5

Variational Monte Carlo

5.1 Quantum Averages and Statistical Samplings

In this chapter, we discuss the general framework in which the variational Monte Carlo methods are defined and few important implementations for interacting systems of bosons and fermions on the lattice. The main advantage of considering this approach relies on the variational principle that has been shown in section 1.4: the energy of a given quantum state is always bounded from below by the exact ground-state one, giving us the route to obtain the best possible solution to the problem. In most cases, a suitable parametrization of the variational wave function allows us to consider a wide range of different quantum phases (e.g., metals, superconductors, and insulators). By performing an optimization of the parameters, we can reach the lowest-energy state, which is expected to capture the correct ground-state behavior. Therefore, variational wave functions represent a flexible and valuable approach to get important insights into the low-energy properties of models that cannot be solved by exact methods. By contrast, the main limitation of this approach is the fact that it is based on a given *Ansatz*, which may contain a relevant bias that cannot be removed within the chosen parametrization.

We also mention that variational wave functions can be easily defined and treated for a wide class of models, irrespective of the range of interactions and the dimension of the local (i.e., single-site) Hilbert space, which can be even infinite and does not need the use of an uncontrolled cutoff to work with a finite-dimensional space. In this respect, variational Monte Carlo is better than other methods, like density-matrix renormalization group or tensor-network methods in which the complexity dramatically increases with both the range of the interaction and the dimension of the local Hilbert space (White, 1992; Schollwöck, 2005, 2011).

Let us start by describing the general framework in which variational Monte Carlo methods are defined. First of all, we fix a complete basis set $\{|x\rangle\}$ in the

https://doi.org/10.1017/9781316417041.006 Published online by Cambridge University Press

Hilbert space, in which (for simplicity) the states are taken to be *orthogonal* and *normalized* such that:

$$\sum_x |x\rangle\langle x| = \mathbb{I}. \tag{5.1}$$

Then, any quantum state $|\Psi\rangle$ can be written as:

$$|\Psi\rangle = \sum_x |x\rangle\langle x|\Psi\rangle = \sum_x \Psi(x)|x\rangle. \tag{5.2}$$

In turn, the expectation value of an operator $\mathcal{O}$ over a given variational wave function $|\Psi\rangle$ takes the following form:

$$\langle\mathcal{O}\rangle = \frac{\langle\Psi|\mathcal{O}|\Psi\rangle}{\langle\Psi|\Psi\rangle} = \frac{\sum_x \langle\Psi|x\rangle\langle x|\mathcal{O}|\Psi\rangle}{\sum_x \langle\Psi|x\rangle\langle x|\Psi\rangle}. \tag{5.3}$$

The main problem in evaluating the expectation value is that the number of configurations in the sum is exponentially large with the number of particles. Although the dimension of the Hilbert space can be slightly reduced by employing few conservation laws (e.g., the conservation of the total number of particles and/or the total spin component along the quantization axis), it still remains exponentially large with the number of particles. Therefore, for large systems, it is impossible to perform an exact enumeration of the configurations to compute $\langle\mathcal{O}\rangle$ exactly. Nevertheless, Eq. (5.3) can be recast into a form that can be easily treated by standard Monte Carlo methods. Indeed, we have that:

$$\langle\mathcal{O}\rangle = \frac{\sum_x |\langle\Psi|x\rangle|^2 \frac{\langle x|\mathcal{O}|\Psi\rangle}{\langle x|\Psi\rangle}}{\sum_x |\langle\Psi|x\rangle|^2} = \frac{\sum_x |\Psi(x)|^2 \mathcal{O}_L(x)}{\sum_x |\Psi(x)|^2}, \tag{5.4}$$

where we have defined the *local estimator* of the operator $\mathcal{O}$:

$$\mathcal{O}_L(x) = \frac{\langle x|\mathcal{O}|\Psi\rangle}{\langle x|\Psi\rangle}. \tag{5.5}$$

The important point is that

$$\mathcal{P}(x) = \frac{|\Psi(x)|^2}{\sum_x |\Psi(x)|^2} \tag{5.6}$$

can be interpreted as a probability, since it is a non-negative quantity for all configurations $|x\rangle$ and is normalized, i.e., $\sum_x \mathcal{P}(x) = 1$. Therefore, the problem of computing a quantum average of the operator $\mathcal{O}$ can be rephrased into the calculation of the average of the random variable $\mathcal{O}_L(x)$ of Eq. (5.5) over the distribution probability $\mathcal{P}(x)$ given by Eq. (5.6). In particular, if we consider

https://doi.org/10.1017/9781316417041.006 Published online by Cambridge University Press

the expectation value of the Hamiltonian, the local estimator corresponds to the so-called *local energy*, which is defined by:

$$e_L(x) = \frac{\langle x|\mathcal{H}|\Psi\rangle}{\langle x|\Psi\rangle}. \tag{5.7}$$

In summary, it is possible to define a stochastic algorithm (e.g., a Markov process) in which a sequence of configurations $\{|x_n\rangle\}$ is generated (for example by using the Metropolis algorithm described in section 3.9). Then, after an equilibration time, they are distributed according to the desired probability $\mathcal{P}(x)$. Then, the quantum expectation value $\langle\mathcal{O}\rangle$ is evaluated from the mean value of the random variable $\mathcal{O}_L(x)$ over the visited configurations:

$$\langle\mathcal{O}\rangle \approx \frac{1}{N}\sum_{n=1}^{N}\mathcal{O}_L(x_n). \tag{5.8}$$

Finally, errorbars can be computed as described in section 3.11.

5.2 The Zero-Variance Property

An important feature of the variational Monte Carlo approach is the *zero-variance property*. Let us suppose that the variational state $|\Psi\rangle$ coincides with an exact eigenstate of $\mathcal{H}$ (not necessarily the ground state), namely $\mathcal{H}|\Psi\rangle = E|\Psi\rangle$. Then, it follows that the local energy $e_L(x)$ is constant:

$$e_L(x) = \frac{\langle x|\mathcal{H}|\Psi\rangle}{\langle x|\Psi\rangle} = E\frac{\langle x|\Psi\rangle}{\langle x|\Psi\rangle} = E. \tag{5.9}$$

Therefore, the random variable $e_L(x)$ does not depend on $|x\rangle$, which immediately implies that its variance is zero, while its mean value E coincides with the exact eigenvalue (in other words, $e_L(x)$ is not a random variable). Clearly, this is an extreme case that is very rare for generic correlated models. However, in general, the variance of $e_L(x)$ will decrease its value whenever the variational state $|\Psi\rangle$ will approach an exact eigenstate. This fact is very important to reduce the statistical fluctuations and improve the numerical efficiency. 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.

Finally, we would like to notice that the average of the square of the local energy corresponds to the quantum average of the Hamiltonian squared:

$$\frac{\langle\Psi|\mathcal{H}^2|\Psi\rangle}{\langle\Psi|\Psi\rangle} = \frac{\sum_x\langle\Psi|\mathcal{H}|x\rangle\langle x|\mathcal{H}|\Psi\rangle}{\sum_x\langle\Psi|x\rangle\langle x|\Psi\rangle} = \frac{\sum_x|\Psi(x)|^2|e_L(x)|^2}{\sum_x|\Psi(x)|^2}. \tag{5.10}$$

https://doi.org/10.1017/9781316417041.006 Published online by Cambridge University Press

Thus, the variance of the random variable $e_L(x)$ is exactly equal to the quantum variance of the Hamiltonian over the variational state $|\Psi\rangle$:

$$\sigma^2_{e_L} = \frac{\langle\Psi|(\mathcal{H}-E)^2|\Psi\rangle}{\langle\Psi|\Psi\rangle}. \tag{5.11}$$

5.3 Jastrow and Jastrow-Slater Wave Functions

In the Monte Carlo evaluation of quantum averages, see Eq. (5.4), we must compute:

- The ratio of probabilities with different configurations, which implies the ratio of overlaps between the given variational state and two configurations of the basis set:

$$\frac{\mathcal{P}(x')}{\mathcal{P}(x)} = \left|\frac{\langle x'|\Psi\rangle}{\langle x|\Psi\rangle}\right|^2, \tag{5.12}$$

  as required in the Metropolis algorithm.
- The local estimator $\mathcal{O}_L(x)$, which, in turn, implies ratios of overlaps and matrix elements of the observable between states of the basis set. For example, when considering the energy, we have:

$$e_L(x) = \frac{\langle x|\mathcal{H}|\Psi\rangle}{\langle x|\Psi\rangle} = \sum_{x'}\langle x|\mathcal{H}|x'\rangle\frac{\langle x'|\Psi\rangle}{\langle x|\Psi\rangle}. \tag{5.13}$$

Naively, the computation of the local estimator looks a tremendously hard task, since it requires a summation over all the states of the many-body Hilbert space; however, thanks to the locality of the Hamiltonian (or, similarly, any other local operator or correlation function), only few terms actually contribute to the sum. Indeed, given the configuration $|x\rangle$, the matrix element $\langle x|\mathcal{H}|x'\rangle$ is non-zero only for $O(L)$ configurations $|x'\rangle$. As an example, let us consider the fermionic Hubbard model: by using the local basis, $|x\rangle$ is connected only to few other configurations that differ for 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 2 (due to the spin). Therefore, the computation of the local estimator only requires a small number of operations, usually proportional to the number of sites/particles.

As we have seen, the building block of the variational Monte Carlo approach is the computation of $\langle x|\Psi\rangle$, which is the amplitude of the variational state over a generic element of the basis set. More precisely, along all the Markov process, only ratios of these overlaps must be computed. This calculation must be done for each configuration that is visited along the Markov process and, therefore, it must be done as fast as possible. This fact imposes some constraint on the form

https://doi.org/10.1017/9781316417041.006 Published online by Cambridge University Press

of the variational wave function. Usually, fermionic states require the calculation of determinants or Pfaffians, while bosonic wave functions require permanents (the definition of the permanent of a matrix $\mathbf{M}$ differs from that of the determinant in the fact that the signatures of the permutations are not taken into account). Fortunately, there are fast (i.e., polynomial) algorithms to evaluate determinants and Pfaffians, thus allowing us to consider these states as variational *Ansätze* for electron systems; by contrast, the calculation of permanents remains an exponentially-hard problem, which strongly limits the number of bosons that can be handled in a reasonable computation time. Nevertheless, the wave functions in which all bosons are condensed in a single state do not need such a calculation and can be considered efficiently in variational Monte Carlo approaches.

Within the variational technique that has been described before, it is possible to treat a large variety of fermionic and bosonic states, which may interpolate between weak and strong correlation regimes. In the following, we would like to give some detailed description on how to construct the variational states and how to compute the building blocks that are necessary along the Markov process. We will consider the Jastrow (bosonic) or Jastrow-Slater (fermionic) states that have been introduced and discussed in Chapter 1:

$$|\Psi_J\rangle = \mathcal{J}|\Phi_0\rangle, \tag{5.14}$$

where $|\Phi_0\rangle$ is a generic uncorrelated (bosonic or fermionic) state and $\mathcal{J}$ is the Jastrow factor, which takes into account the electron correlation. A particularly simple but important case is given by the case where $\mathcal{J}$ contains density-density correlations:

$$\mathcal{J} = \exp\left(-\frac{1}{2}\sum_{i,j} v_{i,j} n_i n_j\right), \tag{5.15}$$

where the $v_{i,j}$'s can be taken as variational parameters, whose total number is L^2. Here, the pseudo-potential $v_{i,j}$ couples densities at different sites i and j and not density fluctuations, as in Eq. (1.65); for systems with conserved number of particles, the two definitions of the Jastrow factors coincide, apart for an irrelevant multiplicative factor. For a translational invariant model, $v_{i,j}$ only depends upon the distance between i and j, thus the number of parameters can be reduced to $O(L)$. We would like to remark that, within the Monte Carlo approach, it is possible to treat exactly (but still having statistical errors) the limit of singular Jastrow factors with $v_{i,i} = \infty$, e.g., the Gutzwiller projector that eliminates all multiply occupied sites (this case being relevant for an infinite Hubbard-U interaction). Indeed, this case can be easily incorporated by building a Markov chain where only configurations $|x\rangle$ that satisfy this constraint are visited.

https://doi.org/10.1017/9781316417041.006 Published online by Cambridge University Press

The advantage of considering the wave function of Eq. (5.14) in the variational Monte Carlo technique is that the calculations can be extremely efficient and fast. Indeed, whenever the Jastrow factor is diagonal in the chosen basis, we have that:

$$\langle x|\Psi_J\rangle = \mathcal{J}(x)\langle x|\Phi_0\rangle, \tag{5.16}$$

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 for the Jastrow term of Eq. (5.15). We emphasize that, in order to have a polynomial algorithm, the Jastrow factor must only contain operators that are diagonal in the basis $|x\rangle$, otherwise $\mathcal{J}|x\rangle$ would generate an exponentially large number of states, ruling out any calculation on large systems. In addition, $\langle x|\Phi_0\rangle$ can be also easily evaluated in $O(L^3)$ operations (for fermions) or $O(1)$ operations (for condensed bosons), see sections 5.5, 5.6, and 5.7.

5.4 The Choice of the Basis Sets

In most cases (but this is not mandatory), the basis set is chosen to be a product state in the real space. On a lattice with L sites, we must first give an ordering to them, $i = 1, \ldots, L$ and then take a basis set on each lattice site $\{|\xi\rangle_i\}$, whose dimension depends upon the model under consideration. For example, in the single-band Hubbard model $|\xi\rangle_i$ can be any of the following four states:

$$|0\rangle_i, \quad c^\dagger_{i,\uparrow}|0\rangle_i, \quad c^\dagger_{i,\downarrow}|0\rangle_i, \quad c^\dagger_{i,\uparrow}c^\dagger_{i,\downarrow}|0\rangle_i, \tag{5.17}$$

namely, the empty state on the site i, the two singly occupied states (with spin up or down along the quantization axis z), or the doubly occupied state.

For the bosonic Hubbard model, the local Hilbert space is infinite with an arbitrary number of bosons on the same site:

$$|0\rangle_i, \quad b^\dagger_i|0\rangle_i, \quad \frac{1}{\sqrt{2!}}(b^\dagger_i)^2|0\rangle_i, \quad \frac{1}{\sqrt{3!}}(b^\dagger_i)^3|0\rangle_i, \quad \ldots \tag{5.18}$$

Finally, for the spin-S Heisenberg model, the local basis can be taken as the $2S+1$ states with $S^z_i = -S, \ldots, S$:

$$|-S\rangle_i, \quad |-S+1\rangle_i, \quad \ldots, \quad |S-1\rangle_i, \quad |S\rangle_i. \tag{5.19}$$

Once the local Hilbert space has been fixed, a generic element of the basis set of the whole lattice $\{|x\rangle\}$ is given by the product state:

$$|x\rangle = |\xi\rangle_1 \ldots |\xi\rangle_L. \tag{5.20}$$

https://doi.org/10.1017/9781316417041.006 Published online by Cambridge University Press

For the subsequent use, we also define the vacuum of the whole lattice as:

$$|0\rangle = |0\rangle_1 \dots |0\rangle_L. \tag{5.21}$$

The advantage of taking a local basis, defined on each site of the lattice, comes from the fact that the Hamiltonian is usually local in space and its matrix elements are easily computed: the interaction terms are often diagonal, like for example the Hubbard-U interaction, or local, like the Hund's coupling in multi-band models, while the hopping terms are usually short-range, resulting in a small number (i.e., proportional to the number of particles) of non-vanishing matrix elements for each many-body configuration $|x\rangle$.

Within the Monte Carlo sampling (i.e., along the Markov process), it is convenient to generate the new (proposed) configuration among the ones that are obtained by applying the Hamiltonian. In this way, the generic configuration $|x_n\rangle$ at step n may differ from the one of the basis set by an overall sign (for fermionic systems). A similar issue is present when computing local estimators $\mathcal{O}_L(x)$, such as the local energy (5.7). This fact does not represent a problem, provided we bookkeep the sign change that may appear. Let us consider N_e spinless electrons on L sites, the configuration

$$|x_n\rangle = c^\dagger_{R_1} \dots c^\dagger_{R_l} \dots c^\dagger_{R_{N_e}} |0\rangle \tag{5.22}$$

can be described by a vector $\boldsymbol{\kappa}$ of length L, whose i-th element, if non-zero, gives the position (from left to right) of the corresponding creation operator in the above string defining the configuration. For example, by taking $N_e = 3$ and $L = 6$, the configuration $c^\dagger_3 c^\dagger_1 c^\dagger_6 |0\rangle$ is associated to the vector $(2, 0, 1, 0, 0, 3)$. Then, the application of the hopping term $c^\dagger_{R'_l} c_{R_l}$ gives:

$$|x_{n+1}\rangle = c^\dagger_{R_1} \dots c^\dagger_{R'_l} \dots c^\dagger_{R_{N_e}} |0\rangle; \tag{5.23}$$

showing that the only changes in the string appear at the position R_l and R'_l. In the example above, by applying $c^\dagger_2 c_1$, we get $c^\dagger_3 c^\dagger_2 c^\dagger_6 |0\rangle$ and the vector $\boldsymbol{\kappa}$ becomes $(0, 2, 1, 0, 0, 3)$. Along the Markov chain, it is useful to store and bookkeep $\boldsymbol{\kappa}$, for the correct definition of the many-body state, see sections 5.6 and 5.7.

5.5 Bosonic Systems

Here, we describe in detail how to evaluate efficiently the ratio of wave functions on different configurations to apply the Metropolis algorithm, see Eq. (5.12). First of all, we consider the computation of the non-interacting part $\langle x|\Phi_0\rangle$, then we discuss the contribution coming from $\mathcal{J}(x)$.

https://doi.org/10.1017/9781316417041.006 Published online by Cambridge University Press

5.5.1 Definition of the Non-Interacting State

The natural choice for a bosonic non-interacting wave function is to condense all the particles in a single state. The simplest example is to consider the single-particle "orbital" state defined by:

$$\phi_\alpha^\dagger|0\rangle = \left(\sum_i V_{i,\alpha} b_i^\dagger\right)|0\rangle, \tag{5.24}$$

where $\{V_{i,\alpha}\}$ are generic amplitudes for having the boson on site i; α is an index that specifies the orbital: for example, the case of a boson in a zero-momentum state is given by $V_{i,\alpha} = 1/\sqrt{L}$. Then, the many-body state with N_b bosons occupying the same orbital is:

$$|\Phi_0\rangle = \frac{1}{\sqrt{N_b!}}\left(\sum_i V_{i,\alpha} b_i^\dagger\right)^{N_b}|0\rangle. \tag{5.25}$$

By expanding the summation, we can rewrite the many-body state over the basis set of Eq. (5.18):

$$|\Phi_0\rangle = \frac{1}{\sqrt{N_b!}}\sum_{m_1,\ldots,m_L}\frac{N_b!}{m_1!\ \ldots\ m_L!}V_{1,\alpha}^{m_1}\ \ldots\ V_{L,\alpha}^{m_L}\,(b_1^\dagger)^{m_1}\ \ldots\ (b_L^\dagger)^{m_L}|0\rangle, \tag{5.26}$$

which allows us to have a very simple expression for the overlap $\langle x|\Phi_0\rangle$, where the configuration $|x\rangle \equiv |n_1,\ldots,n_L\rangle$ contains n_i bosons on site i:

$$|n_1,\ldots,n_L\rangle = \frac{1}{\sqrt{n_1!\ldots n_L!}}(b_1^\dagger)^{n_1}\ \ldots\ (b_L^\dagger)^{n_L}|0\rangle, \tag{5.27}$$

leading to:

$$\langle n_1,\ldots,n_L|\Phi_0\rangle = \sqrt{\frac{N_b!}{n_1!\ \ldots\ n_L!}}\,V_{1,\alpha}^{n_1}\ \ldots\ V_{L,\alpha}^{n_L}. \tag{5.28}$$

5.5.2 Fast Computation of the Non-Interacting State

Given the form of the non-interacting state, for which $\langle n_1,\ldots,n_L|\Phi_0\rangle$ is given by Eq. (5.28), the ratio of two overlaps with different configurations can be obtained by a simple calculation. For example, if we consider the ratio between configurations that only differ by a boson hopping between sites l and k, namely $|x\rangle \equiv |n_1,\ldots,n_k,\ldots,n_l,\ldots,n_L\rangle$ and $|x'\rangle \equiv |n_1,\ldots,n_k+1,\ldots,n_l-1,\ldots,n_L\rangle$, we have:

$$\frac{\langle n_1,\ldots,n_k+1,\ldots,n_l-1,\ldots,n_L|\Phi_0\rangle}{\langle n_1,\ldots,n_k,\ldots,n_l,\ldots,n_L|\Phi_0\rangle} = \sqrt{\frac{n_l}{n_k+1}}\left(\frac{V_{k,\alpha}}{V_{l,\alpha}}\right). \tag{5.29}$$

https://doi.org/10.1017/9781316417041.006 Published online by Cambridge University Press

Therefore, whenever single-boson hopping processes are considered along the Markov chain, the contribution of the non-interacting part of the wave function to the Metropolis ratio (5.12) is very simple, since it requires $O(1)$ operations. The case in which more than one boson is moved can be obtained by generalizing the previous analysis in a straightforward way.

We would like to emphasize that the calculation of $\langle x|\Phi_0\rangle$ becomes dramatically complicated if the bosons are not condensed but occupy different orbitals:

$$|\Phi_0\rangle = \prod_{\alpha=1}^{N_b} \left(\sum_i V_{i,\alpha} b_i^\dagger \right) |0\rangle. \tag{5.30}$$

Indeed, in this case, we obtain that the overlap is given by a permanent of an $N_b \times N_b$ matrix:

$$\langle n_1, \ldots, n_L|\Phi_0\rangle = \mathrm{per}(V_{R_j,\alpha}), \tag{5.31}$$

where the R_j's (with $j = 1, \ldots, N_b$) take the values of the occupied sites. Notice that the values of R_j may appear more than once, according to the occupation of each site, e.g., $R_j = k$ appears n_k times. The calculation of the permanent (or the ratio of two permanents) is exponentially hard, thus preventing us to perform calculations with large values of N_b.

5.5.3 Fast Computation of the Jastrow Factor

Let us now consider the contribution coming from the Jastrow factor. A straightforward calculation of $\mathcal{J}(x)$ requires $O(L^2)$ calculations for the Jastrow factor of Eq. (5.15), thus leading to the same complexity when computing the ratio $\mathcal{J}(x')/\mathcal{J}(x)$ that appears in the Metropolis ratio (5.12). However, whenever the two configurations differ only by few boson hoppings, it is possible to apply a fast computation of the ratio, which involves $O(L)$ operations. Indeed, let us consider: $|x\rangle \equiv |n_1, \ldots, n_k, \ldots, n_l, \ldots, n_L\rangle$ and $|x'\rangle \equiv |n_1, \ldots, n_k + 1, \ldots, n_l - 1, \ldots, n_L\rangle$:

$$\frac{\mathcal{J}(n_1, \ldots, n_k + 1, \ldots, n_l - 1, \ldots, n_L)}{\mathcal{J}(n_1, \ldots, n_k, \ldots, n_l, \ldots, n_L)} = \frac{\exp\left(-\sum_i v_{i,k} n_i\right)}{\exp\left(-\sum_i v_{i,l} n_i\right)} e^{v_{k,l} - v_{k,k}}, \tag{5.32}$$

where we used the fact that the pseudo-potential $v_{i,j}$ is symmetric, i.e., $v_{i,j} = v_{j,i}$ and translational invariant, i.e., $v_{k,k} = v_{l,l}$. The second term in the r.h.s. of Eq. (5.32) does not depend upon the bosonic configuration and can be computed at the beginning of the simulation, once and for all. Instead, the first term depends upon the bosonic configuration, which is sampled along the Markov chain. The computation of the ratio can be done in $O(1)$ operations once we compute and store a vector of dimension L that depends upon $|x\rangle = |n_1, \ldots, n_k, \ldots, n_l, \ldots, n_L\rangle$:

$$T_{\mathrm{Jastrow}}(j) = \sum_i v_{i,j} n_i. \tag{5.33}$$

https://doi.org/10.1017/9781316417041.006 Published online by Cambridge University Press

Then, once the new proposed $|x'\rangle = |n_1, \dots, n_k + 1, \dots, n_l - 1, \dots, n_L\rangle$ is accepted, the vector $T_{\text{Jastrow}}(j)$ must be updated, which requires an $O(L)$ operations. Indeed, since $n_i' = n_i + \delta_{i,k} - \delta_{i,l}$, we have that:

$$T'_{\text{Jastrow}}(j) = T_{\text{Jastrow}}(j) + v_{k,j} - v_{l,j}. \tag{5.34}$$

Therefore, we can compute from scratch $T_{\text{Jastrow}}(j)$ at the beginning of the Markov chain for all the sites $j = 1, \dots, L$ and then update it by using Eq. (5.34) every time a new configuration is accepted along the Markov process. As a safe habit, from time to time, it is recommended to recompute $T_{\text{Jastrow}}(j)$ from scratch, since the rounding error of the fast update can accumulate and give rise to numerical errors.

In summary, by using this scheme for the Jastrow factor, the single-move algorithm scales with $O(L)$; however, by performing these local updates, we need to perform $O(L)$ moves to obtain an almost independent configuration, for systems with a finite correlation time. Therefore, the bosonic code has a quadratic scaling with the number of sites/bosons.

5.6 Fermionic Systems with Determinants

Let us now move to fermionic systems, described by Jastrow-Slater wave functions. Since the part on the calculation of the Jastrow term is similar to the previous bosonic case, we do not repeat it here. The same fast updating can be used also for fermions. Instead, the part involving the calculation of the non-interacting state $\langle x|\Phi_0\rangle$ is totally different from the bosonic case. In the following, we describe how it is possible to devise an efficient algorithm to deal with such an object.

5.6.1 Definition of the Non-Interacting State

Any non-interacting wave function can be obtained as the ground state of a suitable quadratic Hamiltonian $\mathcal{H}_0$. First of all, we contract the spin index σ and the lattice site i into a single index I running from 1 to $2L$:

$$c_{i,\uparrow} \equiv d_i\,, \tag{5.35}$$

$$c_{i,\downarrow} \equiv d_{i+L}. \tag{5.36}$$

Then, we start from the simple case in which the non-interacting Hamiltonian is written as:

$$\mathcal{H}_0 = \sum_{I,J} t_{I,J} d_I^\dagger d_J, \tag{5.37}$$

https://doi.org/10.1017/9781316417041.006 Published online by Cambridge University Press

which contains hopping terms only, also including processes in which the spin along z is not conserved, i.e., the terms with $I \leq L$ and $J > L$ and vice-versa. In a compact form, the non-interacting Hamiltonian of Eq. (5.37) can be written as:

$$\mathcal{H}_0 = \mathbf{d}^\dagger \mathbf{T} \mathbf{d} \ , \tag{5.38}$$

where

$$\mathbf{d}^\dagger = \begin{pmatrix} d_1^\dagger & \dots & d_{2L}^\dagger \end{pmatrix}, \tag{5.39}$$

and

$$\mathbf{T} = \begin{pmatrix} t_{1,1} & \dots & t_{1,2L} \\ \vdots & \ddots & \vdots \\ t_{2L,1} & \dots & t_{2L,2L} \end{pmatrix}. \tag{5.40}$$

Since $\mathcal{H}_0$ commutes with the total number of electrons $N_e = \sum_I d_I^\dagger d_I$, the eigenstates are single-particle orbitals. In practice, the $2L \times 2L$ matrix $\mathbf{T}$ can be easily diagonalized by using standard libraries (e.g., LAPACK routines):

$$\mathcal{H}_0 = \mathbf{d}^\dagger \mathbf{U} \, \mathbf{U}^\dagger \mathbf{T} \mathbf{U} \, \mathbf{U}^\dagger \mathbf{d} \ = \mathbf{\Phi}^\dagger \mathbf{E} \mathbf{\Phi} \ = \sum_\alpha \varepsilon_\alpha \phi_\alpha^\dagger \phi_\alpha, \tag{5.41}$$

where $\mathbf{U}$ is a unitary matrix (that preserves anti-commutation relations of fermionic operators), $\mathbf{E} = \mathrm{diag}(\varepsilon_1, \dots, \varepsilon_{2L})$ is the diagonal matrix containing the $2L$ eigenvalues ε_α of $\mathbf{T}$, and $\mathbf{\Phi}^\dagger = (\phi_1^\dagger, \dots, \phi_{2L}^\dagger)$ is defined in terms of the eigenvectors of $\mathbf{T}$:

$$\phi_\alpha^\dagger = \sum_I U_{I,\alpha} d_I^\dagger. \tag{5.42}$$

Now, the many-body state $|\Phi_0\rangle$ can be constructed by occupying the N_e lowest-energy orbitals:

$$|\Phi_0\rangle = \prod_{\alpha=1}^{N_e} \phi_\alpha^\dagger |0\rangle = \left(\sum_I U_{I,1} d_I^\dagger \right) \cdots \left(\sum_I U_{I,N_e} d_I^\dagger \right) |0\rangle \tag{5.43}$$

Following the arguments of section 5.4, the generic configuration, with N_e electrons, which is visited along the Markov process, reads as:

$$|x\rangle = d_{R_1}^\dagger \dots d_{R_{N_e}}^\dagger |0\rangle, \tag{5.44}$$

where $j = 1, \dots, N_e$ includes both up and down spins and R_j assumes values from 1 to $2L$: the positions of spin-up electrons coincide with the site number, while the

https://doi.org/10.1017/9781316417041.006 Published online by Cambridge University Press

positions of spin-down electrons must be shifted by L. The overlap $\langle x|\Phi_0\rangle$ is then given by:

$$\begin{aligned}\langle x|\Phi_0\rangle &= \langle 0|d_{R_{N_e}}\dots d_{R_1}\left(\sum_I U_{I,1}d_I^\dagger\right)\cdots\left(\sum_I U_{I,N_e}d_I^\dagger\right)|0\rangle \\ &= \langle 0|d_{R_{N_e}}\dots d_{R_1}\left[\sum_p (-1)^p\prod_{\alpha=1}^{N_e}U_{p\{R_j\},\alpha}\right]d_{R_1}^\dagger\dots d_{R_{N_e}}^\dagger|0\rangle, \end{aligned} \tag{5.45}$$

where the sum inside the square bracket is over all the possible permutations of the $\{R_j\}$ in $|x\rangle$; the sign appears because of the anti-commutation relations of fermionic operators. Then, we get:

$$\langle x|\Phi_0\rangle = \det\{U_{R_j,\alpha}\}. \tag{5.46}$$

When constructing the many-body state (5.43), we must pay attention to construct a unique many-body state, i.e., occupy the correct lowest-energy orbitals. When the highest-occupied orbital and the lowest-unoccupied one have different energies (*closed shell* configuration), the choice is unique. Instead, it can also happen that there is a degeneracy that does not allow a unique choice (*open shell* configuration). Whenever the non-interacting Hamiltonian $\mathcal{H}_0$ is diagonalized numerically, the eigenstates provided by standard libraries do not carry definite quantum numbers (like momentum), but are given by generic linear combinations of degenerate orbitals (which erratically depend on the numerical precision used for the computation). This is not a problem whenever *all* the degenerate eigenstates are included in the many-body state (5.43), since taking any linear combination of columns in the matrix $\mathbf{U}$ will not change the value of the determinant in Eq. (5.46). By contrast, not including all degenerate eigenstates will cause a problem, since the determinant will depend on the particular combination of states that is considered. Therefore, whenever a numerical diagonalization is done, we must verify that a closed-shell configuration occurs. Otherwise, the problem of having a vanishing gap can be overcome by constructing suitable orbitals with definite quantum numbers, to obtain a reproducible simulation of the many-body state.

In practice, the diagonalization of the non-interacting Hamiltonian $\mathcal{H}_0$ must be performed at the beginning of the Monte Carlo calculation; then, we need to store the reduced part of the $\mathbf{U}$, obtained by keeping only the N_e columns that correspond to occupied orbitals. This is a $2L \times N_e$ matrix:

$$\mathbf{U} = \begin{pmatrix} U_{1,1} & \dots & U_{1,N_e} \\ \vdots & \ddots & \vdots \\ U_{2L,1} & \dots & U_{2L,N_e} \end{pmatrix}. \tag{5.47}$$

https://doi.org/10.1017/9781316417041.006 Published online by Cambridge University Press

Then, the overlap with a generic electronic configuration is given by the determinant of the matrix obtained by considering only the rows of (5.47) corresponding to the electron positions $\{R_j\}$, thus giving a $N_e \times N_e$ matrix.

Remarkably, the same kind of formalism can be used also if the non-interacting Hamiltonian contains an electron pairing that couples up and down spins. Indeed, let us consider a BCS Hamiltonian described by:

$$\mathcal{H}_0 = \sum_{i,j,\sigma} t_{i,j} c^\dagger_{i,\sigma} c_{j,\sigma} - \mu_0 \sum_{i,\sigma} c^\dagger_{i,\sigma} c_{i,\sigma} + \sum_{i,j} \Delta_{i,j} c^\dagger_{i,\uparrow} c^\dagger_{j,\downarrow} + \text{h.c.}, \tag{5.48}$$

where we have included a chemical potential μ_0, which fixes, on average, the number of electrons. In this case, the total number of particles is not conserved and the concept of single-particle orbitals is not defined. Indeed, the ground state of the BCS Hamiltonian is naturally written in terms of a pairing function (see Chapter 1). Nevertheless, we can perform a particle-hole transformation on the spin-down electrons:

$$c_{i,\uparrow} \to f_{i,\uparrow} \equiv d_i \tag{5.49}$$

$$c_{i,\downarrow} \to f^\dagger_{i,\downarrow} \equiv d^\dagger_{i+L}. \tag{5.50}$$

Then, apart from constant terms, the transformed BCS Hamiltonian has the form of Eq. (5.38), where the matrix $\mathbf{T}$ is given by (i and $j \leq L$):

$$T_{i,i} = -\mu_0 \tag{5.51}$$

$$T_{i+L,i+L} = \mu_0 \tag{5.52}$$

$$T_{i,j} = t_{i,j}, \tag{5.53}$$

$$T_{i+L,j+L} = -t_{i,j}, \tag{5.54}$$

$$T_{i,j+L} = T^*_{j+L,i} = \Delta_{i,j}. \tag{5.55}$$

Since, after the particle-hole transformation, the number of particles (but not the z component of the spin) is conserved, the eigenstates of the BCS Hamiltonian can be expressed into "orbitals," similarly to the ones of Eq. (5.42), but without having a definite spin component along z.

Notice that the particle-hole transformations (5.49) and (5.50) change a spin down into the vacuum and vice-versa, so that:

$$|0\rangle_i \to f^\dagger_{i,\downarrow} |\tilde{0}\rangle_i, \tag{5.56}$$

$$c^\dagger_{i,\uparrow} |0\rangle_i \to f^\dagger_{i,\uparrow} f^\dagger_{i,\downarrow} |\tilde{0}\rangle_i, \tag{5.57}$$

$$c^\dagger_{i,\downarrow} |0\rangle_i \to |\tilde{0}\rangle_i, \tag{5.58}$$

$$c^\dagger_{i,\uparrow} c^\dagger_{i,\downarrow} |0\rangle_i \to f^\dagger_{i,\uparrow} |\tilde{0}\rangle_i, \tag{5.59}$$

https://doi.org/10.1017/9781316417041.006 Published online by Cambridge University Press

where $|\tilde{0}\rangle_i$ is the vacuum of the f electrons, i.e., $f_{i,\uparrow}|\tilde{0}\rangle = f_{i,\downarrow}|\tilde{0}\rangle = 0$. By using the anti-commutation relations of the fermion operators, we have that the occupations $u_{i,\sigma}$ of the new fermionic operators are related to the original ones by:

$$u_{i,\uparrow} = f^\dagger_{i,\uparrow} f_{i,\uparrow} = c^\dagger_{i,\uparrow} c_{i,\uparrow} = n_{i,\uparrow}, \tag{5.60}$$

$$u_{i,\downarrow} = f^\dagger_{i,\downarrow} f_{i,\downarrow} = 1 - c^\dagger_{i,\downarrow} c_{i,\downarrow} = 1 - n_{i,\downarrow}, \tag{5.61}$$

which implies that the relations among densities and magnetizations on site i are given by:

$$u_i = u_{i,\uparrow} + u_{i,\downarrow} = n_{i,\uparrow} - n_{i,\downarrow} + 1 = m^z_i + 1, \tag{5.62}$$

$$w^z_i = u_{i,\uparrow} - u_{i,\downarrow} = n_{i,\uparrow} + n_{i,\downarrow} - 1 = n_i - 1, \tag{5.63}$$

where we have introduced u_i and w^z_i, as the density and magnetization for the transformed operators, respectively. The letters u and w are chosen because they look like an upside down n and m, respectively. Therefore, apart from constant terms, upon the particle-hole transformation of Eqs. (5.49) and (5.50), the density changes into the z component of the spin and vice-versa. As a corollary, we obtain that, whenever the total magnetization of the original particles is vanishing, the total number of the new particles is equal to the number of sites L.

5.6.2 *Fast Computation of the Determinants*

Let us now show how to compute efficiently the ratio of determinants when the two configurations $|x\rangle$ and $|x'\rangle$ differ by one or few electron hoppings. According to Eq. (5.46) the overlap $\langle x|\Phi_0\rangle$ is given by the determinant of $\tilde{U}_{j,\alpha} \equiv U_{R_j,\alpha}$:

$$\tilde{\mathbf{U}} = \begin{pmatrix} U_{R_1,1} & \dots & U_{R_1,N_e} \\ \vdots & \ddots & \vdots \\ U_{R_l,1} & \dots & U_{R_l,N_e} \\ \vdots & \ddots & \vdots \\ U_{R_{N_e},1} & \dots & U_{R_{N_e},N_e} \end{pmatrix}, \tag{5.64}$$

which is obtained taking only the rows corresponding to the occupied sites of the matrix $\mathbf{U}$ of Eq. (5.47). Notice that the actual order of rows in $\tilde{\mathbf{U}}$ is determined by the one of creation operators in $|x\rangle$, see Eq. (5.44). In this regard, the generalization of the vector $\boldsymbol{\kappa}$ (see section 5.4) for the spinful case, can be used to put the rows of $\tilde{\mathbf{U}}$ in the correct order.

Let us start and consider the electronic configurations in which $|x'\rangle$ is obtained from $|x\rangle$ just by hopping the l-th electron from R_l to R'_l, i.e., $|x'\rangle = d^\dagger_{R'_l} d_{R_l}|x\rangle$.

https://doi.org/10.1017/9781316417041.006 Published online by Cambridge University Press

The new matrix $\tilde{\mathbf{U}}'$ will be equal to $\tilde{\mathbf{U}}$, except that the elements of the l-th row will be changed from $U_{R_l,\alpha}$ to $U_{R'_l,\alpha}$:

$$\tilde{\mathbf{U}}' = \begin{pmatrix} U_{R_1,1} & \dots & U_{R_1,N_e} \\ \vdots & \ddots & \vdots \\ U_{R'_l,1} & \dots & U_{R'_l,N_e} \\ \vdots & \ddots & \vdots \\ U_{R_{N_e},1} & \dots & U_{R_{N_e},N_e} \end{pmatrix}. \tag{5.65}$$

Then, the ratio between two configurations that differ only by a single fermion hopping is:

$$\frac{\langle x'|\Phi_0\rangle}{\langle x|\Phi_0\rangle} = \frac{\langle x|d^\dagger_{R_l} d_{R'_l}|\Phi_0\rangle}{\langle x|\Phi_0\rangle} = \frac{\det \tilde{\mathbf{U}}'}{\det \tilde{\mathbf{U}}}. \tag{5.66}$$

By denoting with K the new site of the l-th electron (i.e., $K \equiv R'_l$), the updated matrix elements are given by a compact form:

$$\tilde{U}'_{j,\alpha} = \tilde{U}_{j,\alpha} + \delta_{j,l}(U_{K,\alpha} - \tilde{U}_{l,\alpha}) = \tilde{U}_{j,\alpha} + \delta_{j,l} v^{K,l}_\alpha, \tag{5.67}$$

where we have defined $v^{K,l}_\alpha \equiv U_{K,\alpha} - \tilde{U}_{l,\alpha}$; here the indices K and l are fixed, since they specify the site where the electron is hopping and the electron index, respectively. This equation can be rewritten in the following way:

$$\tilde{U}'_{j,\alpha} = \sum_\beta \tilde{U}_{j,\beta} \left(\delta_{\beta,\alpha} + \tilde{U}^{-1}_{\beta,l} v^{K,l}_\alpha\right) = \sum_\beta \tilde{U}_{j,\beta} Q_{\beta,\alpha}, \tag{5.68}$$

where

$$Q_{\beta,\alpha} = \delta_{\beta,\alpha} + \tilde{U}^{-1}_{\beta,l} v^{K,l}_\alpha. \tag{5.69}$$

Therefore, $\tilde{\mathbf{U}}' = \tilde{\mathbf{U}}\mathbf{Q}$, which implies that the calculation of the ratio of the determinants of $\tilde{\mathbf{U}}'$ and $\tilde{\mathbf{U}}$ is equivalent to the calculation of the determinant of $\mathbf{Q}$:

$$\frac{\det \tilde{\mathbf{U}}'}{\det \tilde{\mathbf{U}}} = \det \mathbf{Q}. \tag{5.70}$$

The great simplification comes from the fact that the determinant of $\mathbf{Q}$ can be easily computed. Indeed, $\mathbf{Q}$ has a particularly simple form that can be written as:

$$Q_{\beta,\alpha} = \delta_{\beta,\alpha} + \mathcal{B}_\beta \mathcal{A}_\alpha, \tag{5.71}$$

where $\mathcal{B}_\beta = \tilde{U}^{-1}_{\beta,l}$ and $\mathcal{A}_\alpha = v^{K,l}_\alpha$. Although the matrix is not Hermitian, the eigenvalues of the matrix $\mathbf{Q}$ can be obtained from the secular equation:

$$\sum_\alpha Q_{\beta,\alpha} \upsilon_\alpha = \lambda \upsilon_\beta; \tag{5.72}$$

https://doi.org/10.1017/9781316417041.006 Published online by Cambridge University Press

by using the explicit form of Eq. (5.71), we obtain:

$$\upsilon_\beta + \mathcal{B}_\beta \sum_\alpha \mathcal{A}_\alpha \upsilon_\alpha = \lambda \upsilon_\beta, \tag{5.73}$$

which implies that all vectors υ_α that are orthogonal to $\mathcal{A}_\alpha$ are eigenvectors with eigenvalue $\lambda = 1$ (there are $N_e - 1$ of such vectors); in addition, $\upsilon_\alpha = \mathcal{B}_\alpha$ is also eigenvector with $\lambda = 1 + \sum_\alpha \mathcal{A}_\alpha \mathcal{B}_\alpha = \sum_\alpha U_{K,\alpha} \tilde{U}^{-1}_{\alpha,l}$. Therefore, we have that:

$$\frac{\det \tilde{\mathbf{U}}'}{\det \tilde{\mathbf{U}}} = \det \mathbf{Q} = \sum_\alpha U_{K,\alpha} \tilde{U}^{-1}_{\alpha,l}. \tag{5.74}$$

Having stored (at the beginning of the simulation) the matrix $\tilde{\mathbf{U}}^{-1}$ for the configuration $|x\rangle$, this calculation requires $O(N_e)$ operations, instead of the $O(N_e^3)$ needed to evaluate a determinant. Then, once the new configuration $|x'\rangle$ is accepted along the Markov process, the matrix $\tilde{\mathbf{U}}^{-1}$ must be updated. This can be done in $O(N_e^2)$ operations. In fact, we have that $(\tilde{\mathbf{U}}')^{-1} = \mathbf{Q}^{-1}\tilde{\mathbf{U}}^{-1}$, the inverse of the matrix $\mathbf{Q}$ being given by (as easily verified):

$$Q^{-1}_{\alpha,\beta} = \delta_{\alpha,\beta} - \frac{1}{\det \mathbf{Q}} \mathcal{B}_\alpha \mathcal{A}_\beta. \tag{5.75}$$

Then, the updated matrix elements of $(\tilde{\mathbf{U}}')^{-1}$ are given by:

$$\tilde{U}^{-1\prime}_{\alpha,j} = \tilde{U}^{-1}_{\alpha,j} - \frac{\tilde{U}^{-1}_{\alpha,l}}{\det \mathbf{Q}} \left(\sum_\beta U_{K,\beta} \tilde{U}^{-1}_{\beta,j} - \delta_{l,j} \right). \tag{5.76}$$

This is a closed equation for updating the matrix $\tilde{\mathbf{U}}^{-1}$.

We would like to emphasize that the previous results for the calculation of the ratio of determinants and the updating can be further simplified. Indeed, at the beginning of the calculation, we can compute and store a $2L \times N_e$ matrix $\mathbf{W}$, whose elements are given by:

$$W_{I,j} = \sum_\alpha U_{I,\alpha} \tilde{U}^{-1}_{\alpha,j}; \tag{5.77}$$

then the ratio of determinants (5.74) costs $O(1)$ operations, since it consists in taking the element corresponding to the new site (row) and the electron performing the hopping process (column):

$$\frac{\det \tilde{\mathbf{U}}'}{\det \tilde{\mathbf{U}}} = W_{K,l}. \tag{5.78}$$

The evaluation of $\mathbf{W}$ requires the knowledge of the full matrix $\mathbf{U}$, which has been computed and stored once for all at the beginning of the simulation (it does not

https://doi.org/10.1017/9781316417041.006 Published online by Cambridge University Press

depend upon the electronic configuration), and $\tilde{\mathbf{U}}^{-1}$, which instead depends upon the configuration $|x\rangle$. Then, a simple updating scheme for $\mathbf{W}$ is possible. In fact, by multiplying both sides of Eq. (5.76) by $U_{I,\alpha}$ and summing over α, we obtain:

$$W'_{I,j} = W_{I,j} - \frac{W_{I,l}}{W_{K,l}} \left(W_{K,j} - \delta_{l,j} \right), \tag{5.79}$$

where, we have used that $\det \mathbf{Q} = W_{K,l}$, according to Eq. (5.78) and the definition of Eq. (5.77). Since each matrix element must be updated with $O(1)$ operations, the total cost is $O(2LN_e)$.

By using a wave function that is constructed from filling single-particle states (also including cases with BCS pairing when a particle-hole transformation is done), the algorithm can be written in terms of the matrix elements of $\mathbf{W}$, which is the equal-time Green's function, see Eq. (5.66):

$$\frac{\langle x | d^\dagger_{R_l} d_K | \Phi_0 \rangle}{\langle x | \Phi_0 \rangle} = W_{K,l}. \tag{5.80}$$

We would like to mention that Eq. (5.79) has a direct interpretation by using the Wick theorem (with different *bra* and *ket* states) for the updated (equal-time) Green's function:

$$W'_{I,j} = \frac{\langle x' | d^\dagger_{R_j} d_I | \Phi_0 \rangle}{\langle x' | \Phi_0 \rangle} = \frac{\langle x | d^\dagger_{R_l} d_K d^\dagger_{R_j} d_I | \Phi_0 \rangle}{\langle x | d^\dagger_{R_l} d_K | \Phi_0 \rangle}; \tag{5.81}$$

in this case, we must consider the contractions without including the anomalous (superconducting) ones, which vanish since both $|x\rangle$ and $|\Phi_0\rangle$ have a fixed number of particles.

Let us finish this part by generalizing the previous formalism to the case where more than one electron hop, i.e., $|x'\rangle = d^\dagger_{R'_{l_1}} d_{R_{l_1}} \dots d^\dagger_{R'_{l_m}} d_{R_{l_m}} |x\rangle$, thus leading to a modification of m rows of the $\tilde{\mathbf{U}}$ matrix; for example, this could be the case for pair-hopping or spin-flip processes. Then, Eq. (5.67) generalizes into:

$$\tilde{U}'_{j,\alpha} = \tilde{U}_{j,\alpha} + \sum_{r=1}^{m} \delta_{j,l_r} \left(U_{K_r,\alpha} - \tilde{U}_{l_r,\alpha} \right) = \tilde{U}_{j,\alpha} + \sum_{r=1}^{m} \delta_{j,l_r} v_\alpha^{K_r,l_r}; \tag{5.82}$$

as before, the indices K_r and l_r (for $r = 1, \dots, m$) are fixed, because they specify the sites where the electrons are hopping and the electron indices, respectively. By performing the same algebra as before, we get:

$$\tilde{U}'_{j,\alpha} = \sum_\beta \tilde{U}_{j,\beta} \left(\delta_{\beta,\alpha} + \sum_{r=1}^{m} \tilde{U}^{-1}_{\beta,l_r} v_\alpha^{K_r,l_r} \right) = \sum_\beta \tilde{U}_{j,\beta} Q_{\beta,\alpha}, \tag{5.83}$$

https://doi.org/10.1017/9781316417041.006 Published online by Cambridge University Press

where now the matrix $\mathbf{Q}$ has the following form:

$$Q_{\beta,\alpha} = \delta_{\beta,\alpha} + \sum_{r=1}^{m} \mathcal{B}_{\beta}^{r} \mathcal{A}_{\alpha}^{r}, \tag{5.84}$$

where $\mathcal{B}_{\beta}^{r} = \tilde{U}_{\beta,l_r}^{-1}$ and $\mathcal{A}_{\alpha}^{r} = v_{\alpha}^{K_r,l_r}$. As before, the determinant of $\mathbf{Q}$ can be easily computed by solving the corresponding eigenvalue problem:

$$\sum_{\alpha} Q_{\beta,\alpha} \upsilon_{\alpha} = \upsilon_{\beta} + \sum_{r=1}^{m} \mathcal{B}_{\beta}^{r} \sum_{\alpha} \mathcal{A}_{\alpha}^{r} \upsilon_{\alpha} = \lambda \upsilon_{\beta}, \tag{5.85}$$

which implies that all vectors that are orthogonal to the subspace defined by the $\mathcal{A}_{\alpha}^{r}$'s are eigenvectors with $\lambda = 1$; moreover, $\upsilon_{\alpha} = \sum_{r=1}^{m} x_r \mathcal{B}_{\alpha}^{r}$ is an eigenvector provided that the coefficients x_r satisfy:

$$\sum_{s=1}^{m} \left(\delta_{r,s} + \sum_{\alpha} \mathcal{A}_{\alpha}^{r} \mathcal{B}_{\alpha}^{s} \right) x_s = \lambda x_r. \tag{5.86}$$

Therefore, the m non-trivial eigenvalues of $\mathbf{Q}$ are given by the ones of the $m \times m$ matrix:

$$C_{r,s} = \delta_{r,s} + \sum_{\alpha} \mathcal{A}_{\alpha}^{r} \mathcal{B}_{\alpha}^{s} = W_{K_r,l_s}. \tag{5.87}$$

The final expression of the ratio of the two determinants is given by:

$$\frac{\det \tilde{\mathbf{U}}'}{\det \tilde{\mathbf{U}}} = \det(W_{K_r,l_s}). \tag{5.88}$$

Also in this case, once the move is accepted, we have to update the matrix $(\tilde{\mathbf{U}}')^{-1} = \mathbf{Q}^{-1}\tilde{\mathbf{U}}^{-1}$. As before, the inverse of the $\mathbf{Q}$ matrix can be obtained:

$$Q_{\alpha,\beta}^{-1} = \delta_{\alpha,\beta} - \sum_{r,s=1}^{m} \mathcal{B}_{\alpha}^{r} C_{r,s}^{-1} \mathcal{A}_{\beta}^{s}. \tag{5.89}$$

Therefore, we obtain:

$$\tilde{U}_{\alpha,j}^{-1\prime} = \tilde{U}_{\alpha,j}^{-1} - \sum_{r,s=1}^{m} \tilde{U}_{\alpha,l_r}^{-1} C_{r,s}^{-1} \left(\sum_{\beta} U_{K_s,\beta} \tilde{U}_{\beta,j}^{-1} - \delta_{l_s,j} \right). \tag{5.90}$$

Then, the updated $\mathbf{W}'$ is obtained by multiplying both sides of the previous equation by $U_{i,\alpha}$ and summing over α:

$$W_{I,j}' = W_{I,j} + \sum_{r=1}^{m} W_{I,l_r} b_j^{(r)}, \tag{5.91}$$

https://doi.org/10.1017/9781316417041.006 Published online by Cambridge University Press

where

$$b_j^{(r)} = -\sum_{s=1}^{m} C_{r,s}^{-1} \left(W_{K_s,j} - \delta_{l_s,j} \right). \tag{5.92}$$

Similarly to the fast update of the Jastrow factor, it is highly recommended, from time to time, to recompute the matrix $\mathbf{W}$ from scratch. This costs $O(2LN_e^2)$, thus this calculation does not affect the complexity of the algorithm if it is done every $O(N_e)$ electron moves.

5.6.3 Delayed Updates

Let us consider the simple update in which only one electron is displaced. The basic operation in the updating is the so-called rank-1 update of a $2L \times N_e$ matrix, see Eq. (5.79) that can be written as:

$$W'_{I,j} = W_{I,j} + a_I b_j, \tag{5.93}$$

where $\mathbf{a}$ and $\mathbf{b}$ are two vectors of length $2L$ and N_e, respectively:

$$a_I = W_{I,l}, \tag{5.94}$$

$$b_j = -\frac{W_{K,j} - \delta_{l,j}}{W_{K,l}}. \tag{5.95}$$

This updating operation can be computationally inefficient, whenever, for large size, the matrix $\mathbf{W}$ cannot be contained in the cache of the processor. A way to overcome this drawback is to delay the update of the matrix $\mathbf{W}$, without loosing its information. This can be done by storing a set of left and right vectors $\mathbf{a}^{(p)}$ and $\mathbf{b}^{(p)}$ with $p = 1, \ldots, m$, as well as the "initial" matrix (denoted by $\mathbf{W}^0$) from which we begin to delay the updates. Then, the matrix $\mathbf{W}$, after m updates is given by:

$$W_{I,j} = W^0_{I,j} + \sum_{p=1}^{m} a_I^{(p)} b_j^{(p)}. \tag{5.96}$$

Every time we accept a new configuration, a new pair of vectors $\mathbf{a}^{(m+1)}$ and $\mathbf{b}^{(m+1)}$ can be computed in few operations in term of $\mathbf{W}^0$ and the previous vectors with $p = 1, \ldots, m$, by substituting Eq. (5.96) in Eqs. (5.94) and (5.95). Once the matrix $\mathbf{W}$ is written in the form of Eq. (5.96), the number of operations required to evaluate the factors in the sum is $O[m(2L + N_e)]$, which is negligible compared to the full update for $m \ll L$.

In this way, we can find an optimal $m_{\max}$, for which we can evaluate the full matrix $\mathbf{W}$ by a standard matrix multiplication:

$$\mathbf{W} = \mathbf{W}^0 + \mathbf{A}\mathbf{B}^T, \tag{5.97}$$

https://doi.org/10.1017/9781316417041.006 Published online by Cambridge University Press

where $\mathbf{A}$ and $\mathbf{B}$ are $2L \times m_{\max}$ and $N_e \times m_{\max}$ matrices, which are made of the $p = 1, \ldots, m_{\max}$ vectors $\mathbf{a}^{(p)}$ and $\mathbf{b}^{(p)}$, respectively. After that, we can continue with a new delayed update with a new $\mathbf{W}^0 = \mathbf{W}$, by initializing to zero the integer m. The advantage of this updating procedure is that after a cycle of $m_{\max}$ steps the bulk of the computation is given by the matrix-matrix product in Eq. (5.97), which is much more efficient (and it is not limited by cache memory) than the $m_{\max}$ rank-1 updates of $\mathbf{W}$ given in Eq. (5.79). For large number of electrons, the delayed update procedure, with the optimal value of $m_{\max}$, allows us to improve the speed of the variational Monte Carlo code by about an order of magnitude.

5.6.4 Backflow Correlations

Here, we would like to discuss how to implement the updating of the determinant part in presence of backflow correlations that have been introduced on lattice problem. We notice that, in the lattice case, a continuous change of coordinates, implementing the standard backflow correlation (1.70), is not possible. Therefore, this *Ansatz* requires to find an expression of the orbital in a generic off-lattice backflow coordinate in terms of the allowed lattice positions, consistently to a Taylor expansion of the orbitals (Tocchio et al., 2008 and Tocchio et al., 2011). In practice, we consider a quadratic Hamiltonian (as the ones that have been discussed in section 5.6), to construct the non-interacting orbitals $\{U_{I,\alpha}\}$, see Eq. (5.42). In the simplest approach, in which the backflow correlations act on holon-doublon (nearest-neighbor) pairs (Tocchio et al., 2008), the overlap between the generic configuration $|x\rangle$ and the backflow wave function $|\Phi_0^b\rangle$ is constructed from the "correlated" orbital with backflow correction:

$$U^b_{I,\alpha} = \eta_0 U_{I,\alpha} + \eta_1 \sum_{\langle j \rangle_i} D_i H_j U_{J,\alpha}, \tag{5.98}$$

where $I = i$ ($I = i + L$) for electrons with spin up (down), and equivalently for J and j; $\langle j \rangle_i$ indicates the sites j that are nearest neighbors of i; D_i (H_i) is the operator that gives 1 if the site i is doubly occupied (empty) and 0 otherwise; finally η_0 and η_1 are variational parameters. Then, the wave function for the given configuration $|x\rangle$ is obtained by taking the determinant of the matrix:

$$\tilde{\mathbf{U}}^b = \begin{pmatrix} U^b_{R_1,1} & \cdots & U^b_{R_1,N_e} \\ \vdots & \ddots & \vdots \\ U^b_{R_{N_e},1} & \cdots & U^b_{R_{N_e},N_e} \end{pmatrix}, \tag{5.99}$$

in which the rows correspond to the sites occupied by the electrons.

The fast update for a ratio between two determinants with configurations $|x\rangle$ and $|x'\rangle$, which differ by one or few electron hoppings, can be computed as before.

https://doi.org/10.1017/9781316417041.006 Published online by Cambridge University Press

Now, the only difference is that, even if there is only one electron hopping, the new matrix $\tilde{\mathbf{U}}^{b\prime}$ will differ not only for a single row, but for few of them. Indeed, by moving one electron, an empty or doubly occupied site can be created, leading to a modification of the "correlated" orbitals of Eq. (5.98). Therefore, in general, we must use the general update in which several rows are changed; however, in this case, it is convenient to work directly with the inverse matrix $\tilde{\mathbf{U}}'$, by updating it according to Eq. (5.90).

In practice, we can consider more general forms for the "correlated" orbitals, by considering further terms (Tocchio et al., 2011), but still remaining in the same spirit of considering a linear combination of non-interacting orbitals depending on the many-body configuration $|x\rangle$ (in the previous case, the linear combination is taken for configurations having holons and doublons at nearest-neighbor sites).

5.7 Fermionic Systems with Pfaffians

In the previous section, we have explained how to construct a correlated wave function starting from non-interacting electrons described by the Hamiltonian of Eq. (5.37), which naturally gives rise to the concept of single-particle orbitals. Here, the uncorrelated part of the many-body state is given by a determinant, see Eq. (5.46). By using a particle-hole transformation on spin-down electrons, the same approach can be also extended to BCS Hamiltonians containing a pairing among electrons with different spins, see Eq. (5.48). By contrast, when the pairing terms couple electrons with the same spin, orbitals cannot be defined, even after the particle-hole transformation and the uncorrelated state is written in terms of a pairing function. In turn, a determinant does not represent the most general uncorrelated anti-symmetric state that, instead, can be defined by means of the Pfaffian of an anti-symmetric matrix. In the following, we will show that, in the most general case, the uncorrelated part of the wave function is given by a Pfaffian, which may describe, as particular cases, states with singlet pairing or Slater determinants. This kind of mean-field wave function was first introduced by Bouchaud et al. (1988) and then used to study small atoms and molecules (Bajdich et al., 2006, 2008). Then, we will describe this kind of state and an efficient algorithm to perform local updates.

5.7.1 Definition of the Pfaffian

Before discussing the non-interacting state, we would like to summarize the definition and few important properties of the Pfaffian. Let us consider a $2n \times 2n$

https://doi.org/10.1017/9781316417041.006 Published online by Cambridge University Press

skew-symmetric matrix $\mathbf{A}$, namely a matrix having $a_{i,j} = -a_{j,i}$. The Pfaffian of $\mathbf{A}$ is defined as the anti-symmetrized product:

$$\mathrm{Pf}\mathbf{A} \equiv \mathcal{A}\left[a_{1,2}\, a_{3,4} \ldots a_{2n-1,2n}\right] = \sum_{\alpha} \mathrm{sign}(\alpha) \prod_{k=1}^{n} a_{i_k,j_k}, \tag{5.100}$$

where the sum runs over all the $(2n-1)!!$ pair partitions defined by $\alpha = \{(i_1,j_1),\ldots,(i_n,j_n)\}$ with $i_k < j_k$ and $i_1 < i_2 < \cdots < i_n$; $\mathrm{sign}(\alpha)$ indicates the parity of the permutation corresponding to the partition α. For example, for $n = 2$:

$$\mathrm{Pf}\begin{pmatrix} 0 & a_{1,2} & a_{1,3} & a_{1,4} \\ -a_{1,2} & 0 & a_{2,3} & a_{2,4} \\ -a_{1,3} & -a_{2,3} & 0 & a_{3,4} \\ -a_{1,4} & -a_{2,4} & -a_{3,4} & 0 \end{pmatrix} = a_{1,2}a_{3,4} - a_{1,3}a_{2,4} + a_{1,4}a_{2,3}. \tag{5.101}$$

Notice that $\det\mathbf{A} = (\mathrm{Pf}\mathbf{A})^2$; therefore, *apart form a global sign*, the calculation of the Pfaffian of a matrix can be reduced to the one of the determinant.

The Pfaffian satisfies the following relations:

$$\mathrm{Pf}\begin{pmatrix} \mathbf{A} & 0 \\ 0 & \mathbf{A}' \end{pmatrix} = \mathrm{Pf}\mathbf{A} \times \mathrm{Pf}\mathbf{A}', \tag{5.102}$$

where $\mathbf{A}$ and $\mathbf{A}'$ are $2n \times 2n$ and $2m \times 2m$ skew-symmetric matrices, respectively. Moreover:

$$\mathrm{Pf}\left(\mathbf{B}\mathbf{A}\mathbf{B}^T\right) = \det\mathbf{B} \times \mathrm{Pf}\mathbf{A}, \tag{5.103}$$

where $\mathbf{B}$ is an arbitrary $2n \times 2n$ matrix.

Finally, we report two important identities that will be used to demonstrate the fast updating of the Pfaffian. Let us consider two invertible skew-symmetric matrices $\mathbf{A}$ ($2n \times 2n$) and $\mathbf{C}$ ($2m \times 2m$) and a real matrix $\mathbf{B}$ ($2n \times 2m$). Then, we have that:

$$\begin{aligned} \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ -\mathbf{B}^T & \mathbf{C}^{-1} \end{pmatrix} &= \begin{pmatrix} \mathbf{1} & \mathbf{B}\mathbf{C} \\ \mathbf{0} & \mathbf{1} \end{pmatrix} \begin{pmatrix} \mathbf{A} + \mathbf{B}\mathbf{C}\mathbf{B}^T & \mathbf{0} \\ \mathbf{0} & \mathbf{C}^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ (\mathbf{B}\mathbf{C})^T & \mathbf{1} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ (\mathbf{A}^{-1}\mathbf{B})^T & \mathbf{1} \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}^{-1} + \mathbf{B}^T\mathbf{A}^{-1}\mathbf{B} \end{pmatrix} \begin{pmatrix} \mathbf{1} & \mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0} & \mathbf{1} \end{pmatrix}. \end{aligned} \tag{5.104}$$

Therefore, by using Eqs. (5.102) and (5.103), we arrive to:

$$\frac{\mathrm{Pf}\left(\mathbf{A} + \mathbf{B}\mathbf{C}\mathbf{B}^T\right)}{\mathrm{Pf}\mathbf{A}} = \frac{\mathrm{Pf}\left(\mathbf{C}^{-1} + \mathbf{B}^T\mathbf{A}^{-1}\mathbf{B}\right)}{\mathrm{Pf}\mathbf{C}^{-1}}. \tag{5.105}$$

https://doi.org/10.1017/9781316417041.006 Published online by Cambridge University Press

5.7.2 Definition of the Non-Interacting State

A case that cannot be brought back to the previous formalism based upon single-particle orbitals is when the original non-interacting Hamiltonian contains a pairing between the same spins or there are both spin-flip hopping and pairing terms. Indeed, here it is not possible to eliminate the pairing terms by performing particle-hole transformations. The most general quadratic Hamiltonian for a fermionic system is given by a generalized BCS Hamiltonian:

$$\mathcal{H}_0 = \sum_{i,j,\sigma,\tau} t_{i,j}^{\sigma,\tau} c_{i,\sigma}^\dagger c_{j,\tau} - \mu_0 \sum_{i,\sigma} c_{i,\sigma}^\dagger c_{i,\sigma} + \sum_{i,j,\sigma,\tau} \Delta_{i,j}^{\sigma,\tau} c_{i,\sigma}^\dagger c_{j,\tau}^\dagger + \text{h.c.}, \tag{5.106}$$

where $t_{i,j}^{\sigma,\tau} = \left(t_{j,i}^{\tau,\sigma}\right)^*$ and $\Delta_{j,i}^{\tau,\sigma} = -\Delta_{i,j}^{\sigma,\tau}$ are hopping and pairing terms, respectively.

The Hamiltonian $\mathcal{H}_0$ can be diagonalized by a generalized Bogoliubov transformation:

$$\Phi_\alpha = \sum_I \left(u_{I,\alpha} d_I + v_{I,\alpha} d_I^\dagger\right), \tag{5.107}$$

where the $2L \times 2L$ matrices $\mathbf{u}$ and $\mathbf{v}$ are determined by imposing:

$$\left[\mathcal{H}_0, \Phi_\alpha\right] = -\varepsilon_\alpha \Phi_\alpha, \tag{5.108}$$

as well as the conditions of orthogonality (that preserve the anti-commutation relations of fermionic operators):

$$\mathbf{u}^\dagger \mathbf{u} + \mathbf{v}^\dagger \mathbf{v} = \mathbf{1}, \tag{5.109}$$

$$\mathbf{u}^T \mathbf{v} + \mathbf{v}^T \mathbf{u} = \mathbf{0}. \tag{5.110}$$

Notice that, Eq. (5.108) allows pairs of solutions, since if Φ_α has eigenvalue ε_α then $\Phi_\alpha^\dagger$ will also be a solution with eigenvalue $-\varepsilon_\alpha$. Therefore, we can limit ourselves to the eigenvectors with non-negative eigenvalues.

In most cases, the ground state of the non-interacting Hamiltonian (5.106) can be written as a generalized BCS wave function:

$$|\Phi_0\rangle = \exp\left(\frac{1}{2} \sum_{I,J} F_{I,J} d_I^\dagger d_J^\dagger\right) |0\rangle, \tag{5.111}$$

where $F_{I,J}$ is the anti-symmetric pairing function, defined by imposing that $|\Phi_0\rangle$ is annihilated by all the operators Φ_α with $\varepsilon_\alpha \geq 0$, which leads to:

$$\mathbf{F}\mathbf{u} = \mathbf{v}. \tag{5.112}$$

As in the case with determinants, also here we must pay attention in defining a unique many-body wave function. Therefore, only the cases without vanishing eigenvalues ε_α must be considered. Moreover, in order to have a non-singular pairing function $\mathbf{F}$, we must also require that the matrix $\mathbf{u}$ is invertible. Notice that

https://doi.org/10.1017/9781316417041.006 Published online by Cambridge University Press

a singular $\mathbf{u}$ corresponds to the existence of occupied single-particle orbitals. In this case, the many-body state must be constructed including these unpaired orbitals, increasing the computational complexity of the algorithm (Bajdich et al., 2006, 2008). In the following, for the fast update, we will limit ourselves to non-singular cases.

The overlap between $|\Phi_0\rangle$ and a generic configuration of Eq. (5.44) is given by:

$$\langle x|\Phi_0\rangle = \frac{1}{(N_e/2)!}\langle 0|d_{R_{R_e}}\dots d_{R_1}\left(\frac{1}{2}\sum_{I,J}F_{I,J}d_I^\dagger d_J^\dagger\right)^{N_e/2}|0\rangle, \tag{5.113}$$

then, by expanding the product of the sums, we get a sum over the $(N_e-1)!!$ pair partitions of the fermions in $|x\rangle$ of $(N_e/2)!$ identical contributions. The ordering of the fermion operators gives rise to the sign appearing in the definition of the Pfaffian, see Eq. (5.100), thus leading to:

$$\langle x|\Phi_0\rangle = \mathrm{Pf}\left\{F_{R_i,R_j}\right\}. \tag{5.114}$$

5.7.3 Fast Computation of the Pfaffian

We now show how to compute efficiently the ratio of two Pfaffians when two configurations $|x\rangle$ and $|x'\rangle$ differ by one electron hopping. In particular, suppose that the l-th electron in R_l changes its position to R'_l, i.e., $|x'\rangle = d^\dagger_{R'_l}d_{R_l}|x\rangle$. The new matrix $\mathbf{F}'$ that is needed to compute the overlap with $|x'\rangle$ differs from the old one $\mathbf{F}$ only in the l-th row and column. According to Eq. (5.114), the overlap $\langle x|\Phi_0\rangle$ is given by the Pfaffian of the matrix:

$$\tilde{\mathbf{F}} = \begin{pmatrix} F_{R_1,R_1} & \dots & F_{R_1,R_l} & \dots & F_{R_1,R_{N_e}} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ F_{R_l,R_1} & \dots & F_{R_l,R_l} & \dots & F_{R_l,R_{N_e}} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ F_{R_{N_e},R_1} & \dots & F_{R_{N_e},R_l} & \dots & F_{R_{N_e},R_{N_e}} \end{pmatrix}, \tag{5.115}$$

which can be constructed from the original matrix $\mathbf{F}$. As for the caste with determinants, the actual order of rows (an columns) is determined by the one of creation operators in the configuration $|x\rangle$, see Eq. (5.44). Then, a generalization of the vector $\boldsymbol{\kappa}$ (see section 5.4) for the spinful case can be used to select the correct order of columns and rows. Then, the overlap $\langle x'|\Phi_0\rangle$ is given by the Pfaffian of:

$$\tilde{\mathbf{F}}' = \begin{pmatrix} F_{R_1,R_1} & \dots & F_{R_1,R'_l} & \dots & F_{R_1,R_{N_e}} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ F_{R'_l,R_1} & \dots & F_{R'_l,R'_l} & \dots & F_{R'_l,R_{N_e}} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ F_{R_{N_e},R_1} & \dots & F_{R_{N_e},R'_l} & \dots & F_{R_{N_e},R_{N_e}} \end{pmatrix}. \tag{5.116}$$

https://doi.org/10.1017/9781316417041.006 Published online by Cambridge University Press

In a compact form, we have:

$$\tilde{F}'_{i,j} = \tilde{F}_{i,j} + \delta_{i,l}\left(F_{R'_l,R_j} - \tilde{F}_{l,j}\right) - \delta_{j,l}\left(F_{R'_l,R_i} - \tilde{F}_{l,i}\right), \tag{5.117}$$

which can be cast in the form of $\tilde{\mathbf{F}}' = \tilde{\mathbf{F}} + \mathbf{BCB}^T$ where $\mathbf{B}$ is a $N_e \times 2$ matrix:

$$\mathbf{B} = \begin{pmatrix} F_{R_1,R'_l} - \tilde{F}_{1,l} & \delta_{1,l} \\ \vdots & \vdots \\ F_{R_{N_e},R'_l} - \tilde{F}_{N_e,l} & \delta_{N_e,l} \end{pmatrix}, \tag{5.118}$$

and $\mathbf{C}$ is a 2×2 matrix:

$$\mathbf{C} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{5.119}$$

Therefore, by using Eq. (5.105), we get a simple expression for the ratio of the Pfaffians:

$$\frac{\langle x'|\Phi_0\rangle}{\langle x|\Phi_0\rangle} = \frac{\langle x|d^\dagger_{R'_l} d_{R'_l}|\Phi_0\rangle}{\langle x|\Phi_0\rangle} = \frac{\mathrm{Pf}\tilde{\mathbf{F}}'}{\mathrm{Pf}\tilde{\mathbf{F}}} = \frac{\mathrm{Pf}(\mathbf{C}^{-1} + \mathbf{B}^T\tilde{\mathbf{F}}^{-1}\mathbf{B})}{\mathrm{Pf}\mathbf{C}}, \tag{5.120}$$

which can be easily computed since it requires the computation of the Pfaffian of a 2×2 matrix, being $\mathrm{Pf}\mathbf{C}^{-1} = -1$. The explicit form is worked out by using Eq. (5.118) for the matrix $\mathbf{B}$:

$$\frac{\mathrm{Pf}\tilde{\mathbf{F}}'}{\mathrm{Pf}\tilde{\mathbf{F}}} = \sum_j F_{R'_l,R_j}\tilde{F}^{-1}_{j,l} \equiv G_{R'_l,l}. \tag{5.121}$$

Having computed the matrix $\tilde{\mathbf{F}}^{-1}$ at the beginning of the simulation, this calculation requires $O(N_e)$ operations. As before, once the new configuration $|x'\rangle$ is accepted along the Markov chain, the matrix $\tilde{\mathbf{F}}^{-1}$ must be updated. From Eq. (5.117), we have that:

$$\tilde{F}'_{i,j} = \sum_m \tilde{F}_{i,m}\left(\delta_{m,j} + \tilde{F}^{-1}_{m,l} v_j^{R'_l,R_l} - w_m^{R'_l,R_l}\delta_{j,l}\right), \tag{5.122}$$

where

$$v_j^{R'_l,R_l} = F_{R'_l,R_j} - \tilde{F}_{l,j}, \tag{5.123}$$

$$w_m^{R'_l,R_l} = \sum_j \tilde{F}^{-1}_{m,j}(F_{R'_l,R_j} - \tilde{F}_{l,j}) = \delta_{m,l} - G_{R'_l,m}. \tag{5.124}$$

Then, the inverse $\tilde{\mathbf{F}}^{-1}$ can be easily found by following the same strategy as in section 5.6:

$$\tilde{F}^{-1\prime}_{i,j} = \tilde{F}^{-1}_{i,j} + \frac{1}{G_{R'_l,l}}\left(\tilde{F}^{-1}_{i,l} w_j^{R'_l,R_l} + w_i^{R'_l,R_l}\tilde{F}^{-1}_{l,j}\right). \tag{5.125}$$

https://doi.org/10.1017/9781316417041.006 Published online by Cambridge University Press

The important difference with the case of section 5.6 is that the updating cannot be written by using the matrix $\mathbf{G}$ alone, but $\tilde{\mathbf{F}}^{-1}$ must be also kept. In fact, by multiplying both sides of Eq. (5.125) by F_{R,R_i} and summing over i, we have:

$$G'_{R,j} - (F_{R,R'_l} - F_{R,R_l})\tilde{F}^{-1\prime}_{l,j} = G_{R,j} + \frac{1}{G_{R'_l,l}}\left(G_{R,l}w_j^{R'_l,R_l} + \sum_i F_{R,R_i}w_i^{R'_l,R_l}\tilde{F}^{-1}_{l,j}\right), \quad (5.126)$$

where we have used the fact that:

$$G'_{R,j} = \sum_i F_{R,R'_i}\tilde{F}^{-1\prime}_{i,j} = \sum_i \left[F_{R,R_i} + \delta_{i,l}\left(F_{R,R'_l} - F_{R,R_l}\right)\right]\tilde{F}^{-1\prime}_{i,j}. \quad (5.127)$$

Then, by using Eq. (5.125) to express $\tilde{F}^{-1\prime}_{l,j}$ in terms of $\tilde{F}^{-1}_{l,j}$, we finally get:

$$G'_{R,j} = G_{R,j} + \frac{1}{G_{R'_l,l}}\left[G_{R,l}w_j^{R'_l,R_l} + \left(F_{R,R'_l} + \sum_k G_{R,k}F_{R'_l,R_k}\right)\tilde{F}^{-1}_{l,j}\right]. \quad (5.128)$$

Therefore, a fast updating can be performed by considering the matrices $\mathbf{G}$ and $\tilde{\mathbf{F}}^{-1}$ and using Eqs. (5.125) and (5.128) every time the proposed new configuration is accepted.

Within the Pfaffian wave function, in addition to the standard (equal-time) Green's function:

$$\frac{\langle x|d^\dagger_{R_l}d_R|\Phi_0\rangle}{\langle x|\Phi_0\rangle} = G_{R,l}, \quad (5.129)$$

also the anomalous (superconducting) ones are present:

$$\frac{\langle x|d^\dagger_{R_l}d^\dagger_{R_k}|\Phi_0\rangle}{\langle x|\Phi_0\rangle} = \tilde{F}^{-1}_{k,l}, \quad (5.130)$$

$$\frac{\langle x|d_R d_{R'}|\Phi_0\rangle}{\langle x|\Phi_0\rangle} = F_{R,R'} + \sum_k G_{R,k}F_{R',R_k}. \quad (5.131)$$

As before, the interpretation of Eqs. (5.125) and (5.128) is given by the Wick theorem (when *bra* and *ket* states are different) for the updated Green's functions $G'_{R,j}$ and $\tilde{F}^{-1\prime}_{i,j}$; in this case, the anomalous contractions are also present, since $|\Phi_0\rangle$ has not a fixed number of particles.

The generalization to the case where more than one electron hop is straightforward by using Eq. (5.105) but it is very cumbersome. However, we can bypass the m-electron update by applying successively single-electron updates.

https://doi.org/10.1017/9781316417041.006 Published online by Cambridge University Press

5.8 Energy and Correlation Functions

The variational energy or any other observable, including correlation functions, can be easily computed by using the same tricks of the fast updates that we have discussed in the previous sections. Indeed, the local energy or other local observables are written in terms of a sum of ratios of wave functions times the matrix elements of the operator $\langle x|\mathcal{O}|x'\rangle$, according to Eq. (5.13). Let us consider a generic n-body operator:

$$\mathcal{O} = \sum_{I_1,\ldots,I_n} \sum_{J_1,\ldots,J_n} \mathcal{O}_{I_1,\ldots,I_n;J_1,\ldots,J_n} d_{I_1}^\dagger d_{J_1} \ldots d_{I_n}^\dagger d_{J_n}, \tag{5.132}$$

where $d_I^\dagger$ (d_I) creates (destroys) an electron on the site $I = 1,\ldots,2L$ (a similar definition holds for a bosonic system). Then, the local estimator of Eq. (5.5) is given by:

$$\mathcal{O}_L(x) = \sum_{I_1,\ldots,I_n} \sum_{J_1,\ldots,J_n} \mathcal{O}_{I_1,\ldots,I_n;J_1,\ldots,J_n} \frac{\langle x|d_{I_1}^\dagger d_{J_1} \ldots d_{I_n}^\dagger d_{J_n}|\Psi\rangle}{\langle x|\Psi\rangle}. \tag{5.133}$$

The ratio in the r.h.s. of this equation can be easily computed by using the results discussed in the sections dedicated to the fast-update algorithms.

5.9 Practical Implementation

Here, we would like to sketch the important steps in a practical implementation of the variational Monte Carlo algorithm.

1. **Initialization** at the beginning of the calculation.
   - Generate a random state of the basis set $|x\rangle$ that is stored into a vector `iconf`(L), whose elements give the local configuration on the site $i = 1,\ldots,L$. For example, for the fermionic (single-band) Hubbard model, `iconf`(i) can assume the values 0 (empty site), ± 1 (one electron with spin up or down), or 2 (doubly-occupied site). In order to change the correct row of the matrix $\tilde{\mathbf{U}}$ (for the determinant) or row and column of $\tilde{\mathbf{F}}$ (for the Pfaffian), it is also important to store a vector `kel`$(2L)$, whose non-zero elements give, for each site, the position of the creation operators in the string defining the sampled configuration $|x\rangle$ of Eq. (5.44). For the bosonic Hubbard model, `iconf`(i) can assume all the non-negative integer values and there is no need to store `kel`$(2L)$.
   - Verify that the initial configuration is not singular, i.e., $\langle x|\Psi_J\rangle \neq 0$. The Jastrow factor usually does not give rise to any problem in this sense and we must only check whether $\langle x|\Phi_0\rangle$ is vanishing or not. Indeed, by working with

https://doi.org/10.1017/9781316417041.006 Published online by Cambridge University Press

fermionic states, it may happen that $\langle x|\Phi_0\rangle = 0$ for some $|x\rangle$. Although this is not a problem along the simulation, since if one of these singular configurations is proposed it will not be accepted by the Metropolis algorithm, it would be a problem to initialize the Markov chain with a singular configuration (because the first acceptance probability in the Metropolis algorithm would have a vanishing denominator). Notice that the typical value for the overlap $\langle x|\Phi_0\rangle$ is exponentially small with the size of the system; however, within the fast update algorithms, we never need to compute the actual values of determinants or Pfaffians but only ratios of them. Therefore, at the beginning of the simulation, we must only require that the matrix $\tilde{\mathbf{U}}$ is invertible without numerical roundoff, in order to construct the equal-time Green's functions.

- Compute the table of Eq. (5.33) to perform the fast update of the Jastrow factor and all the Green's functions that are necessary to perform the fast update. For the determinant case, only the Green's function of Eq. (5.80) is necessary, while for the Pfaffian case both the standard Green's function (5.129) and the anomalous ones (5.130) and (5.131) are needed.

2. **Markov process** with the Metropolis algorithm.
   - Propose a new (random) configuration $|x'\rangle$ by moving one or few particles.
   - Compute the ratio between the new and the old wave functions, see Eqs. (5.78), (5.88), or (5.121).
   - Accept or reject the proposed configuration according to the Metropolis algorithm with Eq. (5.12).
   - If the new configuration is accepted, update the table for the Jastrow factor (5.34) and all the Green's functions, i.e., Eq. (5.79) for determinants or Eqs. (5.125) and (5.128) for Pfaffians.
3. **Computation of observables**.

   Observables (e.g., the variational energy) can be computed every $O(L)$ steps (one step corresponds to propose a move in which one or few particles are moved) in order to have uncorrelated configurations. Of course, the frequency at which the observables are computed depends upon the acceptance ratio, e.g., if the latter one is small, we need to perform more Markov steps before recomputing the observable, in order to decorrelate subsequent measurements. A further binning technique (see section 3.11) can be applied to reduce correlation.

   It is always recommended to write the observables on the hard-disk and then perform the statistical analysis, since the correlation time is not usually known. Indeed, the calculation of observables during the Monte Carlo simulation would require the knowledge of this quantity before starting the simulation. In the unfortunate case in which this is not correctly estimated, we must restart the simulation, which may have a large computational cost. Instead, the post-processing analysis is usually much less expensive, since it only requires simple operations.