf{k}) + \hat{\gamma}^{\dagger}_{\downarrow}(\mathbf{k})\hat{\gamma}_{\downarrow}(\mathbf{k})\right) . \tag{5.40}$$

This Hamiltonian gives a mean-field description of the system as a gas of non-interacting Bogoliubov quasiparticles. The quasiparticles have dispersion $E(\mathbf{k})$, and are created

and annihilated by the operators $\hat{\gamma}^\dagger_\sigma(\mathbf{k})$ and $\hat{\gamma}_\sigma(\mathbf{k})$, respectively. From the dispersion relation (5.37), we can understand why the quantity Δ_{BCS} is called a *gap*: there are no quasiparticle excitations whose energy is below Δ_{BCS}.

Which is the ground state of $\hat{H}_{BCS}$? Since each quasiparticle has positive energy $E(\mathbf{k})$, the state of minimal energy is the state which has zero quasiparticles, i.e., the quasiparticle *vacuum*. In other words, the BCS ground state $|BCS\rangle$ has to be annihilated by the quasiparticle destruction operators:

$$\hat{\gamma}_{\mathbf{k},\sigma}\,|BCS\rangle = 0\,, \quad \forall \mathbf{k}, \sigma\,. \tag{5.41}$$

One can easily verify that this relation is satisfied by the state

$$|BCS\rangle = \prod_{\mathbf{k}} \left(u(\mathbf{k}) + v(\mathbf{k})\hat{\psi}^\dagger_\uparrow(\mathbf{k})\hat{\psi}^\dagger_\downarrow(-\mathbf{k}) \right) |0\rangle\ . \tag{5.42}$$

From this definition of $|BCS\rangle$, we can see that the ground state is constructed by creating pairs of particles with opposite momentum and in a spin singlet state, the so-called Cooper pairs. We observe that for zero interaction, $U = 0$, the gap Δ_{BCS} vanishes, and from Eq. (5.38) follows that $u(\mathbf{k}) = \theta\left(\varepsilon(\mathbf{k}) - \mu_{BCS}\right)$ and $v(\mathbf{k}) = \theta\left(\mu_{BCS} - \varepsilon(\mathbf{k})\right)$, so that the BCS state corresponds in the $U \to 0$ limit to an ideal Fermi gas, μ_{BCS} being the Fermi energy. For $U > 0$, on the other hand, the state $|BCS\rangle$ becomes different from the free gas. It is not even an eigenstate of the particle number $\hat{N} = \sum_{k,\sigma} \psi^\dagger_\sigma(\mathbf{k})\psi_\sigma(\mathbf{k})$, hence the need of introducing a chemical potential μ_{BCS}: it has the role of a Lagrange multiplier that fixes the number of particles $\left\langle BCS \middle| \hat{N} \middle| BCS \right\rangle$ to some value N. Imposing this condition results in the *particle number equation*:

$$N = 2\sum_{\mathbf{k}} v^2(\mathbf{k}) = \sum_{\mathbf{k}} \left(1 - \frac{\varepsilon(\mathbf{k}) - \mu_{BCS}}{\sqrt{\left(\varepsilon(\mathbf{k}) - \mu_{BCS}\right)^2 + \Delta^2_{BCS}}} \right)\,, \tag{5.43}$$

which the BCS ground state has to satisfy in a self-consistent fashion. Morever, the mean-field theory was built under the assumption that the operator $\frac{U}{M}\sum_{\mathbf{p}} \hat{\psi}_\downarrow(-\mathbf{p})\hat{\psi}_\uparrow(\mathbf{p})$ has average value Δ_{BCS}. So there is another self-consistency condition to be fulfilled, the so-called *gap equation*, usually written as

$$1 = \frac{1}{2}\frac{U}{M}\sum_{\mathbf{k}} \frac{1}{\sqrt{\left(\varepsilon(\mathbf{k}) - \mu_{BCS}\right)^2 + \Delta^2_{BCS}}}\,. \tag{5.44}$$

Together, the particle number equation (5.43) and the gap equation (5.44) determine the values of the two parameters of the BCS theory, namely the chemical potential μ_{BCS} and the superfluid gap Δ_{BCS}.

5.6 Generalized random phase approximation (GRPA)

The generalized random phase approximation (GRPA) is an approach which takes the BCS as its starting point, and improves on it to overcome some of the shortcomings of a plain BCS computation of the dynamical correlation functions. The procedure we follow is the one presented by Kelly and Vitali in Ref. [15], which was modeled after Refs. [14, 13].

We want to compute the response functions of our lattice system (5.9) in the linear response regime, since from them it is possible to compute the dynamical structure factors. So let's first briefly review the results of standard linear response theory. We consider a system whose unperturbed Hamiltonian is $\hat{H}_0$, initially found in the ground state of $\hat{H}_0$. We subject it to a time-dependent perturbation $\hat{h}(t)$. We want to know the variation induced by the perturbation in the expectation value of some position-dependent operator $\hat{A}(\mathbf{r})$. We assume that the perturbation couples with another position-dependent operator $\hat{B}(\mathbf{r})$ through the relation

$$\hat{h}(t) = \sum_{\mathbf{r}} \left[h(\mathbf{r},t)\, \hat{B}(\mathbf{r}) + h^*(\mathbf{r},t)\, \hat{B}^\dagger(\mathbf{r}) \right] , \tag{5.45}$$

where $h(\mathbf{r},t)$ is a complex scalar function. This definition, involving the sum of $h(\mathbf{r},t)\hat{B}(\mathbf{r})$ and its Hermitian conjugate, guarantees that the perturbation $\hat{h}(t)$ is Hermitian even when $\hat{B}$ is not. The total, time-dependent Hamiltonian is

$$\hat{H}(t) = \hat{H}_0 + \hat{h}(t) = \hat{H}_0 + \sum_{\mathbf{r}} \left[h(\mathbf{r},t)\, \hat{B}(\mathbf{r}) + h^*(\mathbf{r},t)\, \hat{B}^\dagger(\mathbf{r}) \right] . \tag{5.46}$$

Linear response theory tells us that the variation $\delta \left\langle \hat{A}(\mathbf{r},t) \right\rangle$ in the expectation value of $\hat{A}$ is given by

$$\begin{aligned} \delta \left\langle \hat{A}(\mathbf{r},t) \right\rangle = \sum_{\mathbf{r}'} \int_{-\infty}^{+\infty} dt'\, \chi_{AB}(\mathbf{r},\mathbf{r}',t-t')\, h(\mathbf{r}',t') \\ + \sum_{\mathbf{r}'} \int_{-\infty}^{+\infty} dt'\, \chi_{AB^\dagger}(\mathbf{r},\mathbf{r}',t-t')\, h^*(\mathbf{r}',t') . \end{aligned} \tag{5.47}$$

The response function χ_{AB} is defined by

$$\chi_{AB}(\mathbf{r},\mathbf{r}',t-t') = -\frac{i}{\hbar}\theta(t-t') \left\langle \left[e^{\frac{i}{\hbar}\hat{H}_0 t} \hat{A}(\mathbf{r}) e^{-\frac{i}{\hbar}\hat{H}_0 t}, e^{\frac{i}{\hbar}\hat{H}_0 t'} \hat{B}(\mathbf{r}') e^{-\frac{i}{\hbar}\hat{H}_0 t'} \right] \right\rangle , \tag{5.48}$$

where the expectation is to be evaluated on the ground state of the unperturbed Hamiltonian. The response $\chi_{AB^\dagger}(\mathbf{r},\mathbf{r}',t-t')$ has a similar definition, substituting $\hat{B}$ with $\hat{B}^\dagger$.

We now assume that this ground state is translationally invariant, so that $\chi_{AB}(\mathbf{r},\mathbf{r}',t-t')$ depends on the positions $\mathbf{r}$, $\mathbf{r}'$ only through their difference $\mathbf{r}-\mathbf{r}'$, i.e., $\chi_{AB}(\mathbf{r},\mathbf{r}',t-t') = \chi_{AB}(\mathbf{r}-\mathbf{r}',t-t')$. This allows us to switch to frequency and momentum space through a Fourier transform. The response function in frequency and momentum space is defined as

$$\chi_{AB}(\mathbf{q},\omega) = \sum_{\mathbf{r}} \int_{-\infty}^{+\infty} dt\, e^{i\omega t - i\mathbf{q}\cdot\mathbf{r}}\, \chi_{AB}(\mathbf{r},t) , \tag{5.49}$$

and analogous definitions apply to the Fourier transforms of the other quantities involved[1]. The time-dependent perturbation (5.45) can be written in terms of the spatial Fourier transforms of $\hat{B}$ and h as

$$\hat{h}(t) = \frac{1}{M} \sum_{\mathbf{q}} \left[h(\mathbf{q},t)\, \hat{B}(-\mathbf{q}) + h^*(-\mathbf{q},t)\, \hat{B}^\dagger(\mathbf{q}) \right] . \tag{5.50}$$

[1]Please notice that the Fourier transform of the function $h^*(\mathbf{r},t)$ is $h^*(-\mathbf{q},-\omega)$, i.e., the complex conjugate of the Fourier transform of $h(\mathbf{r},t)$ evaluated in $(-\mathbf{q},-\omega)$. A similar relation holds between the Fourier transforms of $\hat{B}$ and $\hat{B}^\dagger$.

After the transform, Eq. (5.47) becomes

$$\delta \left\langle \hat{A}(\mathbf{q},\omega) \right\rangle = \chi_{AB}(\mathbf{q},\omega)\, h(\mathbf{q},\omega) + \chi_{AB^\dagger}(\mathbf{q},\omega)\, h^*(-\mathbf{q},-\omega)\,, \tag{5.51}$$

which is a particularly convenient expression, since the summations and integrals in Eq. (5.47) have been replaced by products. Introducing a basis of eigenvectors $|\Psi_n\rangle$ of the unperturbed Hamiltonian $\hat{H}_0$, with $\hat{H}_0\,|\Psi_n\rangle = E_n\,|\Psi_n\rangle$, we can write the so-called *spectral representation* of the response function:

$$\chi_{AB}(\mathbf{q},\omega) = \\ \frac{1}{M}\lim_{\eta\to 0^+}\sum_{n>0}\left(\frac{\langle\Psi_0|\hat{B}(-\mathbf{q})|\Psi_n\rangle\langle\Psi_n|\hat{A}(\mathbf{q})|\Psi_0\rangle}{\hbar\omega + (E_n - E_0) + i\eta} - \frac{\langle\Psi_0|\hat{A}(\mathbf{q})|\Psi_n\rangle\langle\Psi_n|\hat{B}(-\mathbf{q})|\Psi_0\rangle}{\hbar\omega - (E_n - E_0) + i\eta}\right). \tag{5.52}$$

An analogous definition holds for $\chi_{AB^\dagger}(\mathbf{q},\omega)$, replacing $\hat{B}(-\mathbf{q})$ with $\hat{B}^\dagger(\mathbf{q})$. So, if we are able to compute the matrix elements of $\hat{A}$ and $\hat{B}$ with respect to the ground state and the excited states of the unperturbed Hamiltonian, we can compute the response function $\chi_{AB}(\mathbf{q},\omega)$ by means of the spectral representation.

Unfortunately, we cannot directly apply this theory to our lattice Hamiltonian $\hat{H}_{\mathcal{L}}$, Eq. (5.9), because this would require exact knowledge of its ground state and its excited states. But we can apply linear response theory to the BCS Hamiltonian $\hat{H}_{BCS}$ (5.40), since we know its ground state $|BCS\rangle$ and its excited states, that are obtained by applying the Bogoliubov operators (5.39) to the BCS ground state to excite quasiparticles. So we consider the BCS Hamiltonian subject to a time-dependent perturbation:

$$\hat{H}(t) = \hat{H}_{BCS} + \hat{h}(t)\,, \tag{5.53}$$

where $\hat{h}(t)$ is a very general perturbation that couples with the spin-resolved particle densities $\hat{n}_\sigma(t)$ and the pairing operator $\hat{\Delta}(\mathbf{r}) = \hat{\psi}_\downarrow(\mathbf{r})\hat{\psi}_\uparrow(\mathbf{r})$. It can be written as

$$\hat{h}(t) = \sum_{\mathbf{r}}\left(h_{\hat{n}_\uparrow}(\mathbf{r},t)\,\hat{n}_\uparrow(\mathbf{r}) + h_{\hat{n}_\downarrow}(\mathbf{r},t)\,\hat{n}_\downarrow(\mathbf{r}) + h_{\hat{\Delta}}(\mathbf{r},t)\,\hat{\Delta}(\mathbf{r}) + h^*_{\hat{\Delta}}(\mathbf{r},t)\,\hat{\Delta}^\dagger(\mathbf{r})\right) \tag{5.54}$$

$$= \frac{1}{M}\sum_{\mathbf{q}}\left(h_{\hat{n}_\uparrow}(\mathbf{q},t)\,\hat{n}_\uparrow(-\mathbf{q}) + h_{\hat{n}_\downarrow}(\mathbf{q},t)\,\hat{n}_\downarrow(-\mathbf{q}) + h_{\hat{\Delta}}(\mathbf{q},t)\,\hat{\Delta}(-\mathbf{q}) + h^*_{\hat{\Delta}}(-\mathbf{q},t)\,\hat{\Delta}^\dagger(\mathbf{q})\right), \tag{5.55}$$

where the "sources" $h_i(\mathbf{r},t)$ are scalar functions of position and time; $h_{\hat{n}_\uparrow}$ and $h_{\hat{n}_\downarrow}$ are real, while $h_{\hat{\Delta}}$ is complex. Applying Eq. (5.52) we can compute the response functions of the operators $\hat{n}_\uparrow(\mathbf{r})$, $\hat{n}_\downarrow(\mathbf{r})$, $\hat{\Delta}(\mathbf{r})$, $\hat{\Delta}^\dagger(\mathbf{r})$ with respect to each other, in the BCS approximation. In other words, we can compute the BCS response-function matrix

$$\chi^{BCS}(\mathbf{q},\omega) = \begin{pmatrix} \chi^{BCS}_{\hat{n}_\uparrow,\hat{n}_\uparrow}(\mathbf{q},\omega) & \chi^{BCS}_{\hat{n}_\uparrow,\hat{n}_\downarrow}(\mathbf{q},\omega) & \chi^{BCS}_{\hat{n}_\uparrow,\hat{\Delta}}(\mathbf{q},\omega) & \chi^{BCS}_{\hat{n}_\uparrow,\hat{\Delta}^\dagger}(\mathbf{q},\omega) \\ \chi^{BCS}_{\hat{n}_\downarrow,\hat{n}_\uparrow}(\mathbf{q},\omega) & \chi^{BCS}_{\hat{n}_\downarrow,\hat{n}_\downarrow}(\mathbf{q},\omega) & \chi^{BCS}_{\hat{n}_\downarrow,\hat{\Delta}}(\mathbf{q},\omega) & \chi^{BCS}_{\hat{n}_\downarrow,\hat{\Delta}^\dagger}(\mathbf{q},\omega) \\ \chi^{BCS}_{\hat{\Delta},\hat{n}_\uparrow}(\mathbf{q},\omega) & \chi^{BCS}_{\hat{\Delta},\hat{n}_\downarrow}(\mathbf{q},\omega) & \chi^{BCS}_{\hat{\Delta},\hat{\Delta}}(\mathbf{q},\omega) & \chi^{BCS}_{\hat{\Delta},\hat{\Delta}^\dagger}(\mathbf{q},\omega) \\ \chi^{BCS}_{\hat{\Delta}^\dagger,\hat{n}_\uparrow}(\mathbf{q},\omega) & \chi^{BCS}_{\hat{\Delta}^\dagger,\hat{n}_\downarrow}(\mathbf{q},\omega) & \chi^{BCS}_{\hat{\Delta}^\dagger,\hat{\Delta}}(\mathbf{q},\omega) & \chi^{BCS}_{\hat{\Delta}^\dagger,\hat{\Delta}^\dagger}(\mathbf{q},\omega) \end{pmatrix}. \tag{5.56}$$

Now we defining the vector of the responses

$$\delta\langle\hat{O}(\mathbf{q},\omega)\rangle = \begin{pmatrix} \delta\langle\hat{n}_\uparrow(\mathbf{q},\omega)\rangle \\ \delta\langle\hat{n}_\downarrow(\mathbf{q},\omega)\rangle \\ \delta\langle\hat{\Delta}(\mathbf{q},\omega)\rangle \\ \delta\langle\hat{\Delta}^\dagger(-\mathbf{q},-\omega)\rangle \end{pmatrix} \tag{5.57}$$

and the vector of the external fields

$$h(\mathbf{q},\omega) = \begin{pmatrix} h_{\hat{n}_\uparrow}(\mathbf{q},\omega) \\ h_{\hat{n}_\downarrow}(\mathbf{q},\omega) \\ h_{\hat{\Delta}}(\mathbf{q},\omega) \\ h^*_{\hat{\Delta}}(-\mathbf{q},-\omega) \end{pmatrix} . \tag{5.58}$$

Using these definitions, we have, according to Eq. (5.51), the following matrix equation:

$$\delta\langle\hat{O}(\mathbf{q},\omega)\rangle = \chi^{BCS}(\mathbf{q},\omega)\, h(\mathbf{q},\omega) . \tag{5.59}$$

The main limitation of the BCS response functions is that the BCS Hamiltonian (5.40) assumes fixed values for the gap Δ_{BCS} and the chemical potential μ_{BCS}. However, when a perturbation is applied, Δ_{BCS} and μ_{BCS} should be able to change in response to it. The BCS responses we found do not take into account this variability of Δ_{BCS} and μ_{BCS} as a consequence of the perturbation.

The generalized random phase approximation (GRPA) goes beyond the BCS treatment by using a clever redefinition of the perturbation, which self-consistently includes the first-order effects of the perturbation on the mean fields. To set the stage for GRPA, we go back to the lattice Hamiltonian $\hat{H}_{\mathcal{L}}$, Eq. (5.9), and perform a full mean-field breakup of the interaction term (5.7), yielding

$$\hat{V} = -U \sum_{\mathbf{r}} \hat{\psi}^\dagger_\uparrow(\mathbf{r})\hat{\psi}_\uparrow(\mathbf{r})\hat{\psi}^\dagger_\downarrow(\mathbf{r})\hat{\psi}_\downarrow(\mathbf{r}) \tag{5.60}$$

$$\begin{aligned} &\simeq const - U \sum_{\mathbf{r}} \Big[\Big\langle \hat{\psi}^\dagger_\uparrow(\mathbf{r})\hat{\psi}_\uparrow(\mathbf{r}) \Big\rangle \hat{\psi}^\dagger_\downarrow(\mathbf{r})\hat{\psi}_\downarrow(\mathbf{r}) + \Big\langle \hat{\psi}^\dagger_\downarrow(\mathbf{r})\hat{\psi}_\downarrow(\mathbf{r}) \Big\rangle \hat{\psi}^\dagger_\uparrow(\mathbf{r})\hat{\psi}_\uparrow(\mathbf{r}) \\ &\qquad + \Big\langle \hat{\psi}^\dagger_\uparrow(\mathbf{r})\hat{\psi}^\dagger_\downarrow(\mathbf{r}) \Big\rangle \hat{\psi}_\downarrow(\mathbf{r})\hat{\psi}_\uparrow(\mathbf{r}) + \Big\langle \hat{\psi}_\downarrow(\mathbf{r})\hat{\psi}_\uparrow(\mathbf{r}) \Big\rangle \hat{\psi}(\mathbf{r})^\dagger_\uparrow\hat{\psi}^\dagger_\downarrow(\mathbf{r}) \Big] \end{aligned} \tag{5.61}$$

$$\simeq const - U \sum_{\mathbf{r}} \Big[\langle\hat{n}_\uparrow(\mathbf{r})\rangle\, \hat{n}_\downarrow(\mathbf{r}) + \langle\hat{n}_\downarrow(\mathbf{r})\rangle\, \hat{n}_\uparrow(\mathbf{r}) + \Big\langle \hat{\Delta}^\dagger(\mathbf{r}) \Big\rangle \hat{\Delta}(\mathbf{r}) + \Big\langle \hat{\Delta}(\mathbf{r}) \Big\rangle \hat{\Delta}^\dagger(\mathbf{r}) \Big] , \tag{5.62}$$

where we neglected the terms of higher order in the fluctuations of the mean fields, which this time include not only $\langle\Delta(\mathbf{r})\rangle$ and $\langle\Delta^\dagger(\mathbf{r})\rangle$, as we did when formulating the BCS theory, but *also* the spin-resolved densities $\langle\hat{n}_\uparrow(\mathbf{r})\rangle$ and $\langle\hat{n}_\downarrow(\mathbf{r})\rangle$. After the switching on of the perturbation $\hat{h}(\mathbf{r},t)$ (5.54), the mean-field average of $\hat{n}_\uparrow(\mathbf{r})$ will assume the form

$$\langle\hat{n}_\uparrow(\mathbf{r},t)\rangle = \langle\hat{n}_\uparrow(\mathbf{r})\rangle_0 + \delta\langle\hat{n}_\uparrow(\mathbf{r},t)\rangle , \tag{5.63}$$

where $\langle\hat{n}_\uparrow(\mathbf{r})\rangle_0$ is the unperturbed value and $\delta\langle\hat{n}_\uparrow(\mathbf{r},t)\rangle$ is the first-order change due to the perturbation. Similarly, all the mean-field averages appearing in Eq. (5.62) will

assume an analogous form. So the full perturbed Hamiltonian will be written as

$$\hat{H}(t) = \hat{H}_{\mathcal{L}} + \hat{h}(t) \tag{5.64}$$

$$= \sum_{\mathbf{k},\sigma} \varepsilon(\mathbf{k})\, \hat{\psi}^{\dagger}_{\sigma}(\mathbf{k})\hat{\psi}_{\sigma}(\mathbf{k}) - U \sum_{\mathbf{r}} \hat{\psi}^{\dagger}_{\uparrow}(\mathbf{r})\hat{\psi}_{\uparrow}(\mathbf{r})\hat{\psi}^{\dagger}_{\downarrow}(\mathbf{r})\hat{\psi}_{\downarrow}(\mathbf{r}) + \hat{h}(t) \tag{5.65}$$

$$\begin{aligned} \simeq \hat{H}_{BCS} + \hat{h}(t) - U \sum_{\mathbf{r}} \Big[& \delta \left\langle \hat{n}_{\uparrow}(\mathbf{r},t) \right\rangle \hat{n}_{\downarrow}(\mathbf{r}) + \delta \left\langle \hat{n}_{\downarrow}(\mathbf{r},t) \right\rangle \hat{n}_{\uparrow}(\mathbf{r}) \\ & + \delta \left\langle \hat{\Delta}^{\dagger}(\mathbf{r},t) \right\rangle \hat{\Delta}(\mathbf{r}) + \delta \left\langle \hat{\Delta}(\mathbf{r},t) \right\rangle \hat{\Delta}^{\dagger}(\mathbf{r}) \Big] , \end{aligned} \tag{5.66}$$

where we denote by $\hat{H}_{BCS}$ the mean-field Hamiltonian, defined as

$$\begin{aligned} \hat{H}_{BCS} = \sum_{\mathbf{k},\sigma} \varepsilon(\mathbf{k})\, \hat{\psi}^{\dagger}_{\sigma}(\mathbf{k})\hat{\psi}_{\sigma}(\mathbf{k}) - U \sum_{\mathbf{r}} \Big[& \left\langle \hat{n}_{\uparrow}(\mathbf{r}) \right\rangle_0 \hat{n}_{\downarrow}(\mathbf{r}) + \left\langle \hat{n}_{\downarrow}(\mathbf{r}) \right\rangle_0 \hat{n}_{\uparrow}(\mathbf{r}) \\ & + \left\langle \hat{\Delta}^{\dagger}(\mathbf{r}) \right\rangle_0 \hat{\Delta}(\mathbf{r}) + \left\langle \hat{\Delta}(\mathbf{r}) \right\rangle_0 \hat{\Delta}^{\dagger}(\mathbf{r}) \Big] , \end{aligned} \tag{5.67}$$

to which we can apply the BCS description[2]. Now we perform the crucial passage of the GRPA approach: looking at Eq. (5.66), we define a new self-consistent perturbation

$$\begin{aligned} \hat{h}_{\text{eff}}(t) = \hat{h}(t) - U \sum_{\mathbf{r}} \Big[& \delta \left\langle \hat{n}_{\uparrow}(\mathbf{r},t) \right\rangle \hat{n}_{\downarrow}(\mathbf{r}) + \delta \left\langle \hat{n}_{\downarrow}(\mathbf{r},t) \right\rangle \hat{n}_{\uparrow}(\mathbf{r}) \\ & + \delta \left\langle \hat{\Delta}^{\dagger}(\mathbf{r},t) \right\rangle \hat{\Delta}(\mathbf{r}) + \delta \left\langle \hat{\Delta}(\mathbf{r},t) \right\rangle \hat{\Delta}^{\dagger}(\mathbf{r}) \Big] , \end{aligned} \tag{5.68}$$

$$\begin{aligned} = \sum_{\mathbf{r}} \Big\{ & \left[h_{\hat{n}_{\uparrow}}(\mathbf{r},t) - U\, \delta \left\langle \hat{n}_{\downarrow}(\mathbf{r},t) \right\rangle \right] \hat{n}_{\uparrow}(\mathbf{r}) + \left[h_{\hat{n}_{\downarrow}}(\mathbf{r},t) - U\, \delta \left\langle \hat{n}_{\uparrow}(\mathbf{r},t) \right\rangle \right] \hat{n}_{\downarrow}(\mathbf{r}) \\ & + \left[h_{\hat{\Delta}}(\mathbf{r},t) - U\, \delta \left\langle \hat{\Delta}^{\dagger}(\mathbf{r},t) \right\rangle \right] \hat{\Delta}(\mathbf{r}) + \left[h^{*}_{\hat{\Delta}}(\mathbf{r},t) - U\, \delta \left\langle \hat{\Delta}(\mathbf{r},t) \right\rangle \right] \hat{\Delta}^{\dagger}(\mathbf{r}) \Big\} , \end{aligned} \tag{5.69}$$

and the corresponding vector of the self-consistent external fields

$$h_{\text{eff}}(\mathbf{q},\omega) = \begin{pmatrix} h_{\hat{n}_{\uparrow}}(\mathbf{q},\omega) \\ h_{\hat{n}_{\downarrow}}(\mathbf{q},\omega) \\ h_{\hat{\Delta}}(\mathbf{q},\omega) \\ h^{*}_{\hat{\Delta}}(-\mathbf{q},-\omega) \end{pmatrix} - U \begin{pmatrix} \delta\langle \hat{n}_{\downarrow}(\mathbf{q},\omega) \rangle \\ \delta\langle \hat{n}_{\uparrow}(\mathbf{q},\omega) \rangle \\ \delta\langle \hat{\Delta}^{\dagger}(-\mathbf{q},-\omega) \rangle \\ \delta\langle \hat{\Delta}(\mathbf{q},\omega) \rangle \end{pmatrix} = h(\mathbf{q},\omega) - U\, \Sigma\, \delta\langle \hat{O}(\mathbf{q},\omega) \rangle , \tag{5.70}$$

where we used the definitions (5.58) and (5.57) and introduced a permutation matrix

$$\Sigma = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} . \tag{5.71}$$

The full perturbed Hamiltonian (5.66) can thus be written as

$$\hat{H}(t) = \hat{H}_{BCS} + \hat{h}_{\text{eff}}(t) , \tag{5.72}$$

[2]Consistency requires that the equilibrium expectation values of the spin-resolved densities $\left\langle \hat{n}_{\uparrow}(\mathbf{r}) \right\rangle_0$ and $\left\langle \hat{n}_{\downarrow}(\mathbf{r}) \right\rangle_0$, multiplied by U, coincide with the chemical potential μ_{BCS}. Similarly, the equilibrium expectation value of the pairing operator $\left\langle \hat{\Delta}(\mathbf{r}) \right\rangle_0$ multiplied by U is equal to the superfluid gap Δ_{BCS}.

and we can apply linear response theory to this modified version of the perturbation on $\hat{H}_{BCS}$, obtaining, in analogy with Eq. (5.59),

$$\delta\langle\hat{O}\rangle = \chi^{BCS}\, h_{\text{eff}} \tag{5.73}$$

$$= \chi^{BCS}\, \left(h - U\,\Sigma\,\delta\langle\hat{O}\rangle\right) \tag{5.74}$$

$$= \chi^{BCS}\, h - U\chi^{BCS}\,\Sigma\,\delta\langle\hat{O}\rangle\,, \tag{5.75}$$

where we omitted for simplicity the dependence of all the quantities on $(\mathbf{q},\omega)$. We rewrite Eq. (5.75) as

$$\left(\mathbb{1} + U\,\chi^{BCS}\,\Sigma\right)\,\delta\langle\hat{O}\rangle = \chi^{BCS}\,h \quad , \tag{5.76}$$

and inverting the matrix $\mathbb{1} + U\,\chi^{BCS}\,\Sigma$, we obtain

$$\delta\langle\hat{O}\rangle = \left(\mathbb{1} + U\,\chi^{BCS}\,\Sigma\right)^{-1}\chi^{BCS}\,h\,. \tag{5.77}$$

So we can write the GRPA response-function matrix in terms of the BCS one, as

$$\chi^{GRPA} = \left(\mathbb{1} + U\,\chi^{BCS}\,\Sigma\right)^{-1}\chi^{BCS}\,. \tag{5.78}$$

The elements of the matrix χ^{GRPA} yield the response functions within the GRPA framework. We can combine these elements to form the density-density (χ^{GRPA}_{nn}) and spin-spin (χ^{GRPA}_{ss}) response functions, as follows. From the definitions of the density operator $\hat{n}(\mathbf{r}) = \hat{n}_\uparrow(\mathbf{r}) + \hat{n}_\downarrow(\mathbf{r})$ and the spin density operator $\hat{n}^s(\mathbf{r}) = \frac{1}{2}\left(\hat{n}_\uparrow(\mathbf{r}) - \hat{n}_\downarrow(\mathbf{r})\right)$, these equalities follow:

$$\chi^{GRPA}_{nn}(\mathbf{q},\omega) = \chi^{GRPA}_{\hat{n}_\uparrow,\hat{n}_\uparrow}(\mathbf{q},\omega) + \chi^{GRPA}_{\hat{n}_\uparrow,\hat{n}_\downarrow}(\mathbf{q},\omega) + \chi^{GRPA}_{\hat{n}_\downarrow,\hat{n}_\uparrow}(\mathbf{q},\omega) + \chi^{GRPA}_{\hat{n}_\downarrow,\hat{n}_\downarrow}(\mathbf{q},\omega)\,, \tag{5.79}$$

$$\chi^{GRPA}_{ss}(\mathbf{q},\omega) = \frac{1}{4}\left[\chi^{GRPA}_{\hat{n}_\uparrow,\hat{n}_\uparrow}(\mathbf{q},\omega) - \chi^{GRPA}_{\hat{n}_\uparrow,\hat{n}_\downarrow}(\mathbf{q},\omega) - \chi^{GRPA}_{\hat{n}_\downarrow,\hat{n}_\uparrow}(\mathbf{q},\omega) + \chi^{GRPA}_{\hat{n}_\downarrow,\hat{n}_\downarrow}(\mathbf{q},\omega)\right]. \tag{5.80}$$

Finally, we write the dynamical structure factor $S(\mathbf{q},\omega)$ by invoking the fluctuation-dissipation relation

$$S(\mathbf{q},\omega) = \lim_{\eta\to 0^+}\left[-\frac{\hbar}{\pi n}\chi''_{nn}(\mathbf{q},\omega+i\eta)\right], \tag{5.81}$$

where $n = N/M$ is the average particle number density in the lattice, $\chi''_{nn} = \Im\chi_{nn}$, and η is a convergence parameter (in actual GRPA numerical implementations, the parameter η has to be set to some small positive number). An analogous relation holds between the spin dynamical structure factor $S^s(\mathbf{q},\omega)$ and the spin-spin response function $\chi_{ss}(\mathbf{q},\omega)$. The dynamical structure factors obtained through the GRPA can be compared with those yielded by the unbiased (but more computationally demanding) AFQMC method.

GRPA also yields the response functions that involve the pairing operator $\hat{\Delta}(\mathbf{r})$ and its Hermitian conjugate $\hat{\Delta}^\dagger(\mathbf{r})$, which are all the more interesting since they cannot be obtained by AFQMC calculations: the pairing operator does not conserve the number of particles, while the AFQMC approach assumes particle number conservation. In turn, from the response functions of $\hat{\Delta}(\mathbf{r})$ and $\hat{\Delta}^\dagger(\mathbf{r})$, we can compute those of the phase and amplitude of the order parameter $\langle\hat{\Delta}(\mathbf{r})\rangle$, as follows. The BCS theory assumes that the equilibrium value of $\langle\hat{\Delta}(\mathbf{r})\rangle$ is a real number, so, in the linear response regime, fluctuations of the real and imaginary part of $\langle\hat{\Delta}(\mathbf{r})\rangle$ correspond to fluctuations of the amplitude and phase of the order parameter, respectively. As a consequence, the response

Table 5.1. Momenta $\mathbf{q}$ for which $F(\mathbf{q},\tau)$ is calculated.

q_x	q_y	$\left(\frac{2\pi}{L}\right)$	$\lvert\mathbf{q}\rvert$	(k_F)
1	0		0.4916	
1	1		0.6952	
2	0		0.9832	
2	1		1.0992	
2	2		1.3904	
3	0		1.4748	
3	1		1.5545	

Note. — List of the momenta $\mathbf{q}$ for which we calculate the intermediate scattering function $F(\mathbf{q},\tau)$ by AFQMC. The x and y components are in units of the reciprocal lattice spacing $2\pi/L$; the modulus is in units of the Fermi momentum k_F.

functions of the amplitude can be obtained from the linear responses of the Hermitian part of the pairing operator, $\frac{1}{2}\left(\Delta(\mathbf{r})+\Delta^\dagger(\mathbf{r})\right)$; similarly, the response functions of the phase are constructed from the operator $\frac{U}{2i\Delta_{BCS}}\left(\Delta(\mathbf{r})-\Delta^\dagger(\mathbf{r})\right)$, which is proportional to the anti-Hermitian part of the pairing. Accordingly, the phase-phase χ_{pp}^{GRPA} and the amplitude-amplitude χ_{aa}^{GRPA} response functions are given by

$$\chi_{pp}^{GRPA}(\mathbf{q},\omega)=\frac{U^2}{4\Delta_{BCS}^2}\left[-\chi_{\hat{\Delta},\hat{\Delta}}^{GRPA}(\mathbf{q},\omega)+\chi_{\hat{\Delta},\hat{\Delta}^\dagger}^{GRPA}(\mathbf{q},\omega)+\chi_{\hat{\Delta}^\dagger,\hat{\Delta}}^{GRPA}(\mathbf{q},\omega)-\chi_{\hat{\Delta}^\dagger,\hat{\Delta}^\dagger}^{GRPA}(\mathbf{q},\omega)\right], \tag{5.82}$$

$$\chi_{aa}^{GRPA}(\mathbf{q},\omega)=\frac{1}{4}\left[\chi_{\hat{\Delta},\hat{\Delta}}^{GRPA}(\mathbf{q},\omega)+\chi_{\hat{\Delta},\hat{\Delta}^\dagger}^{GRPA}(\mathbf{q},\omega)+\chi_{\hat{\Delta}^\dagger,\hat{\Delta}}^{GRPA}(\mathbf{q},\omega)+\chi_{\hat{\Delta}^\dagger,\hat{\Delta}^\dagger}^{GRPA}(\mathbf{q},\omega)\right], \tag{5.83}$$

and the corresponding dynamical structure factors $S^p(\mathbf{q},\omega)$ and $S^a(\mathbf{q},\omega)$ can be retrieved from the imaginary parts of these functions, using the analogous of Eq. (5.81).

Moreover, within GRPA we can also perform mode-coupling analyses. For example, we can study the correlations between the order parameter's amplitude and the spin density by computing the function

$$\chi_{as}^{GRPA}(\mathbf{q},\omega)=\frac{1}{4}\left[\chi_{\hat{\Delta},\hat{n}_\uparrow}^{GRPA}(\mathbf{q},\omega)-\chi_{\hat{\Delta},\hat{n}_\downarrow}^{GRPA}(\mathbf{q},\omega)+\chi_{\hat{\Delta}^\dagger,\hat{n}_\uparrow}^{GRPA}(\mathbf{q},\omega)-\chi_{\hat{\Delta}^\dagger,\hat{n}_\downarrow}^{GRPA}(\mathbf{q},\omega)\right]. \tag{5.84}$$

Mode-coupling analyses are especially useful to explore the possibility, suggested in Ref. [13], of detecting the Higgs mode by probing the spin fluctuations.

5.7 Results

Motivated by Ref. [13], we address two key questions: are there significant discrepancies between the GRPA results and the unbiased AFQMC calculations? If AFQMC supports

the GRPA prediction of a peak in the spin dynamical structure factor, can such a peak be identified with the celebrated Higgs mode, as Ref. [13] suggests? In Subsec. 5.7.1 we compare the GRPA and AFQMC results in order to answer to the first question. Subsec. 5.7.2 complements this discussion with an analysis of the finite-size effects in our data. Finally, in Subsec. 5.7.3 we delve into the second question and we propose an explanation for the origin of the peak in the spin dynamical structure factor.

5.7.1 Comparison between GRPA and AFQMC

We perform AFQMC computations for a spin-balanced system of $N = 26$ atoms, with a regularization lattice of $M = 35 \times 35 = 1225$ sites. Previous AFQMC studies [88] indicate that such a lattice is large enough to achieve accurate estimations in the bulk and in the continuum limit. This is also the largest closed-shell size that we are currently able to simulate. We obtain unbiased estimations of the intermediate scattering functions in imaginary-time (5.10). The intermediate scattering functions $F(\mathbf{q}, \tau)$ are computed in both the density and spin channels, for the list of momenta $\mathbf{q}$ in Table 5.1. These are the seven lowest momenta which are compatible with the regularization lattice, the smallest being $q \simeq 0.5\, k_F$. We perform DEAC analytic continuation of $F(\mathbf{q}, \tau)$ to extract the density dynamical structure factor $S(\mathbf{q}, \omega)$ and the spin dynamical structure factor $S^s(\mathbf{q}, \tau)$. We study three different regimes of interaction strength, going from the weakly correlated BCS regime with $\log(k_F a) = 2.5$ to the more correlated, though still not molecule-dominated, $\log(k_F a) = 1.0$. The choices of the values of the interaction were dictated by the intent to compare with the results in Ref. [13]. We also perform GRPA calculations of the same dynamical structure factors, for two different system sizes: the one we investigated by AFQMC, as well as a much larger size at the same particle density, precisely $N = 5850$ atoms on a regularization lattice of $M = 525 \times 525 = 275625$ sites. This larger system, too computationally demanding to be accessible by state-of-the-art AFQMC calculations, allows for the investigation of a denser set of momenta, thus providing a clearer picture. Moreover, it gives important information about finite-size effects. In the GRPA calculation, we need to choose a non-zero value for the convergence parameter η, see Eq. (5.81) and the subsequent discussion. We tune the values empirically, and in particular we choose $\eta \simeq 0.06\, \varepsilon_F$ for 26 atoms and $\eta \simeq 0.03\, \varepsilon_F$ for 5850 atoms.

Fig. 5.1 shows a comparison of our AFQMC results (circles with vertical bars) and GRPA predictions (color plots) for the dynamical structure factors in both the density (left column) and spin (right column) channels. The circles mark the maxima of the AFQMC structure factors, and the vertical bars represent the widths at half heights. The background color plots display the GRPA results for the larger system of 5850 atoms at the same density, from which one can appreciate the much higher momentum resolution that is accessible at this size. The horizontal lines mark the threshold of 2Δ, Δ being the superfluid gap, so that 2Δ is the energy needed to excite a pair of quasiparticles. The red dash-dotted lines show the value of 2Δ according to BCS theory, see Δ_{BCS} in Eq. (5.33), and the white dashed lines display the value of 2Δ according to the AFQMC results from Ref. [88]. The BCS value of the gap informs the GRPA calculations, since the GRPA approach is built on the basis of the BCS theory. The remaining figures show the spectra in more detail: in Fig. 5.2 we plot the AFQMC density and spin dynamical structure factors, for selected momenta and interaction strengths $\log(k_F a) = 1.5$ and 1.0; Fig. 5.3 compare the GRPA and AFQMC results at the same system size, $N = 26$ atoms, for interaction strength $\log(k_F a) = 1.5$; Figs. 5.4 and 5.5 compare the AFQMC and GRPA results for the two different system sizes, to investigate the finite-size effect.

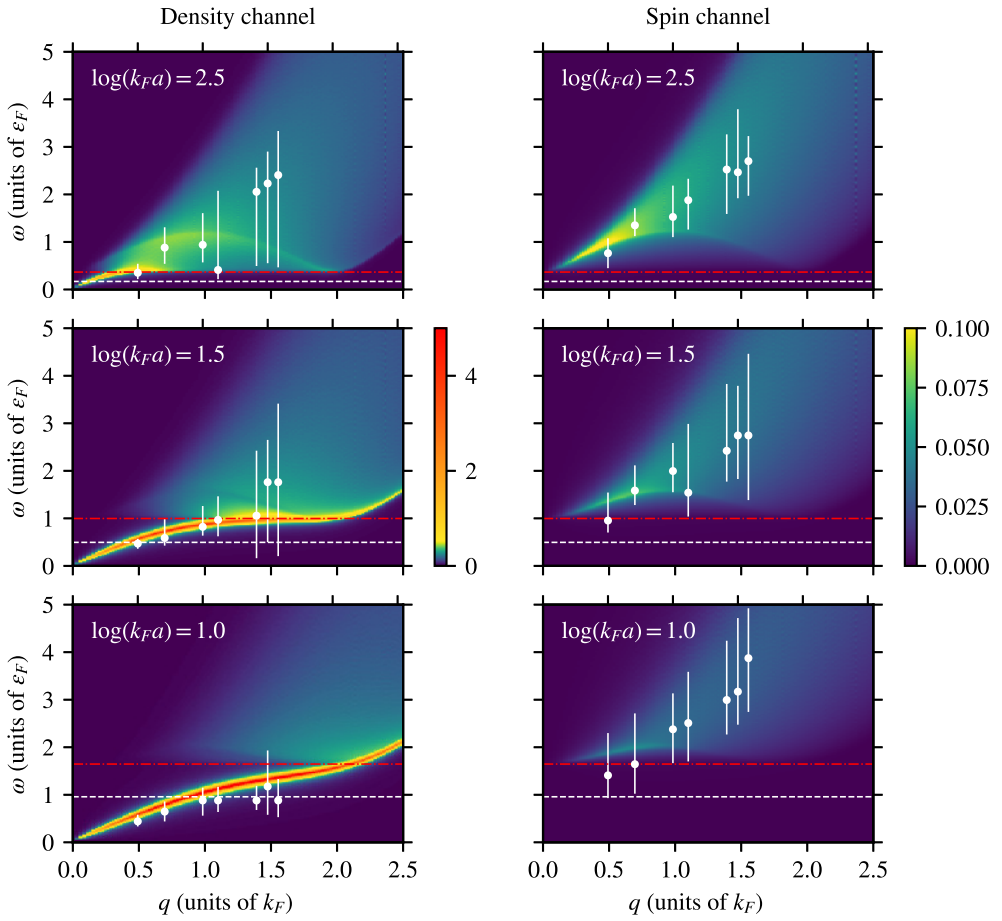

Figure 5.1: Dynamical structure factors (arbitrary units) in the density (left) and spin (right) channels for interaction strengths $\log(k_F a) = 2.5$ (top), 1.5 (middle), and 1.0 (bottom). Dots: maxima of the structure factors obtained by QMC for a system of 26 fermions, with vertical bars representing their full width at half maximum. Color plots: dynamical structure factors (arbitrary units) obtained by GRPA for a larger system of $N = 5850$ fermions at the same density. Dashed horizontal lines: twice the quasiparticle gap for the system of $N = 26$ fermions; for interaction strength $\log(k_F a) = 2.5$, it is calculated from the theory by Gor'kov and Melik-Barkhudarov [89, 90], which was shown to be accurate in this weak-interaction condition [88]; for interaction strengths $\log(k_F a) = 1.5$ and 1.0, it is calculated with QMC from Ref. [88]. Dash-dotted horizontal lines: twice the quasiparticle gap predicted by BCS theory for the system of $N = 5850$ fermions.

Density channel

First of all, we observe that our GRPA results displayed in Fig. 5.1, both in the density and in the spin channels, appear to be consistent with the GRPA results of Ref. [13], despite the different regularization employed. Focusing now on the density spectra, the left panels of Fig. 5.1 show that GRPA predicts a sharp Nambu-Goldstone mode for low energy $\hbar\omega < 2\Delta_{BCS}$. For higher energy $\hbar\omega > 2\Delta_{BCS}$, where the breaking of Cooper pairs is possible, a quasiparticle pair continuum emerges, in which the Nambu-

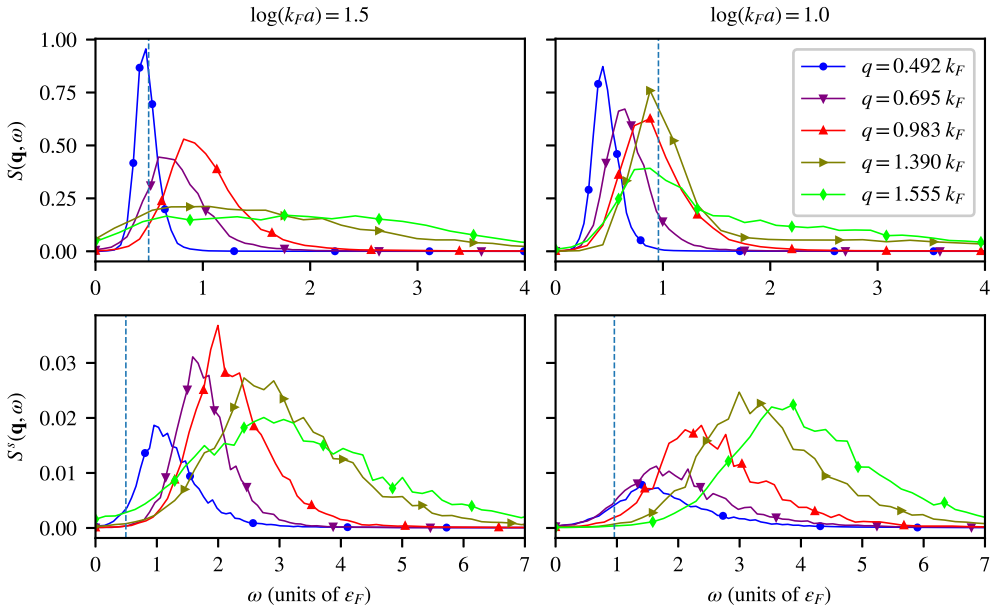

Figure 5.2: QMC results for the dynamical structure factors (arbitrary units) in the density (top panels) and spin (bottom panels) channels, for interaction strengths $\log(k_F a) = 1.5$ (left panels) and 1.0 (right panels), and various values of the wave number q. The solid lines represent our results, with markers indicating selected data points to improve readability. Dashed vertical lines: twice the quasiparticle gap, calculated with QMC from Ref. [88].

Goldstone mode is heavily damped.

Turning to the corresponding AFQMC results, we observe that the chosen regularization lattice supercell is large enough (and the accessible momenta are, accordingly, low enough) to explore the behavior for energies below the threshold of 2Δ. This is especially visible in the regimes with stronger interaction (and higher Δ), of which the upper panels of Fig. 5.2 give a more detailed view. The results confirm the sharp Nambu-Goldstone mode, but with a renormalized dispersion with respect to GRPA: the Goldstone peak calculated by AFQMC lies at a lower energy than the GRPA one at the same momentum, consistently with the smaller QMC superfluid gap Δ. As discussed in Subsec. 5.7.2, this difference is not a finite-size effect, so we conclude that it is due to the many-body correlations, that are not fully captured by an approximation like GRPA, but are well described by an exact approach like AFQMC. Going to higher momentum, the AFQMC results show that the Goldstone peak is gradually damped and the spectrum broadens, consistently with the appearance of a quasiparticle continuum above 2Δ with which the Goldstone mode eventually merges.

Up to now we compared the AFQMC results for $N = 26$ atoms with the GRPA results for $N = 5850$ atoms. The upper panels of Fig. 5.3 compare the QMC density spectra with the GRPA ones at the same size, $N = 26$, and interaction strength $\log(k_F a) = 1.5$. We only point out that, taking into account the renormalization of the superfluid gap, the two behaviors are qualitatively the same.

Regarding the possibility of detecting the Higgs mode in the density structure factor, in Figs. 5.7, 5.8 and 5.9 we show in more detail the GRPA peaks in the range $\hbar\omega > 2\Delta_{BCS}$,

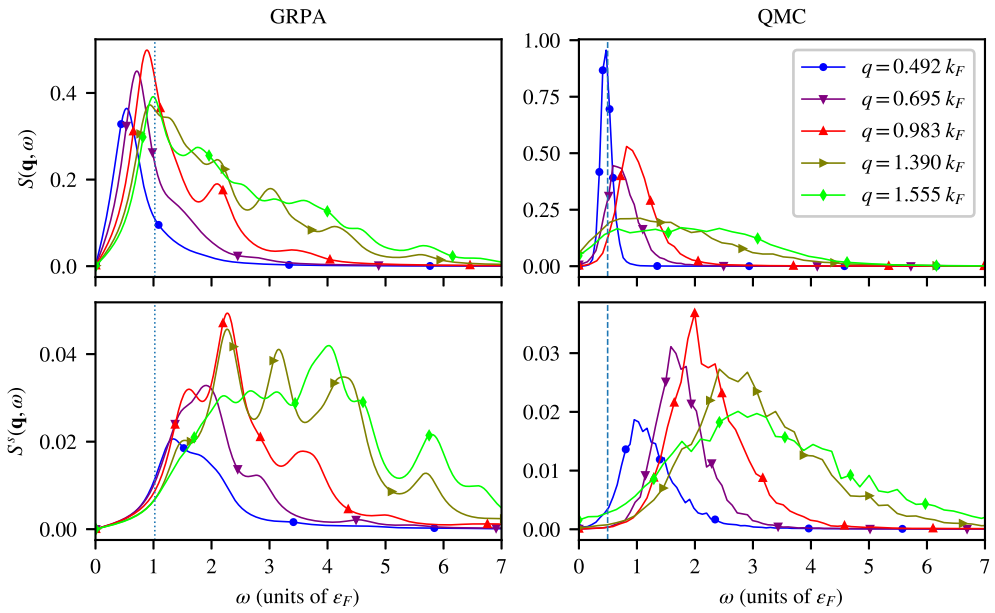

Figure 5.3: GRPA (left panels) and QMC (right panels) results for the dynamical structure factors (arbitrary units) in the density (top panels) and spin (bottom panels) channels, for a system of 26 atoms with interaction strength $\log(k_F a) = 1.5$ and various values of the wave number q. The solid lines represent our results, with markers indicating selected data points to improve readability. Dashed vertical lines: twice the quasiparticle gap, calculated with QMC from Ref. [88]. Dotted vertical lines: twice the quasiparticle gap, as predicted by BCS theory.

for interactions strengths $\log(k_F a) = 2.5$, 1.5, and 1.0, respectively. At the GRPA level we observe, as expected, a peak which can be interpreted as the Higgs mode, as we will discuss later. On the other hand, at the AFQMC level, the dominant Goldstone peak and the broadening due to the quasiparticle pair continuum make the requirements for the detection of the amplitude mode beyond the resolution we can currently achieve with AFQMC and analytic continuation.

Spin channel

Focusing on the spin spectra in the right panels of Fig. 5.1, the GRPA predicts no excitation for low energy $\hbar\omega < 2\Delta_{\mathrm{BCS}}$, as expected, while for $\hbar\omega > 2\Delta_{BCS}$ a sharp peak appears at low momentum and broadens at higher momenta. This is the peak that Zhao et al. [13] interpret as the celebrated Higgs mode, with a dispersion of the form $\omega_s(q) = 2\Delta_{\mathrm{BCS}} + \alpha q^2$ as $q/k_F \to 0$. We will discuss the interpretation of this peak later. Comparing different interaction strengths, we observe that at stronger interaction (i.e., lower $\log(k_F a)$) the spin spectrum gets flatter and fainter. This behavior is explained by the strongly-attractive fermions forming tightly bound pairs, or molecules, that constitute a Bose-Einstein condensate in which the spin degrees of freedom are suppressed. We also notice that the widths of the spin spectra increase with the interaction strength, and they do so faster in the AFQMC results than within GRPA.

Considering now the spin dynamical structure factor results from AFQMC, we are

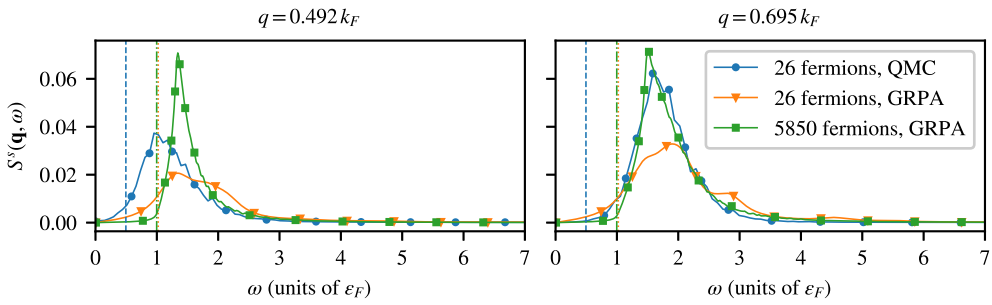

Figure 5.4: Spin dynamical structure factor $S^s(\mathbf{q},\omega)$ (arbitrary units) for interaction strength $\log(k_F a) = 1.5$ and wave numbers $q = 0.492\,k_F$ (left panel) and $0.695\,k_F$ (right panel). The solid lines represent our results, with markers indicating selected data points to improve readability. Blue dots: QMC results for 26 fermions. Orange triangles: GRPA results for 26 fermions. Green squares: GRPA results for 5850 fermions. Dashed vertical line: twice the quasiparticle gap for 26 fermions, calculated with QMC from Ref. [88]. Dotted vertical line: twice the quasiparticle gap predicted by BCS theory for 26 fermions. Dash-dotted vertical line: twice the quasiparticle gap predicted by BCS theory for 5850 fermions.

limited by the system size to momenta not smaller than $q \simeq 0.5 k_F$. The spectra for the lowest accessible momenta do exhibit a clear peak, of decreasing intensity with increasing interaction strength, as shown in the more detailed plots of Fig. 5.2 (lower panels). Its energy is compatible with a mode converging to 2Δ for $q \to 0$, where Δ is the AFQMC estimate of the superfluid gap. This peak is not very sharp, especially in comparison with the Nambu-Goldstone peak in the density channel. Nevertheless, as we argue in Subsec. 5.7.2, this broadening is probably a finite-size effect and we expect the peak to become sharp as the size of the system increases.

The lower panels of Fig. 5.3 compare the QMC spin spectra with the GRPA ones at the same size, $N = 26$, and interaction strength $\log(k_F a) = 1.5$. Like we already observed for the density spectra, the qualitative behavior is the same, with the energies shifted by the renormalization of the superfluid gap.

5.7.2 Discussion of the finite-size effects

We claimed that the low momentum peak visible in the AFQMC results for the spin channel may be broadened by finite-size effects, and that it probably becomes sharper for larger system sizes. To corroborate this statement, in Fig. 5.4 we show the spin dynamical structure factor $S^s(\mathbf{q},\omega)$ for interaction strength $\log(k_F a) = 1.5$, focusing on the two smallest wave numbers $q = 0.492\,k_F$ and $0.695\,k_F$. The AFQMC spectra are computed for 26 atoms, as always. The GRPA ones are calculated for 26 and 5850 atoms at the same density. All of them show a peak that moves to higher frequency as q increases from $0.492\,k_F$ to $0.695\,k_F$. The GRPA spectra exhibit maxima at approximately the same positions for the two system sizes, but the larger system shows much sharper peaks. The AFQMC results have a width which is compatible with the corresponding GRPA result for the same system size of $N = 26$ fermions, but the peaks are shifted to lower frequency, as we expect from the lower QMC superfluid gap. If we assume that the significant sharpening observed in the GRPA spectra for the larger system is not an artifact of the GRPA approach, we can suggest that the same would be observed also

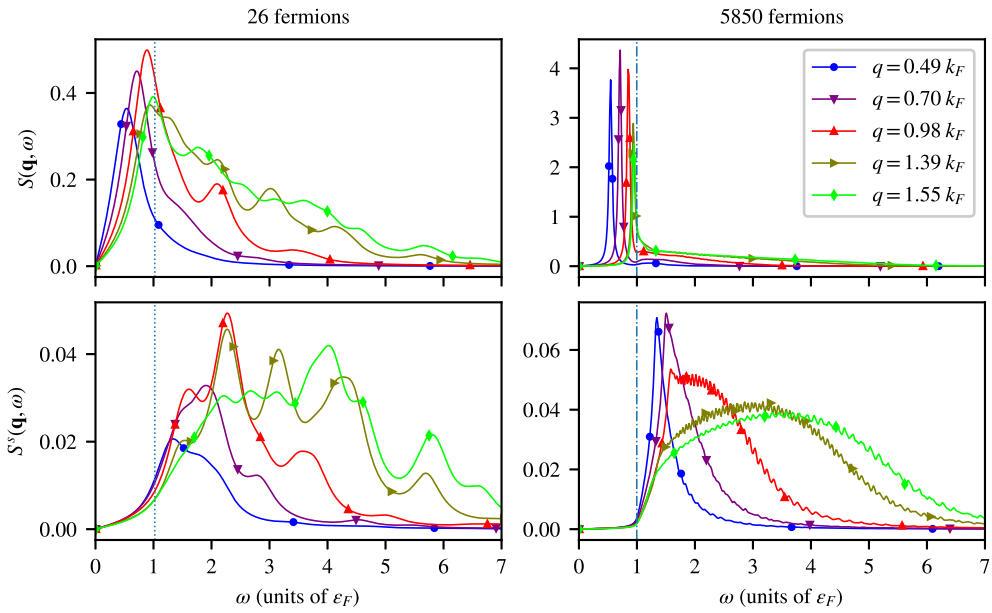

Figure 5.5: GRPA results for the dynamical structure factors (arbitrary units) in the density (top panels) and spin (bottom panels) channels of a system of 26 (left panels) and 5850 (right panels) fermions, for interaction strength $\log(k_F a) = 1.5$ and various values of the wave number q. The solid lines represent our results, with markers indicating selected data points to improve readability. Dotted vertical lines: twice the quasiparticle gap for 26 fermions, as predicted by BCS theory. Dash-dotted vertical lines: twice the quasiparticle gap for 5850 fermions, as predicted by BCS theory.

within AFQMC, if a much larger system could be investigated.

In Fig. 5.5, we take a closer look at the size dependence within GRPA, comparing the spectra for interaction $\log(k_F a) = 1.5$ and for the two different system sizes: $N = 26$ and $N = 5850$ atoms. The qualitative behavior is the same for the two sizes, and the peaks do not significantly change position. On the other hand, we again observe significantly sharper peaks in the larger system, for both the density and the spin channels. This makes the collective modes easier to tell apart from the quasiparticle pair continuum.

On one hand, this discussion implies that the dynamical structure factors of the 26-atoms system are still relevantly affected by finite-size effects, encouraging to invest efforts in advancing the AFQMC methodologies to access larger sizes; at the same time this suggests that, under the assumption that GRPA provides a qualitatively correct description of the size effects, a sharp peak does indeed exist at the QMC level in the spin dynamical structure factor.

5.7.3 Interpretation of the peak in the spin dynamical structure factor

In Subsections 5.7.1 and 5.7.2 we observed that, within the resolution of our calculations, the GRPA predictions are qualitatively confirmed by the AFQMC results. The main observable difference lies in the value of the superfluid gap, which is made significantly smaller by the many-body correlations that are better captured by the unbiased AFQMC

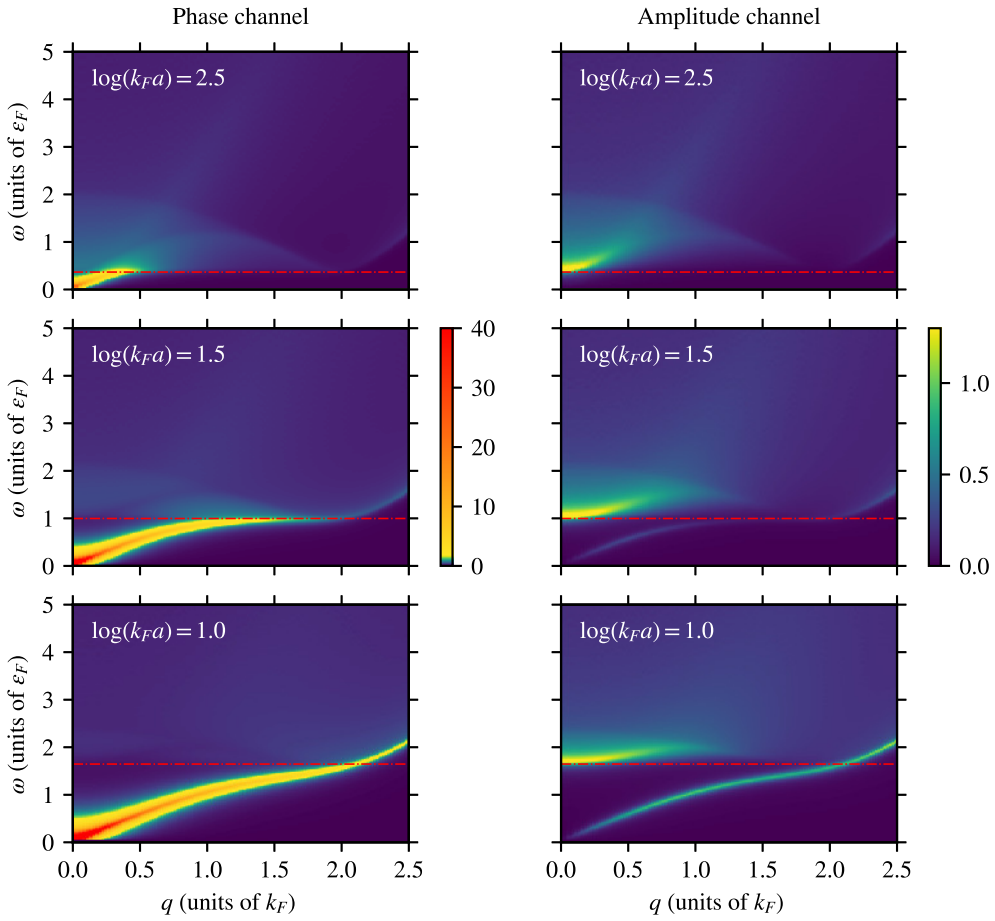

Figure 5.6: Dynamical structure factors (arbitrary units) in the order parameter's phase (left) and amplitude (right) channels for interaction strengths $\log(k_F a) = 2.5$ (top), 1.5 (middle), and 1.0 (bottom). The spectra are obtained by GRPA for a system of $N = 5850$ fermions. Dash-dotted horizontal lines: twice the quasiparticle gap predicted by BCS theory for the system of $N = 5850$ fermions.

approach. Importantly, we were able to reproduce within an exact QMC method the peak in the spin dynamical structure factor that was originally highlighted by Ref. [13]. In this Subsection, we deepen our GRPA study of the dynamical correlations to address the claim that such a peak is the Higgs, or amplitude, mode.

First of all, we remark that, within GRPA, there is no direct coupling between the dynamics of the order parameter and the spin excitations, as it is evident from Eq. 14 in Ref. [13] and the analysis in Ref. [91] (Eq. 120 and following). We explicitly test this fact by computing, within GRPA, the matrix $\chi_{\alpha,\beta}(\mathbf{q},\omega)$ of linear response functions, with labels $\alpha, \beta = n, s, p, a$, where n denotes particle density, s denotes spin density, p denotes the phase of the order parameter, and a denotes its amplitude. Within GRPA, we rigorously find that $\chi_{n,s}(\mathbf{q},\omega) = \chi_{p,s}(\mathbf{q},\omega) = \chi_{a,s}(\mathbf{q},\omega) = 0$. This means that a perturbation coupled with the spin density (a magnetic field) cannot excite modulations in the particle

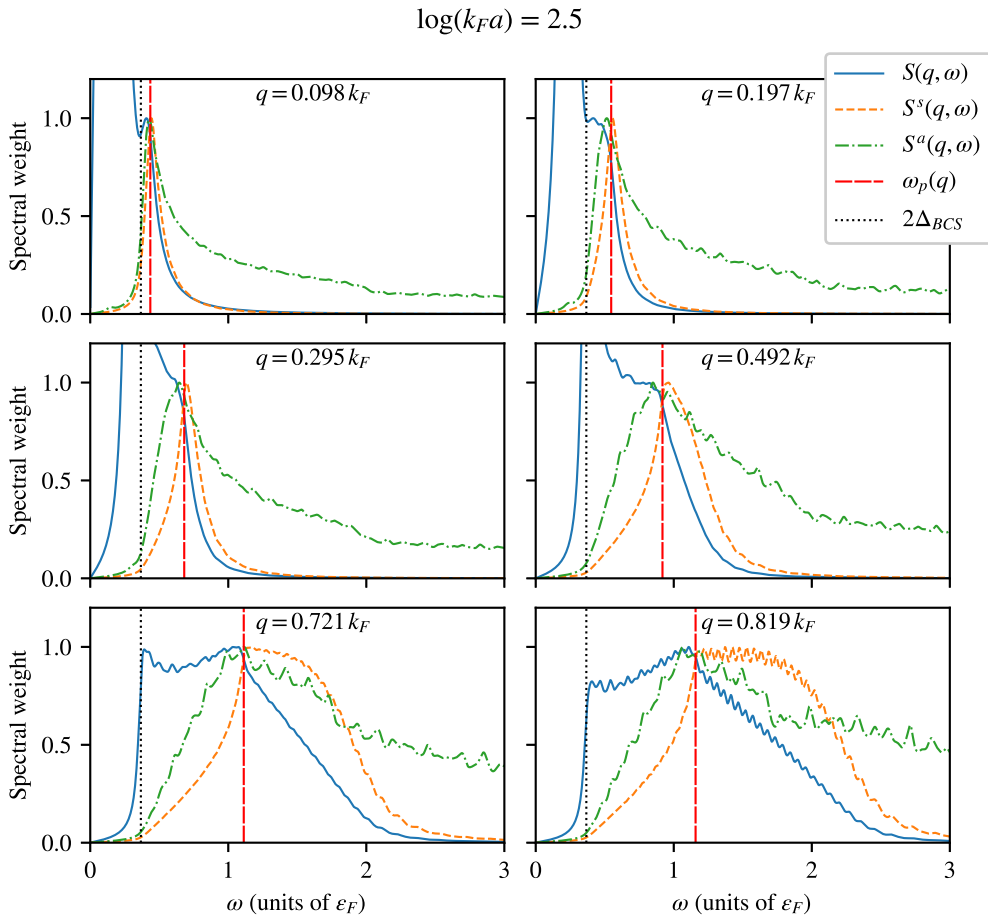

Figure 5.7: GRPA dynamical structure factors (arbitrary units) in the density (S^n), spin (S^s) and amplitude (S^a) channels, for 5850 fermions at interaction strength $\log(k_F a) = 2.5$. Each panel shows the results for a different wave number q. For ease of comparison, the dynamical structure factors were scaled so that their peaks in the range $\omega > 2\Delta_{BCS}$ have the same height, conventionally set to 1. The vertical dotted line marks twice the BCS quasiparticle gap. The vertical dashed line marks $\omega_p(q)$, the maximum of the density of pair quasiparticle states of total momentum q.

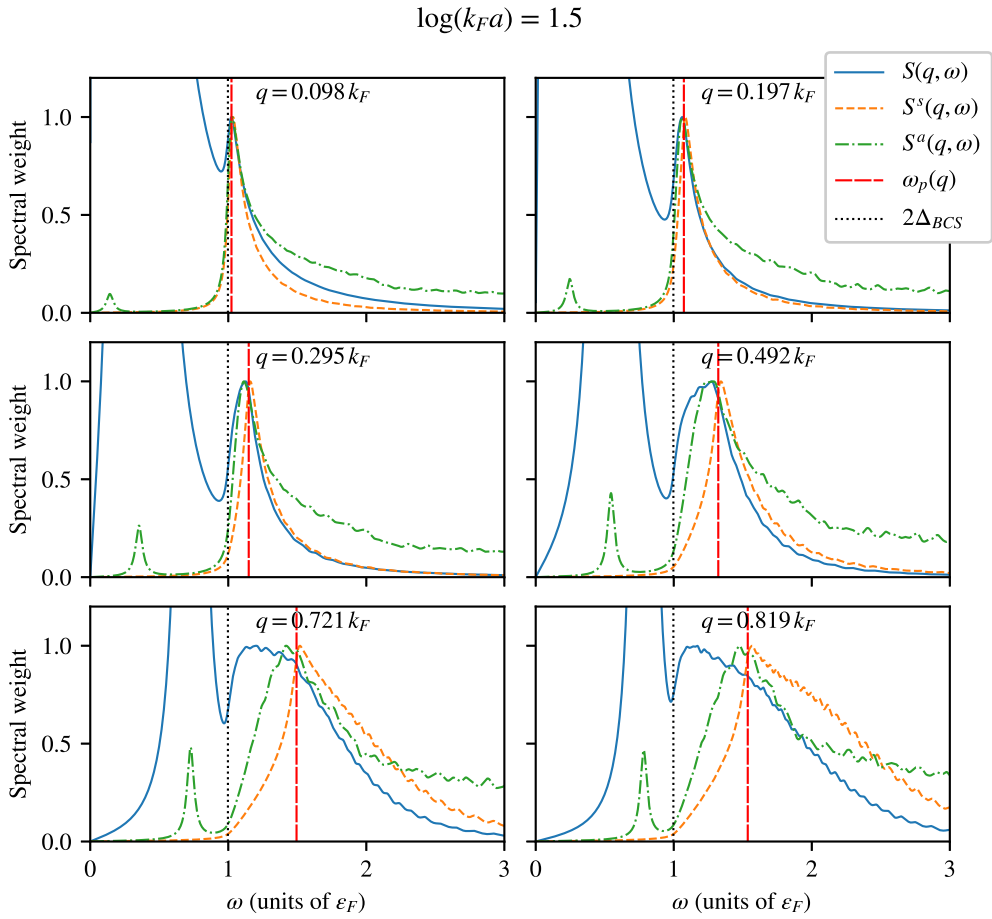

Figure 5.8: GRPA dynamical structure factors (arbitrary units) in the density (S^n), spin (S^s) and amplitude (S^a) channels, for 5850 fermions at interaction strength $\log(k_F a) = 1.5$. Each panel shows the results for a different wave number q. For ease of comparison, the dynamical structure factors were scaled so that their peaks in the range $\omega > 2\Delta_{BCS}$ have the same height, conventionally set to 1. The vertical dotted line marks twice the BCS quasiparticle gap. The vertical dashed line marks $\omega_p(q)$, the maximum of the density $\rho_p(\mathbf{q}, \omega)$ of pair quasiparticle states of total momentum q.

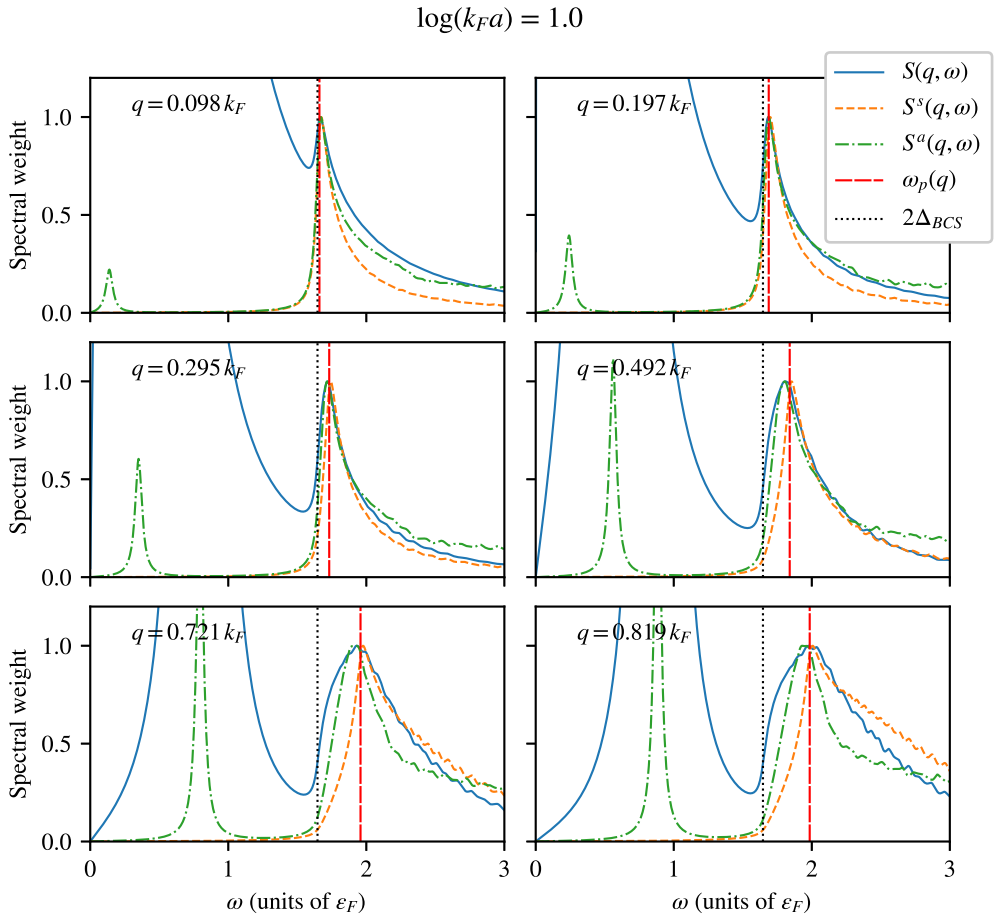

Figure 5.9: GRPA dynamical structure factors (arbitrary units) in the density (S^n), spin (S^s) and amplitude (S^a) channels, for 5850 fermions at interaction strength $\log(k_F a) = 1.0$. Each panel shows the results for a different wave number q. For ease of comparison, the dynamical structure factors were scaled so that their peaks in the range $\omega > 2\Delta_{BCS}$ have the same height, conventionally set to 1. The vertical dotted line marks twice the BCS quasiparticle gap. The vertical dashed line marks $\omega_p(q)$, the maximum of the density $\rho_p(\mathbf{q}, \omega)$ of pair quasiparticle states of total momentum q.

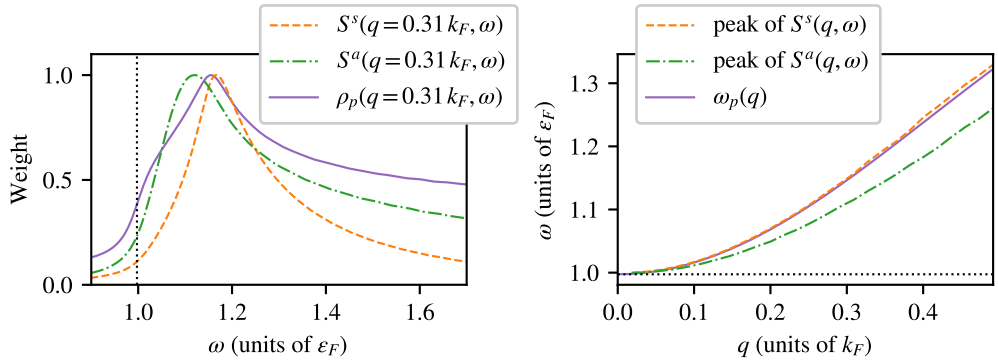

Figure 5.10: Left panel: GRPA dynamical structure factors (arbitrary units) in the density spin (S^s) and amplitude (S^a) channels, compared with the density $\rho_p(q,\omega)$ of pair quasiparticle states at momentum $q = 0.31\,k_F$, for 18954 fermions at interaction strength $\log(k_F a) = 1.5$; the vertical dotted line marks twice the BCS quasiparticle gap; for ease of comparison, the curves were scaled so that their peaks have the same height, conventionally set to 1. Right panel: Dispersions of the peaks in $S^s(q,\omega)$ (dashed line) and $S^a(q,\omega)$ (dash-dotted line) for 18954 atoms and interaction $\log(k_F a) = 1.5$; the solid line shows $\omega_p(q)$, the maximum of $\rho_p(\mathbf{q},\omega)$; the horizontal dotted line marks twice the BCS quasiparticle gap.

density or in the order parameter, at least at the GRPA level.

Keeping this in mind, the relations

$$\int_0^{+\infty} d\omega\, \chi''_{p,s}(-\mathbf{q},\omega) \propto \langle\Psi_{BCS}|\left(\hat{\Delta}(\mathbf{q}) - \hat{\Delta}^\dagger(-\mathbf{q})\right)\hat{n}^s(-\mathbf{q})|\Psi_{BCS}\rangle\,, \tag{5.85}$$

$$\int_0^{+\infty} d\omega\, \chi''_{a,s}(-\mathbf{q},\omega) \propto \langle\Psi_{BCS}|\left(\hat{\Delta}(\mathbf{q}) + \hat{\Delta}^\dagger(-\mathbf{q})\right)\hat{n}^s(-\mathbf{q})|\Psi_{BCS}\rangle \tag{5.86}$$

$$\tag{5.87}$$

(where, again, χ'' denotes the imaginary part of χ) imply that the quantum state $\hat{n}^s(-\mathbf{q})|\Psi_{BCS}\rangle$ ($|\Psi_{BCS}\rangle$ being the BCS ground state) must be orthogonal to the quantum states of the form $\hat{\Delta}^\dagger(\mathbf{q})|\Psi_{BCS}\rangle$ and $\hat{\Delta}(-\mathbf{q})|\Psi_{BCS}\rangle$; in simple words, if we induce a spin modulation on top of the BCS ground state we obtain a quantum state that has no overlap with states which are obtained by modulating the order parameter (Nambu-Goldstone and Higgs excitations).

To deepen our analysis, we compute the GRPA dynamical structure factors in the order parameter's phase and amplitude channel for the 5850-particles system. For the convergence parameter η, we empirically choose the value $\eta \simeq 0.06\,\varepsilon_F$. Fig. 5.6 shows the color plots for the resulting $S^p(\mathbf{q},\omega)$ (left panels) and $S^a(\mathbf{q},\omega)$ (right panels). The phase dynamical structure factor $S^p(\mathbf{q},\omega)$ is dominated by the Nambu-Goldstone mode, that presents the same properties we observed in the density channel. Besides the Goldstone, $S^p(\mathbf{q},\omega)$ also presents a faint signal above the threshold of $2\Delta_{BCS}$, linked to the quasiparticle excitations. The amplitude dynamical structure factor $S^a(\mathbf{q},\omega)$ exhibits a clear peak whose energy is $2\Delta_{BCS}$ as $q \to 0$ and increases quadratically with the momentum magnitude, eventually fading at high momenta. Because of the very definition of $S^a(\mathbf{q},\omega)$ as the spectrum of amplitude fluctuations, it is natural to identify this peak as the Higgs mode. Its dispersion becomes flatter at higher interaction strength (and lower $\log(k_F a)$). Moreover, in $S^a(\mathbf{q},\omega)$ for $\log(k_F a) = 1.5$ and 1.0 it can also be seen a signal

of the Goldstone mode, fainter than the Higgs, that becomes stronger as the interaction strength increases. This is coherent with the well-known fact that the phase-amplitude coupling, vanishing in the weak-interaction limit, becomes stronger at higher interaction strength.

The Figures 5.7, 5.8 and 5.9 show the comparison, for interaction strengths $\log(k_F a) = 2.5, 1.5$ and 1.0 respectively, of the density dynamical structure factor $S(\mathbf{q}, \omega)$, spin dynamical structure factor $S^s(\mathbf{q}, \omega)$, as well as the amplitude dynamical structure factor $S^a(\mathbf{q}, \omega)$. Each panel focuses on a different wave number, and the dynamical structure factors are scaled in order to make it easy to compare the spectra in the energy range $\hbar\omega > 2\Delta_{BCS}$. For $\log(k_F a) = 1.5$ and 1.0, for the smallest wave numbers shown, we observe that both the density (blue solid line) and the amplitude (green dashed-dotted line) dynamical correlations display the Goldstone mode below $2\Delta_{BCS}$ (black dotted vertical line) and higher-energy peaks at exactly the same energy above $2\Delta_{BCS}$. As we already discussed, we naturally identify this peak as the Higgs mode, that is thus visible also in the density channel, as it is known. This secondary peak in the density channel is also visible for $\log(k_F a) = 2.5$ at the smallest momenta shown, but its position is skewed by the presence of the Goldstone peak, which in this regime is very close and prominent. At all interaction strengths, as the magnitude of the wave vector increases, the secondary peak in the density channel broadens due to the quasiparticle pair continuum, to the point that for high momenta it cannot be identified with the Higgs any more. Also the spin dynamical structure factor (orange dashed line) shows a peak, as we have already extensively discussed. However, it is located at a slightly higher energy with respect to the Higgs, with a discrepancy increasing with the magnitude of the wave vector, as it is clear from the dispersions of the two peaks (amplitude and spin), shown in the right panel of Fig. 5.10 for an even larger system of 18954 atoms (always at the same density). Our conclusion is thus that, while a peak exists in the spin channel, it cannot be interpreted as a manifestation of the amplitude (Higgs) mode of the order parameter.

On the other hand, we propose a different origin for the spin "mode", kinematic in nature. When a probe acts on the system with a momentum $\mathbf{q}$ an energy $\hbar\omega$ within the quasiparticle pair continuum, we expect the formation of pairs of quasiparticles with momenta $-\mathbf{k}$ and $\mathbf{k} + \mathbf{q}$ for all possible values of $\mathbf{k}$. At the mean-field level, the energy of such a pair is $E(-\mathbf{k}) + E(\mathbf{k} + \mathbf{q})$, where $E(\mathbf{k}) = \sqrt{(\varepsilon(\mathbf{k}) - \mu_{BCS})^2 + \Delta_{BCS}^2}$ is the BCS dispersion. The (suitably η-regularized) function

$$\rho_p(\mathbf{q}, \omega) = \sum_{\mathbf{k}} \delta(\hbar\omega - E(-\mathbf{k}) - E(\mathbf{k} + \mathbf{q})) \tag{5.88}$$

counts the number of such quasiparticle pair states of momentum $\mathbf{q}$ and energy $\hbar\omega$. We denote by $\omega_p(\mathbf{q})$ the position of the maximum of $\rho_p(\mathbf{q}, \omega)$ for a given wave vector $\mathbf{q}$. In the Figures 5.7, 5.8 and 5.9 the red vertical dashed lines mark the positions of $\omega_p(q)$ for the chosen momenta, and we observe that the peak in the spin dynamical structure factor follows very closely the value of $\omega_p(q)$. The left panel of Fig. 5.10 compares the peaks of S^s, S^a, and ρ_p at momentum $q = 0.31\, k_F$ and interaction $\log(k_F a) = 1.5$ for a very large system of 18954 atoms, while the right panel displays the positions of the peaks of these three quantities as a function of the wave number q. Again, we observe that, while the positions of the spin and the amplitude peaks have a significant discrepancy, the position of the spin peak is very close to the position $\omega_p(q)$ of the peak of $\rho_p(q, \omega)$. This observation suggests that the spin dynamical structure factor is entirely governed by the quasiparticle pair continuum, which narrows at small momentum, giving rise to a peak in the response function.

We remark that, in order to resolve the difference between the spin and the amplitude peaks, we had to study a very large system, composed of 5850 fermions. For smaller systems, as we extensively discussed in Subsec. 5.7.2, finite-size effects broaden the structures (see Fig. 5.4), making it impossible to achieve the required resolution. So, it is not realistic, within our current capabilities, to investigate this discrepancy with AFQMC. While we cannot exclude *a priori* that many-body correlations beyond GRPA might produce a coupling between the amplitude and the spin channel, we notice that any such coupling, if it even exists, would be well below the resolution that is currently within the reach of AFQMC calculations.

Conclusions and future directions

The research work we presented in this Thesis is centered on the computational investigation of the dynamical and excitation properties of strongly-interacting many-body systems. We focused in particular on the time-dependent variational Monte Carlo (t-VMC) method, which generalizes the usual variational Monte Carlo (VMC) to perform simulations of the quantum dynamics using a parametrized wave function Ansatz.

We proposed a technique to compute the dynamical structure factor of a quantum system using the t-VMC method, that was developed independently but is substantially equivalent to the technique published in 2023 by T. Mendes-Santos, M. Schmitt, and M. Heyl [23]. Due to time constraints, we did not implement this method in our code yet, but an interesting future direction would be the application of this technique to different systems, other than the one considered in Ref. [23]. It would be especially interesting to apply it to continuous systems of particles, and investigate how it performs compared to the method based on a weak probe pulse described in Ref. [38].

The main novelty of our work is the extension of the t-VMC technique to shadow wave functions (SWF). This extension comes at the cost of having to perform the Monte Carlo sampling of two shadow configurations in addition to the real configuration. It also involves a reweighting procedure in the computation of the time-dependent expectation values, which can become problematic if the normalization $\langle \cos\theta \rangle(t)$ becomes close to zero. However, this allows a greater flexibility in the choice of the architecture of the wave function, that we exploit to introduce the very concise Baerswyl-Shadow wave function (BSWF), on the basis of the physical insight coming from the path-integral formalism applied to the transverse-field Ising (TFI) model.

We developed from scratch a C++ code to perform t-VMC simulations with various variational Ansatzë for lattice systems. Our imaginary-time simulations show that the BSWF achieves a description of the TFI ground state that is competitive with the translation-invariant RBMs, despite using only three parameters. The real-time simulations of a translation-invariant quench show that for short times the BSWF is very accurate, but it loses accuracy as the time progresses. We attribute this result to the lack of long-range correlation terms in the definition of the BSWF Ansatz. However, we find that the BSWF can significantly improve other Ansatzë, if used in conjunction with them as a multiplicative factor. The real-time simulation of the localized quantum quench shows that the non-translation-invariant BSWF achieves a quantitatively accurate description for short times, then it becomes less consistent with the exact evolution, but it still captures the qualitative features of the evolution.

We generalized the BSWF architecture and introduced the multilayer BSWF and the long-range BSWF. They both achieve a very accurate description of the TFI ground state,

outperforming the BSWF. The real-time evolutions of these Ansatzë are more consistent with the exact evolution, but suffer from the explosion of noise due to the reweighting procedure, in which the normalization $\langle \cos\theta \rangle$ becomes very close to zero. The problem is more severe in the multilayer BSWF than in the LRBSWF.

We also tested the two-dimensional (2D) BSWF in imaginary time for a small system that is tractable by exact diagonalization, and compared the results with the non-translation-invariant RBM with density 1. We found that it is very accurate, considering the small number of variational parameters. Further investigations are needed to explore the effectiveness of the BSWF in 2D as a multiplicative factor to improve other Ansatzë.

It would also be interesting to test our ideas on other models, other than the TFI. The promising results obtained for the ground state suggest that the BSWF or similarly-designed wave functions could describe accurately the ground state properties of spin systems that are difficult to tackle with other techniques. Further research could also address the problem of noise explosion in the real-time simulations: different reweighting strategies could mitigate this difficulty and allow better descriptions of the dynamics.

In a separate project, we computed the dynamical correlations of a two-dimensional dilute gas of attractive fermions, motivated by the suggestion in Ref. [13], based on generalized random phase approximation (GRPA), that the celebrated Higgs mode could be detectable in the spectrum of spin fluctuations. We used GRPA to compute the dynamical structure factors in the channels of the particle density, the spin density, and the order parameter's phase and amplitude. We checked the approximate density and spin spectra yielded by GRPA against the results obtained by the unbiased auxiliary-field quantum Monte Carlo (AFQMC) method, coupled with analytic continuation. The comparison of the AFQMC results with those yielded by GRPA for a system of 26 particles show good accordance if we take into account the renormalization of the superfluid gap Δ due to many-body correlations: this is related to a renormalization of the Nambu-Goldstone mode dispersion, and an energy shift in the spectrum structures that are located above 2Δ. At the lowest accessible wave vectors for this system size, we observe a maximum in the spin spectrum above 2Δ that is not very sharp. However, comparison of the GRPA results for 26 and 5850 fermions show that the peak is likely broadened by size effects. So, if we assume that the GRPA gives a qualitatively correct description of the size effects, we can conclude that the peak observed in the QMC spectrum is compatible with the sharp peak that the Authors of Ref. [13] observed and interpreted as a manifestation of the Higgs mode.

To test the correctness of this interpretation, we compared the GRPA dynamical structure factors for a system of 5850 particles in the channels of the density, the spin density, and the order parameter's amplitude. We found that the density and amplitude channels at low momentum exhibit a peak above 2Δ at the same energy, which we naturally interpret as the Higgs mode. The resolution needed to detect this secondary peak in the density dynamical structure factor is beyond our capability with AFQMC. The peak in the spin channel is found at a slightly higher energy, and this discrepancy increases with the momentum magnitude. Therefore, we conclude that the spin peak first observed in Ref. [13] cannot be identified, at least at the GRPA level, with the Higgs mode. On the other hand, we propose a kinematic origin of the spin peak: we show that it is located extremely close to a maximum in the density of quasiparticle pair states. This observation suggests that the spin dynamical structure factor is entirely governed by the quasiparticle pair continuum, which narrows at small momentum, giving rise to a peak in the response function.

This study thus contributes to the fundamental understanding of the dynamical properties of two-dimensional Fermi superfluids. Moreover, our unbiased QMC results, al-

beit affected by finite-size effects, can serve as a benchmark for many-body theories and can provide insight to experimental researchers in the challenge of measuring collective modes in superfluids.

Bibliography

[1] Jae yoon Choi, Sebastian Hild, Johannes Zeiher, Peter Schauß, Antonio Rubio-Abadal, Tarik Yefsah, Vedika Khemani, David A. Huse, Immanuel Bloch, and Christian Gross. Exploring the many-body localization transition in two dimensions. *Science*, 352(6293):1547–1552, 2016.

[2] Michael Schreiber, Sean S. Hodgman, Pranjal Bordia, Henrik P. Lüschen, Mark H. Fischer, Ronen Vosk, Ehud Altman, Ulrich Schneider, and Immanuel Bloch. Observation of many-body localization of interacting fermions in a quasirandom optical lattice. *Science*, 349(6250):842–845, 2015.

[3] J Smith, A Lee, P Richerme, B Neyenhuis, P W Hess, P Hauke, M Heyl, D A Huse, and C Monroe. Many-body localization in a quantum simulator with programmable random disorder. *Nat. Phys.*, 12(10):907–911, October 2016.

[4] P. Jurcevic, H. Shen, P. Hauke, C. Maier, T. Brydges, C. Hempel, B. P. Lanyon, M. Heyl, R. Blatt, and C. F. Roos. Direct Observation of Dynamical Quantum Phase Transitions in an Interacting Many-Body System. *Phys. Rev. Lett.*, 119:080501, Aug 2017.

[5] J Zhang, G Pagano, P W Hess, A Kyprianidis, P Becker, H Kaplan, A V Gorshkov, Z-X Gong, and C Monroe. Observation of a many-body dynamical phase transition with a 53-qubit quantum simulator. *Nature*, 551(7682):601–604, November 2017.

[6] N Fläschner, D Vogel, M Tarnowski, B S Rem, D-S Lühmann, M Heyl, J C Budich, L Mathey, K Sengstock, and C Weitenberg. Observation of dynamical vortices after quenches in a system with topology. *Nat. Phys.*, 14(3):265–268, March 2018.

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

[8] Giuseppe Carleo. *Spectral and dynamical properties of strongly correlated systems*. PhD thesis, SISSA, October 2011.

[9] Federico Becca and Sandro Sorella. *Time-Dependent Variational Monte Carlo*, page 156–163. Cambridge University Press, 2017.

[10] Silvio Vitiello, Karl Runge, and M. H. Kalos. Variational Calculations for Solid and Liquid ^{4}He with a "Shadow" Wave Function. *Phys. Rev. Lett.*, 60:1970–1972, May 1988.

[11] S. A. Vitiello, K. J. Runge, G. V. Chester, and M. H. Kalos. Shadow wave-function variational calculations of crystalline and liquid phases of ^{4}He. *Phys. Rev. B*, 42:228–239, Jul 1990.

[12] L. Reatto and G. L. Masserini. Shadow wave function for many-boson systems. *Phys. Rev. B*, 38:4516–4522, Sep 1988.

[13] Huaisong Zhao, Xiaoxu Gao, Wen Liang, Peng Zou, and Feng Yuan. Dynamical structure factors of a two-dimensional Fermi superfluid within random phase approximation. *New J. Phys.*, 22(9):093012, sep 2020.

[14] R. Ganesh, A. Paramekanti, and A. A. Burkov. Collective modes and superflow instabilities of strongly correlated Fermi superfluids. *Phys. Rev. A*, 80:043612, Oct 2009.

[15] Patrick Kelly and Ettore Vitali. On the Accuracy of Random Phase Approximation for Dynamical Structure Factors in Cold Atomic Gases. *Atoms*, 9(4), 2021.

[16] Yoshihiro Yunomae, Daisuke Yamamoto, Ippei Danshita, Nobuhiko Yokoshi, and Shunji Tsuchiya. Instability of superfluid Fermi gases induced by a rotonlike density mode in optical lattices. *Phys. Rev. A*, 80:063627, Dec 2009.

[17] R. Combescot, M. Yu. Kagan, and S. Stringari. Collective mode of homogeneous superfluid Fermi gases in the BEC-BCS crossover. *Phys. Rev. A*, 74:042717, Oct 2006.

[18] S. Zhang. Auxiliary-Field Quantum Monte Carlo for Correlated Electron Systems. In E. Pavarini, E. Koch, and U. Schollwöck, editors, *Emergent Phenomena in Correlated Matter: Modeling and Simulation Vol. 3*. Verlag des Forschungszentrum Jülich, Jülich, Germany, 2013.

[19] Ettore Vitali, Hao Shi, Mingpu Qin, and Shiwei Zhang. Computation of dynamical correlation functions for many-fermion systems with auxiliary-field quantum Monte Carlo. *Phys. Rev. B*, 94:085140, Aug 2016.

[20] M. Motta, D. E. Galli, S. Moroni, and E. Vitali. Imaginary time correlations and the phaseless auxiliary field quantum Monte Carlo. *J. Chem. Phys.*, 140(2):024107, 01 2014.

[21] Nathan S. Nichols, Paul Sokol, and Adrian Del Maestro. Parameter-free differential evolution algorithm for the analytic continuation of imaginary time correlation functions. *Phys. Rev. E*, 106:025312, Aug 2022.

[22] Nathan S. Nichols. DEAC paper code repository, 2021. `https://github.com/DelMaestroGroup/papers-code-DEAC`.

[23] Tiago Mendes-Santos, Markus Schmitt, and Markus Heyl. Highly Resolved Spectral Functions of Two-Dimensional Systems with Neural Quantum States. *Phys. Rev. Lett.*, 131:046501, Jul 2023.

[24] Giuseppe Carleo, Federico Becca, Marco Schiró, and Michele Fabrizio. Localization and Glassy Dynamics Of Many-Body Quantum Systems. *Scientific Reports*, 2(1):243, February 2012.

[25] A.D. McLachlan. A variational solution of the time-dependent Schrodinger equation. *Molecular Physics*, 8(1):39–44, January 1964.

[26] Mathias Gartner. Monte Carlo simulations of non-equilibrium dynamics in bosonic many-body systems. Master's thesis, Johannes Kepler Universität Linz, 2023.

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

[28] Rocco Martinazzo and Irene Burghardt. Local-in-Time Error in Variational Quantum Dynamics. *Phys. Rev. Lett.*, 124:150601, Apr 2020.

[29] Christian Lubich. *From Quantum to Classical Molecular Dynamics: Reduced Models and Numerical Analysis*. European Mathematical Society, September 2008.

[30] Federico Becca and Sandro Sorella. *Optimization of Variational Wave Functions*, page 131–155. Cambridge University Press, 2017.

[31] Peter Kramer and Marcos Saraceno. *Geometry of the Time-Dependent Variational Principle in Quantum Mechanics*, volume 140 of *Lecture Notes in Physics*. Springer, Berlin,

Heidelberg, 1981.

[32] Damian Hofmann, Giammarco Fabiani, Johan H. Mentink, Giuseppe Carleo, and Michael A. Sentef. Role of stochastic noise and generalization error in the time propagation of neural-network quantum states. *SciPost Phys.*, 12:165, 2022.

[33] Markus Schmitt and Markus Heyl. Quantum Many-Body Dynamics in Two Dimensions with Artificial Neural Networks. *Physical Review Letters*, 125(10):100503, September 2020.

[34] Tao Pang. Spin Jastrow state. *Phys. Rev. B*, 43:3362–3365, Feb 1991.

[35] R. P. Feynman and Michael Cohen. Energy Spectrum of the Excitations in Liquid Helium. *Phys. Rev.*, 102:1189–1204, Jun 1956.

[36] Yongkyung Kwon, D. M. Ceperley, and Richard M. Martin. Effects of backflow correlation in the three-dimensional electron gas: Quantum Monte Carlo study. *Phys. Rev. B*, 58:6800–6806, Sep 1998.

[37] Yongkyung Kwon, D. M. Ceperley, and Richard M. Martin. Effects of three-body and backflow correlations in the two-dimensional electron gas. *Phys. Rev. B*, 48:12037–12046, Oct 1993.

[38] Mathias Gartner, Ferran Mazzanti, and Robert E. Zillich. Time-dependent variational Monte Carlo study of the dynamic response of bosons in an optical lattice. *SciPost Phys.*, 13:025, 2022.

[39] D. Baeriswyl. Variational schemes for many-electron systems. In Alan R. Bishop, David K. Campbell, Pradeep Kumar, and Steven E. Trullinger, editors, *Nonlinearity in Condensed Matter*, pages 183–193, Berlin, Heidelberg, 1987. Springer Berlin Heidelberg.

[40] Michael Dzierzawa, Dionys Baeriswyl, and Michela Di Stasio. Variational wave functions for the Mott transition: The 1/r Hubbard chain. *Phys. Rev. B*, 51:1993–1996, Jan 1995.

[41] B. Valenzuela, S. Fratini, and D. Baeriswyl. Charge and spin order in one-dimensional electron systems with long-range Coulomb interactions. *Phys. Rev. B*, 68:045112, Jul 2003.

[42] Michele Ruggeri, Saverio Moroni, and Markus Holzmann. Nonlinear Network Description for Many-Body Quantum Systems in Continuous Space. *Phys. Rev. Lett.*, 120:205302, May 2018.

[43] Richard P. Fynman and Albert R. Hibbs. *Quantum Mechanics and Path Integrals*. McGraw-Hill, New York, 1965.

[44] Francesco Calcavecchia, Francesco Pederiva, Malvin H. Kalos, and Thomas D. Kühne. Sign problem of the fermionic shadow wave function. *Phys. Rev. E*, 90:053304, Nov 2014.

[45] Pierre Pfeuty. The one-dimensional Ising model with a transverse field. *Annals of Physics*, 57(1):79–90, 1970.

[46] Glen Bigan Mbeng, Angelo Russomanno, and Giuseppe E. Santoro. The quantum Ising chain for beginners, 2020.

[47] E. M. Inack, G. E. Santoro, L. Dell'Anna, and S. Pilati. Projective quantum monte carlo simulations guided by unrestricted neural network states. *Phys. Rev. B*, 98:235145, Dec 2018.

[48] S. Pilati, E. M. Inack, and P. Pieri. Self-learning projective quantum Monte Carlo simulations guided by restricted Boltzmann machines. *Phys. Rev. E*, 100:043301, Oct 2019.

[49] Ulli Wolff. Collective Monte Carlo Updating for Spin Systems. *Phys. Rev. Lett.*,

62:361–364, Jan 1989.

[50] Mark E. J. Newman and Gerard T. Barkema. *Monte Carlo Methods in Statistical Physics*, chapter 4. Oxford University Press, 1999.

[51] S F Edwards and P W Anderson. Theory of spin glasses. *Journal of Physics F: Metal Physics*, 5(5):965, may 1975.

[52] Vl.S. Dotsenko, W. Selke, and A.L. Talapov. Cluster Monte Carlo algorithms for random Ising models. *Physica A: Statistical Mechanics and its Applications*, 170(2):278–281, 1991.

[53] Mark E. J. Newman and Gerard T. Barkema. *Monte Carlo Methods in Statistical Physics*, chapter 6. Oxford University Press, 1999.

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

[55] David Pekker and C.M. Varma. Amplitude/Higgs modes in condensed matter physics. *Annual Review of Condensed Matter Physics*, 6(1):269–297, March 2015.

[56] Ryo Shimano and Naoto Tsuji. Higgs Mode in Superconductors. *Annual Review of Condensed Matter Physics*, 11(1):103–124, March 2020.

[57] Ryo Shimano and Naoto Tsuji. Higgs Mode in Superconductors. *Annu. Rev. Condens. Matter Phys.*, 11(1):103–124, March 2020.

[58] David Pekker and C.M. Varma. Amplitude/Higgs Modes in Condensed Matter Physics. *Annu. Rev. Condens. Matter Phys.*, 6(1):269–297, 2015.

[59] Ulf Bissbort, Sören Götze, Yongqiang Li, Jannes Heinze, Jasper S. Krauser, Malte Weinberg, Christoph Becker, Klaus Sengstock, and Walter Hofstetter. Detecting the Amplitude Mode of Strongly Interacting Lattice Bosons by Bragg Scattering. *Phys. Rev. Lett.*, 106(20):205303, May 2011.

[60] Manuel Endres, Takeshi Fukuhara, David Pekker, Marc Cheneau, Peter Schauβ, Christian Gross, Eugene Demler, Stefan Kuhr, and Immanuel Bloch. The Higgs amplitude mode at the two-dimensional superfluid/Mott insulator transition. *Nature*, 487(7408):454–458, July 2012.

[61] Julian Léonard, Andrea Morales, Philip Zupancic, Tobias Donner, and Tilman Esslinger. Monitoring and manipulating Higgs and Goldstone modes in a supersolid quantum gas. *Science*, 358(6369):1415–1418, December 2017.

[62] T. Cea, C. Castellani, G. Seibold, and L. Benfatto. Nonrelativistic Dynamics of the Amplitude (Higgs) Mode in Superconductors. *Phys. Rev. Lett.*, 115(15):157002, October 2015.

[63] H. Kurkjian, J. Tempere, and S. N. Klimin. Linear response of a superfluid Fermi gas inside its pair-breaking continuum. *Sci. Rep.*, 10(1):11591, July 2020.

[64] A. Behrle, T. Harrison, J. Kombe, K. Gao, M. Link, J.-S. Bernier, C. Kollath, and M. Köhl. Higgs Mode in a Strongly Interacting Fermionic Superfluid. *Nat. Phys.*, 14(8):781–785, August 2018.

[65] Luca Bayha, Marvin Holten, Ralf Klemt, Keerthan Subramanian, Johannes Bjerlin, Stephanie M. Reimann, Georg M. Bruun, Philipp M. Preiss, and Selim Jochim. Observing the Emergence of a Quantum Phase Transition Shell by Shell. *Nature*, 587(7835):583–587, November 2020.

[66] P. Dyke, S. Musolino, H. Kurkjian, D. J. M. Ahmed-Braun, A. Pennings, I. Herrera, S. Hoinka, S. J. J. M. F. Kokkelmans, V. E. Colussi, and C. J. Vale. Higgs oscillations in a unitary Fermi superfluid, 2023.

[67] Andreas Kell, Moritz Breyer, Daniel Eberz, and Michael Köhl. Exciting the Higgs mode in a strongly-interacting Fermi gas by interaction modulation, 2024.

[68] R. G. Scott, F. Dalfovo, L. P. Pitaevskii, and S. Stringari. Rapid ramps across the BEC-BCS crossover: A route to measuring the superfluid gap. *Phys. Rev. A*, 86:053604, Nov 2012.

[69] S. Hannibal, P. Kettmann, M. D. Croitoru, A. Vagov, V. M. Axt, and T. Kuhn. Quench dynamics of an ultracold Fermi gas in the BCS regime: Spectral properties and confinement-induced breakdown of the Higgs mode. *Phys. Rev. A*, 91:043630, Apr 2015.

[70] Jun Tokimoto, Shunji Tsuchiya, and Tetsuro Nikuni. Excitation of Higgs Mode in Superfluid Fermi Gas in BCS–BEC Crossover. *J. Phys. Soc. Jpn.*, 88(2):023601, February 2019.

[71] Guitao Lyu, Kui-Tian Xi, Sukjin Yoon, Qijin Chen, and Gentaro Watanabe. Exciting the long-lived Higgs mode in superfluid Fermi gases with particle removal. *Phys. Rev. A*, 107(2):023321, February 2023.

[72] Andrea Barresi, Antoine Boulet, Gabriel Wlazłowski, and Piotr Magierski. Generation and decay of Higgs mode in a strongly interacting Fermi gas. *Sci. Rep.*, 13(1):11285, July 2023.

[73] Dan Phan and Andrey V. Chubukov. Following the Higgs mode across the BCS-BEC crossover in two dimensions. *Phys. Rev. B*, 107(13):134519, April 2023.

[74] Xinloong Han, Boyang Liu, and Jiangping Hu. Observability of Higgs mode in a system without Lorentz invariance. *Phys. Rev. A*, 94(3):033608, September 2016.

[75] J. Bjerlin, S. M. Reimann, and G. M. Bruun. Few-Body Precursor of the Higgs Mode in a Fermi Gas. *Phys. Rev. Lett.*, 116(15):155302, April 2016.

[76] S. Hoinka, M. Lingham, M. Delehaye, and C. J. Vale. Dynamic Spin Response of a Strongly Interacting Fermi Gas. *Phys. Rev. Lett.*, 109(5):050403, August 2012.

[77] Chris J. Vale and Martin Zwierlein. Spectroscopic Probes of Quantum Gases. *Nat. Phys.*, 17(12):1305–1315, December 2021.

[78] Sascha Hoinka, Paul Dyke, Marcus G. Lingham, Jami J. Kinnunen, Georg M. Bruun, and Chris J. Vale. Goldstone mode and pair-breaking excitations in atomic Fermi superfluids. *Nat. Phys.*, 13(10):943–946, October 2017.

[79] Hauke Biss, Lennart Sobirey, Niclas Luick, Markus Bohlen, Jami J. Kinnunen, Georg M. Bruun, Thomas Lompe, and Henning Moritz. Excitation Spectrum and Superfluid Gap of an Ultracold Fermi Gas. *Phys. Rev. Lett.*, 128(10):100401, March 2022.

[80] Lennart Sobirey, Hauke Biss, Niclas Luick, Markus Bohlen, Henning Moritz, and Thomas Lompe. Observing the Influence of Reduced Dimensionality on Fermionic Superfluids. *Phys. Rev. Lett.*, 129(8):083601, August 2022.

[81] Gianluca Bertaina. Two-dimensional short-range interacting attractive and repulsive Fermi gases at zero temperature. *Eur. Phys. J. Spec. Top.*, 217(1):153–162, March 2013.

[82] Hao Shi, Simone Chiesa, and Shiwei Zhang. Ground-state properties of strongly interacting Fermi gases in two dimensions. *Phys. Rev. A*, 92:033603, Sep 2015.

[83] Félix Werner and Yvan Castin. General relations for quantum gases in two and three dimensions: Two-component fermions. *Phys. Rev. A*, 86:013626, Jul 2012.

[84] E. Vitali, M. Rossi, L. Reatto, and D. E. Galli. Ab initio low-energy dynamics of superfluid and solid ^{4}He. *Phys. Rev. B*, 82(17):174510, November 2010.

[85] G. Bertaina, D. E. Galli, and E. Vitali. Statistical and computational intelligence approach to analytic continuation in Quantum Monte Carlo. *Adv. Phys.: X*, 2(2):302–323, March 2017.

[86] David Pines and Chia-Wei Woo. Sum Rules, Structure Factors, and Phonon Dispersion in Liquid He^4 at Long Wavelengths and Low Temperatures. *Phys. Rev. Lett.*, 24:1044–1045, May 1970.

[87] Allen Miller, David Pines, and Philippe Nozières. Elementary Excitations in Liquid Helium. *Phys. Rev.*, 127:1452–1464, Sep 1962.

[88] Ettore Vitali, Hao Shi, Mingpu Qin, and Shiwei Zhang. Visualizing the BEC-BCS crossover in a two-dimensional Fermi gas: Pairing gaps and dynamical response functions from ab initio computations. *Phys. Rev. A*, 96:061601(R), Dec 2017.

[89] D. S. Petrov, M. A. Baranov, and G. V. Shlyapnikov. Superfluid transition in quasi-two-dimensional Fermi gases. *Phys. Rev. A*, 67:031601(R), Mar 2003.

[90] L. P. Gor'kov and T. K. Melik-Barkhudarov. Contribution to the theory of superfluidity in an imperfect Fermi gas. *Sov. Phys. JETP*, 13:1018, November 1961.

[91] Lianyi He. Dynamic density and spin responses of a superfluid Fermi gas in the BCS–BEC crossover: Path integral formulation and pair fluctuation theory. *Ann. Phys.*, 373:470–511, 2016.

Acknowledgments

First and foremost, I would like to express my deepest gratitude to my Supervisor, Prof. Davide E. Galli, for his unwavering support, guidance, and encouragement throughout my doctoral studies. Always ready to share his expertise and insight, he has been both a competent guide and a friend.

My heartfelt thanks go also to my Co-Supervisor, Dr. Gianluca Bertaina, for his scrupulous scientific pursuit. His lucid and sharp observations have enriched my work and taught me more than he may know.

I would also like to acknowledge my institution, the Università degli Studi di Milano, for providing a welcoming environment in which to grow and for the high-performance computing resources through the INDACO platform.

I am incredibly grateful to the students that I co-advised: Davide Piccioni, Marco Tesoro, Francesco Tognocchi, and Giorgia Iori. Their enthusiasm, dedication, and skills have been invaluable. I also want to mention the new entry, Raffaele Salioni; I look forward to the work we will be doing together.

I cannot thank enough my "LCP mates," Jacopo D'Alberto and Michele Caresana. Thank you, Jacopo, for the research we did together, and thank you, Michele, for putting up with both Jacopo and me. Without your company and friendship, these years would not have been the same.

I would also like to thank Prof. Ettore Vitali for being a valuable guide in the world of Fermi superfluidity. Additionally, the work he did with his students Patrick Kelly, Annette Lopez, and Kaelyn Dauer has been essential.

I also want to thank some mentors whose help is not directly related to the work presented in this Thesis but was priceless during my doctoral studies.

I am deeply thankful to Prof. Sandro Sorella for his enthusiastic teaching and invaluable expertise that he shared with me during our short collaboration. Though he is no longer with us, his legacy continues to inspire and guide young researchers.

I also wish to sincerely thank Prof. Giuseppe Santoro for his assistance during a particularly challenging time. When Davide Piccioni and I were feeling disoriented and uncertain, his guidance and support provided us with the clarity and direction we needed. During that period, the help of Prof. Federico Becca, Dr. Guglielmo Mazzola, and Prof. Alberto Parola was also inestimable.

My heartfelt thanks go to my family for their endless support and encouragement. To my parents, Fabio and Cinzia, thank you for always believing in me and for your unconditional love and support, even when I would not have put up with myself. To my brother Federico and my sister Aurora, thank you for being a source of joy and inspiration.

Thank you all for helping me reach this milestone. This accomplishment would not have been possible without your support and friendship.