

Variational Monte Carlo for the Hubbard Model

*Original*

Variational Monte Carlo for the Hubbard Model / Tocchio, L.F. - In: Correlated Materials: Methods and Applications / Eva Pavarini and Erick Koch. - STAMPA. - [s.l.] : Forschungszentrum Juelich GmbH, 2026. - ISBN 978-3-95806-908-4. - pp. 13.1-13.27

*Availability:*

This version is available at: 11583/3016150 since: 2026-09-30T10:34:55Z

*Publisher:*

Forschungszentrum Juelich GmbH

*Published*

DOI:

*Terms of use:*

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

*Publisher copyright*

(Article begins on next page)

# 13 Variational Monte Carlo for the Hubbard Model

Luca F. Tocchio  
Politecnico di Torino  
Corso Duca degli Abruzzi, 24 10129-Turin, Italy

## Contents

|          |                                                                                   |           |
|----------|-----------------------------------------------------------------------------------|-----------|
| <b>1</b> | <b>Introduction</b>                                                               | <b>2</b>  |
| <b>2</b> | <b>The Hubbard model</b>                                                          | <b>3</b>  |
| 2.1      | The single-orbital Hubbard model . . . . .                                        | 3         |
| 2.2      | The multi-orbital Hubbard model . . . . .                                         | 3         |
| <b>3</b> | <b>Variational Monte Carlo</b>                                                    | <b>4</b>  |
| 3.1      | The Metropolis algorithm . . . . .                                                | 6         |
| <b>4</b> | <b>Variational wave functions</b>                                                 | <b>8</b>  |
| 4.1      | Jastrow-Slater wave functions for the single-orbital Hubbard model . . . . .      | 8         |
| 4.2      | Computing ratios for the Jastrow-Slater wave function . . . . .                   | 10        |
| 4.3      | The backflow terms . . . . .                                                      | 13        |
| 4.4      | Jastrow-Slater wave functions for the multi-orbital Hubbard model . . . . .       | 16        |
| <b>5</b> | <b>Optimizing the variational state: Stochastic reconfiguration</b>               | <b>17</b> |
| <b>6</b> | <b>Correlation functions in variational Monte Carlo: Definitions and examples</b> | <b>19</b> |
| <b>7</b> | <b>Final remarks</b>                                                              | <b>24</b> |

# 1 Introduction

One of the reasons why the study of correlated Hamiltonians in two dimensions is one of the most active and debated fields in modern condensed matter physics is the absence of a clear-cut method for tackling non-exactly solvable Hamiltonians, like the Hubbard model. Variational Monte Carlo (VMC) is one of the few methods that allows us to address Hubbard models, including its multi-orbital extensions, directly in two dimensions, describing non-local correlations and ordering, including non-local superconductivity [1]. We remark that VMC is a nonperturbative approach for which performance is not significantly affected by the dimensionality of the system or by the size of the local Hilbert space. In this respect, VMC might be more efficient than other methods, like density-matrix renormalization group or tensor-network methods, in which the complexity increases with both the range of the interaction and the dimension of the local Hilbert space [2]. It can also provide complementary information with respect to methods like dynamical mean-field theory [3], which can capture accurate dynamical local correlations, but does not include nonlocal spatial correlations beyond mean field.

The VMC methods rely on two main ingredients. The first one is the construction of a variational wave function that approximates the correct ground-state. This approach relies on the variational principle, stating that the energy of a given quantum state is always bounded from below by the exact ground-state energy, thus giving us guidance to obtain the best solution to the problem. Moreover, a suitable parametrization of the ground state allows us to consider a wide range of different quantum phases (superconductivity, magnetic and charge order, Mott insulators...). By contrast, the main limitation of our approach is the fact that it is based on a given Ansatz, which may contain a relevant bias.

The second one is the Monte Carlo sampling: Computing expectation values of observables over the variational estimate of the ground-state wave function requires a numerical Monte Carlo sampling since the wave function is correlated. Moreover, the variational state contains many parameters to be optimized at the same time. This is a highly non-trivial and computationally demanding task that is made possible thanks to the Stochastic Reconfiguration method, which allows us to find the optimal set for hundreds of independent parameters [4]. Once the optimal variational state is obtained, the presence of, for instance, spin order, charge order, and superconductivity can be detected with the computation of appropriate correlation functions. These physical observables are computed on finite lattice sizes with periodic boundary conditions and might require size scaling to the thermodynamic limit.

In Sec. 2, we briefly introduce the single-orbital Hubbard model and its multi-orbital extension. In this lecture, we will focus only on electronic models, even though the Hubbard model can also be extended to bosonic particles. In Sec. 3, we introduce the general scheme under which variational Monte Carlo works. Then, in Sec. 4, we present the variational wave functions that can properly describe superconductive and magnetic states in single-orbital and multi-orbital Hubbard models. We also detail how to perform Monte Carlo calculations for the given wave functions. In Sec. 5, we describe how to optimize a variational wave function with the stochastic reconfiguration method. In Sec. 6, we present a few results for correlation functions that can be computed in variational Monte Carlo. Finally, in Sec 7, we draw the conclusions.

## 2 The Hubbard model

### 2.1 The single-orbital Hubbard model

In this section, we introduce the single-orbital Hubbard model on the square lattice, a prototype model for studying the role of electronic correlations [5, 6]. In particular, this model is the simplest model able to capture the essential characteristics of the cuprates' phase diagram. This Hamiltonian comprises an interaction term  $\mathcal{H}_{\text{int}}$  that contains the on-site electron-electron repulsion  $U$  and a kinetic part  $\mathcal{H}_0$  that includes the nearest neighbor hopping integral  $t$ , and the next-nearest neighbor hopping integral  $t'$ . This last term is directly connected with the critical temperature for superconductivity [7]. The Hamiltonian is defined as

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_{\text{int}} = -t \sum_{\langle R, R' \rangle, \sigma} c_{R\sigma}^\dagger c_{R'\sigma} - t' \sum_{\langle\langle R, R' \rangle\rangle, \sigma} c_{R\sigma}^\dagger c_{R'\sigma} + \text{H.c} + U \sum_R n_{R\uparrow} n_{R\downarrow}, \quad (1)$$

where  $c_{R\sigma}^\dagger$  and  $c_{R\sigma}$  denote the creation and annihilation operators of an electron with spin  $\sigma$  on site  $R$ , respectively, while  $n_{R\sigma} = c_{R\sigma}^\dagger c_{R\sigma}$  is the electron density per spin  $\sigma$  on site  $R$ . The symbols  $\langle \dots \rangle$  and  $\langle\langle \dots \rangle\rangle$  represent the summation over nearest- and next-nearest-neighbor sites, respectively. By calling  $N_e$  the total number of electrons and  $L$  the total number of sites, the electron density is  $n = N_e/L$ . At half filling, i.e., when  $n = 1$ , the model interpolates between a metal at  $U = 0$  and an insulator driven by electronic correlations, the so-called Mott insulator, at  $t = 0$ . A metal-insulator transition will occur at a critical value of  $U/t$  that depends on the lattice geometry.

Despite its simplicity, this model cannot be exactly solved, except in a few limiting cases, like in one spatial dimension with nearest-neighbor hopping, where the so-called Bethe Ansatz provides us with exact results [8], and in the half-filled case for any value of  $U/t$  at  $t' = 0$  on the square lattice, where the auxiliary-field Monte Carlo approach becomes numerically exact, with no sign problem [9]. In all other cases, we need to rely on approximate numerical methods.

### 2.2 The multi-orbital Hubbard model

When extending the Hubbard model from a single relevant orbital to a multi-orbital scheme, we need to include new interaction terms. In addition to the intra-orbital on-site Coulomb interaction  $U$ , we shall also include the inter-orbital on-site Coulomb repulsion  $U'$  and the Hund's coupling  $J$ . The latter term tends to maximize the total spin on a single site, as a consequence of the intra-atomic exchange energy. The interaction term of the Hubbard Hamiltonian assumes then a more complicated form, also known as the Kanamori Hamiltonian [10, 11]

$$\begin{aligned} \mathcal{H}_{\text{int}} = & U \sum_R \sum_{\alpha} n_{R,\alpha,\uparrow} n_{R,\alpha,\downarrow} + U' \sum_R \sum_{\alpha \neq \beta} n_{R,\alpha,\uparrow} n_{R,\beta,\downarrow} + (U' - J) \sum_R \sum_{\alpha > \beta} \sum_{\sigma} n_{R,\alpha,\sigma} n_{R,\beta,\sigma} \\ & - J \sum_R \sum_{\alpha \neq \beta} c_{R,\alpha,\uparrow}^\dagger c_{R,\alpha,\downarrow}^\dagger c_{R,\beta,\downarrow} c_{R,\beta,\uparrow} + J \sum_R \sum_{\alpha \neq \beta} c_{R,\alpha,\uparrow}^\dagger c_{R,\alpha,\downarrow}^\dagger c_{R,\beta,\downarrow} c_{R,\beta,\uparrow}, \end{aligned} \quad (2)$$

where  $\alpha$  and  $\beta$  are orbital indices,  $c_{R,\alpha,\sigma}^\dagger$  ( $c_{R,\alpha,\sigma}$ ) is the fermionic operator that creates (annihilates) an electron in orbital  $\alpha$ , site  $R$  and spin  $\sigma$ , and  $n_{R,\alpha,\sigma} = c_{R,\alpha,\sigma}^\dagger c_{R,\alpha,\sigma}$  is the density operator on site  $R$ , orbital  $\alpha$  and spin  $\sigma$ .

The first three terms refer to the on-site density-density interaction between electrons with opposite spins in the same orbital ( $U$ ), opposite spins in different orbitals ( $U'$ ), and parallel spins in different orbitals ( $U' - J < U'$ ). In the latter term, the interaction constant is lower because of the spin alignment effect. The fourth is the spin-flip process ( $J$ ), and the last term describes pair hopping, i.e., an up-down spin pair of electrons that moves from one orbital to another one ( $J$ ). The preservation of rotational symmetry is guaranteed for  $U' = U - 2J$ , which will be used throughout the discussion.

The kinetic part of the multi-orbital Hubbard model depends on the specific hopping scheme that characterizes the model under investigation. Here, we consider iron-based superconductors as an example. In this case, a three Fe orbital ( $d_{xz}$ ,  $d_{yz}$ , and  $d_{xy}$ ) kinetic Hamiltonian  $\mathcal{H}_0$  [12, 13] is able to reproduce qualitatively the generic shape and the orbital content of the Fermi surfaces of iron-based superconductors, while remaining sufficiently light to allow for accurate numerical simulations. The tight-binding Hamiltonian is then defined as

$$\mathcal{H}_0 = \sum_k \sum_{\alpha,\beta} \sum_{\sigma} c_{k,\alpha,\sigma}^\dagger T_{\alpha,\beta}(k) c_{k,\beta,\sigma}, \quad (3)$$

where  $\alpha$  and  $\beta$  are orbital indices ( $1=xz$ ,  $2=yz$ , and  $3=xy$ ), while  $c_{k,\alpha,\sigma}^\dagger$  ( $c_{k,\alpha,\sigma}$ ) is the fermionic operator that creates (annihilates) an electron in orbital  $\alpha$ , with momentum  $k$  and spin  $\sigma$ . The intra- and inter-orbital hoppings are given as

$$\begin{aligned} T_{1,1}(k) &= 2t_2 \cos k_x + 2t_1 \cos k_y + 4t_3 \cos k_x \cos k_y, \\ T_{2,2}(k) &= 2t_1 \cos k_x + 2t_2 \cos k_y + 4t_3 \cos k_x \cos k_y, \\ T_{3,3}(k) &= 2t_5(\cos k_x + \cos k_y) + 4t_6 \cos k_x \cos k_y + \varepsilon_{xy}, \\ T_{1,2} &= T_{2,1}^* = 4t_4 \sin k_x \sin k_y, \\ T_{1,3} &= T_{3,1}^* = 2it_7 \sin k_x + 4it_8 \sin k_x \cos k_y, \\ T_{2,3} &= T_{3,2}^* = 2it_7 \sin k_y + 4it_8 \sin k_y \cos k_x. \end{aligned}$$

The hopping parameters (in units of eV) are set as:  $t_1 = 0.02$ ,  $t_2 = 0.06$ ,  $t_3 = 0.03$ ,  $t_4 = -0.01$ ,  $t_5 = 0.1$ ,  $t_6 = 0.15$ ,  $t_7 = -0.1$ ,  $t_8 = -t_7/2$  [13]. The two orbitals  $d_{xz}$  and  $d_{yz}$  are degenerate, while the orbital  $d_{xy}$  has a crystal field  $\varepsilon_{xy} = 0.2$ .

### 3 Variational Monte Carlo

Let us now introduce the general scheme under which variational Monte Carlo is defined. Let us first fix a complete orthonormalized basis in the Hilbert space  $\{|x\rangle\}$

$$\sum_x |x\rangle\langle x| = \mathbb{I}. \quad (4)$$

In order to clarify this statement with an example, an element of the basis for the single-orbital Hubbard model is a vector that contains the electronic occupation of each lattice site.

Given the aforementioned basis, each quantum state can be written as:

$$|\Psi\rangle = \sum_x |x\rangle \langle x|\Psi\rangle = \sum_x \Psi(x) |x\rangle. \quad (5)$$

In this way, the expectation value of a generic operator  $\mathcal{O}$  over a given variational state can be written as

$$\langle \mathcal{O} \rangle = \frac{\langle \Psi | \mathcal{O} | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \frac{\sum_x \langle \Psi | x \rangle \langle x | \mathcal{O} | \Psi \rangle}{\sum_{x'} \langle \Psi | x' \rangle \langle x' | \Psi \rangle}. \quad (6)$$

The main problem in evaluating this expectation value lies in the number of configurations in the sum, which becomes exponentially large with the number of sites and particles. Therefore, for large sizes, it is impossible to perform an exact enumeration of the configurations to compute  $\mathcal{O}$  exactly. However, Eq. (6) can be recast into a form that can be easily treated by standard Monte Carlo methods

$$\langle \mathcal{O} \rangle = \frac{\sum_x |\langle \Psi | x \rangle|^2 \frac{\langle x | \mathcal{O} | \Psi \rangle}{\langle x | \Psi \rangle}}{\sum_{x'} |\langle \Psi | x' \rangle|^2} = \frac{\sum_x |\Psi(x)|^2 O_L(x)}{\sum_{x'} |\Psi(x')|^2}, \quad (7)$$

where we have defined the local estimator

$$O_L(x) = \frac{\langle x | \mathcal{O} | \Psi \rangle}{\langle x | \Psi \rangle} = \sum_{x'} \langle x | \mathcal{O} | x' \rangle \frac{\langle x' | \Psi \rangle}{\langle x | \Psi \rangle}. \quad (8)$$

At first sight, the computation of the local estimator looks like a hard task, since it requires a summation over all the states of the many-body Hilbert space; however, when considering operators like the Hamiltonian, only a few terms contribute to the sum, thanks to the locality of the Hubbard Hamiltonian. Indeed, if we use the local basis, whose elements are the electronic configurations in the lattice, an element  $|x\rangle$  is connected only to a few other configurations that differ in the hopping of one electron from a given site to one of its neighbors; then, the maximum number of such processes is equal to the number of sites  $L$  times the number of bonds times the spin degeneracy. Therefore, the computation of the local estimator requires only a limited number of operations, usually proportional to the number of sites or particles in the system.

The quantity

$$P(x) = \frac{|\Psi(x)|^2}{\sum_{x'} |\Psi(x')|^2} \quad (9)$$

can be interpreted as a probability distribution, since it is non-negative and normalized. This property guarantees that variational Monte Carlo is not affected by the sign problem that is common to other Monte Carlo approaches.

Thus, the problem of computing a quantum average of the operator  $\mathcal{O}$  can be rewritten as the average of the random variable  $O_L(x)$ , see Eq. (8), over the probability distribution  $P(x)$ , see Eq. (9). In particular, if we consider the expectation value of the Hamiltonian, the local estimator is named local energy and is defined by

$$e_L(x) = \frac{\langle x | \mathcal{H} | \Psi \rangle}{\langle x | \Psi \rangle}. \quad (10)$$

It is then possible to define a stochastic algorithm (e.g., a Markov process) in which a sequence of electronic configurations  $\{|x_n\rangle\}$  is generated. One can use, for example, the Metropolis algorithm, which is a customary algorithm in Monte Carlo simulations, for generating states according to the given probability distribution. With the Metropolis algorithm, after an initial equilibration time, configurations will be distributed according to the desired probability  $P(x)$ . Then, the quantum expectation value  $\langle O \rangle$  is evaluated as the mean value of the random variable  $O_L(x)$  over the visited configurations

$$\langle O \rangle \approx \frac{1}{N} \sum_{n=1}^N O_L(x_n), \quad (11)$$

where  $N$  is the total number of visited configurations, after equilibration.

An important feature of VMC is the zero-variance property: if the variational state  $|\Psi\rangle$  coincides with an eigenstate of the Hamiltonian with eigenvalue  $E$ , the local energy  $e_L(x)$  is a constant

$$e_L(x) = \frac{\langle x | \mathcal{H} | \Psi \rangle}{\langle x | \Psi \rangle} = E \frac{\langle x | \Psi \rangle}{\langle x | \Psi \rangle} = E. \quad (12)$$

Therefore, the variance of the random variable  $e_L(x)$  is zero. Even if this is an extreme case, it teaches us that the variance of  $e_L(x)$  decreases when the variational state approaches an exact eigenstate. This fact is important to reduce statistical fluctuations and improve the numerical efficiency. Notice that the zero-variance property is a feature that exists only for quantum expectation values, while it is absent in classical calculations, where observables have thermal fluctuations.

### 3.1 The Metropolis algorithm

In this lecture, we do not discuss the details of Markov chains and the conditions that guarantee that the configurations  $|x_n\rangle$  generated along the Markov chain are distributed according to the probability distribution  $P(x)$  of Eq. (9), for large values of  $n$ . More details can be found, for instance, in Ref. [1]. Instead, we just recall the Metropolis algorithm that defines a transition probability for generating a new electronic configuration  $|x'\rangle$ , starting from an old one  $|x\rangle$ . Once a move is proposed to generate the new configuration  $|x'\rangle$  starting from the old configuration  $|x\rangle$ , the acceptance probability  $A(x'|x)$  is defined as

$$A(x'|x) = \min \left\{ 1, \frac{T(x|x')P(x')}{T(x'|x)P(x)} \right\}, \quad (13)$$

where  $T(x'|x)$  is the trial probability for generating  $|x'\rangle$  from  $|x\rangle$ . If the trial probability is symmetric ( $T(x'|x) = T(x|x')$ ), Eq. (13) further simplifies to:

$$A(x'|x) = \min \left\{ 1, \frac{P(x')}{P(x)} \right\}, \quad (14)$$

where

$$\frac{P(x')}{P(x)} = \frac{|\Psi(x')|^2}{|\Psi(x)|^2}, \quad (15)$$

thus making it unnecessary to know the normalization of the wave function.

Recalling that each configuration  $|x\rangle$  is the electron occupation of each lattice site (with a fixed  $z$  component of the spin), we set in the following a generic procedure to run a Metropolis Monte Carlo algorithm for fermions:

- Generate a random number to identify the sites  $R$  and  $R'$  among which the electron (with fixed spin  $\sigma$ ) will hop, defining the trial probability  $T(x'|x)$ . In strongly correlated configurations (i.e., large values of  $U/t$  at half filling), an additional trial move can be proposed, as a spin flip between two electrons with opposite spin on singly occupied sites.
- If the electron can perform the move (specifically, it must not violate the Pauli principle), the algorithm computes the acceptance ratio as  $A(x'|x)$ ; otherwise, the move is rejected.
- Generate a random uniform number  $\eta \in (0, 1]$  to compare with the acceptance ratio. The new configuration  $|x'\rangle$  is accepted for  $A(x'|x) \geq \eta$ ; otherwise, the move is rejected.
- Once thermalization is reached, i.e., when configurations are generated according to the desired probability distribution, observables, such as the local energy  $E_L(x)$ , can be computed every  $O(L)$  steps, where  $L$  is the lattice size, to have uncorrelated configurations. As a rule of thumb, thermalization time is similar to correlation time, that is, the necessary time to run a simulation from a starting configuration  $|x\rangle$  to obtain an independent configuration  $|x'\rangle$ .

If the configurations are uncorrelated, the error bar associated with the variational estimate of an observable is given by

$$\sigma_{\langle \mathcal{O} \rangle} = \sqrt{\frac{1}{N(N-1)} \left( \sum_{i=1}^N O_L^2(x_i) - \frac{1}{N} \left( \sum_{i=1}^N O_L(x_i) \right)^2 \right)}. \quad (16)$$

However, the usual acceptance probability for VMC in the Hubbard model is quite small, implying that many moves are rejected, and the electronic configurations stored during the Monte Carlo sampling could be correlated. Consequently, the error bar will be underestimated if we use Eq. (16). A simple procedure used to solve the problem of successive correlated configuration, in the evaluation of the error bar, is the *binning technique*, in which  $L_b$  measures are grouped in a bin, in order to compute a partial average within the bin  $b$ :

$$\bar{O}_b = \frac{1}{L_b} \sum_{i=(b-1)L_b+1}^{bL_b} O_L(x_i); \quad (17)$$

here,  $b = 1, \dots, N_{\text{bin}}$  such that  $N = \sum_{b=1}^{N_{\text{bin}}} L_b$ . Whenever  $L_b$  is larger than the correlation time, the configurations are almost uncorrelated. Based on the binning technique, the variational energy and its error bar are given by

$$\langle \mathcal{O} \rangle = \frac{1}{N_{\text{bin}}} \sum_{b=1}^{N_{\text{bin}}} \bar{O}_b \quad \text{and} \quad \sigma_{\langle \mathcal{O} \rangle} = \sqrt{\frac{1}{N_{\text{bin}}(N_{\text{bin}}-1)} \sum_{b=1}^{N_{\text{bin}}} (\bar{O}_b - \langle \mathcal{O} \rangle)^2}. \quad (18)$$

## 4 Variational wave functions

### 4.1 Jastrow-Slater wave functions for the single-orbital Hubbard model

The Jastrow-Slater wave function can be generically written as

$$|\Psi_J\rangle = \mathcal{J}|\Phi_0\rangle, \quad (19)$$

where  $\mathcal{J}$  is a Jastrow factor that embeds correlation in the wave function and  $|\Phi_0\rangle$  is the ground state of a noninteracting quadratic auxiliary Hamiltonian  $\mathcal{H}_{\text{aux}}$ , according to the definition

$$|\Phi_0\rangle = \prod_{k=1}^{N_e} \varphi_k^\dagger |0\rangle, \quad (20)$$

where  $N_e$  is the total number of electrons and  $\varphi_k$  are the eigenstates of  $\mathcal{H}_{\text{aux}}$ . Indeed, we can easily diagonalize any Hamiltonian that can be written in a simple quadratic form as

$$\mathcal{H}_{\text{aux}} = \mathbf{d}^\dagger \mathbf{T} \mathbf{d}, \quad (21)$$

where  $\mathbf{d}^\dagger = (\mathbf{d}_\uparrow^\dagger, \mathbf{d}_\downarrow^\dagger)$  is a vector with  $2L$  components ( $\mathbf{d}_\sigma^\dagger = d_{1,\sigma}^\dagger, \dots, d_{L,\sigma}^\dagger$ ),  $\mathbf{T}$  is a  $2L \times 2L$  matrix, and  $L$  is the number of lattice sites. By introducing a unitary matrix  $\mathbf{U}$  (that preserves the anticommutation rules of fermions), we can diagonalize  $\mathcal{H}_{\text{aux}}$

$$\mathcal{H}_{\text{aux}} = \mathbf{d}^\dagger \mathbf{U} \mathbf{U}^\dagger \mathbf{T} \mathbf{U} \mathbf{U}^\dagger \mathbf{d} = \mathbf{\Phi}^\dagger \mathbf{E} \mathbf{\Phi} = \sum_k \varepsilon_k \varphi_k^\dagger \varphi_k, \quad (22)$$

where  $\mathbf{E}$  is a diagonal matrix that contains the eigenvalues and  $\mathbf{\Phi} = (\varphi_1, \dots, \varphi_{2L})$  is defined in terms of the eigenvectors.

The choice of the noninteracting quadratic Hamiltonian is a careful aspect, since we need to include in it the relevant physical terms that we want to study in our model. For instance, if one would be interested in studying the emergence of superconductivity in the Hubbard model, when varying the electronic density, and its competition with antiferromagnetic Néel order, one could define the following auxiliary Hamiltonian

$$\mathcal{H}_{\text{aux}} = \mathcal{H}_0 + \mathcal{H}_{AF} + \mathcal{H}_{BCS}. \quad (23)$$

The first term consists of the *kinetic energy* of the electrons in the Hubbard model

$$\mathcal{H}_0 = -t \sum_{\langle R, R' \rangle, \sigma} c_{R\sigma}^\dagger c_{R'\sigma} - \tilde{t}' \sum_{\langle\langle R, R' \rangle\rangle, \sigma} c_{R\sigma}^\dagger c_{R'\sigma} + \text{H.c.} \quad (24)$$

The second term includes antiferromagnetism by *standard Néel order*

$$\mathcal{H}_{AF} = \Delta_{AF} \sum_R (-1)^{x+y} \left( c_{R\uparrow}^\dagger c_{R\uparrow} - c_{R\downarrow}^\dagger c_{R\downarrow} \right), \quad (25)$$

while the last one introduces the *BCS electron pairing*

$$\mathcal{H}_{BCS} = \sum_{R, \eta=x, y} \Delta_\eta \left( c_{R,\uparrow}^\dagger c_{R+\eta,\downarrow}^\dagger - c_{R,\downarrow}^\dagger c_{R+\eta,\uparrow}^\dagger \right) + \text{H.c.} - \mu \sum_{R, \sigma} c_{R\sigma}^\dagger c_{R\sigma}, \quad (26)$$

where  $\mu$  is a chemical potential and  $\Delta_x = -\Delta_y$  represents the usual  $d$ -wave symmetry that is realized in the single-band Hubbard model. Further developments can be included to describe the striped modulations of charge and spin that are a common feature in the ground state of the single-orbital Hubbard model on the square lattice [14–16]. The parameters  $\tilde{t}'$ ,  $\Delta_{AF}$ ,  $\Delta\eta$ ,  $\mu$  are variational parameters that can be optimized, as described in Sec. 5, while we keep  $t = 1$  to fix the energy scale.

The reader could now wonder how a BCS state can be written as a Slater determinant. This is indeed possible, via a convenient particle-hole canonical transformation of the annihilation and creation operators for down spins, i.e.

$$d_{R,\uparrow}^\dagger \equiv c_{R,\uparrow}^\dagger \quad d_{R,\downarrow} \equiv c_{R,\downarrow}^\dagger. \quad (27)$$

The products of creation and annihilation operators that build up the auxiliary Hamiltonian are then rewritten as

$$\begin{aligned} c_{R\uparrow}^\dagger c_{R'\uparrow} &= d_{R\uparrow}^\dagger d_{R'\uparrow} & c_{R\downarrow}^\dagger c_{R'\downarrow} &= -d_{R'\downarrow}^\dagger d_{R\downarrow} \\ c_{R,\uparrow}^\dagger c_{R+\eta,\downarrow}^\dagger &= c_{R,\uparrow}^\dagger c_{R+\eta,\downarrow} & c_{R,\downarrow}^\dagger c_{R+\eta,\uparrow}^\dagger &= -c_{R+\eta,\uparrow}^\dagger c_{R,\downarrow}. \end{aligned} \quad (28)$$

In this way, the eigenstates of the BCS Hamiltonian can be expressed as “orbitals”, similarly to the ones that can be obtained by diagonalizing the kinetic term.

It should be noted that the particle-hole transformation changes the total number of electrons per spin species  $N_\sigma^c$ , according to the following relations

$$\begin{aligned} N_\uparrow^c &= \sum_R c_{R,\uparrow}^\dagger c_{R,\uparrow} = \sum_R d_{R,\uparrow}^\dagger d_{R,\uparrow} = N_\uparrow^d \\ N_\downarrow^c &= \sum_R c_{R,\downarrow}^\dagger c_{R,\downarrow} = L - \sum_R d_{R,\downarrow}^\dagger d_{R,\downarrow} = L - N_\downarrow^d, \end{aligned} \quad (29)$$

where  $L$  is the total number of sites.

Whenever the total magnetization is zero, i.e.,  $N_\uparrow^c = N_\downarrow^c$ , the total number of electrons after the particle-hole transformation, given by  $N_\uparrow^d + N_\downarrow^d$ , is equal to  $L$ , that is, the system, after the particle-hole transformation, is always half-filled.

The *Jastrow factor*  $\mathcal{J}$  is a long-range correlator that, on the lattice, takes the following form

$$\mathcal{J} = \exp \left[ -\frac{1}{2} \sum_{R,R'} v_{R,R'} (n_R - n)(n_{R'} - n) \right], \quad (30)$$

where  $v_{R,R'}$  is a pseudo-potential for density-density correlations in the variational state,  $n_R$  is the electronic density on site  $R$ , and  $n$  is the average electronic density. For translationally invariant models, like the Hubbard Hamiltonian of Eq. (1),  $v_{R,R'}$  only depends on the relative distance of the two sites  $|\mathbf{R} - \mathbf{R}'|$ . The Jastrow pseudo-potential can be optimized for all possible distances, which are  $O(L)$  in translationally invariant systems, via the stochastic reconfiguration method of Sec. 5. The role of the long-range tail of the Jastrow factor is to create a bound state between doubly occupied sites and empty sites, thus not allowing for conduction, but still

allowing local density fluctuations. Indeed, it has been shown that the Jastrow term may turn a non-interacting metallic state  $|\Phi_0\rangle$  at half filling into a Mott insulator, that is, an insulator driven by electronic correlation, even in the absence of magnetic order [17]. We mention the existence of a deep connection between a Mott insulator at half-filling and a superconductor at finite doping in the Resonant Valence Bond (RVB) theory of high-temperature superconductivity proposed by Anderson [18]. He suggested that a superconducting phase may emerge when doping a Mott insulator with “preformed” singlet pairs. In fact, the RVB state describes an insulator that is a liquid of spin singlets, which become mobile when the system is doped, thus forming actual superconducting pairs.

We remark that replacing the Jastrow factor with a much simpler term that penalizes the formation of double occupancies is not able to reproduce a Mott insulator unless  $U = \infty$ . This simplified interaction term, called the Gutzwiller factor, is defined as [19]

$$|\Psi_G\rangle = P_G|\Phi_0\rangle \quad P_G = \exp \left[ -\frac{g}{2} \sum_R (n_R - n)^2 \right] \quad (31)$$

and corresponds to the  $R = R'$  case of the Jastrow factor. The effect of the Gutzwiller factor becomes clear once the variational state is expanded in a basis set whose elements  $\{|x\rangle\}$  represent configurations with electron occupations of each lattice site. Since the Gutzwiller factor is diagonal in this basis, we have that:

$$\langle x|\Psi_G\rangle = P_G(x)\langle x|\Phi_0\rangle, \quad (32)$$

where  $P_G(x) \leq 1$  is a number that depends on how many doubly-occupied sites are present in a given electronic configuration. Therefore, the amplitude of the non-interacting state  $|\Phi_0\rangle$  is renormalized by  $P_G(x)$ . At half filling, density excitations are represented by doublons (doubly-occupied sites) and holons (empty sites). In the absence of electronic correlation, these objects are free to move and are then responsible for conductivity (for example, in the fermionic model, a doublon is negatively charged with respect to the average background, while the holon is positively charged). The effect of the Gutzwiller factor is to penalize the formation of such excitations; however, once created, doublons and holons are no longer correlated, thus being free to move independently. Only when the energetic penalty is infinite, an insulator is obtained. On the contrary, the long-range tail of the Jastrow factor allows to correlate holons and doublons and keeps them bound together. This fact can be understood via a classical mapping to a model of charged particles, where holons and doublons represent positive and negative charges [1].

## 4.2 Computing ratios for the Jastrow-Slater wave function

Both in the Metropolis algorithm of Eq. (15) and in the evaluation of the local observables of Eq. (8), we need to compute ratios of amplitudes of the variational state with respect to a generic element of the basis set, i.e.,  $\langle x'|\Psi\rangle/\langle x|\Psi\rangle$ .

The advantage of considering Jastrow-Slater wave functions in the variational Monte Carlo technique is that the calculations can be efficient and fast. Indeed, since the Jastrow factor is

diagonal in the chosen basis, we have that

$$\langle x|\Psi\rangle = \mathcal{J}(x)\langle x|\Phi_0\rangle, \quad (33)$$

where  $\mathcal{J}(x)$  is the value of the Jastrow operator computed for the configuration  $|x\rangle$ , i.e.,  $\mathcal{J}|x\rangle = \mathcal{J}(x)|x\rangle$ . Therefore, given the electronic configuration,  $\mathcal{J}(x)$  is a number that can be evaluated in  $O(L^2)$  operations. Moreover,  $\langle x|\Phi_0\rangle$  can be evaluated in  $O(L^3)$  operations since it is a Slater determinant.

Let us now briefly discuss how to compute ratios for the two cases. First, let us focus on the ratio of Jastrow factors  $\mathcal{J}(x')/\mathcal{J}(x)$  and show that it requires only  $O(L)$  operations. We focus on a simplified form for the Jastrow factor

$$\mathcal{J} = \exp \left[ -\frac{1}{2} \sum_{R,R'} v_{R,R'} n_R n_{R'} \right], \quad (34)$$

which is equivalent to the one of Eq. (30) when the particle number is conserved (apart from a multiplicative factor).

When  $|x'\rangle$  is close to  $|x\rangle$ , i.e.

$$|x\rangle \equiv |n_1, \dots, n_k, \dots, n_l, \dots, n_L\rangle \quad |x'\rangle \equiv |n_1, \dots, n_k+1, \dots, n_l-1, \dots, n_L\rangle, \quad (35)$$

it is possible to define a fast computation

$$\frac{\mathcal{J}(x')}{\mathcal{J}(x)} = \frac{\mathcal{J}(n_1, \dots, n_k+1, \dots, n_l-1, \dots, n_L)}{\mathcal{J}(n_1, \dots, n_k, \dots, n_l, \dots, n_L)} = \frac{\exp(-\sum_R v_{R,k} n_R)}{\exp(-\sum_R v_{R,l} n_R)} e^{v_{k,l}-v_{k,k}}, \quad (36)$$

where we used the fact that the pseudo-potential is symmetric and translationally invariant. The term  $e^{v_{k,l}-v_{k,k}}$  in Eq. (36) does not depend upon the electronic configuration and can be computed at the beginning of the simulation, once for all. Instead, the ratio of the exponentials depends on the fermionic configuration, which is sampled along the Markov chain. However, the computation of the ratio can be done in  $O(1)$  operations once we compute and store a vector of dimension  $L$  that depends on  $|x\rangle$ , defined as

$$T_{\text{Jastrow}}(R') = \sum_R v_{R,R'} n_R. \quad (37)$$

Then, updating  $T_{\text{Jastrow}}(R')$  to the new configuration  $|x'\rangle$  costs  $O(L)$ :

$$n'_R = n_R + \delta_{k,R} - \delta_{l,R} \Rightarrow T'_{\text{Jastrow}}(R') = T_{\text{Jastrow}}(R') + v_{k,R'} - v_{l,R'}. \quad (38)$$

Therefore, we can compute from scratch  $T_{\text{Jastrow}}(R')$  at the beginning of the Markov chain for all the sites  $R' = 1, \dots, L$  and then update it every time a new configuration is accepted along the Markov chain. From time to time, it is recommended to recompute Eq. (37) from scratch, since the rounding error of the update can accumulate and give rise to numerical errors.

Let us now move to the  $\langle x|\Phi_0\rangle$  term, and show that it indeed corresponds to a Slater determinant that can be computed in  $O(L^3)$  operations with standard linear algebra routines. First of all, let

us contract the site and spin indices of the auxiliary quadratic Hamiltonian  $\mathcal{H}_{\text{aux}}$  into a single index  $I$  from 1 to  $2L$ :  $c_{R,\uparrow} \equiv d_R$ ,  $c_{R,\downarrow} \equiv d_{R+L}$ . Then, we recall that  $\mathcal{H}_{\text{aux}}$  can be diagonalized and its eigenstates are the single-particle orbitals  $\varphi_k^\dagger|0\rangle$

$$\mathcal{H}_{\text{aux}} = \sum_k \varepsilon_k \varphi_k^\dagger \varphi_k, \quad (39)$$

with  $\varepsilon_k$  the single-particle energies and  $\varphi_k^\dagger = \sum_I U_{I,k} d_I^\dagger$  (where  $\mathbf{U}$  is the unitary matrix that diagonalizes  $\mathcal{H}_{\text{aux}}$ ). Then, the many-particle state  $|\Phi_0\rangle$  can be constructed by occupying the  $N_e$  lowest single-particle orbitals, where  $N_e$  is the number of electrons

$$|\Phi_0\rangle = \prod_{k=1}^{N_e} \varphi_k^\dagger |0\rangle = \prod_{k=1}^{N_e} \left( \sum_I U_{I,k} d_I^\dagger \right) |0\rangle. \quad (40)$$

A generic electronic configuration for  $N_e$  electrons in the lattice, visited along the Markov chain, can be written as

$$|x\rangle = d_{R_1}^\dagger \dots d_{R_{N_e}}^\dagger |0\rangle, \quad (41)$$

where a generic  $R_j$  assumes values from 1 to  $2L$ : the positions of spin-up electrons coincide with the site number, while the positions of spin-down electrons must be shifted by  $L$ . The overlap  $\langle x|\Phi_0\rangle$  can then be written as

$$\begin{aligned} \langle x|\Phi_0\rangle &= \langle 0|d_{R_{N_e}} \dots d_{R_1} \left( \sum_I U_{I,1} d_I^\dagger \right) \dots \left( \sum_I U_{I,N_e} d_I^\dagger \right) |0\rangle = \\ &= \langle 0|d_{R_{N_e}} \dots d_{R_1} \left[ \sum_p (-1)^p \prod_{k=1}^{N_e} U_{p\{R_j\},k} \right] d_{R_1}^\dagger \dots d_{R_{N_e}}^\dagger |0\rangle. \end{aligned} \quad (42)$$

The sum in  $[\dots]$  is over all the permutations of the  $\{R_j\}$  in  $|x\rangle$  (with the sign from anticommutations), thus implying that  $\langle x|\Phi_0\rangle = \det\{U_{R_j,k}\}$ . In practice, the diagonalization of the auxiliary Hamiltonian can be done at the beginning of the simulation, obtaining the matrix of the eigenvectors  $\mathbf{U}$ , of which we store only the  $N_e$  columns corresponding to the occupied orbitals. Then, the overlap  $\langle x|\Phi_0\rangle$  with the generic configuration  $|x\rangle$  is given by the determinant of the matrix  $\tilde{\mathbf{U}}$

$$\tilde{\mathbf{U}} = \begin{pmatrix} U_{R_1,1} & \dots & U_{R_1,N_e} \\ \vdots & \ddots & \vdots \\ U_{R_l,1} & \dots & U_{R_l,N_e} \\ \vdots & \ddots & \vdots \\ U_{R_{N_e},1} & \dots & U_{R_{N_e},N_e} \end{pmatrix}, \quad (43)$$

obtained by taking only the rows corresponding to the occupied sites of the matrix  $\mathbf{U}$ . Now, if we consider a new electronic configuration  $|x'\rangle$ , obtained from  $|x\rangle$  after the hopping of the  $l$ -th electron from position  $R_l$  to position  $R_{l'}$ , i.e.,  $|x'\rangle = d_{R_{l'}}^\dagger d_{R_l} |x\rangle$ , we obtain a new matrix  $\tilde{\mathbf{U}}'$  that

will differ from  $\tilde{\mathbf{U}}$  for the  $l$ -th row

$$\tilde{\mathbf{U}}' = \begin{pmatrix} U_{R_1,1} & \cdots & U_{R_1,N_e} \\ \vdots & \ddots & \vdots \\ U_{R_l',1} & \cdots & U_{R_l',N_e} \\ \vdots & \ddots & \vdots \\ U_{R_{N_e},1} & \cdots & U_{R_{N_e},N_e} \end{pmatrix}. \quad (44)$$

Then, the ratio between two configurations that differ only by the hopping of a single electron is given by

$$\frac{\langle x' | \Phi_0 \rangle}{\langle x | \Phi_0 \rangle} = \frac{\langle x | d_{R_l}^\dagger d_{R_l'} | \Phi_0 \rangle}{\langle x | \Phi_0 \rangle} = \frac{\det \tilde{\mathbf{U}}'}{\det \tilde{\mathbf{U}}}. \quad (45)$$

It is then possible to show that this ratio of determinants can be computed in only  $O(L^2)$  operations, thus speeding up the simulation with respect to a full calculation of a determinant from scratch [1].

### 4.3 The backflow terms

While the Jastrow-Slater wave function is able to capture Mott insulating and superconductive states, it has a strong intrinsic limitation: Electronic correlations are introduced only via a multiplicative factor, while the nodes of the variational state are the ones set by the uncorrelated state  $|\Phi_0\rangle$ . In order to overcome this limitation, we have introduced a novel variational state that redefines the single-particle orbitals of the uncorrelated state, taking into account the electronic repulsion, the so-called backflow correlations [20, 21].

Let us consider the single-particle eigenstates  $\varphi_k(R, \sigma)$  of the auxiliary Hamiltonian  $\mathcal{H}_{\text{aux}}$ , written in the basis of lattice sites and  $z$  component of the spins. Backflow correlations modify them, according to the electronic configuration on the lattice

$$\varphi_k^b(R, \sigma) \equiv \tilde{\varepsilon} \varphi_k(R, \sigma) + \eta_1 \sum_{R' \text{ n.n. } R} D_R H_{R'} \varphi_k(R', \sigma) + \eta_2 \sum_{R' \text{ n.n.n. } R} D_R H_{R'} \varphi_k(R', \sigma), \quad (46)$$

where  $\tilde{\varepsilon} = \varepsilon$  if the site  $R$  is doubly occupied and is surrounded by at least one empty site, while  $\tilde{\varepsilon} = 1$  in all the other cases. Here  $\varepsilon, \eta_l$ , ( $l = 1, 2$ ) are variational parameters to be optimized, according to the method defined in Sec. 5. The operators  $D_R = n_{R,\uparrow} n_{R,\downarrow}$  and  $H_R = h_{R,\uparrow} h_{R,\downarrow}$ , with  $h_{R,\sigma} = 1 - n_{R,\sigma}$ , are non-zero only if the site  $R$  is doubly occupied or empty, respectively. Thus, they are the operators that allow us to include many-body correlations inside the single-particle orbitals. Moreover, the shorthand notations n.n. and n.n.n. indicate nearest- and next-nearest-neighbor sites, respectively. In this way, already the determinant part of the wave function includes correlation effects, strongly improving the accuracy of the many-body state. This is a substantial improvement with respect to Jastrow factors, where electron-electron correlation is included only via a multiplicative term.

The terms in Eq. (46) are the dominant contributions to the backflow operator, having a qualitative influence on the properties of the correlated wave functions. In addition, we can also take

into account further terms that are useful in the intermediate-coupling regime and correspond to all possible hopping processes

$$\begin{aligned}
\varphi_k^b(R, \sigma) \equiv & \tilde{\varepsilon} \varphi_k(R, \sigma) + \eta_1 \sum_{R' \text{ n.n. } R} D_R H_{R'} \varphi_k(R', \sigma) \\
& + \eta_2 \sum_{R' \text{ n.n.n. } R} D_R H_{R'} \varphi_k(R', \sigma) \\
& + \beta_1 \sum_{R' \text{ n.n. } R} n_{R, \sigma} h_{R, -\sigma} n_{R', -\sigma} h_{R', \sigma} \varphi_k(R', \sigma) \\
& + \beta_2 \sum_{R' \text{ n.n.n. } R} n_{R, \sigma} h_{R, -\sigma} n_{R', -\sigma} h_{R', \sigma} \varphi_k(R', \sigma) \\
& + \gamma_1 \sum_{R' \text{ n.n. } R} (D_R n_{R', -\sigma} h_{R', \sigma} + n_{R, \sigma} h_{R, -\sigma} H_{R'}) \varphi_k(R', \sigma) \\
& + \gamma_2 \sum_{R' \text{ n.n.n. } R} (D_R n_{R', -\sigma} h_{R', \sigma} + n_{R, \sigma} h_{R, -\sigma} H_{R'}) \varphi_k(R', \sigma), \quad (47)
\end{aligned}$$

where, in addition to  $\varepsilon$  and  $\eta_l, \beta_l, \gamma_l$  ( $l=1, 2$ ) are variational parameters to be optimized. We mention that the computation of  $\langle x | \Phi_0 \rangle$  and of the determinant ratios introduced in the previous subsection can be extended to the presence of backflow correlations with a similar degree of complexity [1].

Let us now present a few numerical results that show the relevance of backflow correlations in describing electronic correlations. We focus on the half-filled case, where the effect of electronic correlations is maximal, and we consider the simpler single-orbital Hubbard model. Moreover, we work in the so-called “nonmagnetic sector”, that is, neglecting any possible magnetic order in the variational wave function. In this way, we focus on the appearance of Mott insulating states that are driven solely by electronic correlations.

First, let us discuss the range of backflow correlations. We present the results in Fig. 1, left panels, for a 2D square lattice with both nearest neighbor hopping  $t$  and next-nearest neighbor hopping  $t'$ . We report that backflow parameters are irrelevant for all distances larger than the ones connected by the hopping amplitude. Most importantly, a remarkable gain in energy is obtained only by a suitable optimization of the backflow parameters up to second neighbors, while poor results are obtained when there is no backflow (or its range is smaller than the one of the hopping). Moreover, in Fig. 1, right panel, we present the effect of backflow correlations in the large- $U$  limit. We consider a one-dimensional Hubbard model, which is the same as the one introduced in Eq. (1), just denoting the nearest-neighbor hopping with  $t_1$  and the next-nearest-neighbor hopping with  $t_2$ . Only the presence of backflow correlations in the wave function allows for a proper extrapolation to the infinite- $U$  limit of the Hubbard model, that is, the Heisenberg model. Here, the energy of the Heisenberg model is provided by density-matrix renormalization group (DMRG) calculations, which are numerically exact for 1D models. We also compare the variational results with the exact ones on the 18-site cluster. In Fig. 2, we show the accuracy of the variational state (with and without backflow correlations) and the overlap with the exact ground state for two hopping values, i.e.,  $t'/t = 0$  and  $t'/t = 0.7$ . The backflow term is able to highly improve the accuracy both for weak and strong couplings. Moreover,

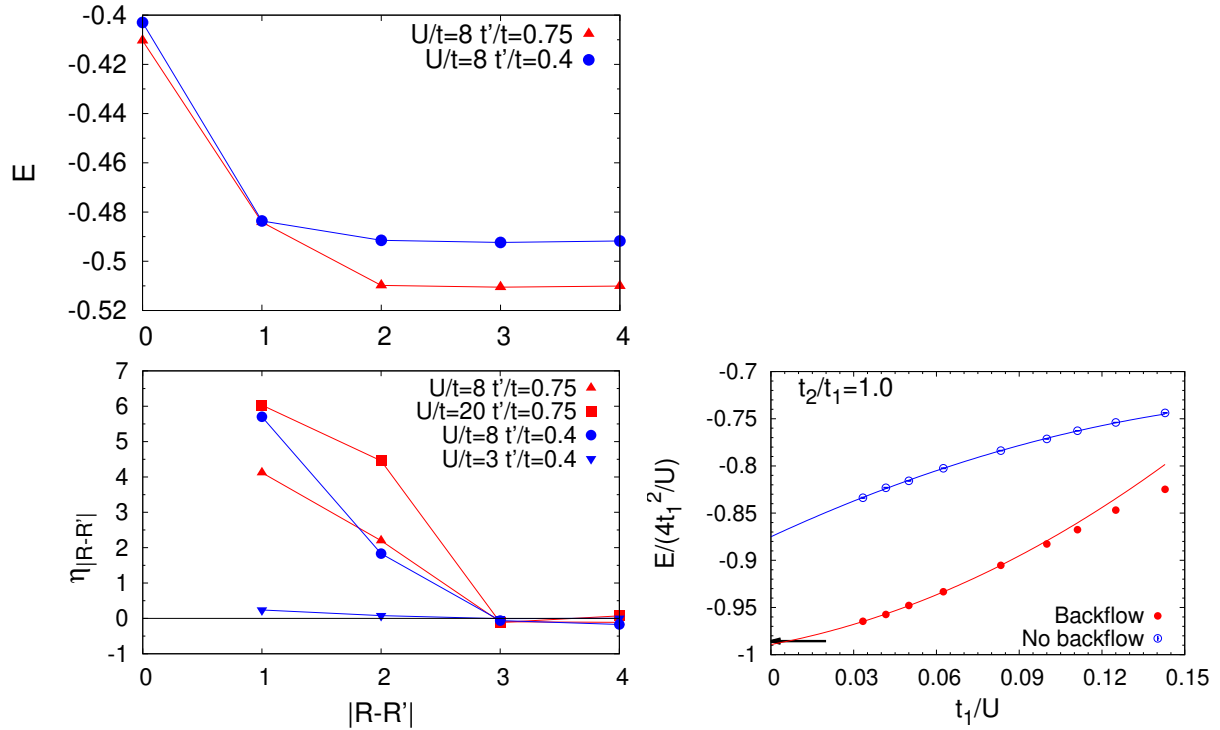


**Fig. 1:** Upper left panel: Ground-state energy as a function of the range of backflow parameters on the 2D square lattice for  $t'/t = 0.4$  and  $0.75$  at  $U/t = 8$ . A range equal to zero corresponds to no backflow correlations in the wave function. Lower left panel: Optimized backflow parameters  $\eta_{|R-R'|}$  as a function of the distance  $|R-R'|$  between lattice sites. Results for  $U/t = 3$  and  $t'/t = 0.4$  correspond to a metallic state, the others to Mott insulating states. Data are shown for a lattice with  $L = 98$  sites [21]. Right panel: Energy (in units of  $4t_1^2/U$ ), as a function of  $t_1/U$ , for the 1D Hubbard model with  $t_2/t_1 = 1$ . Filled (empty) circles denote the results with (without) backflow correlations. The DMRG energy of the corresponding Heisenberg model with  $J_2/J_1 = 1$  is shown by an arrow [22]. Data are presented for an  $L = 120$  chain [21].

the overlap between the exact ground state and the backflow state remains very high, even for large values of  $U/t$ , and the improvement with respect to the state without backflow is crucial, especially when a finite value of  $t'$  is included.

These results suggest that backflow correlations are indeed necessary to properly describe electronic correlations. In particular, backflow correlations turned out to be particularly useful in the presence of frustration [20, 23]. We do not discuss frustration in this lecture, but let us just say that it is related to the presence of competing hopping parameters in Mott insulators, like the case of square lattices with both nearest neighbor and next-nearest neighbor hopping parameters. The presence of frustration may destroy magnetic orders and lead to the formation of spin liquid phases, i.e., Mott insulating states without long-range magnetic order. Moreover, the backflow wave function has been the starting point for constructing a new class of wave functions, named neural network backflow, which uses a neural network to learn the optimal transformation via variational Monte Carlo calculations [24].

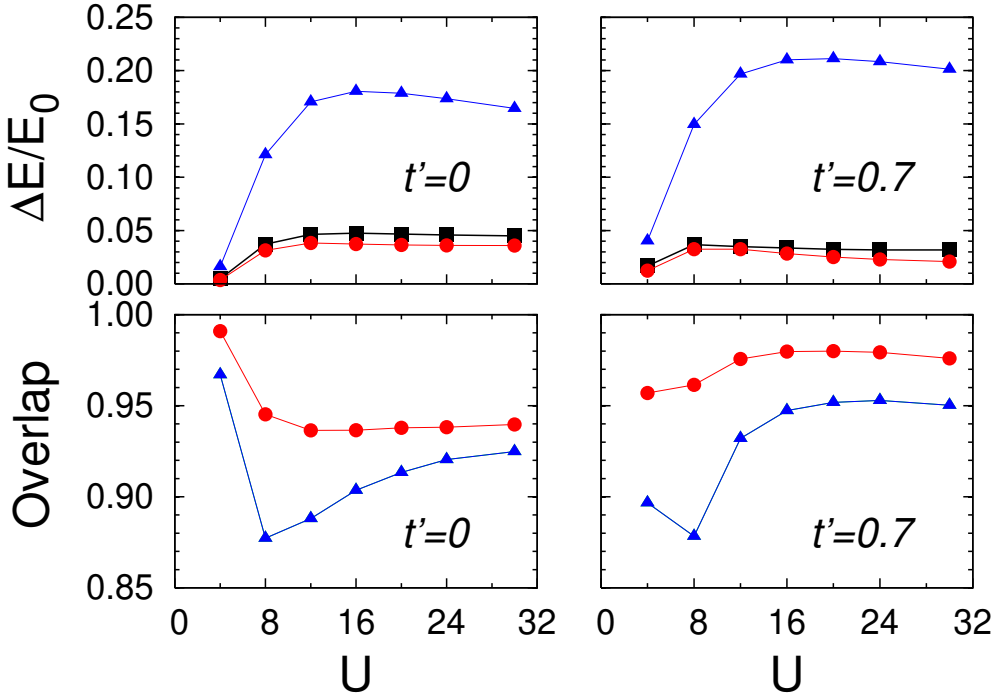


**Fig. 2:** Results for 18 electrons on 18 sites as a function of  $U/t$ . Upper panels: Accuracy of energy  $\Delta E = (E_v - E_0)$ ,  $E_v$  and  $E_0$  being the variational and the exact values, respectively. Lower panels: Overlap between the exact ground state and the variational wave functions. The variational state without backflow is denoted by blue triangles, the variational state with backflow correlations is denoted by red circles [20].

#### 4.4 Jastrow-Slater wave functions for the multi-orbital Hubbard model

The Jastrow-Slater wave function for a multi-orbital Hubbard model can be generalized in the following form [25, 26]

$$|\Psi\rangle = \mathcal{J}_U \mathcal{J}_{U'} \mathcal{J}_J |\Phi_0\rangle. \quad (48)$$

The Jastrow factor has been generalized as

$$\begin{aligned} \mathcal{J}_U &= \exp \left( -\frac{1}{2} \sum_{\alpha} \sum_{R,R'} v_{R,R'}^{\alpha,\alpha} n_{\alpha,R} n_{\alpha,R'} \right) \\ \mathcal{J}_{U'} &= \exp \left( -\frac{1}{2} \sum_{\alpha \neq \beta} \sum_{R,R'} v_{R,R'}^{\alpha,\beta} n_{\alpha,R} n_{\beta,R'} \right). \end{aligned} \quad (49)$$

When  $\alpha = \beta$ , the Jastrow factor is intra-orbital with a pseudopotential  $v_{R,R'}^{\alpha,\alpha}$ , whereas for  $\alpha \neq \beta$  the electronic correlations are inter-orbital and its pseudopotential is written as  $v_{R,R'}^{\alpha,\beta}$ . The Jastrow factors mimic the Coulomb interaction present in the Hubbard-Kanamori Hamiltonian of Eq. (2). Specifically,  $\mathcal{J}_U$  and  $\mathcal{J}_{U'}$  take into account the presence of the intra-orbital repulsion  $U$  and the inter-orbital repulsion  $U'$ , respectively. In order to take into account the presence of the Hund's coupling in the wave function, we define a spin-spin Jastrow factor that favors configu-

rations where electron spins in different orbitals (and same site) are aligned

$$\mathcal{J}_J = \exp \left( -\frac{1}{2} \sum_{\alpha \neq \beta} u^{\alpha, \beta} \sum_R S_{\alpha, R}^z S_{\beta, R}^z \right). \quad (50)$$

Here, the correlator involves the  $z$  component of the spin  $S_{\alpha, R}^z$  only for different orbitals on site  $R$ , as suggested by the definition of the Hund's coupling.

The uncorrelated state  $|\Phi_0\rangle$  is now the ground state of a new quadratic auxiliary Hamiltonian, which is appropriate for the multi-orbital model. Focusing only on the investigation of superconductivity in the model and neglecting any possible magnetic order, the auxiliary Hamiltonian can be defined as

$$\mathcal{H}_{\text{aux}} = \mathcal{H}_0 + \mathcal{H}_{BCS}, \quad (51)$$

where  $\mathcal{H}_0$  is the same as the tight-binding Hamiltonian of Eq. (3), while  $\mathcal{H}_{BCS}$  is now defined as

$$\mathcal{H}_{BCS} = - \sum_{R, \alpha, \sigma} \mu_\alpha c_{R, \alpha, \sigma}^\dagger c_{R, \alpha, \sigma} + \sum_{R, \alpha, \delta} [\Delta_{\alpha, \delta} (c_{R, \alpha, \uparrow} c_{R+\delta, \alpha, \downarrow} - c_{R, \alpha, \downarrow} c_{R+\delta, \alpha, \uparrow}) + \text{h.c.}], \quad (52)$$

where  $\delta = x, y, x+y$ , and  $x-y$  indicates nearest and next-nearest neighbors of the site  $R$  and  $\mu_\alpha$  defines the chemical potential of orbital  $\alpha$ . The (singlet) pairing amplitudes  $\Delta_{\alpha, \delta}$  and  $\mu_\alpha$  are optimized, while the hopping parameters in  $\mathcal{H}_0$  are usually kept fixed to the bare Hamiltonian ones. Inter-orbital pairing amplitudes are found to be negligible in the optimal wave function, in agreement with [27] and therefore not included in the auxiliary Hamiltonian.

## 5 Optimizing the variational state: Stochastic reconfiguration

We have introduced the functional form of the wave function, which depends on many variational parameters, ranging from the pseudopotentials in the Jastrow factor to the BCS coupling or the hopping parameters in the auxiliary Hamiltonian. This freedom in the choice of the variational wave function is meaningful only if we can construct a procedure that allows the optimization of many variational parameters at the same time, in order to find the best approximation of the ground state. One possible way to optimize the set of parameters is to use the zero-variance property, minimizing the variance of the energy in the variational state [28]. In contrast, the *stochastic reconfiguration* exploits the minimization of the variational energy; unlike the minimization of variance, where the minimum value is zero, the ground state energy is unknown a priori [4]. However, for a given form of the trial wave function, energy-minimized wave functions yield more accurate values of expectation values than wave functions whose variance is minimized [29].

Let us explain the stochastic reconfiguration assuming a generic functional form for the wave function  $|\Psi(\{p_k\})\rangle$  with  $k = 1, \dots, N_p$ , which depends on a set of  $N_p$  variational parameters

$\{p_k\}$ . The variational energy of the state can be decreased by the application of the projection operator  $(\mathcal{H}-\Lambda)$ , where  $\Lambda$  is an appropriate constant energy shift

$$|\Psi_\Lambda\rangle = (\mathcal{H}-\Lambda) |\Psi^0\rangle, \quad (53)$$

where  $|\Psi^0\rangle = |\Psi(\{p_k^0\})\rangle$  is the starting state, depending on an initial set of variational parameters  $\{p_k^0\}$ . However, in general, the new state  $|\Psi_\Lambda\rangle$  will not be in the Jastrow-Slater form. Therefore, the idea is to find the quantum state  $|\Psi'\rangle = |\Psi(\{p'_k\})\rangle$ , which gives the best possible approximation of  $|\Psi_\Lambda\rangle$ . If the change of the variational parameters is taken to be small, we can approximate the new parameters as  $p'_k = p_k^0 + \delta p_k$ , and the new state  $|\Psi'\rangle$  can be expanded at first order in  $\{\delta p_k\}$

$$|\Psi'\rangle = \delta p_0 |\Psi^0\rangle + \sum_{k=1}^{N_p} \delta p_k \frac{\partial |\Psi^0\rangle}{\partial p_k} + O(\delta p_k^2). \quad (54)$$

Here, the series is truncated to the second order, and  $\delta p_0$  is considered to match the normalization condition. Then, we define the local operator  $\mathcal{O}_k$  (with  $k = 1, \dots, N_p$ ) on the configuration  $|x\rangle$  as the logarithmic derivative with respect to the variational parameter  $p_k$

$$\mathcal{O}_k = \frac{1}{\Psi^0(x)} \frac{\partial \Psi^0(x)}{\partial p_k} = \frac{\partial \ln \langle x | \Psi^0 \rangle}{\partial p_k}. \quad (55)$$

Imposing  $\mathcal{O}_0 = 1$ ,  $|\Psi'\rangle$  can be expressed by using the logarithmic derivative operators

$$|\Psi'\rangle = \sum_{k=0}^{N_p} \delta p_k \mathcal{O}_k |\Psi^0\rangle. \quad (56)$$

Then, the stochastic reconfiguration method imposes that the projection of the shifted state  $|\Psi_\Lambda\rangle$  and of the new state  $|\Psi'\rangle$  on the subspace  $\mathcal{O}_j |\Psi^0\rangle$  is equal for  $j = 0 \dots N_p$

$$\langle \Psi^0 | \mathcal{O}_j | \Psi' \rangle = \langle \Psi^0 | \mathcal{O}_j | \Psi_\Lambda \rangle. \quad (57)$$

This ensures that the energy computed on  $|\Psi'\rangle$  is lower than the one computed on  $|\Psi^0\rangle$ . Then, by inserting the definition of  $|\Psi_\Lambda\rangle$  from Eq. (53) and the one of  $|\Psi'\rangle$  from Eq. (56), the stochastic reconfiguration condition becomes

$$\sum_{k=0}^{N_p} \delta p_k \langle \mathcal{O}_j \mathcal{O}_k \rangle = \langle \mathcal{O}_j (\Lambda - \mathcal{H}) \rangle. \quad (58)$$

The symbol  $\langle \dots \rangle$  stands for  $\langle \Psi^0 | \dots | \Psi^0 \rangle$ . Subsequently, the conditions for  $j=0$  and  $j \neq 0$  are separated, and the  $k=0$  term is extracted from the summation, leading to the following relations

$$\delta p_0 = - \sum_{k=1}^{N_p} \delta p_k \langle \mathcal{O}_k \rangle + \Lambda - \langle \mathcal{H} \rangle \quad \text{and} \quad \delta p_0 \langle \mathcal{O}_j \rangle + \sum_{k=1}^{N_p} \delta p_k \langle \mathcal{O}_j \mathcal{O}_k \rangle = \Lambda \langle \mathcal{O}_j \rangle - \langle \mathcal{O}_j \mathcal{H} \rangle. \quad (59)$$

The parameter  $\delta p_0$ , related to normalization, does not affect any physical observable; indeed, the two conditions can be unified in a compact form, in the unknowns  $\{\delta p_k\}$  as

$$\sum_{k=1}^{N_p} \delta p_k S_{jk} = f_j, \quad (60)$$

where  $f_j = \langle \mathcal{O}_j \rangle \langle \mathcal{H} \rangle - \langle \mathcal{O}_j \mathcal{H} \rangle$  is the generalized force and  $S_{jk} = \langle \mathcal{O}_j \mathcal{O}_k \rangle - \langle \mathcal{O}_j \rangle \langle \mathcal{O}_k \rangle$  is the covariance matrix of logarithmic derivatives of  $N_p \times N_p$  dimension (symmetric and semi-positive definite). During the Monte Carlo procedure, the expectation values are computed by stochastic sampling, the linear system is numerically solved, and, successively, parameters are updated according to the following scheme

$$\delta p_k = \sum_{j=1}^{N_p} S_{jk}^{-1} f_j, \quad p'_k = p_k^0 + \tau \delta p_k. \quad (61)$$

$\tau$  is the acceleration parameter that determines the convergence of the algorithm (in principle, each parameter could have a different  $\tau_k$ ). A large  $\tau$  guarantees rapid convergence, although it could threaten the stability of the algorithm. Hence,  $\tau$  is chosen sufficiently small to generate a variational energy that can be expanded in series

$$E(\Psi') = E(\Psi_0) + \tau \sum_{k=0}^{N_p} \frac{\partial E(\Psi_0)}{\partial p_k} \delta p_k + O(\tau^2). \quad (62)$$

Therefore, by using the definition of  $\delta p_k$  in Eq. (61) and the fact that the derivative of the energy with respect to a parameter is equal to minus the generalized force ( $\frac{\partial E(\Psi_0)}{\partial p_k} = -f_k$ ), we can rewrite the change in the variational energy as

$$\delta E = E(\Psi') - E(\Psi_0) = -\tau \sum_{j,k=1}^{N_p} S_{jk}^{-1} f_j f_k + O(\tau^2). \quad (63)$$

Here, the decrease of the variational energy ( $\delta E \leq 0$ ) is ensured by the semi-positive definition of the covariance matrix of the logarithmic derivatives  $S_{jk}$ .

Examples of calculations of logarithmic derivative operators for Jastrow-Slater wave functions can be found in Refs. [1, 30]

## 6 Correlation functions in variational Monte Carlo: Definitions and examples

In this section, we introduce a few relevant correlators, which expectation values over the optimal variational state allow us to describe the properties of our variational approximation of the ground-state wave function.

The first one is the *static structure factor*  $N(\mathbf{q})$  defined in reciprocal space as:

$$N(\mathbf{q}) = \langle n_{\mathbf{q}} n_{-\mathbf{q}} \rangle = \frac{1}{L} \sum_{R,R'} \langle n_R n_{R'} \rangle e^{i\mathbf{q} \cdot (\mathbf{R} - \mathbf{R}')}, \quad (64)$$

where  $\langle \dots \rangle$  indicates the expectation value over the variational wave function. In particular, the presence of a peak at a given  $\mathbf{q}$  vector, diverging with the system size, denotes the presence of true charge order in the system. A non-diverging peak signals the presence of short-range charge correlations that do not lead to long-range order.

Moreover, the static structure factor allows us to assess the metallic or insulating nature of the ground state by considering just ground-state expectation values, without directly calculating energy differences. Indeed, the behavior of the static structure factor at small momenta (or, equivalently, at large wavelengths) reveals the nature of the state associated with that wave function. To show this, let us recall the Feynman single-mode approximation, originally formulated in the context of collective excitations in liquid Helium [31] and then applied to fermionic systems [32]. In this context, a variational Ansatz of an excited state  $|\Psi_{\mathbf{q}}\rangle$ , with a given momentum  $\mathbf{q}$ , can be obtained by applying  $n_{\mathbf{q}}$  to the ground state wave function  $|\Psi\rangle$  (or one approximation for it), namely

$$|\Psi_{\mathbf{q}}\rangle = n_{\mathbf{q}}|\Psi\rangle. \quad (65)$$

The variational estimator of the excitation energy is then given by

$$E_{\mathbf{q}} - E_0 = \frac{\langle\Psi_{\mathbf{q}}|\mathcal{H} - E_0|\Psi_{\mathbf{q}}\rangle}{\langle\Psi_{\mathbf{q}}|\Psi_{\mathbf{q}}\rangle} = \frac{\langle\Psi|n_{-\mathbf{q}}[\mathcal{H}, n_{\mathbf{q}}]|\Psi\rangle}{\langle\Psi_{\mathbf{q}}|\Psi_{\mathbf{q}}\rangle} = \frac{\langle\Psi|[n_{-\mathbf{q}}\mathcal{H}], n_{\mathbf{q}}|\Psi\rangle}{\langle\Psi_{\mathbf{q}}|\Psi_{\mathbf{q}}\rangle}, \quad (66)$$

where  $\mathcal{H}$  is the Hubbard Hamiltonian. The sum of both commutators  $(n_{-\mathbf{q}}\mathcal{H}n_{\mathbf{q}} - n_{-\mathbf{q}}n_{\mathbf{q}}\mathcal{H}) + (n_{-\mathbf{q}}\mathcal{H}n_{\mathbf{q}} - \mathcal{H}n_{-\mathbf{q}}n_{\mathbf{q}})$  is equivalent to the double commutator

$$[n_{-\mathbf{q}}, [\mathcal{H}, n_{\mathbf{q}}]] = n_{-\mathbf{q}}(\mathcal{H}n_{\mathbf{q}} - n_{\mathbf{q}}\mathcal{H}) - (\mathcal{H}n_{\mathbf{q}} - n_{\mathbf{q}}\mathcal{H})n_{-\mathbf{q}} \quad (67)$$

due to the inversion symmetry  $\mathbf{q} \leftrightarrow -\mathbf{q}$ . Therefore, the excitation energy is given by

$$E_{\mathbf{q}} - E_0 = \frac{1}{2} \frac{\langle\Psi|[n_{-\mathbf{q}}, [\mathcal{H}, n_{\mathbf{q}}]]|\Psi\rangle}{N_{\mathbf{q}}}, \quad (68)$$

where  $N(\mathbf{q}) = \langle\Psi|n_{-\mathbf{q}}n_{\mathbf{q}}|\Psi\rangle$  is the static structure factor for the ground state. The double commutator  $[n_{-\mathbf{q}}, [\mathcal{H}, n_{\mathbf{q}}]]$  can be evaluated and involves the kinetic term only, since the potential term of the Hamiltonian contains density-density interactions that commute with  $n_{\mathbf{q}}$  [21]

$$[n_{-\mathbf{q}}, [\mathcal{H}, n_{\mathbf{q}}]] = \frac{1}{L} \sum_{\mathbf{k}, \sigma} (\varepsilon_{\mathbf{k}+\mathbf{q}} + \varepsilon_{\mathbf{k}-\mathbf{q}} - 2\varepsilon_{\mathbf{k}}) c_{\mathbf{k}, \sigma}^{\dagger} c_{\mathbf{k}, \sigma}, \quad (69)$$

where  $\varepsilon_{\mathbf{k}}$  is the dispersion relation of the kinetic part of the Hubbard Hamiltonian. Then, in the limit of small momenta, one gets

$$E_{\mathbf{q}} - E_0 \propto \lim_{q \rightarrow 0} \frac{|\mathbf{q}|^2}{N(\mathbf{q})}. \quad (70)$$

Then, whenever  $N(\mathbf{q})$  is linear in  $\mathbf{q}$ , the energy spectrum is gapless, suggesting that the system is metallic; by contrast, if  $N(\mathbf{q}) \sim \mathbf{q}^2$ , this construction gives a finite gap, leading to an insulator. As an example, we show in Fig. 3 the behavior of  $N(\mathbf{q})$  across the metal-insulator transition that occurs in the single-orbital Hubbard model as a function of  $U/t$  at  $t'/t = 0.75$ . In order to highlight the location of the metal-insulator transition we plot  $N(\mathbf{q})/|\mathbf{q}|$ : If the system is metallic,  $N(\mathbf{q})/|\mathbf{q}|$  extrapolates to a finite value for  $|\mathbf{q}| \rightarrow 0$ ; conversely, if the system is insulating,  $N(\mathbf{q})/|\mathbf{q}|$  extrapolates to zero for  $|\mathbf{q}| \rightarrow 0$ ; see Eq. (70).

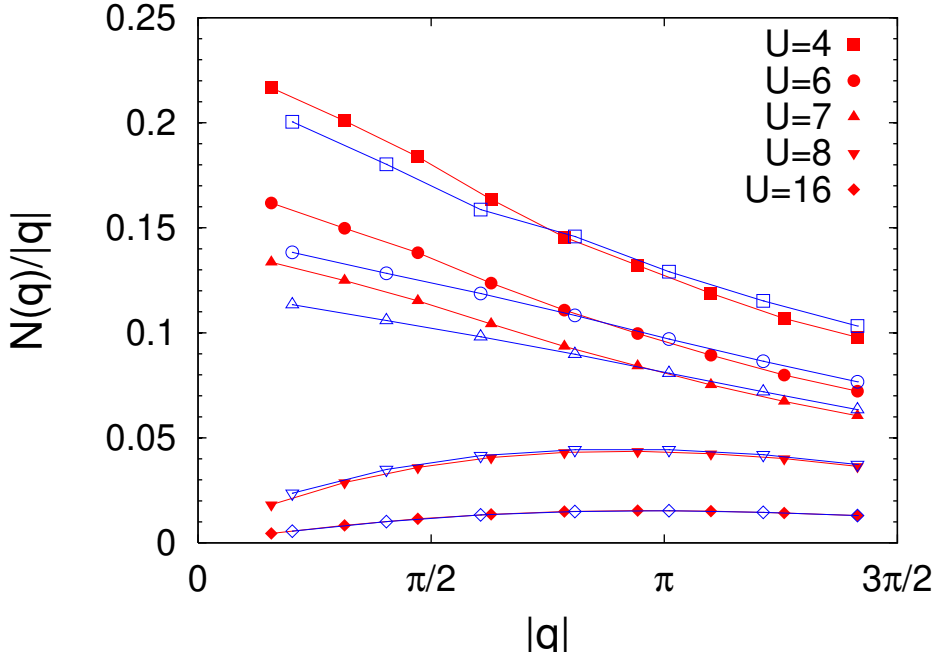


**Fig. 3:** Variational results for  $N(\mathbf{q})$  divided by  $|\mathbf{q}|$ , as a function of  $|\mathbf{q}|$ , for  $L = 98$  (empty symbols) and for  $L = 162$  (full symbols) at  $t'/t = 0.75$ . The metal-insulator transition takes place between  $U/t = 7$  and  $U/t = 8$ , that is, where  $N(\mathbf{q})$  changes its behavior from linear to quadratic for  $|\mathbf{q}| \rightarrow 0$  [20].

The second one is the spin-spin correlator. Analogously to the case of charge order, spin order in the system is associated to a peak in the *spin-spin correlations* defined as

$$S(\mathbf{q}) = \frac{1}{L} \sum_{R,R'} \langle S_R^z S_{R'}^z \rangle e^{i\mathbf{q} \cdot (\mathbf{R} - \mathbf{R}')}, \quad (71)$$

where  $S_R^z$  is the spin operator along the  $z$  direction, i.e.,  $S_R^z = 1/2(c_{R,\uparrow}^\dagger c_{R,\uparrow} - c_{R,\downarrow}^\dagger c_{R,\downarrow})$ . Also, the presence of a spin gap can be detected by looking at the behavior of  $S(\mathbf{q})$  at small momenta. As an example, we show the behavior of  $S(\mathbf{q})$  for the one dimensional Hubbard model. The model is the same as the one we introduced in Eq. (1), just denoting nearest neighbor hopping as  $t_1$  and next-nearest neighbor hopping as  $t_2$ . Even if in one dimension there is no long-range order, short-range spin correlations are present in the insulating phase. They can be detected as a non-diverging peak in  $S(\mathbf{q})$ . The peak remains commensurate at  $Q = \pi$  up to  $t_2/t_1 \sim 0.75$  and then shifts, as a function of the ratio  $t_2/t_1$ , to the value of  $Q = \pi/2$  for large values of  $t_2/t_1$ . Results can be seen in Fig. 4.

Then, another crucial task in ground-state simulations is the detection of superconductivity in the model. The presence of finite BCS parameters in the variational state after optimization is a necessary condition for the presence of superconductivity. It is not a sufficient condition because either the system can be an insulator, induced by the long-range tail of the Jastrow factor in the variational state, or finite BCS parameters can be a finite-size effect, disappearing in the thermodynamic limit, as it occurs in the  $U = 0$  limit of the Hubbard model. Hence, to reveal the presence of true superconductivity in the state, we can compute the so-called *singlet*

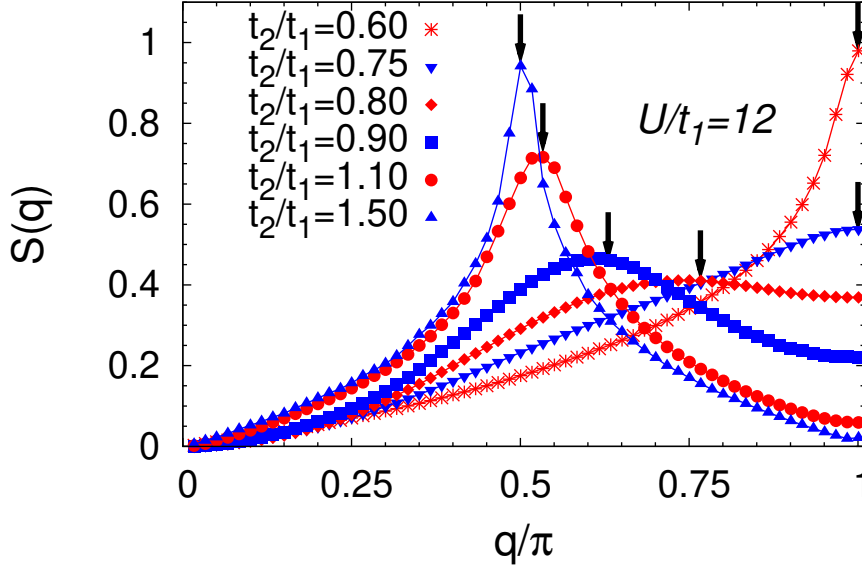


**Fig. 4:** Spin-spin correlations  $S(\mathbf{q})$  at  $U/t_1 = 12$  for different values of the hopping ratio  $t_2/t_1$ . Arrows indicate the location of the peaks. Data are shown on a  $L = 120$  lattice size [33].

pairing correlation function in real space, that we restrict here to be intra-orbital

$$D_\alpha(r) = \frac{1}{L} \sum_R \langle P_{\alpha,R} P_{\alpha,R+r}^\dagger \rangle, \quad (72)$$

where  $\langle \dots \rangle$  means the usual expectation value on the variational ground state,  $\alpha$  is the orbital index, and the singlet operator is defined as:  $P_{\alpha,R} = c_{\alpha,R+\eta,\downarrow} c_{\alpha,R,\uparrow} - c_{\alpha,R+\eta,\uparrow} c_{\alpha,R,\downarrow}$ . In practice, the  $P_{\alpha,R}$  operator destroys a singlet oriented along the  $\eta$  direction (for instance,  $\eta = x$  or  $\eta = y$ ) in position  $R$ , while  $P_{\alpha,R+r}^\dagger$  creates a singlet with the same orientation in position  $R+r$ . Then, the system is superconductive when the following condition is met:

$$\phi_\alpha^2 = \lim_{r \rightarrow \infty} D_\alpha(r) \text{ is finite.} \quad (73)$$

The symbol  $\phi_\alpha$  denotes the superconductive order parameter, that, for  $|\phi_\alpha| \neq 0$  indicates that the system is superconductive.

From a practical point of view, on a finite lattice, we evaluate  $D_\alpha(r)$  at the maximum distance, i.e.,  $r = l/2$  (if the lattice size is defined as  $L = l \times l$ ) because of the periodic boundary conditions. Then, we consider different lattice sizes and make an extrapolation of  $\phi_\alpha$  to the thermodynamic limit ( $l \rightarrow \infty$ ), in order to assess the actual presence of superconductivity. Moreover, it might be useful to compare the superconductive correlation functions for finite values of  $U$  with the ones obtained at  $U = 0$ . In the latter case, finite-size kinetic contributions from the hopping are present. Hence, superconductivity emerges only when we observe an increase of  $\phi_\alpha$  in the presence of interaction, with respect to the  $U = 0$  case.

As an example, we show in Fig. 5, left panels, the results for  $\phi_\alpha^2$  as a function of  $U$  in the three-orbital Hubbard-Kanamori Hamiltonian of Eqs. (2) and (3). Results are presented for two relevant electronic densities of  $n = 3 + 1/3$  and  $n = 3 + 2/3$ . In both cases, superconductivity

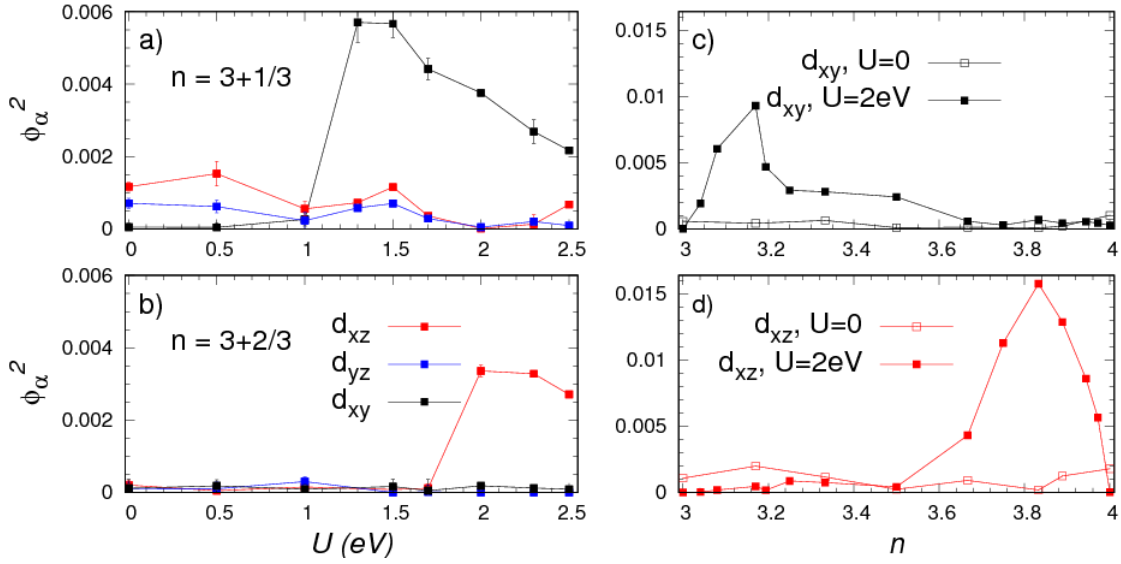


**Fig. 5:** Square of the superconducting order parameter  $\phi_\alpha^2$  in each orbital as a function of  $U$  at  $n = 3 + 1/3$  [panel a)] and  $n = 3 + 2/3$  [panel b)]. Data are for the  $L = 18 \times 18$  cluster. Square of the superconducting order parameter in the  $d_{xy}$  [panel c)] and  $d_{xz}$  [panel d)] orbitals as a function of  $n$  at  $U = 2$  eV (full squares) and at  $U = 0$  (empty squares). Data are for the  $L = 12 \times 12$  cluster. All data are computed at  $J = 0.2 U$  [25].

develops only in one of the three orbitals, while the other two orbitals, instead, do not show any enhancement of the superconducting correlations with respect to the  $U = 0$  case. We also show the behavior of  $\phi_\alpha^2$  as a function of density  $n$  in Fig. 5, right panels, for a value of  $U$  where superconductivity is present in one of the orbitals. Here, we also report the non-interacting values at  $U = 0$ , to emphasize the relevance of the Hubbard interaction to generate the electron pairing. Even if not shown here, we mention that superconductivity is not present when the Hund's coupling is set to zero.

As a comparison, in Fig. 6, we report the intra-orbital pairing correlations  $D_\alpha(r)$  at densities  $n = 3 + 1/3$  and  $n = 3 + 2/3$ , both within the orbital-selective phase and at  $U = 0$ . In both cases, there is one orbital where pairing correlations at large distances are strongly enhanced going from  $U = 0$  to large values of  $U$ . On the contrary, in the other two orbitals, pairing correlations remain similar or are even suppressed when electron-electron correlation is switched on.

We mention, that in addition to the correlators introduced above, one can also compute expectation values of simpler single-particle observables. For instance, in a multi-orbital model, one could compute the electronic density per orbital, defined as  $n_\alpha = \frac{1}{L} \langle \sum_{R,\sigma} n_{R,\alpha,\sigma} \rangle$ . This observable allows us to identify orbital selective phenomena, that are related to superconductivity. In Fig. 7, we observe a differentiation in the orbital occupation of the two degenerate orbitals ( $d_{xz}$  and  $d_{yz}$ ) above a critical value of  $U$  that corresponds to the onset of superconductivity in the most occupied orbital. We remark that the onset of superconductivity and orbital selectivity is related to the presence of a long-range tail in the Jastrow term of the variational wave function, that allows for a proper treatment of spatial correlations.

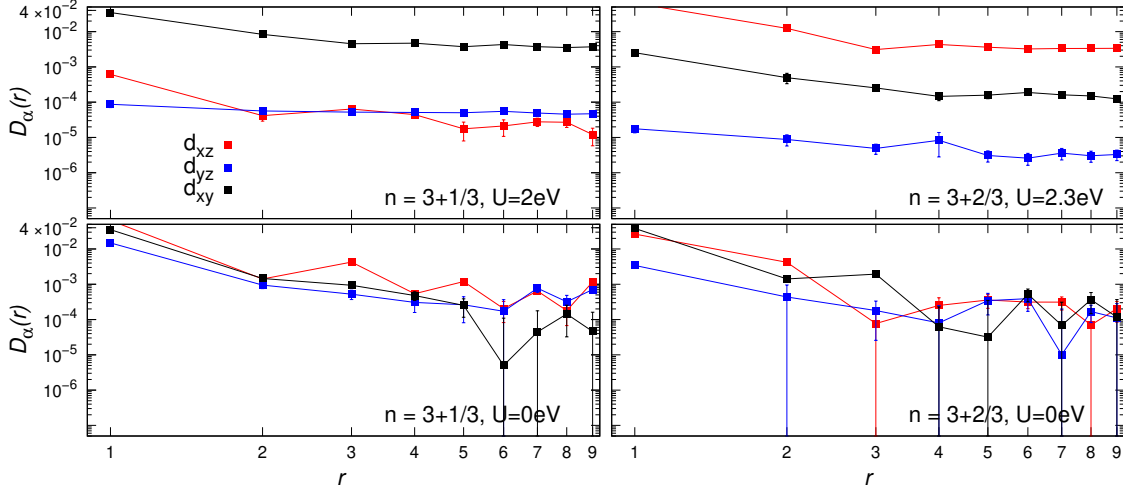


**Fig. 6:** Upper panels:  $D_\alpha(r)$  in all orbitals, for a value of  $U$  where superconductivity is present. Data are shown at electronic densities  $n = 3+1/3$  (left column) and  $n = 3+2/3$  (right column). Lower panels: The same as above, but for  $U = 0$ . Data are computed on a  $L = 18 \times 18$  cluster at  $J = 0.2U$  [25].

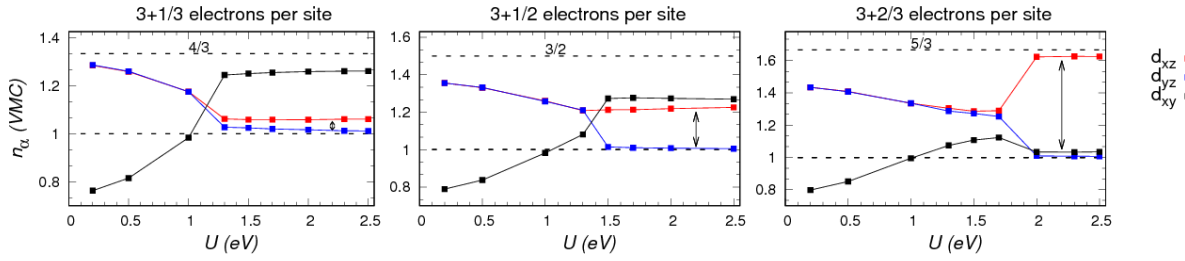


**Fig. 7:** Electronic density per orbital  $n_\alpha$  as a function of  $U$  for three different total electronic densities. We fix  $J/U = 0.2$  and compute data for the  $L = 12 \times 12$  cluster. The arrows highlight the size of the differentiation between the  $d_{xz}$  and the  $d_{yz}$  orbitals [25].

## 7 Final remarks

In this lecture, we introduced the basics of the variational Monte Carlo method in general and in relation with the Hubbard model, both in its simpler single-orbital version and in its multi-orbital extension.

Variational Monte Carlo is a conceptually simple approach, where the computation of physical observables over a variational approximation of the ground state wave function is rewritten as averages of local observables over the usual quantum mechanical probability distribution. This definition ensures that VMC has no sign problem, since we sample a nonnegative quantity.

The core of the variational Monte Carlo approach is to define accurate variational states that can reproduce the correct ground state physics; for this reason, we devoted Sec. 4 to the definition of different variational states. We described the Jastrow-Slater wave function for the single-orbital Hubbard model and how it can be improved with the backflow terms. We also presented an extension of the Jastrow-Slater wave function to the multi-orbital case.

Then, once a functional form of the variational wave function is built, it is crucial to be able to optimize the variational parameters that are contained within the variational state. The concomitant optimization of many variational parameters is made possible with the stochastic re-configuration method, described in Sec. 5.

Finally, we introduced several many-body correlators that allow us to characterize the physics of the optimal variational approximation of the ground state. Among the other, we discussed how to discriminate whether the system is metallic or insulating, detect the presence of charge or spin order and the emergence of superconductivity. All these features lead to a complete understanding of the ground-state properties of our correlated Hubbard Hamiltonian.

We close this section by commenting on which paths might variational Monte Carlo take in the future.

A first one is based on the application of relatively simple variational states, of the Jastrow-Slater form, to complicated multi-orbital Hamiltonians where approaches that are nonperturbative in electronic correlations and can address spatial correlations are lacking. These ideas have been already discussed in these lecture notes, in the field of iron-based superconductors, and can have a relevant impact, for instance, in the field of Kagome metals. Here, electronic correlations are a necessary ingredient, while the complicated kinetic electronic model is a challenge for all the contemporary many-body methods [34].

A second one is the improvement of variational wave functions with respect to the scheme introduced in this lecture. In this respect, neural-network quantum states, based on the transformer architecture, show excellent accuracy and appear as promising tools for solving the intricacies of competing orders in the Hubbard model. However, the domain of applicability of these more evolved variational wave functions is, at the moment, limited to the single-orbital Hubbard model on lattices with a simple geometry [35–37].

Finally, a third one that goes away from the target of these notes is to consider not only ground-state properties, but to provide VMC calculations of dynamics and finite-temperature properties, that involve also excited states. This is a wide and growing field, based on the time-dependent variational principle, which might be the subject of another set of lecture notes.

## References

- [1] F. Becca and S. Sorella: *Quantum Monte Carlo Approaches for Correlated Systems* (Cambridge University Press, 2017)
- [2] U. Schollwöck, *Ann. Phys.* **326**, 96 (2011)
- [3] A. Georges, G. Kotliar, W. Krauth, and M.J. Rozenberg, *Rev. Mod. Phys.* **68**, 13 (1996)
- [4] S. Sorella, *Phys. Rev. B* **71**, 241103 (2005)
- [5] J.P.F. LeBlanc *et al.*, *Phys. Rev. X* **5**, 041041 (2015)
- [6] M. Qin, T. Schäfer, S. Andergassen, P. Corboz, and E. Gull, *Annu. Rev. Condens. Matter Phys.* **13**, 275 (2022)
- [7] E. Pavarini, I. Dasgupta, T. Saha-Dasgupta, O. Jepsen, and O. Andersen, *Phys. Rev. Lett.* **87**, 047003 (2001)
- [8] E.H. Lieb, and F.Y. Wu, *Phys. Rev. Lett.* **20**, 1445 (1968)
- [9] S. Zhang, J. Carlson, and J.E. Gubernatis, *Phys. Rev. Lett.* **74**, 3652 (1995)
- [10] J. Kanamori, *Progr. Theor. Phys.* **30**, 275 (1963)
- [11] L. Georges, L. de Medici and J. Mravlje, *Annu. Rev. Condens. Matter Phys.* **4**, 137 (2013)
- [12] M. Daghofer, A. Nicholson, A. Moreo, and E. Dagotto, *Phys. Rev. B* **81**, 014511 (2010)
- [13] L. Fanfarillo, A. Valli, and M. Capone, *Phys. Rev. Lett.* **125**, 177001 (2020)
- [14] B.-X. Zheng *et al.*, *Science* **358**, 1155 (2017)
- [15] A.S. Darmawan, Y. Nomura, Y. Yamaji, and M. Imada, *Phys. Rev. B* **98**, 205132 (2018)
- [16] V. Marino, F. Becca, and L.F. Tocchio, *SciPost Phys.* **12**, 180 (2022)
- [17] M. Capello, F. Becca, M. Fabrizio, S. Sorella, and E. Tosatti, *Phys. Rev. Lett.* **94**, 026406 (2005)
- [18] P.W. Anderson, *Science* **235**, 1196 (1987)
- [19] M.C. Gutzwiller, *Phys. Rev. Lett.* **10**, 159 (1963)
- [20] L.F. Tocchio, F. Becca, A. Parola, and S. Sorella, *Phys. Rev. B* **78**, 041101(R) (2008)
- [21] L.F. Tocchio, F. Becca, and C. Gros, *Phys. Rev. B* **83**, 195138 (2011)
- [22] R. Chitra, S. Pati, H.R. Krishnamurty, D. Sen, and S. Ramasesha, *Phys. Rev. B* **52**, 6581 (1995)

- [23] L.F. Tocchio, H. Feldner, F. Becca, R. Valenti, and C. Gros, Phys. Rev. B **87**, 035143 (2013)
- [24] D. Luo and B.K. Clark, Phys. Rev. Lett. **122**, 226401 (2019)
- [25] V. Marino, A. Scazzola, F. Becca, M. Capone, and L.F. Tocchio, Phys. Rev. Lett. **134**, 196502 (2025)
- [26] V. Marino: *Superconductivity and Orbital Selectivity in a three-orbital Hubbard model for the iron-based superconductors – Variational Monte Carlo analysis* (Doctoral Thesis, 2025) <http://iris.polito.it/handle/11583/3001278>
- [27] T. Misawa and M. Imada, Nat. Comm. **5**, 5738 (2014)
- [28] C.J. Umrigar, K.G. Wilson, and J.W. Wilkins, Phys. Rev. Lett. **60**, 1719 (1988)
- [29] M. Snajdr and S.M. Rothstein, J. Chem. Phys. **112**, 4935(2000)
- [30] S. Yunoki and S. Sorella, Phys. Rev. B **74**, 014408 (2006)
- [31] R.P. Feynman, Phys. Rev. **94**, 262 (1954)
- [32] A.W. Overhauser, Phys. Rev. B **3**, 1888 (1971)
- [33] L.F. Tocchio, F. Becca, and C. Gros, Phys. Rev. B **81**, 205109 (2010)
- [34] D. Di Sante, T. Neupert, G. Sangiovanni, R. Thomale, R. Comin, I. Zeljkovic, J.G. Checkelsky, and S.D. Wilson, Rev. Mod. Phys. **98**, 015002 (2026)
- [35] J.R. Moreno, G. Carleo, A. Georges, and J. Stokes, Proc. Natl. Acad. Sci. USA **119**, e2122059119 (2022)
- [36] R. Rende, A. Nikolaenko, L.L. Viteritti, S. Sachdev, and Y. -Hui Zhang, arXiv:2603.02316 (2026)
- [37] L.L. Viteritti, R. Rende, C. Roth, A. Sengupta, G. Carleo, and A. Georges, arXiv:2604.21978 (2026)