8

Green's Function Monte Carlo

8.1 Basic Notions and Formal Derivations

In this Chapter, we describe both the basic principles and the more advanced details that are necessary to implement the so-called Green's function Monte Carlo (GFMC) technique. Within this approach, the ground-state wave function of a given Hamiltonian is stochastically sampled by using the power method that has been discussed in section 1.7. As already done in Chapter 5 for the variational Monte Carlo technique, we fix a complete basis set $\{|x\rangle\}$ in the Hilbert space:

$$\sum_x |x\rangle\langle x| = \mathbb{I}, \tag{8.1}$$

in which the states are taken to be *orthogonal* and *normalized*. Moreover, we assume that, given the Hamiltonian $\mathcal{H}$, the matrix elements $\mathcal{H}_{x',x} = \langle x'|\mathcal{H}|x\rangle$ can be computed efficiently for each $|x\rangle$ and $|x'\rangle$. For local Hamiltonians, although the dimension of the Hilbert space increases exponentially with the system size L, for each $|x\rangle$, the number of the non-zero elements $\mathcal{H}_{x',x}$ scales with L; therefore, given the state of the basis $|x\rangle$, all matrix elements $\mathcal{H}_{x',x}$ can be computed with a reasonable computational effort. By using this property, it is possible to define a stochastic algorithm that allows us to perform the power method (see section 1.7) in a statistical way (Ceperley and Alder, 1980; Reynolds et al., 1982; Foulkes et al., 2001). Here, the ground-state wave function $|\Upsilon_0\rangle$ is filtered out from an initial state $|\Psi_0\rangle$, by using a suitable projection operator:

$$\lim_{n\to\infty} (\Lambda - \mathcal{H})^n |\Psi_0\rangle \propto |\Upsilon_0\rangle, \tag{8.2}$$

where Λ is a diagonal operator with $\Lambda_{x,x} = \lambda$. The necessary requirements to approach the ground state $|\Upsilon_0\rangle$ is to take a sufficiently large value of λ and

https://doi.org/10.1017/9781316417041.009 Published online by Cambridge University Press

choose an initial state $|\Psi_0\rangle$ such that $\langle \Upsilon_0 | \Psi_0 \rangle \neq 0$. In practice, Eq. (8.2) can be implemented iteratively:

$$|\Psi_{n+1}\rangle = (\Lambda - \mathcal{H})|\Psi_n\rangle. \tag{8.3}$$

By expanding over the given basis set, we have:

$$\Psi_{n+1}(x') = \sum_x \mathcal{G}_{x',x} \Psi_n(x), \tag{8.4}$$

where $\Psi_n(x) = \langle x | \Psi_n \rangle$ and

$$\mathcal{G}_{x',x} = \langle x' | (\Lambda - \mathcal{H}) | x \rangle \tag{8.5}$$

is the so-called Green's function (even though it does not correspond to a Green's function). On the lattice, the GFMC technique is particularly simple and has been used for the first time to study the ground-state properties of the Heisenberg model on the square lattice (Trivedi and Ceperley, 1989, 1990).

Within a statistical implementation of the power method, we would be tempted to interpret Eq. (8.4) as a Master equation for a stochastic variable where $\Psi_n(x)$ represents the probability distribution at the iteration n and $\mathcal{G}_{x',x}$ is the transition probability. However, some important points must be elucidated in order to reach a final statistical interpretation of this quantum evolution. First of all, within this approach, *all* the matrix elements of the Green's function $\mathcal{G}_{x',x}$ have to be non-negative, as required for defining a transition probability. As far as the diagonal elements $\mathcal{G}_{x,x}$ are concerned, there is no problem, since we can always define a sufficiently large and positive value of λ that gives $\mathcal{G}_{x,x} \geq 0$. By contrast, the requirement that off-diagonal elements are non-negative is highly non trivial and, indeed, is satisfied only by certain models; if $\mathcal{G}_{x',x} < 0$ for some couples (x', x), we say that we are in presence of the *sign problem*, which will be discussed in Chapter 10. In the following, we will assume that the Hamiltonian is such that $\mathcal{G}_{x',x} \geq 0$ for all the couples (x, x'). Furthermore, even when the Green's function is non-negative, we generically have that:

$$\sum_{x'} \mathcal{G}_{x',x} \neq 1, \tag{8.6}$$

and thus $\mathcal{G}_{x',x}$ cannot be interpreted as a transition probability, since it is not normalized. Notice that Eq. (8.6) implies that the normalization of the wave function is not conserved along the projection technique of Eq. (8.4). Nevertheless, we can always split up the Green's function into the product of two factors: a stochastic matrix $p_{x',x}$ that represents a *bona fide* transition probability, and a real number b_x, which is the normalization of the Green's function:

$$\mathcal{G}_{x',x} = p_{x',x} b_x, \tag{8.7}$$

https://doi.org/10.1017/9781316417041.009 Published online by Cambridge University Press

where b_x is given by:

$$b_x = \sum_{x'} \mathcal{G}_{x',x}, \tag{8.8}$$

and then the matrix $p_{x',x}$ can be obtained by:

$$p_{x',x} = \frac{\mathcal{G}_{x',x}}{b_x}. \tag{8.9}$$

We now want to devise a simple Markov process from which the evolution of the wave function of Eq. (8.4) can be obtained. Since the Green's function cannot be directly interpreted as a transition probability, a description with only the configuration x_n (that identifies the quantum state $|x\rangle$ at the iteration step n) is not sufficient. The simplest way of considering the presence of the scale factor b_x is to add a weight w_n in the statistical description of the projection method. Therefore, the Markov process is defined by a dyad (x_n, w_n); its evolution, according to the decomposition of Eqs. (8.8) and (8.9) is described by:

$$\text{1) generate } x_{n+1} = x' \text{ with probability } p_{x',x_n}, \tag{8.10}$$

$$\text{2) update the weight with } w_{n+1} = w_n b_x. \tag{8.11}$$

Notice that here the new configuration x_{n+1} can be the same as x_n, because it is selected among all the ones that have $p_{x',x_n} \neq 0$, including also x_n when $p_{x_n,x_n} \neq 0$ (this is the case when $\mathcal{G}_{x,x} = \lambda - \mathcal{H}_{x,x} \neq 0$). The Markov process describes a *diffusion* of the configuration x_n (as well as its weight w_n) with the following transition probability:

$$K(x', w'|x, w) = p_{x',x}\, \delta(w' - w b_x), \tag{8.12}$$

which defines the new dyad (x', w') given the old one (x, w). Thus, the Master equation corresponding to the probability density $\mathcal{P}_n(x, w)$ is:

$$\mathcal{P}_{n+1}(x', w') = \sum_x \int dw\; K(x', w'|x, w)\; \mathcal{P}_n(x, w). \tag{8.13}$$

Hereafter, the integration limits over the variable w are assumed to run from $-\infty$ to $+\infty$, the probability density $\mathcal{P}_n(x, w)$ being zero for the values of w that are not reached along the Markov chain (e.g., for $w < 0$). Within this formalism, the wave function $\Psi_n(x)$ can be obtained from integrating over the weight the probability density $\mathcal{P}_n(x, w)$ multiplied by w. Indeed, from the Master equation (8.13) and the definition of the transition probability of Eq. (8.12), we have that:

$$\mathcal{P}_{n+1}(x', w') = \sum_x p_{x',x} \int dw\; \delta(w' - w b_x)\; \mathcal{P}_n(x, w); \tag{8.14}$$

https://doi.org/10.1017/9781316417041.009 Published online by Cambridge University Press

by multiplying both sides by w' and then integrating over w', we have:

$$\int dw'\ w'\ \mathcal{P}_{n+1}(x', w') = \sum_x p_{x',x}\ b_x \int dw\ w\ \mathcal{P}_n(x, w)$$
$$= \sum_x \mathcal{G}_{x',x} \int dw\ w\ \mathcal{P}_n(x, w), \tag{8.15}$$

which is just the original iterative procedure of Eq. (8.4), once the following identification is performed:

$$\Psi_n(x) \equiv \int dw\ w\ \mathcal{P}_n(x, w). \tag{8.16}$$

Therefore, after a large number n of iterations, $\mathcal{P}_n(x, w)$ converges to an equilibrium distribution $\mathcal{P}_{\text{eq}}(x, w)$, which determines the ground-state wave function $\Upsilon_0(x)$:

$$\Upsilon_0(x) \equiv \int dw\ w\ \mathcal{P}_{\text{eq}}(x, w). \tag{8.17}$$

In the following, we will describe a practical implementation of the Markov process, which allows us to compute observables over the ground state $|\Upsilon_0\rangle$, once the high-energy components present in $|\Psi_0\rangle$ are filtered out according to the iterative procedure of Eq. (8.4).

8.2 Single Walker Technique

Here, we explain how to implement the Markov process of Eqs. (8.10) and (8.11) within the simplest approach. To this purpose, we define the basic element of this stochastic process, the so-called *walker*. A walker is determined by an index x, labelling the configuration $|x\rangle$, and a weight w, which is associated to the amplitude of the wave function, as given by Eq. (8.16). The walker changes its configuration and weight by performing a Markovian process with a discrete iteration time n: the dyad (x_n, w_n), which denotes the walker at the time n, is distributed according to $\mathcal{P}_n(x, w)$. Most importantly, the walker determines, in a statistical sense, the quantum state $\Psi_n(x)$:

$$\Psi_n(x) \equiv \int dw\ w\ \mathcal{P}_n(x, w) \approx \langle\langle w_n \delta_{x,x_n} \rangle\rangle, \tag{8.18}$$

where $\langle\langle \dots \rangle\rangle$ denotes the statistical average over (infinitely) many independent realizations of the Markov chain with configurations (x_n, w_n). We would like to mention that the actual determination of the wave function $\Psi_n(x)$ is rarely pursued in the actual practice, as in a many-body system the Hilbert space is exponentially large and the information on the wave function cannot be even stored in the computer memory. Moreover, in order to have an adequate errorbar on the configurations, each of them must be visited many times, implying an exponentially

https://doi.org/10.1017/9781316417041.009 Published online by Cambridge University Press

large computation time. Remarkably, quantum Monte Carlo approaches are based on the fact that it is not necessary to have an accurate statistical information on the wave function to obtain a reliable estimation of its energy or correlation functions.

In practice, we can use a single walker and start from an initial condition with x_0, corresponding to a given many-body configuration $|x_0\rangle$, and a weight $w_0 = 1$, so to have:

$$\mathcal{P}_0(x, w) = \delta(w - 1)\delta_{x,x_0}. \tag{8.19}$$

Then, we evolve the walker by using Eq. (8.10) and (8.11). The Markov process can be very easily implemented for generic Hamiltonians on a lattice, since the number of non-zero entries in the stochastic matrix p_{x',x_n}, for given x_n, is small, and moderately growing with the number of lattice sites L. Thus, in order to define x_{n+1}, it is enough to divide the interval $[0, 1)$ into smaller intervals for all possible $\{x'\}$ connected to x_n with non-zero probability p_{x',x_n} and generate a random number ξ between 0 and 1. As discussed in section 3.5, ξ will lie in one of the above defined intervals, with a probability of hitting the interval corresponding to a certain x' exactly equal to p_{x',x_n}, see Fig. 8.1. The weight w_{n+1} is just obtained by multiplying w_n by b_{x_n}.

Let us finally discuss the evolution of the marginal probability of the configuration x alone:

$$\Pi_n(x) = \int dw\, \mathcal{P}_n(x, w). \tag{8.20}$$

Since, within the single-walker implementation, the evolution of the configuration x_n is independent from its weight, as given by Eqs. (8.10) and (8.11), it is easy to write down the Master equation for $\Pi_n(x)$:

$$\Pi_{n+1}(x') = \sum_x p_{x',x}\Pi_n(x). \tag{8.21}$$

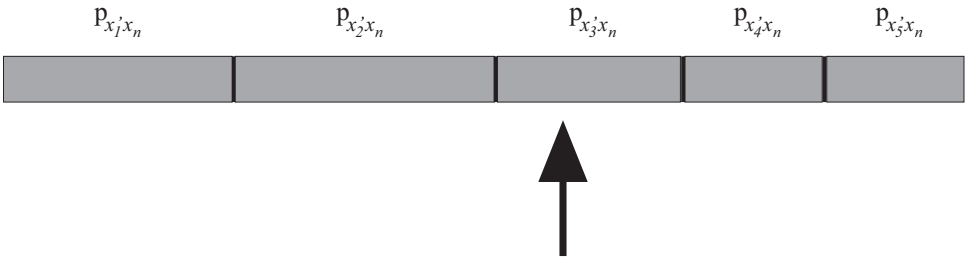

Figure 8.1 The interval $[0, 1)$ is divided into sub-intervals of length equal to the probabilities p_{x_i',x_n}, for all i's labeling the possible configurations x' with non-zero p_{x_i',x_n} (here, we show the case with 5 entries). Then, a random number $0 \leq \xi < 1$ is generated and the new configuration x_{n+1} is selected (here $i = 3$), as ξ (here denoted by the arrow) lies in the corresponding interval, see section 3.5.

https://doi.org/10.1017/9781316417041.009 Published online by Cambridge University Press

Therefore, at equilibrium, the marginal probability is given by the right eigenvector of the stochastic matrix $p_{x',x}$ with eigenvalue $\lambda_0 = 1$, which is given by:

$$\Pi_{\text{eq}}(x) = \frac{b_x}{\sum_{x'} b_{x'}}. \tag{8.22}$$

Indeed, we have that:

$$\sum_x p_{x',x} b_x = \sum_x \mathcal{G}_{x',x} = \sum_x \mathcal{G}_{x,x'} = b_{x'}, \tag{8.23}$$

where we have just used the fact that the matrix $\mathcal{G}_{x',x}$ is symmetric. Then, for an ergodic Hamiltonian, which implies an ergodic matrix $p_{x',x}$, the probability $\Pi_n(x)$ converges (with an exponential rate) to $\Pi_{\text{eq}}(x)$, no matter what is the initial condition, see section 3.8.

8.2.1 Ground-State Energy

Let us now discuss how to compute physical observables and start by considering the ground-state energy E_0. Since for large enough n, $\mathcal{H}\mathcal{G}^n|\Psi_0\rangle \propto E_0|\Upsilon_0\rangle$, we have that:

$$E_0 \approx \frac{\sum_{x'} \langle x'|\mathcal{H}\mathcal{G}^n|\Psi_0\rangle}{\sum_{x'} \langle x'|\mathcal{G}^n|\Psi_0\rangle} = \frac{\sum_{x,x'} \mathcal{H}_{x',x}\langle x|\mathcal{G}^n|\Psi_0\rangle}{\sum_{x'} \langle x'|\mathcal{G}^n|\Psi_0\rangle}, \tag{8.24}$$

which, by using Eq. (8.18) for $\Psi_n(x) = \langle x|\mathcal{G}^n|\Psi_0\rangle$, gives:

$$E_0 \approx \frac{\langle\langle e_L(x_n) w_n\rangle\rangle}{\langle\langle w_n\rangle\rangle}, \tag{8.25}$$

where we have defined the local energy as:

$$e_L(x) = \sum_{x'} H_{x',x} = \lambda - b_x. \tag{8.26}$$

Therefore, the ground-state energy can be obtained by considering several independent calculations with n steps (n being large enough to filter out the ground-state wave function from the initial state $|\Psi_0\rangle$) and then averaging them together.

A more efficient and straightforward way of performing the numerical calculation is to do a single simulation of length $N_{\text{sim}} \gg n$; after a thermalization time, the configurations x along the Markov process will be equilibrated according to $\Pi_{\text{eq}}(x)$ of Eq. (8.22); then, we can imagine to start, from *each* step $n - p$, a projection technique of length p. In this sense, the initial probability is no longer given by Eq. (8.19), but instead by:

$$\mathcal{P}_0(x, w) = \delta(w - 1)\Pi_{\text{eq}}(x). \tag{8.27}$$

https://doi.org/10.1017/9781316417041.009 Published online by Cambridge University Press

After p steps, the accumulated weight for each projection process is given by:

$$G_n^p = \prod_{j=1}^{p} b_{x_{n-j}}. \tag{8.28}$$

Then, the ground-state energy can be estimated by taking a sufficiently large value of p in the accumulated weight:

$$E_0 \approx \frac{\sum_n G_n^p e_L(x_n)}{\sum_n G_n^p}. \tag{8.29}$$

In practice, in order to avoid numerical overflows/underflows, it is useful to store (e.g., in the hard disk) the factor b_{x_n} for each Markov step, or even better $b_{x_n}/\overline{b}$ (where $\overline{b}$ is an estimation of the average value of b_{x_n}), without performing many multiplications. In the post-processing analysis, we can compute G_n^p for the smallest possible value of p that gives a converged energy and evaluate E_0 through Eq. (8.29).

This simple procedure gives an exact estimation of the ground-state energy. However, few remarks should be done. First of all, when computing the error-bars, we must keep in mind that, by taking each iteration as a starting point for the projection technique, we have correlated measures, which require a binning approach, as described in section 3.11. Most importantly, according to Eq. (8.26), the calculation of the energy does not satisfy the zero-variance property because the random quantity $e_L(x)$ does not depend on any variational guess. In this sense, the inclusion of the importance sampling is fundamental to reduce the statistical fluctuations.

8.3 Importance Sampling

Let us now describe how it is possible to implement importance sampling within the GFMC technique. This can be done by introducing the so-called *guiding function* $\Psi_G(x)$, which must be real (see below) and non-vanishing for all configurations $|x\rangle$. Usually, it is chosen to be the best variational *Ansatz* or very close to it. By multiplying both sides of Eq. (8.4) by $\Psi_G(x')$, we obtain:

$$\Psi_G(x')\Psi_{n+1}(x') = \sum_x \mathcal{G}_{x',x} \frac{\Psi_G(x')}{\Psi_G(x)} \Psi_G(x)\Psi_n(x). \tag{8.30}$$

Then, by defining the Green's function with importance sampling:

$$\tilde{\mathcal{G}}_{x',x} = \mathcal{G}_{x',x} \frac{\Psi_G(x')}{\Psi_G(x)}, \tag{8.31}$$

https://doi.org/10.1017/9781316417041.009 Published online by Cambridge University Press

we have that the wave function

$$\tilde{\Psi}_n(x) = \Psi_G(x)\Psi_n(x) \tag{8.32}$$

satisfies the same evolution as Eq. (8.4) with $\tilde{\mathcal{G}}$ replacing $\mathcal{G}$:

$$\tilde{\Psi}_{n+1}(x') = \sum_x \tilde{\mathcal{G}}_{x',x}\tilde{\Psi}_n(x). \tag{8.33}$$

We would like to mention the important fact that Eq. (8.31) represents a similarity transformation (also used when discussing the Markov chains in section 3.8). Such transformation does not modify the spectrum of the matrix, i.e., $\mathcal{G}$ and $\tilde{\mathcal{G}}$ have the same eigenvalues, while it trivially changes the eigenvectors, i.e., if $\upsilon(x)$ is an eigenvector of $\mathcal{G}$, then $\tilde{\upsilon}(x) = \Psi_G(x)\upsilon(x)$ is a (right) eigenvector of $\tilde{\mathcal{G}}$. Notice that, in general, $\tilde{\mathcal{G}}_{x',x}$ is no longer symmetric.

Whenever $\tilde{\mathcal{G}}_{x',x}$ is real and non-negative for all couples (x', x), it is possible to apply the same decomposition of Eq. (8.7) in terms of a conditional probability and a real coefficient (the fact of having a real Green's function implies that the guiding function must be also real):

$$\tilde{\mathcal{G}}_{x',x} = \tilde{p}_{x',x}\tilde{b}_x, \tag{8.34}$$

where:

$$\tilde{b}_x = \sum_{x'} \tilde{\mathcal{G}}_{x',x}, \tag{8.35}$$

$$\tilde{p}_{x',x} = \frac{\tilde{\mathcal{G}}_{x',x}}{\tilde{b}_x}. \tag{8.36}$$

Then, the corresponding Markov process can be devised, in analogy to Eqs. (8.10) and (8.11).

Two important facts can be achieved by using the importance sampling transformation. First of all, the exact value of the ground-state energy E_0 is obtained without statistical fluctuations whenever the guiding wave function $\Psi_G(x)$ is the correct ground-state wave function (see below). Moreover, the guiding function can be used as a remedy to the sign problem: there are cases where the original Green's function $\mathcal{G}_{x',x}$ does not satisfy the non-negativity condition for all off-diagonal matrix elements, and this fact can be adjusted by including a properly chosen $\Psi_G(x)$. The antiferromagnetic nearest-neighbor Heisenberg model on the square lattice represents the simplest example in which the Green's function can be negative, i.e., $\mathcal{G}_{x',x} = -J/2$ for configurations $|x\rangle$ and $|x'\rangle$ having two nearest-neighbor spins with opposite orientations. Once considering a guiding function with the so-called Marshall sign (Marshall, 1955):

$$\text{Sign}[\Psi_G(x)] = (-1)^{N_{A,\uparrow}(x)}, \tag{8.37}$$

https://doi.org/10.1017/9781316417041.009 Published online by Cambridge University Press

where $N_{A,\uparrow}(x)$ is the number of up spin in one sub-lattice in the configuration $|x\rangle$, the Green's function with importance sampling is non-negative for all couples of configurations.

In presence of importance sampling, all the derivations of section 8.2 can be repeated. For example, Eq. (8.14) is replaced by:

$$\tilde{\mathcal{P}}_{n+1}(x', \tilde{w}') = \sum_x \tilde{p}_{x',x} \int d\tilde{w}\ \delta(\tilde{w}' - \tilde{w}\tilde{b}_x)\ \tilde{\mathcal{P}}_n(x, \tilde{w}); \tag{8.38}$$

moreover, Eq. (8.18) becomes:

$$\tilde{\Psi}_n(x) = \Psi_G(x)\Psi_n(x) \equiv \int d\tilde{w}\ \tilde{w}\ \tilde{\mathcal{P}}_n(x, \tilde{w}) \approx \langle\langle \tilde{w}_n \delta_{x,x_n} \rangle\rangle. \tag{8.39}$$

Finally, we can show that the right eigenvector of the stochastic matrix $\tilde{p}_{x',x}$ (with unit eigenvalue) is given by:

$$\tilde{\Pi}_{\text{eq}}(x) = \frac{\tilde{b}_x \Psi_G^2(x)}{\sum_{x'} \tilde{b}_{x'} \Psi_G^2(x')}. \tag{8.40}$$

8.3.1 Ground-State Energy and Correlation Functions

We are now in the position to compute the ground-state energy in presence of the importance sampling. Again, for large enough n, we have that:

$$E_0 \approx \frac{\langle \Psi_G | \mathcal{H} \mathcal{G}^n | \Psi_0 \rangle}{\langle \Psi_G | \mathcal{G}^n | \Psi_0 \rangle} = \frac{\sum_x \langle \Psi_G | \mathcal{H} | x \rangle \langle x | \mathcal{G}^n | \Psi_0 \rangle}{\sum_x \langle \Psi_G | x \rangle \langle x | \mathcal{G}^n | \Psi_0 \rangle}. \tag{8.41}$$

By using the fact that:

$$\tilde{\Psi}_n(x) = \langle x | \tilde{\mathcal{G}}^n | \Psi_0 \rangle = \Psi_G(x)\Psi_n(x) = \langle x | \Psi_G \rangle \langle x | \mathcal{G}^n | \Psi_0 \rangle, \tag{8.42}$$

together with Eq. (8.39) and considering that the guiding function is real, we have:

$$E_0 \approx \frac{\langle\langle \tilde{e}_L(x_n) \tilde{w}_n \rangle\rangle}{\langle\langle \tilde{w}_n \rangle\rangle}, \tag{8.43}$$

where we have defined the local energy with importance sampling:

$$\tilde{e}_L(x) = \frac{\langle \Psi_G | \mathcal{H} | x \rangle}{\langle \Psi_G | x \rangle} = \sum_{x'} \mathcal{H}_{x',x} \frac{\Psi_G(x')}{\Psi_G(x)}. \tag{8.44}$$

Notice that if $|\Psi_G\rangle$ is an eigenvector of the Hamiltonian $\mathcal{H}$ with eigenvalue E, then $\tilde{e}_L(x) = E$ for all the configurations $|x\rangle$, which implies that there are no statistical fluctuations. Of course, this is a very unrealistic situation for a generic correlated Hamiltonian; however, the advantage of considering importance sampling is that, whenever the guiding function is close to an eigenstate (e.g., the exact ground state),

https://doi.org/10.1017/9781316417041.009 Published online by Cambridge University Press

the statistical fluctuations are strongly reduced with respect to the case without importance sampling.

In practice, as before, we can perform a single simulation of length $N_{\rm sim} \gg n$ and, after a thermalization time, imagine to start the projection technique of p steps from the configurations that are distributed according to $\tilde{\Pi}_{\rm eq}(x)$ of Eq. (8.40). In this case, the accumulated weight from iteration $n-p$ for p steps is:

$$\tilde{G}_n^p = \prod_{j=1}^{p} \tilde{b}_{x_{n-j}}, \tag{8.45}$$

and the ground-state energy is given by:

$$E_0 \approx \frac{\sum_n \tilde{G}_n^p \tilde{e}_L(x_n)}{\sum_i \tilde{G}_n^p}. \tag{8.46}$$

By using the so-called *forward-walking* technique, the GFMC method can also efficiently compute expectation values of local operators $\mathcal{O}$, which are diagonal in the basis set $\{|x\rangle\}$:

$$\mathcal{O}|x\rangle = \mathcal{O}(x)|x\rangle, \tag{8.47}$$

where $\mathcal{O}(x)$ is the eigenvalue corresponding to the configuration $|x\rangle$. Indeed, for large n and m, the true expectation value over the ground state can be written as:

$$\frac{\langle \Upsilon_0|\mathcal{O}|\Upsilon_0\rangle}{\langle \Upsilon_0|\Upsilon_0\rangle} \approx \frac{\langle \Psi_0|\tilde{\mathcal{G}}^m \mathcal{O} \tilde{\mathcal{G}}^{n-m}|\Psi_0\rangle}{\langle \Psi_0|\tilde{\mathcal{G}}^n|\Psi_0\rangle}. \tag{8.48}$$

For these operators, the Markov chain can be easily modified in order to account the application of the operator $\mathcal{O}$ at a selected iteration $n-m$. In fact, this is exactly equivalent, in the statistical sense, to modify the weight:

$$\tilde{w}_{n-m} \to \mathcal{O}(x_{n-m})\tilde{w}_{n-m}, \tag{8.49}$$

in such a way that:

$$\frac{\langle \Upsilon_0|\mathcal{O}|\Upsilon_0\rangle}{\langle \Upsilon_0|\Upsilon_0\rangle} \approx \frac{\langle\langle \mathcal{O}(x_{n-m})\tilde{w}_n\rangle\rangle}{\langle\langle \tilde{w}_n\rangle\rangle}. \tag{8.50}$$

As before, the quantity on the r.h.s. can be computed in a single run by using the accumulated weight of Eq. (8.45):

$$\frac{\langle \Upsilon_0|\mathcal{O}|\Upsilon_0\rangle}{\langle \Upsilon_0|\Upsilon_0\rangle} \approx \frac{\sum_n \tilde{G}_n^p \mathcal{O}(x_{n-m})}{\sum_n \tilde{G}_n^p}, \tag{8.51}$$

which gives the ground-state expectation value for $m = p/2$ and large enough p. The important point here is that both numerator and denominator of Eq. (8.48) are evaluated in a single simulation, i.e., by using the same Markov chain.

https://doi.org/10.1017/9781316417041.009 Published online by Cambridge University Press

An alternative method to compute averages of general operators is obtained by performing two independent simulations for the numerator and the denominator of Eq. (8.48). The advantage of this approach is the possibility to measure not only diagonal operators but also general ones with $\mathcal{O}_{x',x} = \langle x'|\mathcal{O}|x\rangle \neq 0$ for $|x'\rangle \neq |x\rangle$. In this case, first of all, we carry out a simulation for evaluating the denominator. Then, we take the equilibrated walker configurations $(x_n, \tilde{w}_n)$ and apply the operator $\mathcal{O}$ at a given iteration. Whenever the operator has off-diagonal elements, this is not simply equivalent to scale the weight as in Eq. (8.49). Nevertheless, we can use a stochastic approach to select only one configuration among the ones that are connected to x with non-zero matrix elements $\mathcal{O}_{x',x}$. Finally, starting from the selected configuration, we perform the propagation for further m steps.

Whenever $\mathcal{O}_{x',x} > 0$, this stochastic approach can be implemented in few steps:

1. As far as the denominator of Eq. (8.51) is concerned, the original Markov chain with n steps is used to compute the accumulated weight $\tilde{G}_n^p$.
2. As far as the numerator of Eq. (8.51) is concerned:
   - Consider the original Markov chain at the iteration $n-m$ with the corresponding configuration $x_{n-m} \equiv x$ and weight $\tilde{G}_{n-m}^p$.
   - Compute $\mathcal{O}(x_{n-m}) \equiv \mathcal{O}(x)$:

$$\mathcal{O}(x) = \sum_{x'} \mathcal{O}_{x',x} \frac{\Psi_G(x')}{\Psi_G(x)}. \tag{8.52}$$

   - Select a random new configuration x' with a probability proportional to $\mathcal{O}_{x',x}\Psi_G(x')/\Psi_G(x)$.
   - Finally, start an independent Markov chain from x' for m steps and keep accumulating the weight from $\tilde{G}_{n-m}^p$ to $\left(\tilde{G}_n^p\right)'$. Notice that in this case the weights on numerator and denominator are different, given the forward propagation with m steps.

We would like to stress the important point that the general operators $\mathcal{O}$, with arbitrary signs in the matrix elements, can be always cast as a difference $\mathcal{O} = \mathcal{O}^+ - \mathcal{O}^-$ of two operators with positive definite matrix elements and, therefore, the method just described can be applied to $\mathcal{O}^+$ and $\mathcal{O}^-$ separately.

Finally, we mention that, by using Eq. (8.48) for $m = 0$, the so-called *mixed average* $\mathcal{O}_{\mathrm{MA}}$ is obtained, which is a biased estimator of the exact quantum average:

$$\mathcal{O}_{\mathrm{MA}} = \frac{\langle \Psi_G|\mathcal{O}|\Upsilon_0\rangle}{\langle \Psi_G|\Upsilon_0\rangle}. \tag{8.53}$$

The calculation of this mixed average is possible for any type of operator, not only the local ones. For operators that are defined on the ground state, i.e., $\mathcal{O}|\Upsilon_0\rangle = \mathcal{O}_0|\Upsilon_0\rangle$ ($\mathcal{O}_0$ being the eigenvalue), the mixed average estimator is exact (as we

https://doi.org/10.1017/9781316417041.009 Published online by Cambridge University Press

have seen for the ground-state energy). For all other operators, an approximated scheme to evaluate the ground-state expectation value can be used, which is valid whenever the state $|\Psi_G\rangle$ is close to $|\Upsilon_0\rangle$. Indeed, if $|\Psi_G\rangle = |\Upsilon_0\rangle + \epsilon|\Upsilon'\rangle$ (where $|\Upsilon'\rangle$ is normalized and orthogonal to $|\Upsilon_0\rangle$), the following relation holds up to $O(\epsilon^2)$:

$$\frac{\langle\Upsilon_0|\mathcal{O}|\Upsilon_0\rangle}{\langle\Upsilon_0|\Upsilon_0\rangle} \approx 2\frac{\langle\Psi_G|\mathcal{O}|\Upsilon_0\rangle}{\langle\Psi_G|\Upsilon_0\rangle} - \frac{\langle\Psi_G|\mathcal{O}|\Psi_G\rangle}{\langle\Psi_G|\Psi_G\rangle}. \tag{8.54}$$

8.4 The Continuous-Time Limit

Up to now, we have seen a statistical implementation of the power method, which is based upon the Markov process of Eqs. (8.10) and (8.11). As we discussed in section 8.1, the Green's function must be non-negative, given the probabilistic nature of the GFMC approach. Therefore, the constant λ in Eq. (8.2) has to be taken large enough to have all the diagonal elements $\mathcal{G}_{x,x} \geq 0$. Here, to simplify the notation, we do not put a tilde over the various quantities, assuming that we are using importance sampling. The requirement of having non-negative diagonal elements, often needs a very large shift λ, which also increases with the size of the cluster; in addition, in some cases, this shift cannot be given *a priori* (as for example when the fixed-node approximation is considered, see Chapter 10). If, for the chosen λ, a negative diagonal element $\mathcal{G}_{x,x}$ is found, we need to start a new simulation from scratch with a larger value for λ, with a considerable waste of computational time. In order to avoid this problem, we could work with an exceedingly large value of λ; however, this choice slows down the efficiency of the algorithm, since the probability to remain in the same configuration

$$p_d = \frac{p_{x,x}}{b_x} = \frac{\lambda - \mathcal{H}_{x,x}}{\lambda - e_L(x)} \tag{8.55}$$

becomes very close to one, thus implying a very large correlation time.

In the following, we show that the problem of working with a large λ can be solved without loss of efficiency. Indeed, given p_d, the probability $\mathcal{Q}(k)$ that the walker remains in the same configuration $|x\rangle$ for k steps (i.e., it makes k diagonal moves) and then changes into $|x'\rangle \neq |x\rangle$ is simply given by:

$$\mathcal{Q}(k) = p_d^k(1 - p_d). \tag{8.56}$$

As we described in section 3.3, we can sample this discrete distribution by extracting a random number ξ that is uniformly distributed in $[0, 1)$; then the number of diagonal moves k is determined by the inequalities:

$$(1 - p_d)\sum_{i=0}^{k-1} p_d^i \leq \xi < (1 - p_d)\sum_{i=0}^{k} p_d^i, \tag{8.57}$$

https://doi.org/10.1017/9781316417041.009 Published online by Cambridge University Press

which leads to

$$k \leq \frac{\ln(1-\xi)}{\ln p_d} < k+1. \tag{8.58}$$

These relations are equivalent to take:

$$k = \left[\frac{\ln(1-\xi)}{\ln p_d}\right], \tag{8.59}$$

where $[\dots]$ indicates the integer part. Here, we assume that $0 \leq \xi < 1$, such that k is always finite.

The idea is to consider a set of M elementary iterations as a single Markov step, for which $\mathcal{Q}(k)$ is given by Eq. (8.56) for $k = 0, \dots, M-1$ and $\mathcal{Q}(M) = p_d^M$ (denoting the possibility that there are no off-diagonal moves during M trials). Then, along the M iterations, we can sample $\mathcal{Q}(k)$ with:

$$k = \min\left(M_{\text{left}}, \left[\frac{\ln(1-\xi)}{\ln p_d}\right]\right), \tag{8.60}$$

where M_{left} is the number of iterations that are left to complete the Markov step. At the beginning, $M_{\text{left}} = M$, then we can iteratively apply Eq. (8.60) and bookkeep M_{left}. The weight w of the walker is updated according to k diagonal moves:

$$w \to w b_x^k; \tag{8.61}$$

if $k < M_{\text{left}}$ a new configuration is extracted randomly according to the transition probability $t_{x',x}$ defined by:

$$t_{x',x} = \begin{cases} \frac{p_{x',x}}{1-p_{x,x}} & \text{for } x \neq x', \\ 0 & \text{for } x = x'. \end{cases} \tag{8.62}$$

Finally, if $k < M_{\text{left}}$, M_{left} is changed into $M_{\text{left}} - (k+1)$ (k diagonal moves and the off-diagonal one), so that we can continue to use Eq. (8.60) until $M_{\text{left}} = 0$, where the Markov step is concluded, see Fig. 8.2. We stress the fact that the Markov step is composed of M elementary steps; therefore, in this case, the appropriate factors to compute G_n^p in Eq. (8.45) are replaced by the accumulated weights along the M steps.

It is interesting to observe that this method can be readily generalized for $\lambda \to \infty$ by increasing M with λ, namely for $M = \lambda\beta$, where β represents the imaginary-time evolution of the exact propagator $\exp(-\beta\mathcal{H})$ applied statistically:

$$(\Lambda - \mathcal{H})^M = \lambda^M \left(\mathbb{I} - \frac{\mathcal{H}}{\lambda}\right)^{\lambda\beta} \approx \exp(-\beta\mathcal{H}). \tag{8.63}$$

https://doi.org/10.1017/9781316417041.009 Published online by Cambridge University Press

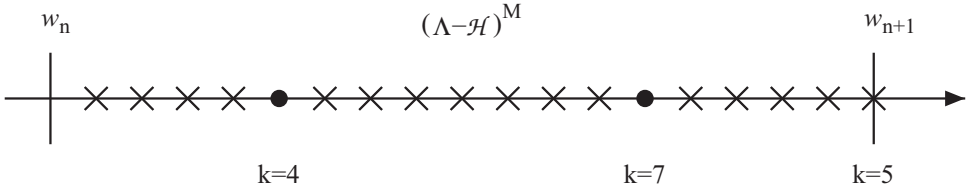

Figure 8.2 Picture of the Markov process based upon sampling the number of diagonal moves of Eq. (8.60). Here, the number of elementary steps is $M = 18$: the steps in which the configuration does not change are marked by crosses, while the (two) steps in which the configuration is changed are marked by full dots. In order to compute observables, only the weights at the end of the Markov step (i.e., after 18 elementary processes) are written in the file, to construct the accumulated weight G_n^p of Eq. (8.45). When $\lambda \to \infty$, a similar picture holds, where the weights are stored only at the end of the Markov step described by the evolution in imaginary time τ.

To this purpose, it is enough to bookkeep the corresponding time β_{left} remaining to complete the imaginary-time propagation of length β. For $\lambda \to \infty$, p_d becomes:

$$p_d \approx 1 + \frac{e_L(x) - \mathcal{H}_{x,x}}{\lambda}. \tag{8.64}$$

Then, Eq. (8.60) becomes:

$$\frac{k}{\lambda} \equiv \tau = \min\left(\beta_{\text{left}}, \frac{\ln(1-\xi)}{e_L(x) - \mathcal{H}_{x,x}}\right), \tag{8.65}$$

where we have used that $\ln p_d \approx [e_L(x) - \mathcal{H}_{x,x}]/\lambda$. Along the imaginary time β, we iteratively apply this formula and bookkeep the remaining time β_{left} that is left to complete the Markov step of length β. At the beginning, $\beta_{\text{left}} = \beta$, then τ is extracted using Eq. (8.65); the weight is updated according to:

$$w \to w \exp\left[-\tau e_L(x)\right]; \tag{8.66}$$

if $\tau < \beta_{\text{left}}$ the new configuration is taken according to Eq. (8.62) and the remaining time is changed into $\beta_{\text{left}} - \tau$. The Markov step is completed when $\beta_{\text{left}} = 0$.

The important aspect of this approach is that the evolution with the imaginary-time propagator $\exp(-\beta\mathcal{H})$ is done *without* any Trotter error (Trotter, 1959; Suzuki, 1976a,b), which is possible on the lattice.

8.5 Many Walkers Formulation

In practice, the single-walker formulation described in the previous sections is unstable when the number of projection p in Eq. (8.45) is too large. Indeed, the

https://doi.org/10.1017/9781316417041.009 Published online by Cambridge University Press

accumulated weight G_n^p is a product of p random variables, thus having huge fluctuations as seen in section 2.7:

$$\frac{\sqrt{\langle (G_n^p)^2\rangle - \langle G_n^p\rangle^2}}{\langle G_n^p\rangle} \approx \exp\left(\frac{p}{2}\sigma^2\right), \tag{8.67}$$

where σ^2 is the variance of $\ln b_x$. (Also here we prefer to use a simplified notation without a tilde over the different quantities, but still assuming that we are using importance sampling.)

In order to overcome the problem of large fluctuations, we discuss a way to propagate a set of N_w walkers simultaneously, defined by their weights $w_{\alpha,n}$ and configurations $x_{\alpha,n}$, for each Markov iteration n and $\alpha = 1,\ldots,N_\mathrm{w}$. Given N_w walkers, we indicate the corresponding configurations and weights with a dyad of vectors $(\mathbf{x}_n, \mathbf{w}_n)$. By evolving them independently, no improvement is obtained for the aforementioned large fluctuations, (e.g., we would just reduce the variance of the accumulated weight by a factor $1/N_\mathrm{w}$):

$$\langle (G_n^p)^2\rangle - \langle G_n^p\rangle^2 \to \frac{1}{N_\mathrm{w}}\left[\langle (G_n^p)^2\rangle - \langle G_n^p\rangle^2\right]. \tag{8.68}$$

Instead, by performing an iterative *reconfiguration* of the walkers, before that the variance of the weights becomes too large, it is possible to keep the weights of all the walkers approximately equal during the simulation and then obtain a considerable reduction of statistical errors of each single factor appearing in the evaluation of G_n^p, implying that:

$$\sigma^2 \to \frac{\sigma^2}{N_\mathrm{w}}, \tag{8.69}$$

which leads an exponential reduction of the fluctuations in Eq. (8.67). Therefore, by taking $N_\mathrm{w} \approx p$, this procedure allows us a stable propagation for a projection of p steps.

Let us now discuss the reconfiguration scheme, which goes under the name of *branching*. The Master equation (8.38) can be easily generalized to the case of many *independent* walkers. Indeed, if the evolution of the probability distribution $\mathcal{P}_n(\mathbf{x},\mathbf{w})$ is done without any restriction, each walker is uncorrelated from any other one, and we have:

$$\mathcal{P}_n(\mathbf{x},\mathbf{w}) = \prod_\alpha \mathcal{P}_n(x_\alpha, w_\alpha). \tag{8.70}$$

Similarly to the previous case of Eq. (8.39), we can define the state evolved at iteration n by:

$$\Psi_G(x)\Psi_n(x) \equiv \int \mathbf{dw} \sum_{\mathbf{x}} \left(\frac{\sum_\alpha w_\alpha \delta_{x,x_\alpha}}{N_\mathrm{w}}\right) \mathcal{P}_n(\mathbf{x},\mathbf{w}) \approx \left\langle\!\left\langle \frac{1}{N_\mathrm{w}} \sum_{\alpha=1}^{N_\mathrm{w}} w_{\alpha,n}\delta_{x,x_{\alpha,n}} \right\rangle\!\right\rangle. \tag{8.71}$$

https://doi.org/10.1017/9781316417041.009 Published online by Cambridge University Press

Table 8.1. *Reconfiguration scheme. Here $\overline{w}_n$ is defined in Eq. (8.72) and the new walkers are chosen according to the probability $p_{\alpha,n}$ of Eq. (8.73), see Fig. 8.3.*

Old Walkers		New Walkers	
$(x_{1,n}, w_{1,n})$	$\longrightarrow$	$x'_{1,n} = x_{j(1),n}$	$w'_{1,n} = \overline{w}_n$
$(x_{2,n}, w_{2,n})$	$\longrightarrow$	$x'_{2,n} = x_{j(2),n}$	$w'_{2,n} = \overline{w}_n$
$(x_{3,n}, w_{3,n})$	$\longrightarrow$	$x'_{3,n} = x_{j(3),n}$	$w'_{2,n} = \overline{w}_n$
$\vdots$	$\vdots$	$\vdots$	
$(x_{N_w,n}, w_{N_w,n})$	$\longrightarrow$	$x'_{N_w,n} = x_{j(N_w),n}$	$w'_{N_w,n} = \overline{w}_n$

The crucial point is that the physical information is contained in the first moment (with respect to the weight) of the distribution $\mathcal{P}_n(\mathbf{x}, \mathbf{w})$, which gives the wave function (e.g., $\Psi_G(x)\Psi_n(x)$ with importance sampling). Therefore, we can define a reconfiguration process that changes the probability distribution *without* changing the statistical average of Eq. (8.71). This can be obtained by a Markovian step that changes the configurations $(\mathbf{x}_n, \mathbf{w}_n)$ into $(\mathbf{x}'_n, \mathbf{w}'_n)$. In particular, the following reconfiguration does a proper job (Calandra Buonaura and Sorella, 1998):

1. Take the new weights equal to the average of the old ones:

$$w'_{\alpha,n} = \overline{w}_n \equiv \frac{1}{N_w} \sum_{\beta} w_{\beta,n}. \tag{8.72}$$

2. Choose the new configurations among the old ones with a probability $p_{\alpha,n}$ that is proportional to the weight of the old configurations:

$$p_{\alpha,n} \equiv \frac{w_{\alpha,n}}{\sum_{\beta} w_{\beta,n}}. \tag{8.73}$$

The branching procedure is schematically shown in Table 8.1 and Fig. 8.3. In principle, the selection of the new configurations can be done by extracting N_w independent random numbers z_α, uniformly distributed in $[0, 1)$, and apply the strategy of section 3.5. However, a better reconfiguration scheme is obtained by extracting N_w *correlated* random numbers $z_\alpha = (\alpha + \xi - 1)/N_w$, where ξ is a single random number uniformly distributed in $[0, 1)$ (and $\alpha = 1, \ldots, N_w$). This set of numbers is then used to select the new configurations, giving an unbiased result despite their correlation. Notice that, when all weights $p_{\alpha,n}$ are equal, this improved scheme does not modify the population of walkers, while, in the general case, the number of surviving walkers is optimal. After this reconfiguration, the new set of N_w walkers have the same weights $\overline{w}_n$, and most of the irrelevant walkers with small weights have dropped out, since they will not be selected according to $p_{\alpha,n}$.

https://doi.org/10.1017/9781316417041.009 Published online by Cambridge University Press

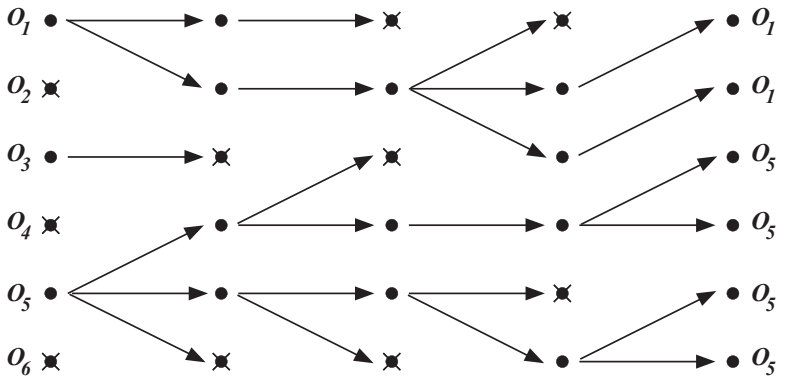

Figure 8.3 Schematic picture of the branching process for $N_{\rm w} = 6$ walkers. Here, every configuration $x_{\alpha,n}$ (denoted with a full dot) may be killed (denoted by crosses) or continue its evolution (denoted by arrows). In the latter case, the configuration may give rise to one or more copies of it, according to the probability of Eq. (8.73). Reconfigurations are done every $n_{\rm bra}$ Markov steps (here, the intermediate steps are not shown). The forward-walking technique is also shown: at the leftmost step, a (diagonal) operator is applied, whose values are $\mathcal{O}_\alpha$ for $\alpha = 1, \ldots, 6$ (the index of the Markov step is omitted); at the rightmost step the forward walking is finished (after $m = 4$ steps) and the values of the operators are changed according to the branching procedure.

This reconfiguration plays the same stabilization effect of the conventional branching scheme (Ceperley and Alder, 1980; Reynolds et al., 1982; Foulkes et al., 2001), but with the advantage that the number of walkers is kept fixed along the simulation. Indeed, one of the problems of the traditional branching scheme is to control the walker population and the fluctuations of the weights, which may induce some bias in the simulation.

Within the reconfiguration of Eqs. (8.72) and (8.73), some kind of correlation among the walkers is introduced, however, we can rigorously prove that the statistical average of Eq. (8.71) over the probability $\mathcal{P}_n(\mathbf{x}, \mathbf{w})$ is equal to the one performed with the new probability $\mathcal{P}'_n(\mathbf{x}', \mathbf{w}')$, namely we have that:

$$\Psi_G(x)\Psi'_n(x) = \Psi_G(x)\Psi_n(x). \tag{8.74}$$

In fact, the branching process is described by:

$$\mathcal{P}'_n(\mathbf{x}', \mathbf{w}') = \int \mathbf{dw} \sum_{\mathbf{x}} T_{\rm B}(\mathbf{x}', \mathbf{w}'|\mathbf{x}, \mathbf{w})\ \mathcal{P}_n(\mathbf{x}, \mathbf{w}), \tag{8.75}$$

where the transition probability $T_{\rm B}(\mathbf{x}', \mathbf{w}'|\mathbf{x}, \mathbf{w})$ is given by:

$$T_{\rm B}(\mathbf{x}', \mathbf{w}'|\mathbf{x}, \mathbf{w}) = \prod_\beta \left(\frac{\sum_\alpha w_\alpha \delta_{x'_\beta, x_\alpha}}{\sum_\alpha w_\alpha} \right) \delta \left(w'_\beta - \frac{\sum_\alpha w_\alpha}{N_{\rm w}} \right). \tag{8.76}$$

https://doi.org/10.1017/9781316417041.009 Published online by Cambridge University Press

Therefore, we have that the new wave function, after the branching process, is given by:

$$\Psi_G(x)\Psi'_n(x) = \int \mathbf{dw}' \sum_{\mathbf{x}'} \left(\frac{\sum_\alpha w'_\alpha \delta_{x,x'_\alpha}}{N_\mathrm{w}} \right) \mathcal{P}'_n(\mathbf{x}', \mathbf{w}')$$
$$= \int \mathbf{dw}' \sum_{\mathbf{x}'} \left(\frac{\sum_\alpha w'_\alpha \delta_{x,x'_\alpha}}{N_\mathrm{w}} \right) \int \mathbf{dw} \sum_{\mathbf{x}} T_\mathrm{B}(\mathbf{x}', \mathbf{w}'|\mathbf{x}, \mathbf{w}) \mathcal{P}_n(\mathbf{x}, \mathbf{w}). \tag{8.77}$$

Here, it is convenient to single out each term of the sum, i.e., $w'_\gamma \delta_{x,x'_\gamma}$ and sum (integrate) over all the other variables x'_α (w'_α) with $\alpha \neq \gamma$. The contribution of this term (indicated by $\left[\Psi_G(x)\Psi'_n(x)\right]_\gamma$) is given by:

$$\left[\Psi_G(x)\Psi'_n(x)\right]_\gamma = \int dw'_\gamma \sum_{x'_\gamma} \left(\frac{w'_\gamma \delta_{x,x'_\gamma}}{N_\mathrm{w}} \right) \int \mathbf{dw} \sum_{\mathbf{x}} \left(\frac{\sum_\alpha w_\alpha \delta_{x'_\gamma, x_\alpha}}{\sum_\alpha w_\alpha} \right)$$
$$\times \delta \left(w'_\gamma - \frac{\sum_\alpha w_\alpha}{N_\mathrm{w}} \right) \mathcal{P}_n(\mathbf{x}, \mathbf{w}). \tag{8.78}$$

Then, the factor $\sum_\alpha w_\alpha$, appearing in the previous equation, cancels out after integrating over w'_γ and summing over x'_γ, due to the delta-function that imposes $w'_\gamma = 1/N_\mathrm{w} \sum_\alpha w_\alpha$:

$$\left[\Psi_G(x)\Psi'_n(x)\right]_\gamma = \frac{1}{N_\mathrm{w}} \int \mathbf{dw} \sum_{\mathbf{x}} \left(\frac{\sum_\alpha w_\alpha \delta_{x,x_\alpha}}{N_\mathrm{w}} \right) \mathcal{P}_n(\mathbf{x}, \mathbf{w}), \tag{8.79}$$

namely, $\left[\Psi_G(x)\Psi'_n(x)\right]_\gamma = \Psi_G(x)\Psi_n(x)/N_\mathrm{w}$. By summing over γ we have proven the statement of Eq. (8.74).

In practice, the reconfiguration scheme of Eq. (8.75) is performed every n_bra power iterations. We would like to emphasize that taking a large n_bra implies a rather inefficient branching process, since the walkers will have considerably different weights and only few of them will survive the reconfiguration (here the inefficiency comes from the fact that we loose time in evolving walkers that will not contribute to the final averages). By contrast, a too small n_bra is also not optimal, since in this case all the walkers will be unaffected and we spend time in performing a reconfiguration that does not change the distribution $\mathcal{P}_n(\mathbf{x}, \mathbf{w})$.

It is useful, after each reconfiguration, to store the average weight $\overline{w}_n$ and put $w'_{\alpha,n} = 1$ for all walkers $\alpha = 1, \ldots, N_\mathrm{w}$ (instead of taking $w'_{\alpha,n} = \overline{w}$). The value of the accumulated weight G_n^p can be recovered by following the evolution of the weights in the p reconfiguration processes and reads:

$$G_n^p = \prod_{j=1}^{p} \overline{w}_{n-j}, \tag{8.80}$$

which corresponds to the application of $p \times n_\mathrm{bra}$ iterations of the power method.

https://doi.org/10.1017/9781316417041.009 Published online by Cambridge University Press

The continuous-time approach of section 8.4 can be also applied; in this case the branching scheme can be done after an evolution of length τ_{bra}, i.e., after a propagation given by $\exp(-\tau_{\text{bra}}\mathcal{H})$.

Once the branching process is done, it is no longer possible to obtain an explicit form of the equilibrium distribution of the marginal probability:

$$\Pi_n(x) = \int \mathbf{dw} \sum_{\mathbf{x}} \left(\frac{\sum_\alpha \delta_{x,x_\alpha}}{N_{\text{w}}} \right) \mathcal{P}_n(\mathbf{x}, \mathbf{w}). \tag{8.81}$$

Indeed, by repeating the same steps as before, we can show that the new marginal probability after the reconfiguration is given by:

$$\Pi'_n(x) = \int \mathbf{dw} \sum_{\mathbf{x}} \left(\frac{\sum_\alpha w_\alpha}{N_{\text{w}}} \right)^{-1} \left(\frac{\sum_\alpha w_\alpha \delta_{x,x_\alpha}}{N_{\text{w}}} \right) \mathcal{P}_n(\mathbf{x}, \mathbf{w}), \tag{8.82}$$

which depends upon the number of walkers N_{w} and recovers the wave function of Eq. (8.77) when increasing N_{w}, i.e., when the term in the first parenthesis approaches a constant value (by the central limit theorem, see section 2.7).

8.5.1 Ground-State Energy and Correlation Functions

As in the case of the single-walker technique, the ground-state energy can be computed by using a single simulation, with configurations that are equilibrated according to $\Pi_{\text{eq}}(x)$ (not explicitly known) and considering the accumulated weight of Eq. (8.80):

$$E_0 \approx \frac{\sum_n G_n^p e_L(x_n)}{\sum_n G_n^p}; \tag{8.83}$$

where n denotes the n-th reconfiguration process; the local energy can be evaluated either immediately after the reconfiguration process, when all the walkers have the same weight:

$$e_L(x_n) \equiv \frac{1}{N_{\text{w}}} \sum_\alpha e_L(x_{\alpha,n}), \tag{8.84}$$

or, for a better statistical error, just before the reconfiguration, taking properly into account the weight of each single walker:

$$e_L(x_n) \equiv \frac{\sum_\alpha w_{\alpha,n} e_L(x_{\alpha,n})}{\sum_\alpha w_{\alpha,n}}. \tag{8.85}$$

An example of the convergence for two values of N_{w} is reported in Fig. 8.4, for the continuous-time algorithm of section 8.4. Here, we show the case of the

https://doi.org/10.1017/9781316417041.009 Published online by Cambridge University Press

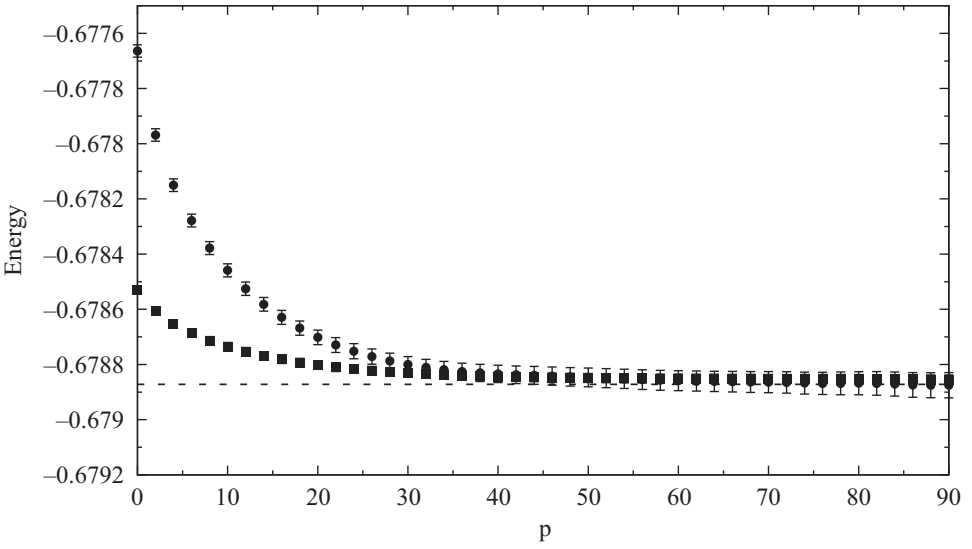

Figure 8.4 Energy per site for the Heisenberg model on the 6×6 cluster (square lattice) as a function of the projection index p, see Eq. (8.83). Two cases with $N_w = 10$ (circles) and $N_w = 50$ (squares) are shown. The continuous-time approach with $\tau = 0.1$ is used. The dashed line indicates the exact result obtained by Lanczos diagonalization.

ground-state energy of the Heisenberg model on the square lattice for the 6×6 cluster (where the exact solution is available by using Lanczos diagonalizations). Notice that the starting point with $p = 0$ is different for $N_w = 10$ and 50, the latter case being closer to the exact result.

Let us finish this part by discussing how to compute correlation functions. For local operators $\mathcal{O}$ that satisfy the condition of Eq. (8.47), we can perform the calculation by using a single Markov chain for both the numerator and the denominator of Eq. (8.48), as discussed in section 8.3 for the single-walker technique. However, in order to control the forward walking technique, the set of measured values $O_{\alpha,n}$, for $\alpha = 1, \ldots, N_w$ at the iteration n, must be modified after each reconfiguration process occurring in the forward direction. Indeed, within the Markov process with the reconfiguration of Eq. (8.76), the denominator can be easily computed by constructing the accumulated weight G_n^p; instead, for the numerator, the branching scheme must be corrected by the fact that the application of the operator at the iteration n would change the reconfigurations at the next iterations, since it modifies the weights. In practice, after each branching it is important to bookkeep only the values of the observables that correspond to the walkers that survive the reconfiguration. This can be done by storing (after each branching) the value $\mathcal{O}_{\alpha,n}$ of the operator for each walker and a table $\beta = j_n(\alpha)$ that gives the old walker β from which the

https://doi.org/10.1017/9781316417041.009 Published online by Cambridge University Press

new one α has been generated. Then, the contribution after m reconfiguration steps is given by the recursive application of the function $j_n(\alpha)$:

$$\mathcal{O}'_{\alpha,n+m} = \mathcal{O}_{j_{n+1}(j_{n+2}(\ldots(j_{n+m}(\alpha)))),n}. \tag{8.86}$$

It must be stressed that the forward walking technique may be unstable when few walkers are taken and a large propagation is considered: in fact, only few values of the observable $\mathcal{O}_{\alpha,n}$ survive the forward walk and eventually, for $m \to \infty$, only one value will remain, see Fig. 8.3. Therefore, this approach gives reliable results only for small enough propagation, before some instability appears.

Examples of the convergence as a function of the reconfiguration steps are shown in Fig. 8.5 for $N_w = 50$, by using the continuous-time algorithm of section 8.4. Here, we show the case of the spin-spin correlations in real space for the Heisenberg model on the square lattice (again for the 6×6 cluster, where the exact solution is possible):

$$S(r) = \frac{1}{L} \sum_i S^z_{i+r} S^z_i, \tag{8.87}$$

where L is the number of sites and S^z_i is the z-component of the spin operator at the site i.

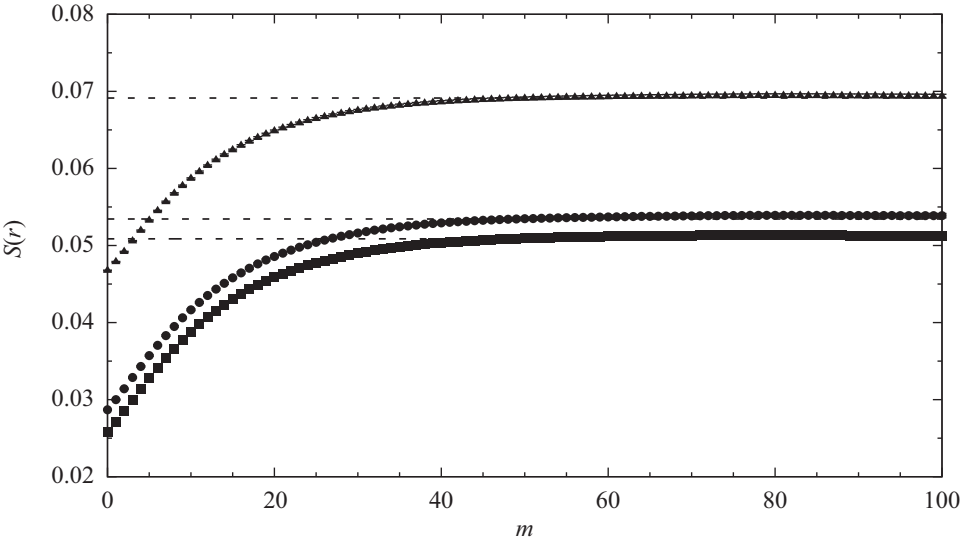

Figure 8.5 Spin-spin correlations $S(r)$ of Eq. (8.87) for the Heisenberg model on the 6×6 cluster (square lattice) as a function of the reconfiguration step m, see Eq. (8.86). The three independent distances along the diagonal of the cluster are considered (with increasing distance from top to bottom). Here, $N_w = 50$ and the continuous-time approach with $\tau = 0.1$ is used. The dashed lines indicate the exact results obtained by Lanczos diagonalization.

https://doi.org/10.1017/9781316417041.009 Published online by Cambridge University Press

8.6 Practical Implementation

We finally give a brief summary of the important steps in a practical implementation of the Green's function Monte Carlo algorithm.

1. **Initialization** at the beginning of the calculation.
   - Generate N_w random configurations $|x_\alpha\rangle$ and initialize the weights $w_\alpha = 1$. As in the variational Monte Carlo method, configurations can be stored into a set of vectors $\texttt{iconf}(L, N_w)$, whose elements give the local state on the site $i = 1, \ldots, L$ for each walker $\alpha = 1, \ldots, N_w$, and $\texttt{kel}(2L, N_w)$, 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).
   - Verify that the initial configurations are not singular, i.e., $\langle x_\alpha | \Psi_J \rangle \neq 0$, similarly to what has been discussed in the variational Monte Carlo method.
   - As for the variational Monte Carlo approach, compute the table of Eq. (5.33) and all the Green's functions that are necessary to compute $\tilde{\mathcal{G}}$ with importance sampling and perform the time evolution. For the determinant case, only the static 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. **GFMC projection (with the continuous-time approach)**.
   - For each walker α independently, select the new configuration $|x'_\alpha\rangle$ among the ones that are connected to $|x_\alpha\rangle$ through the stochastic matrix $t_{x',x}$ of Eq. (8.62).
   - For each walker α, update the weight factor w_α by using Eq. (8.66).
   - 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.
   - After a time propagation of length τ_{bra}, perform a reconfiguration of the walkers, according to Eqs. (8.72) and (8.73), see also Table 8.1.
3. **Computation of observables**.

   Observables can be computed every time a branching reconfiguration is performed (either just before or after it). For the calculation of the ground-state energy, it is recommended to write the average weight $\overline{w}$ of Eq. (8.72) and the local energy of Eq. (8.84) (before the branching) or Eq. (8.85) (after the branching) on the hard disk and then perform the calculations of the accumulated weight of Eq. (8.45), with the smallest value of p. For the calculation of local observables, it is recommended to write all the configurations $\texttt{iconf}(L, N_w)$ and the table $\beta = j_n(\alpha)$ (obtained at each n branching procedure), in order to perform the reshuffling of Eq. (8.86) that is needed within the forward-walker technique.

https://doi.org/10.1017/9781316417041.009 Published online by Cambridge University Press