The fermion Monte Carlo revisited

The fermion Monte Carlo revisited

Roland Assaraf[1]**, Michel Caffarel**[2] **and Anatole Khelif**[3]

[1] Laboratoire de Chimie Théorique CNRS-UMR 7616, Université Pierre et Marie Curie,
4 Place Jussieu, 75252 Paris, France
[2] Laboratoire de Chimie et Physique Quantiques CNRS-UMR 5626, IRSAMC,
Université Paul Sabatier, 118 route de Narbonne 31062 Toulouse Cedex, France
[3] Laboratoire de Logique Mathématique, Université Denis Diderot, 2 Place Jussieu, 75251 Paris,
France

Abstract
In this work we present a detailed study of the fermion Monte Carlo algorithm
(FMC), a recently proposed stochastic method for calculating fermionic
ground-state energies. A proof that the FMC method is an exact method is
given. In this work the stability of the method is related to the difference
between the lowest (bosonic-type) eigenvalue of the FMC diffusion operator
and the exact Fermi energy. It is shown that within a FMC framework the lowest
eigenvalue of the new diffusion operator is no longer the bosonic ground-state
eigenvalue as in standard exact diffusion Monte Carlo (DMC) schemes but a
modified value which is strictly greater. Accordingly, FMC can be viewed
as an exact DMC method built from a correlated diffusion process having
a *reduced* Bose–Fermi gap. As a consequence, the FMC method is more
stable than any transient method (or nodal release-type approaches). It is
shown that the most recent ingredient of the FMC approach (Kalos M H and
Pederiva F 2000 *Phys. Rev. Lett.* **85** 3547), namely the introduction of non-
symmetric guiding functions, does not necessarily improve the stability of the
algorithm. We argue that the stability observed with such guiding functions is in
general a finite-size population effect disappearing for a very large population
of walkers. The counterpart of this stability is a control population error
which is different in nature from the standard diffusion Monte Carlo algorithm
and which is at the origin of an uncontrolled approximation in FMC. We
illustrate the various ideas presented in this work with calculations performed
on a very simple model having only nine states but a full 'sign problem'.
Already for this toy model it is clearly seen that FMC calculations are inherently
uncontrolled.

PACS numbers: 02.70.Ss, 05.30.Fk

1. Introduction

In theory quantum Monte Carlo (QMC) techniques [1] are capable of giving an exact estimate of the energy with an evaluation of the error: the statistical error. Unfortunately, such an ideal situation is not realized in practice. Exact results with a controlled finite statistical error are only achieved for bosonic systems. For fermionic systems, we do not have at our disposal an algorithm which is both exact and stable (statistical fluctuations going to zero in the large simulation time regime). This well-known problem is usually referred to as the 'sign problem'. The usual solution to cope with this difficulty consists in defining a stable algorithm based on an uncontrolled approximation, the so-called fixed-node approximation [2–7]. In practice, the fixed-node error on the energy is small when one uses good trial wavefunctions and, thus, QMC methods can be considered today as reference methods to compute ground-state energies as shown by a large variety of applications [5, 8–13]. However, the accuracy of the results is never known from the calculation; it is known only *a posteriori*, for example by comparison with experimental data. Exact methods, which are basically transient methods [14–17] including the nodal release method [14], have been applied with success only to very specific models (small or homogeneous systems) for which the sign instability is not too severe (small Bose–Fermi energy gap).

Recently, Kalos and co-workers [18] have proposed a method presented as curing the sign problem, the so-called fermion Monte Carlo method (FMC). This work makes use of two previously introduced ingredients, a cancellation process between 'positive' and 'negative' walkers introduced by Arnow *et al* [19] and a modified process correlating explicitly the dynamics of the walkers of different signs [20]. The new feature introduced in [18] is the introduction of non-symmetric guiding functions. The method has been tested on various simple systems including free fermions and interacting systems such as the ^{3}He fluid [18, 21]. The results are found to be compatible with the assumption that the method is stable and not biased. However, this conclusion is not clear at all because of the presence of large error bars. The purpose of this work is to present a detailed analysis of the algorithm and a definitive answer to this assumption.

The content of this paper is as follows. In section 2, we give a brief presentation of the 'sign problem'. The sign instability in an exact DMC approach comes from the blowing up in imaginary time of the undesirable bosonic component associated with the lowest mathematical eigenstate of the Hamiltonian (a wavefunction which is positive and symmetric with respect to the exchange of particles). The fluctuations of the transient energy estimator grows like $e^{t(E_0^F - E_0^B)}$, where E_0^F is the fermionic ground-state energy (here and in what follows, the superscript F stands for 'fermionic') and E_0^B, the lowest mathematical eigenvalue of the Hamiltonian (the superscript 'B' standing for Bosonic). In section 3, we briefly recall the main elements of the standard DMC method, the basic stochastic algorithm simulating the imaginary-time evolution of the Hamiltonian. In section 4, we describe the FMC method as a generalization of the DMC method and we show that *FMC is an exact method*, that is, no systematic bias is introduced. This section is made of two parts. In the first part we introduce the notion of positive and negative walkers to represent a signed wavefunction in DMC. This will help us to view the FMC method as a generalization of the DMC method, the FMC introducing two important modifications with respect to DMC: a correlated dynamics for positive and negative walkers and a cancellation process for such pairs whenever they meet.

In section 5, we study the stability of the algorithm. We prove that the *FMC method is in general not stable*; the fluctuations of the transient estimator of the energy growing exponentially like $e^{t(E_0^F - \tilde{E}_0^B)}$, where $\tilde{E}_0^B$ is the lowest bosonic-like eigenvalue of the generalized diffusion process operator associated with FMC whose expression is given explicitly. It

is shown that *FMC is more stable than the standard nodal release DMC method* because $\tilde{E}_0^B > E_0^B$. In section 6, we illustrate our theoretical results using a toy model, a 'minimal' quantum system having a genuine sign problem (two coupled oscillators on a finite lattice). The different aspects of the FMC method are highlighted in this application. In the case of non-symmetric wavefunctions introduced recently in [18] it is shown that, in contrast with DMC where the population control error decays linearly as a function of the population size, the FMC decay displays a much more slower power law. As a consequence, one needs a very large population of walkers to remove the control population error and to observe the instability of FMC. Let us emphasize that observing such a subtle finite population effect for a genuine many-fermion system is actually impossible. Here, to illustrate this important point numerically, we have been led to consider a very simple system having only a few states. For this system, a large population of walkers—eventually much larger than the dimension of the quantum Hilbert space itself—can be considered. The results obtained in that regime confirm our theoretical findings, in particular the fact that the stability of the algorithm presented elsewhere [18] is only apparent. As an important conclusion, we emphasize that the FMC control population error is an uncontrolled approximation for realistic fermion systems.

2. Fermion instability in quantum Monte Carlo

We consider a Schrödinger operator for a system of N fermions:

$$H = -\tfrac{1}{2}\nabla^2 + V(\mathbf{R}), \tag{1}$$

where we note $\mathbf{R} = (\mathbf{r}_1, \ldots, \mathbf{r}_N)$ the $3N$ coordinates of the N particles in the three-dimensional space. In this expression, the first term is the kinetic energy (∇^2 is the Laplacian operator in the space of the $3N$ coordinates). The second term, V, is the potential. DMC techniques are based upon the evolution of the Hamiltonian in imaginary time. We express this evolution using the spectral decomposition

$$\mathrm{e}^{-t(H-E_T)} = \sum_i \mathrm{e}^{-t(E_i-E_T)}|\phi_i\rangle\langle\phi_i|, \tag{2}$$

where ϕ_i are the (normalized) eigenfunctions of H, E_i are the corresponding eigenvalues and E_T is a so-called reference energy. The fundamental property of operator (2) is to filter the lowest eigenstate ϕ_0. To understand how it works, we consider the time evolution of a wavefunction $f_0(\mathbf{R})$:

$$|f_t\rangle \equiv \mathrm{e}^{-t(H-E_T)}|f_0\rangle \tag{3}$$

and calculate the overlap with an eigenstate ϕ_i

$$\langle\phi_i|f_t\rangle = \mathrm{e}^{-t(E_i-E_T)}\langle\phi_i|f_0\rangle. \tag{4}$$

From this expression it is easily seen that the component on the lowest eigenstate ϕ_0 is growing exponentially faster than the higher components. In DMC methods the function f_t is generated using random walks. For t large enough the lowest eigenstate

$$f_t \sim \mathrm{e}^{-t(E_0-E_T)}\phi_0 \tag{5}$$

is produced. This asymptotic behaviour makes DMC an accurate method for computing the properties of ϕ_0, in particular the energy E_0. Unfortunately, for fermionic systems the physical ground state ϕ_0^F, which is antisymmetric with respect to the exchange of particles, is not ϕ_0, but some 'mathematically' excited state because ϕ_0 is positive and symmetric for a Schrödinger operator (bosonic ground state). For the sake of clarity, we shall denote from now on this bosonic ground state as ϕ_0^B and its energy, E_0^B. The necessity of extracting an exponentially

small component to evaluate E_0^F or any fermionic property is at the origin of the fermion instability in quantum Monte Carlo.

We now give a quantitative analysis of this instability. Since the asymptotic behaviour of f_t is not useful to compute the fermionic energy, we consider the transient behaviour of the evolution of f_t. Basically, exact methods [14] filter ϕ_0^F by projecting the transient behaviour of f_t on the antisymmetric space. Introducing the antisymmetrization operator $\mathcal{A}$, one obtains the physical ground state for t large enough:

$$\mathcal{A}|f_t\rangle \sim e^{-t(E_0^F - E_T)}\phi_0^F, \tag{6}$$

since the components of f_t over the higher antisymmetric eigenstates are decreasing exponentially with respect to the component on ϕ_0^F. In practice, the fermionic energy E_0^F is calculated using an antisymmetric function ψ_T and using the fact that, at large enough t, one has

$$E_0^F = \frac{\langle \psi_T | H | f_t \rangle}{\langle \psi_T | f_t \rangle}, \tag{7}$$

up to an exponentially small correction. In the long-time regime the stochastic estimation of the RHS of equation (7) is unstable. In essence, the signal, the antisymmetric component of f_t, decreases exponentially fast with respect to f_t, equation (5). The signal-over-noise ratio behaves like $e^{-t(E_0^F - E_0^B)}$ and, thus, an exponential growth of the relative fluctuations of the DMC estimator, equation (7), appears.

Now, let us give a more quantitative analysis by writing this estimator and evaluating the variance. In a standard diffusion Monte Carlo calculation, the time-dependent distribution f_t is generated by a random walk over a population of walkers $\{\mathbf{R}_i\}$. Formally, f_t is given in the calculation as an average at time t over Dirac functions centred on the walkers and weighted by some positive function ψ_G:

$$f_t(\mathbf{R}) = \frac{1}{\psi_G(\mathbf{R})}\left\langle \sum_i \delta(\mathbf{R} - \mathbf{R}_i) \right\rangle. \tag{8}$$

In this expression, $\langle \cdots \rangle$ denotes the average over populations of walkers $\{\mathbf{R}_i\}$ obtained at the given time t. The function ψ_G is usually called the importance or guiding function. Replacing f_t in (7) by its expression (8), the estimator of the energy reads for large enough t:

$$E_0^F = \frac{\left\langle \sum_i \frac{H\psi_T}{\psi_G}(\mathbf{R}_i) \right\rangle}{\left\langle \sum_i \frac{\psi_T}{\psi_G}(\mathbf{R}_i) \right\rangle}. \tag{9}$$

In practice, both numerator and denominator are computed as an average over the N_S walkers produced by the algorithm at time t. As a consequence, the energy is obtained as

$$E_0^F = \frac{\mathcal{N}}{\mathcal{D}} \equiv \frac{\frac{1}{N_S}\sum_{i=1}^{N_S} \frac{H\psi_T}{\psi_G}(\mathbf{R}_i)}{\frac{1}{N_S}\sum_{i=1}^{N_S} \frac{\psi_T}{\psi_G}(\mathbf{R}_i)}. \tag{10}$$

The ratio $\frac{\mathcal{N}}{\mathcal{D}}$ is an estimator of the energy for N_S large enough, when the numerator and denominator have small fluctuations around their average. Now, let us evaluate the fluctuations of this ratio in the limit of a large population N_S:

$$\sigma^2\left(\frac{\mathcal{N}}{\mathcal{D}}\right) = \frac{\left\langle \left(\mathcal{N} - E_0^F \mathcal{D}\right)^2 \right\rangle}{\langle \mathcal{D} \rangle^2}. \tag{11}$$

Here, we have used the hypothesis that the fluctuations of the denominator and the numerator are very small. Using the fact that $\mathcal{N}$ and $\mathcal{D}$ are statistical averages over independent random variables with the same distribution, one obtains

$$\sigma^2 \left(\frac{\mathcal{N}}{\mathcal{D}} \right) = \frac{1}{N_S} \frac{\left\langle \left[\frac{(H-E_0^{\mathrm{F}})\psi_T}{\psi_G}(\mathbf{R}_i) \right]^2 \right\rangle}{\left\langle \frac{\psi_T}{\psi_G}(\mathbf{R}_i) \right\rangle^2}. \tag{12}$$

We can replace these averages by integrals over the distribution $f_t \psi_G$, equation (8),

$$\sigma^2 \left(\frac{\mathcal{N}}{\mathcal{D}} \right) = \frac{1}{N_S} \frac{\left\langle \frac{[(H-E_0^{\mathrm{F}})\psi_T]^2}{\psi_G} \big| \phi_0^{\mathrm{B}} \right\rangle \left\langle \psi_G \big| \phi_0^{\mathrm{B}} \right\rangle}{\left\langle \psi_T \big| \phi_0^{\mathrm{F}} \right\rangle^2} \mathrm{e}^{2t(E_0^{\mathrm{F}}-E_0^{\mathrm{B}})} \propto \frac{1}{N_S} \mathrm{e}^{2t(E_0^{\mathrm{F}}-E_0^{\mathrm{B}})}, \tag{13}$$

which confirms quantitatively that the statistical error grows exponentially in time. Let us emphasize that this problem is particularly severe because the Bose–Fermi energy gap, $\Delta_{\mathrm{B-F}} \equiv E_0^{\mathrm{F}} - E_0^{\mathrm{B}}$, usually grows faster than linearly as a function of the number of particles. In practice, one has to find a trade-off between the systematic error coming from short projection times t and the large fluctuations arising at large projection times.

3. The diffusion Monte Carlo method

The FMC method is a generalization of the well-known DMC method. Presenting this algorithm is a useful preparation for the following section about fermion Monte Carlo. The DMC method generates the function f_t following the imaginary time dynamics of H:

$$f_t \equiv \mathrm{e}^{-t(H-E_T)} f_0, \tag{14}$$

where f_0 is a positive function. The imaginary time dynamics is produced by iterating many times the short-time Green function $\mathrm{e}^{-\tau(H-E_T)}$, where τ is a small time step. The distribution $f_{t'}$ at the time $t' \equiv t + \tau$ is then obtained from f_t as follows:

$$f_{t'} = \mathrm{e}^{-\tau(H-E_T)} f_t, \tag{15}$$

where the density f_t is sampled by the population of walkers $\{\mathbf{R}_i\}$. Using the Dirac ket notation, f_t given by equation (8) is rewritten as

$$f_t = \frac{1}{\psi_G} \left\langle \sum_i |\mathbf{R}_i\rangle \right\rangle, \tag{16}$$

where ψ_G is some positive function, the so-called guiding function. Let us show how the density $f_{t'}$ is generated from the distribution (16). Replacing in (15) the function f_t by the RHS of (16) one has

$$f_{t'} = \mathrm{e}^{-\tau(H-E_T)} \frac{1}{\psi_G} \left\langle \sum_i |\mathbf{R}_i\rangle \right\rangle \tag{17}$$

$$= \left\langle \sum_i \mathrm{e}^{-\tau(H-E_T)} \frac{1}{\psi_G} |\mathbf{R}_i\rangle \right\rangle \tag{18}$$

$$= \frac{1}{\psi_G} \left\langle \sum_i \mathrm{e}^{\tau L} |\mathbf{R}_i\rangle \right\rangle, \tag{19}$$

where we have introduced the operator

$$L \equiv -\psi_G(H - E_T)\frac{1}{\psi_G}. \tag{20}$$

For a Schrödinger Hamiltonian (1) the operator L takes the form

$$L = -\psi_G(H - E_L)\frac{1}{\psi_G} - (E_L - E_T) \tag{21}$$

$$= \frac{1}{2}\nabla^2 - \nabla[\mathbf{b}.] \tag{22}$$

$$-(E_L - E_T), \tag{23}$$

where we have introduced the so-called drift vector

$$\mathbf{b} \equiv \frac{\nabla\psi_G}{\psi_G} \tag{24}$$

and the local energy of the guiding function ψ_G:

$$E_L \equiv \frac{H\psi_G}{\psi_G}. \tag{25}$$

The operator L is the sum of the so-called Fokker Planck operator (22) and a local operator (23). Using this decomposition, the vector $e^{\tau L}|\mathbf{R}_i\rangle$ appearing in the average (19) can be rewritten for small enough time step τ as follows:

$$e^{\tau L}|\mathbf{R}_i\rangle = e^{\tau\left[\frac{1}{2}\nabla^2 - \nabla[\mathbf{b}.]\right]} \tag{26}$$

$$\times e^{-\tau(E_L - E_T)}|\mathbf{R}_i\rangle. \tag{27}$$

The action of $e^{\tau L}$ on $|\mathbf{R}_i\rangle$ is sampled as follows. The short-time dynamics of the Fokker Planck operator (26) is performed by the way of a Langevin process:

$$R_i'^{\mu} = R_i^{\mu} + b_i^{\mu}\tau + \sqrt{\tau}\eta_i^{\mu}, \tag{28}$$

where μ runs over the $3N$ coordinates (three space coordinate for each fermion), and η_i^{μ} are independent Gaussian random variables centred and normalized:

$$\langle\eta_i^{\mu}\eta_j^{\nu}\rangle = \delta^{\mu\nu}\delta_{ij}. \tag{29}$$

The averaged Langevin process (28) is equivalent to apply the short-time dynamics of the Fokker Planck operator (26):

$$\langle|\mathbf{R}_i'\rangle\rangle = e^{\tau\left[\frac{1}{2}\nabla^2 - \nabla[\mathbf{b}.]\right]}|\mathbf{R}_i\rangle. \tag{30}$$

The factor

$$w_i \equiv e^{-\tau(E_L(\mathbf{R}_i) - E_T)} \tag{31}$$

being a normalization term, called the branching term. The new walker $\mathbf{R}_i'$ is duplicated (branched) a number of times equal to w_i in average. This process is a birth–death process since some walkers can be duplicated and some can be removed. The population of walkers fluctuating, one has to resort to control population techniques [22–24]. With these two processes, diffusion and branching a new population of walkers $\{\mathbf{R}_i'\}$ is produced, representing in average the desired result:

$$\left\langle\sum_i|\mathbf{R}_i'\rangle\right\rangle = \left\langle\sum_i e^{\tau L}|\mathbf{R}_i\rangle\right\rangle. \tag{32}$$

From (19) and (32) one can see that the distribution $f_{t'}$ is sampled by the new population of walkers $\{\mathbf{R}'_i\}$ according to (16)

$$f_{t'} = \frac{1}{\psi_G}\langle|\mathbf{R}'_i)\rangle. \tag{33}$$

In summary, by iterating these two simple operations, namely the Langevin and branching processes, the DMC method allows us to simulate the imaginary time dynamics of the Hamiltonian, thus producing a sample of f_t, equation (16). Various properties of the system can be computed from this sample, e.g. ground-state bosonic energies [2, 3, 25], excited-state energies [16, 17] and various observables [26–32].

For the vast majority of the DMC simulations on fermionic systems, only an approximation of the exact fermionic ground-state energy is computed, namely the so-called fixed-node energy [3, 5, 8, 33–35]. In a fixed-node DMC calculation the guiding function is chosen as $\psi_G = |\psi_T|$ where ψ_T is some fermionic antisymmetric trial wavefunction. With this choice, the guiding function vanishes at the nodes (zeroes) of the trial wavefunction and the drift vector diverges. As a consequence, the walkers cannot cross the nodes of ψ_T and are confined within the nodal regions of the configuration space. It can be shown that the resulting DMC stationary state is the best variational solution having the same nodes as ψ_T. In other words, the 'fixed-node' energy obtained from the RHS of (9) or (10) is an upper bound of the exact Fermi energy, $E_0^{FN} > E_0^F$ [33, 34]. Note that, in practice, E_0^{FN} is in general a good approximation of the true energy [8, 35].

In the present work, we are considering 'exact' DMC approaches for which the exact fermionic energy calculated from expression (9) is searched for. As discussed in the previous section, such exact DMC calculations are fundamentally unstable. A famous example of an exact DMC approach is the nodal release method of Ceperley and Alder [14, 34]. Basically, nodal release methods are standard DMC methods where the fixed-node distribution (sampled with a standard fixed-node DMC) is chosen as initial distribution f_0. In exact methods the guiding function ψ_G is strictly positive everywhere, so that the walkers can cross the nodes of the trial wavefunction. Exact Fermi methods are efficient in practice only when the convergence of the estimator is fast enough, that is, when it occurs before the blowing up of fluctuations, equation (13). In practice, two conditions are to be satisfied. First, the fixed-node wavefunction ϕ_0^{FN} must be already close enough to the exact solution ϕ_0^F. For this reason the choice of the trial function ψ_T (quality of the nodes of ψ_T) is crucial. Second, the Bose–Fermi gap, $\Delta_{B-F} = E_0^F - E_0^B$, which drives the asymptotic behaviour of the fluctuations, equation (13), must be small. The quantity $E_0^F - E_0^B$ depends only on the Hamiltonian at hand; there is no freedom in the nodal release method to modify the asymptotic behaviour. We will define in the following section the fermion Monte Carlo method (FMC) as a generalization of the DMC method and show in sections 5 and 6 that, in contrast with the nodal release method, the FMC method can improve substantially the asymptotic behaviour, equation (13).

4. The FMC method

4.1. Preliminary: Introducing positive and negative walkers in DMC

In fermion Monte Carlo a dynamics on a signed function f_t is performed. In what follows, we show that DMC can be easily generalized to the case of a signed distribution f_t and, thus, FMC can be viewed as a simple generalization of DMC. If f_t carries a sign, it can be written as the difference of two positive functions,

$$f_t = f_t^+ - f_t^-, \tag{34}$$

both satisfying the following equations of evolution:

$$f_t^+ = e^{-t(H-E_T)} f_0^+ \tag{35}$$

$$f_t^- = e^{-t(H-E_T)} f_0^- . \tag{36}$$

To sample these expressions, two independent DMC calculations can be carried out. The positive part f_t^+ is then sampled by a population of walkers $\{\mathbf{R}_i^+\}$ (called 'positive' walkers). The distribution f_t^+ is related to $\{\mathbf{R}_i^+\}$ as in equation (16)

$$f_t^+ = \frac{1}{\psi_G^+} \left\langle \sum_i |\mathbf{R}_i^+\rangle \right\rangle , \tag{37}$$

and the negative part is similarly sampled by a population of 'negative' walkers $\{\mathbf{R}_i^-\}$:

$$f_t^- = \frac{1}{\psi_G^-} \left\langle \sum_i |\mathbf{R}_i^-\rangle \right\rangle . \tag{38}$$

Note that we consider here the general case where the guiding functions associated with the positive and negative walkers, ψ_G^+ and ψ_G^-, are different. Finally, the dynamics of positive and negative walkers is described by the DMC-like diffusion operators:

$$L^\pm \equiv -\psi_G^\pm (H - E_T) \frac{1}{\psi_G^\pm} \tag{39}$$

$$= \frac{1}{2} \nabla^2 - \nabla[\mathbf{b}^\pm .] \tag{40}$$

$$- (E_L^\pm - E_T), \tag{41}$$

where the drift vectors are given by

$$\mathbf{b}^\pm \equiv \frac{\nabla \psi_G^\pm}{\psi_G^\pm} , \tag{42}$$

and the local energies of the guiding functions $\psi_G^\pm$ by

$$E_L^\pm \equiv \frac{H \psi_G^\pm}{\psi_G^\pm} . \tag{43}$$

In actual calculations, two Langevin dynamics on the positive and negative walkers are performed

$$R_i'^{\mu\pm} = R_i^{\mu\pm} + b_i^{\mu\pm} \tau + \sqrt{\tau} \eta_i^{\mu\pm} \tag{44}$$

and positive and negative walkers are branched according to their respective weight:

$$W^\pm (\mathbf{R}_i^\pm) \equiv e^{-\tau (E_L{}^\pm (\mathbf{R}_i^\pm) - E_T)} . \tag{45}$$

4.2. The detailed rules of FMC

In a few words, the FMC method is similar to a DMC method on a signed function, except that the positive and negative walkers are correlated and can annihilate whenever they meet. The Langevin processes are correlated in such a way that positive walkers and negative walkers meet as much as possible. We will see in the following section that the cancellation procedure is at the origin of an improved stability of the algorithm. In this section, we give a complete

description of the algorithm and show that this algorithm does not introduce any systematic bias.

In FMC, the guiding functions ψ_G^+ and ψ_G^- are not arbitrary; they are related under any permutation P of two particles as follows:

$$\psi_G^+(\mathbf{R}) = \psi_G^-(P\mathbf{R}). \tag{46}$$

Various choices are possible for the guiding functions. Here, we consider the form proposed by Kalos and Pederiva [18], namely

$$\psi_G^\pm = \sqrt{\psi_S^2 + c^2\psi_T^2} \pm c\psi_T, \tag{47}$$

where ψ_S is a symmetric (Bose-like) function, ψ_T an antisymmetric trial wavefunction and c some positive mixing parameter allowing to introduce some antisymmetric component into $\psi_G^\pm$.

Having in mind this choice for the guiding functions we will show how the two DMC processes over the two populations of walkers $\{\mathbf{R}_i^+\}$ and $\{\mathbf{R}_i^-\}$ can be replaced by a diffusion process over a population of *pairs* of walkers $\{(\mathbf{R}_i^+, \mathbf{R}_i^-)\}$. We first show how the DMC process can be modified to maintain as many positive and negative walkers during the simulation. In the DMC dynamics, the branching terms associated with the positive and negative walkers are in general different. As a consequence, the number of positive walkers N_S^+, can be different from the number of negative walkers N_S^-. At time t the DMC density f_t reads

$$f_t = \left\langle \sum_{i=1}^{N_S^+} \frac{1}{\psi_G^+} |\mathbf{R}_i^+\rangle - \sum_{i=1}^{N_S^-} \frac{1}{\psi_G^-} |\mathbf{R}_i^-\rangle \right\rangle. \tag{48}$$

This formula is obtained by introducing in equation (34) expressions (37) and (38) for f_t^+ and f_t^-, respectively. If N_S^+ and N_S^- are different, we will replace f_t (48) by a new function g_t sampled with an equal number of positive and negative walkers. Such an operation does not introduce any bias if the antisymmetric components of the future evolution of f_t and g_t are identical. Indeed, only the antisymmetric component of f_t contributes to the estimator of the energy, equation (7). At time $t' > t$ the two densities are given by

$$f_{t'} = e^{-(t'-t)(H-E_T)} f_t \tag{49}$$

$$g_{t'} = e^{-(t'-t)(H-E_T)} g_t. \tag{50}$$

Let us write that the antisymmetric components of $f_{t'}$ and $g_{t'}$ must be equal:

$$\mathcal{A} e^{-(t'-t)(H-E_T)} f_t = \mathcal{A} e^{-(t'-t)(H-E_T)} g_t. \tag{51}$$

Using the fact that the antisymmetrization operator $\mathcal{A}$ commutes with the evolution operator and regrouping all the terms one finally finds

$$e^{-(t'-t)(H-E_T)} \mathcal{A}(f_t - g_t) = 0. \tag{52}$$

This condition is satisfied whenever $\mathcal{A}(f_t - g_t) = 0$. The important conclusion is that one can replace f_t by any function g_t such that the difference $g_t - f_t$ is orthogonal to the space of antisymmetric functions. Let us now show how this property can be used to impose a common number of walkers in the positive and negative populations.

Let us consider the case where there are more positive walkers than negative walkers, $N_S^+ > N_S^-$. In this case one can subtract from f_t, equation (48), the following vector:

$$\frac{1}{\psi_G^+} |\mathbf{R}_i^+\rangle + P \frac{1}{\psi_G^+} |\mathbf{R}_i^+\rangle, \tag{53}$$

where $\mathbf{R}_i^+$ is a positive walker and P is a two-particle permutation. Such an operation is allowed since the application of the antisymmetrizer to the vector (53) gives zero [a direct consequence of $\mathcal{A}(1 + P) = 0$]. Now, because of equation (46) the vector (53) can also be written as

$$\frac{1}{\psi_G^+} |\mathbf{R}_i^+\rangle + \frac{1}{\psi_G^-} P |\mathbf{R}_i^+\rangle. \tag{54}$$

Subtracting this vector from f_t removes the contribution $\frac{1}{\psi_G^+} |\mathbf{R}_i^+\rangle$ from (48) and adds the contribution $\frac{1}{\psi_G^-} P |\mathbf{R}_i^+\rangle$ to (48). In other words, the positive walker $\mathbf{R}_i^+$ has been removed and the negative walker

$$\mathbf{R}_i^- = P\mathbf{R}_i^+ \tag{55}$$

has been created. Similarly, one can remove a negative walker $\mathbf{R}_i^-$ and create a positive walker

$$\mathbf{R}_i^+ = P\mathbf{R}_i^-. \tag{56}$$

This possibility of transferring one walker from one population to the other one allows us to keep an identical number of walkers for the two populations at each step. Now, thanks to this possibility, we can interpret the two populations consisting of the N_S positive walkers $\mathbf{R}_i^+$ and the N_S negative walkers $\mathbf{R}_i^-$ ($i \in [1, \ldots, N_S]$) as a unique population of N_S pairs of walkers $\{(\mathbf{R}_i^+, \mathbf{R}_i^-)\}$. Following this interpretation, the density f_t (48) can be then rewritten as an average over a population of pairs of walkers:

$$f_t = \left\langle \sum_{i=1}^{N_S} \left(\frac{1}{\psi_G^+} |\mathbf{R}_i^+\rangle - \frac{1}{\psi_G^-} |\mathbf{R}_i^-\rangle \right) \right\rangle, \tag{57}$$

and the energy can be computed as a ratio of averages performed on the population of pairs:

$$E_0^F = \frac{\left\langle \sum_i \left(\frac{H\psi_T}{\psi_G^+}(\mathbf{R}_i^+) - \frac{H\psi_T}{\psi_G^-}(\mathbf{R}_i^-) \right) \right\rangle}{\left\langle \sum_i \left(\frac{\psi_T}{\psi_G^+}(\mathbf{R}_i^+) - \frac{\psi_T}{\psi_G^-}(\mathbf{R}_i^-) \right) \right\rangle}, \tag{58}$$

where equation (7) has been rewritten by replacing f_t using equation (57). Now, everything is in order to detail the short-time dynamics of FMC. The FMC dynamics consists of three steps (Langevin, branching and cancellation steps).

(i) *Langevin step*. The Langevin processes (44) are simulated as in DMC, except that the Gaussian random variables of the positive and negative walkers are no longer independent. The positive walker $\mathbf{R}_i^+$ and the negative walker $\mathbf{R}_i^-$ are moved according to equation (44):

$$R_i'^{\mu\pm} = R_i^{\mu\pm} + b_i^{\mu\pm}\tau + \sqrt{\tau}\eta_i^{\mu\pm}, \tag{59}$$

where $\eta_i^{\mu\pm}$ are Gaussian centred random variables verifying

$$\langle \eta_i^{\mu\pm} \eta_j^{\mu\pm} \rangle = \delta_{ij}. \tag{60}$$

Such a move insures that the density of positive and negative walkers obey the Fokker Planck equation:

$$\langle |\mathbf{R}_i'^{\pm}\rangle \rangle = e^{\tau(\frac{1}{2}\nabla^2 - \nabla[\mathbf{b}^{\pm}.])} |\mathbf{R}_i^{\pm}\rangle. \tag{61}$$

However, the Gaussian random variables are no more independent; they are correlated within a pair

$$c_i^{\mu\nu} \equiv \langle \eta_i^{\mu+} \eta_i^{\nu-} \rangle \neq 0. \tag{62}$$

Different ways of correlating positive and negative walkers can be considered. We shall employ here the approach used in [18, 20] which consists in obtaining the vector η_i^-, representing the $3N$ coordinates $\eta_i^{\nu-}$, from η_i^+ by reflexion with respect to the hyperplane perpendicular to the vector $\mathbf{R}_i^+ - \mathbf{R}_i^-$:

$$\eta_i^- = \eta_i^+ - 2\frac{(\mathbf{R}_i^+ - \mathbf{R}_i^-) \cdot \eta_i^+}{(\mathbf{R}_i^+ - \mathbf{R}_i^-)^2}(\mathbf{R}_i^+ - \mathbf{R}_i^-). \tag{63}$$

This relation between the Gaussian random variables makes the move deterministic along the direction $\mathbf{R}_i^+ - \mathbf{R}_i^-$. Such a construction insures that the walkers within a pair will meet each other in a finite time (even in large-dimensional spaces). This aspect will be illustrated numerically in the last section. Formally, the two correlated Langevin processes can be seen as one Langevin process in the space of pairs of walkers:

$$\left(\mathbf{R}_i^{+\prime}, \mathbf{R}_i^{-\prime}\right) = \left(\mathbf{R}_i^+, \mathbf{R}_i^-\right) + \left(\mathbf{b}^+(\mathbf{R}_i^+), \mathbf{b}^-(\mathbf{R}_i^-)\right)\tau + \sqrt{\tau}\left(\eta_i^+, \eta_i^-\right), \tag{64}$$

where η_i^+ and η_i^- are related via (63).

(ii) *Branching step.* As we have already noticed, the branching of the negative and the positive walker is different within a pair:

$$w_i^+ \equiv e^{-\tau(E_L^+(\mathbf{R}_i^+)-E_T)} \neq w_i^- \equiv e^{-\tau(E_L^-(\mathbf{R}_i^-)-E_T)}. \tag{65}$$

Taking into account their respective weights, the two walkers of a pair give the following contribution to the density f_t:

$$w_i^+ \frac{1}{\psi_G^+}|\mathbf{R}_i^+\rangle - w_i^- \frac{1}{\psi_G^-}|\mathbf{R}_i^-\rangle. \tag{66}$$

If for example $w_i^+ > w_i^-$, this vector can be written as

$$w_i^- \left[\frac{1}{\psi_G^+}|\mathbf{R}_i^+\rangle - \frac{1}{\psi_G^-}|\mathbf{R}_i^-\rangle\right] \tag{67}$$

$$+\left(w_i^+ - w_i^-\right)\frac{1}{\psi_G^+}|\mathbf{R}_i^+\rangle. \tag{68}$$

This density is the sum of two contributions. The first contribution, equation (67), comes from a pair of walkers $\left(\mathbf{R}_i^+, \mathbf{R}_i^-\right)$ and carries the weight w_i^-. The second, equation (68), comes from a single positive walker $\mathbf{R}_i^+$ and carries the weight $w_i^+ - w_i^-$. This single walker $\mathbf{R}_i^+$ can be replaced by a pair as follows. First, this single walker can be replaced by two positive walkers $\mathbf{R}_i^+$ with half of the weight, $\frac{1}{2}\left(w_i^+ - w_i^-\right)$. One of these two positive walkers carrying half of the weight can be transferred to the population of negative walkers by exchanging two particles. Finally, this single walker $\mathbf{R}_i^+$ can be replaced by a pair $\left(\mathbf{R}_i^+, P\mathbf{R}_i^+\right)$ carrying the weight $\frac{1}{2}\left(w_i^+ - w_i^-\right)$. The resulting process just described is a branching of the pair $\left(\mathbf{R}_i^+, \mathbf{R}_i^-\right)$ with the weight w_i^- and the creation of the pair $\left(\mathbf{R}_i^+, P\mathbf{R}_i^+\right)$ with the weight $\frac{1}{2}\left(w_i^+ - w_i^-\right)$. Of course, if one has $w_i^+ < w_i^-$, then, the pair $\left(\mathbf{R}_i^+, \mathbf{R}_i^-\right)$ is branched with the weight w_i^+ and the pair $\left(P\mathbf{R}_i^-, \mathbf{R}_i^-\right)$ is created with the weight $\frac{1}{2}\left(w_i^- - w_i^+\right)$.

Both cases, $w_i^- < w_i^+$ or $w_i^- > w_i^+$, can be summarized as follows. The pair of walkers $(\mathbf{R}_i^+, \mathbf{R}_i^-)$ is branched with the weight

$$\min\left(w_i^+, w_i^-\right) = e^{-\tau(\max(E_L^+, E_L^-) - E_T)}, \tag{69}$$

and the pairs $(\mathbf{R}_i^+, P\mathbf{R}_i^+)$ and $(P\mathbf{R}_i^-, \mathbf{R}_i^-)$ are created with their respective weights

$$\frac{1}{2}\left(w_i^+ - \min\left(w_i^+, w_i^-\right)\right) = -\frac{\tau}{2}\left[E_L^+ - \max\left(E_L^+, E_L^-\right)\right] + O(\tau^2) \tag{70}$$

and

$$\frac{1}{2}\left(w_i^- - \min\left(w_i^+, w_i^-\right)\right) = -\frac{\tau}{2}\left[E_L^- - \max\left(E_L^+, E_L^-\right)\right] + O(\tau^2). \tag{71}$$

(iii) *Cancellation step.* The third step is a cancellation procedure performed whenever a positive and a negative walkers meet. When $\mathbf{R}_i{}^+ = \mathbf{R}_i{}^-$, the contribution of the pair to the density can be simplified as follows:

$$\left[1 - \frac{\psi_G^+}{\psi_G^-}(\mathbf{R}_i^+)\right]\frac{1}{\psi_G^+}(\mathbf{R}_i^+)|\mathbf{R}_i^+\rangle. \tag{72}$$

If the term in brackets is positive, this contribution comes from one single positive walker $\mathbf{R}_i^+$ with multiplicity $\left[1 - \frac{\psi_G^+}{\psi_G^-}(\mathbf{R}_i^+)\right]$. One can transform this single walker into a pair of positive and negative walkers $(\mathbf{R}_i^+, P\mathbf{R}_i^+)$ with the new multiplicity:

$$\frac{1}{2}\left[1 - \frac{\psi_G^+}{\psi_G^-}(\mathbf{R}_i^+)\right]. \tag{73}$$

If the term in brackets is negative, the pair $(P\mathbf{R}_i^+, \mathbf{R}_i^+)$ is drawn and the multiplicity is given by

$$\frac{1}{2}\left[1 - \frac{\psi_G^-}{\psi_G^+}(\mathbf{R}_i^+)\right]. \tag{74}$$

This is a cancellation procedure because a pair $(\mathbf{R}_i^+, \mathbf{R}_i^+)$ with a multiplicity 1 has been transformed into a pair with a multiplicity smaller than 1. Note that, when $\psi_G^+ = \psi_G^-$, the multiplicities (73) or (74) reduce both to zero. In other words, there is a total cancellation of the pair whenever the walkers meet. As we shall see in the following section, the cancellation step is at the origin of the improved stability. The basic reason is that this procedure removes pairs which do not contribute to the signal but only to the statistical noise. A rigorous analysis of this point is provided in the following section.

5. Stability of the FMC method

5.1. Criterium for stability

We have just seen that the FMC method is a generalization of the DMC approach, and we have shown that FMC preserves the evolution of the antisymmetric component of the sampled density. Now, having shown that FMC is an exact method, it is necessary to study the stability of the method. For that purpose, we consider the estimator of the energy, equation (58), in the large-time regime:

$$E_0^F = \frac{\mathcal{N}}{\mathcal{D}} = \frac{\frac{1}{N_S}\sum_i \frac{H\psi_T}{\psi_G^+}(\mathbf{R}_i^+) - \frac{H\psi_T}{\psi_G^-}(\mathbf{R}_i^-)}{\frac{1}{N_S}\sum_i \frac{\psi_T}{\psi_G^+}(\mathbf{R}_i^+) - \frac{\psi_T}{\psi_G^-}(\mathbf{R}_i^-)}, \tag{75}$$

where N_S is the population size at time t. In the same way as done for the estimator (10), we can evaluate the variance of E_0^F by supposing that N_S is large enough so that both the numerator and the denominator have small fluctuations around their average:

$$\sigma^2 \left(\frac{\mathcal{N}}{\mathcal{D}} \right) = \frac{\left\langle \left(\mathcal{N} - E_0^F \mathcal{D} \right)^2 \right\rangle}{\langle \mathcal{D} \rangle^2}. \tag{76}$$

Let us begin with the denominator. Using identity (57), the average of the random variable $\mathcal{D}$ defined in (75) is nothing but

$$\langle \mathcal{D} \rangle = \frac{1}{N_S} \langle \psi_T | f_t \rangle. \tag{77}$$

Replacing f_t by its asymptotic behaviour, equation (6), we finally find that the denominator of (76) behaves for large t as

$$\langle \mathcal{D} \rangle^2 = \frac{1}{N_S^2} \, \mathrm{e}^{-2t(E_0^F - E_T)} \langle \psi_T | \phi_0^F \rangle^2. \tag{78}$$

Now, let us compute the numerator of (76). This numerator can be written as the variance of a sum of random variables defined over the pairs of walkers

$$\left\langle \left(\mathcal{N} - E_0^F \mathcal{D} \right)^2 \right\rangle = \left\langle \left[\frac{\sum_i \Gamma\left(\mathbf{R}_i^+, \mathbf{R}_i^-\right)}{N_S} \right]^2 \right\rangle, \tag{79}$$

where we have introduced the function $\Gamma(\mathbf{R}^+, \mathbf{R}^-)$:

$$\Gamma(\mathbf{R}^+, \mathbf{R}^-) \equiv \frac{\left(H - E_0^F\right)\psi_T}{\psi_G^+}(\mathbf{R}^+) - \frac{\left(H - E_0^F\right)\psi_T}{\psi_G^-}(\mathbf{R}^-). \tag{80}$$

Using the fact that the pairs of walkers have the same distribution and supposing that they are independent we finally find

$$\left\langle \left(\mathcal{N} - E_0^F \mathcal{D} \right)^2 \right\rangle = \frac{1}{N_S} \sigma^2 \left(\Gamma(\mathbf{R}^+, \mathbf{R}^-) \right). \tag{81}$$

If the pairs of walkers are not independent this expression is only modified by a correlation factor independent on the time and the population size N_S provided that N_S and t are large enough. Finally, up to a multiplicative factor, the variance of the energy estimator has the following asymptotic behaviour:

$$\sigma^2 \left(\frac{\mathcal{N}}{\mathcal{D}} \right) \propto \frac{1}{N_S(t)} C(t), \tag{82}$$

where the coefficient $C(t)$ is given by

$$C(t) = N_S(t)^2 \, \mathrm{e}^{2t(E_0^F - E_T)}, \tag{83}$$

and where $N_S(t)$, the number of pairs depends on time t due to the birth–death process. In expression (83) we note that the behaviour of (76) at large times t is related to the value of E_0^F and the asymptotic behaviour of the number of pairs of walkers $N_S(t)$. We can already understand physically the interest of the cancellation process: this process limits the growth of, $N_S(t)$, the number of pairs of walkers, thus limiting the growth of $C(t)$ (83) and the variance (82). Let us now precise this criterion more rigorously by evaluating the asymptotic behaviour of $N_S(t)$. For that purpose we introduce the density of pairs $\Pi_t\left(\mathbf{R}_i^+, \mathbf{R}_i^-\right)$; this density obeys a diffusion equation

$$\frac{\partial \Pi_t}{\partial t} = -(D_{\mathrm{FMC}} - E_T)\Pi_t, \tag{84}$$

where we have introduced the diffusion operator $-(D_{FMC} - E_T)$. We will give its expression later; for the present purpose we just need the asymptotic behaviour of Π_t given by

$$\Pi_t = e^{-t(\tilde{E}_0^B - E_T)} \Pi_S, \tag{85}$$

where Π_S is the stationary density of the process, namely the lowest eigenstate of the operator D_{FMC}, and $\tilde{E}_0^B$ is the corresponding eigenvalue. The number of pairs $N_S(t)$ behaves as the normalization of Π_t, and consequently grows like $e^{-t(\tilde{E}_0^B - E_T)}$. Note that in practice one adjusts the reference energy E_T to $\tilde{E}_0^B$ during the simulation to keep a constant population of average size $\bar{N}_S$ along the dynamics. Such a procedure is referred to as a control population technique [24] and will be discussed later. Finally, the asymptotic behaviour of the variance of the FMC estimator of the energy is

$$\sigma^2 \left(\frac{\mathcal{N}}{\mathcal{D}} \right) \propto \frac{1}{\bar{N}_S} e^{2t(E_0^F - \tilde{E}_0^B)}. \tag{86}$$

This expression is analogous to (13) except that the lowest energy of the Hamiltonian operator H has been replaced by the lowest energy of the operator D_{FMC}. In conclusion the stability of the algorithm is related to the lowest eigenvalue of the FMC diffusion operator, $\tilde{E}_0^B$. It is clear from (86) that the higher this eigenvalue is, the more stable the simulation will be. In the following section, we will discuss the allowed values of $\tilde{E}_0^B$. This will prove that FMC is not a stable method in general, but is more stable than any standard transient method.

5.2. Stability of the fermion Monte Carlo algorithm

In this section, we prove that the lowest eigenvalue of the FMC operator, $\tilde{E}_0^B$, has the following upper and lower bounds:

$$\tilde{E}_0^B \leqslant E_0^F \tag{87}$$

$$\tilde{E}_0^B > E_0^B \tag{88}$$

From the expression of the variance, equation (86), one can easily understand the meaning of these two inequalities. The first inequality indicates that FMC is not a stable method, the stability being achieved only in the limit $\tilde{E}_0^B = E_0^F$. Note that, even for very simple systems, this stability is in general not obtained. This important point will be illustrated in the following section. The second inequality shows that FMC is more stable than any standard transient DMC method (nodal release method). Indeed, the exponent associated with the explosion of fluctuations, equation (86), is smaller than in the standard case, equation (13). Before giving a mathematical proof of these two inequalities, let us first present some intuitive arguments in their favour. The first inequality, equation (87), takes its origin in the fact that the signal—the antisymmetric component of f_t, equation (6)—is extracted from the population of pairs of walkers, equation (57), and, consequently, cannot grow faster than the population of pairs itself, equation (85). The second inequality can be understood as follows. Without the cancellation process, the FMC method reduces to two correlated DMC algorithms. The number of walkers grows as in a standard DMC, namely $\sim e^{-(E_0^B - E_T)t}$. The cancellation process obviously reduces the growth of the population of walkers, equation (85), and, thus, we should expect that $\tilde{E}_0^B > E_0^B$. Now, let us give some more rigorous proofs. For that purpose we compare the fermion Monte Carlo operators with and without cancellation process. Without the cancellation process the FMC diffusion operator reads

$$D - E_T \equiv \psi_G^+ \left(H^+ - E_L^+ \right) \frac{1}{\psi_G^+} + \psi_G^- (H^- - E_L^-) \frac{1}{\psi_G^-} \tag{89}$$

$$-\frac{1}{2}\sum_{\mu\neq\nu}\frac{\partial^2}{\partial R^+{}_\mu\partial R^-{}_\nu}[c_{\mu\nu}.] \tag{90}$$

$$+\max\left(E_L^+, E_L^-\right) - E_T \tag{91}$$

$$+\frac{1}{2}\left[E_L^+ - \max\left(E_L^+, E_L^-\right)\right]e^{(P\mathbf{R}^+ - \mathbf{R}^-)\cdot\nabla^-}. \tag{92}$$

$$+\frac{1}{2}\left[E_L^- - \max\left(E_L^+, E_L^-\right)\right]e^{(P\mathbf{R}^- - \mathbf{R}^+)\cdot\nabla^+}, \tag{93}$$

where the operators $H^\pm$ are both identical to H, except that H^+ and H^- act on the space of positive and negative configurations, respectively:

$$H^\pm \equiv H(\mathbf{R}^\pm) = -\frac{1}{2}\sum_\mu\frac{\partial^2}{\partial R_\mu^{\pm 2}} + V(\mathbf{R}^\pm). \tag{94}$$

The coefficients $c_{\mu\nu}$ are real coefficients and will be defined below. To justify that this operator is the diffusion operator corresponding to FMC with no cancellation process, we need to check that the short-time dynamics described by

$$\frac{\partial\Pi_t}{\partial t} = -(D - E_T)\Pi_t \tag{95}$$

is indeed realized via the two first steps of the FMC algorithm (Langevin and branching steps). In the expression of D, the two operators appearing in equation (89) define a Fokker Planck operator in analogy to equation (40). This operator is the diffusion operator associated with the Langevin process, equation (64). The term in equation (90) is a coupling term between the moves of positive and negative walkers taking into account the correlation of the Gaussian random variables η_μ^+ and η_μ^-. The quantities $c_{\mu\nu}(\mathbf{R}^+, \mathbf{R}^-)$ introduced in equation (90) are nothing but the covariance of these variables:

$$c_{\mu\nu}(\mathbf{R}^+, \mathbf{R}^-) = \langle\eta_\mu^+\eta_\nu^-\rangle. \tag{96}$$

The three last contributions describe the branching processes at work in FMC. One recognizes in equation (91) the branching of a pair, equation (69). The two following contributions, equations (92) and (93), correspond to the creation of pairs $(\mathbf{R}^+, P\mathbf{R}^+)$ and $(P\mathbf{R}^-, \mathbf{R}^-)$, with the respective weights given by (70) and (71). Note that in equations (92) and (93) the operator $e^{(P\mathbf{R}^- - \mathbf{R}^+)\cdot\nabla^+}$ is written in a symbolic form representing a translation of the vector $\mathbf{R}^+$ to $P\mathbf{R}^-$, the action of this operator on the pair $(\mathbf{R}^+, \mathbf{R}^-)$ being indeed to create the pair $(P\mathbf{R}^-, \mathbf{R}^-)$.

Now, let us prove that the lowest eigenvalue of D is E_0^B (bosonic ground state). For that purpose, it is convenient to introduce the operator $\mathcal{R}$ which transforms a distribution of pairs of walkers into a distribution of walkers and then to define the following reduced density:

$$\mathcal{R}\Pi_t(\mathbf{R}) \equiv \int dR'\left[\frac{\Pi_t(\mathbf{R}, \mathbf{R}')}{\psi_G^+(\mathbf{R})} + \frac{\Pi_t(\mathbf{R}', P\mathbf{R})}{\psi_G^-(\mathbf{R})}\right]. \tag{97}$$

The density $\mathcal{R}\Pi_t$ represents the sum of the distributions sampled by each type of walkers when the contribution of the other type of walkers is integrated out. Using the explicit expression of D it is a simple matter of algebra to verify that

$$\mathcal{R}D\Pi_t = H\mathcal{R}\Pi_t. \tag{98}$$

Using equations (98) and (95), one can also write

$$\frac{\partial\mathcal{R}\Pi_t}{\partial t} = -(H - E_T)\mathcal{R}\Pi_t, \tag{99}$$

which means that the reduced density evolves under the dynamics of H. In other words, the set of positive and negative walkers sample the same distribution as a standard diffusion Monte Carlo algorithm. Now, suppose that λ_S is the lowest eigenvalue of D and Π_S the corresponding eigenstate (the stationary density of the process described by equation (95)):

$$D\Pi_S = \lambda_S \Pi_S. \tag{100}$$

Applying $\mathcal{R}$ on both sides of this identity and using relation (98) one gets

$$H\mathcal{R}\Pi_S = \lambda_S \mathcal{R}\Pi_S. \tag{101}$$

In other words, the reduced density $\mathcal{R}\Pi_S$ is a positive eigenstate of H with eigenvalue λ_S. The bosonic state being non-degenerate, we can conclude that $\lambda_S = E_0^{\mathrm{B}}$. This ends our proof.

Let us now consider the genuine FMC diffusion operator including the cancellation process, D_{FMC}. To simplify the notations let us suppose that we are in the symmetric case for which $\psi_G^+ = \psi_G^- = \psi_G$ (the common guiding function is symmetric under permutation of particles). This particular case is much simple because when walkers meet, there is a full cancellation and no residual branching. From an operatorial point of view the cancellation step consists in introducing a projection operator, P_c, at each step of the dynamics:

$$P_c \equiv \left[1 - \int \mathrm{d}R \, |R^+ = R, R^- = R\rangle\langle R^+ = R, R^- = R| \right], \tag{102}$$

where $|R^+, R^-\rangle$ denotes the usual tensorial product. The full FMC diffusion operator can thus be written as

$$D_{\mathrm{FMC}} \equiv P_c D. \tag{103}$$

It is important to realize that the D_{FMC} operator defined via P_c and D (equations (102) and (93)) represents indeed an equivalent operatorial description of the stochastic rules of FMC described in section 4.2 (Langevin, branching and cancellation steps). Note also that using expression (103) of D_{FMC}, we have a simple alternative way of recovering the proof just presented above that FMC is a bias-free approach. Since this is an important point of this work, let us present this alternative proof. The action of the projection operator P_c is to remove from the sample components of the form $|R^+ = R, R^- = R\rangle$ for which the antisymmetric component of the reduced density is zero:

$$\mathcal{A}\mathcal{R}|R^+ = R, R^- = R\rangle = \mathcal{A}\frac{1}{\psi_G(\mathbf{R})}(|R\rangle + |PR\rangle) = 0. \tag{104}$$

In other words one has the following algebraic identity:

$$\mathcal{A}\mathcal{R}P_c = \mathcal{A}\mathcal{R}. \tag{105}$$

In the general case where $\psi_G^+ \neq \psi_G^-$, the cancellation procedure still corresponds to define a new operator written as in equation (103) with P_c satisfying the same identity as in (105). Applying $A\mathcal{R}$ to the LHS and RHS of equation (84), one has

$$\frac{\partial \mathcal{A}\mathcal{R}\Pi_t}{\partial t} = -\mathcal{A}\mathcal{R}D_{\mathrm{FMC}}\Pi_t = -\mathcal{A}(H - E_T)\mathcal{R}\Pi_t. \tag{106}$$

This equation indicates that the evolutions of the antisymmetric component of the reduced density under the dynamics of D_{FMC} and H are identical. This confirms that the energy estimator, equation (57), or any observable estimator not coupling directly positive and negative walkers, is not biased. In the case of the energy, the estimator can be written as a function of the reduced density as follows:

$$E_0^{\mathrm{F}} = \frac{\langle\psi_T|H\mathcal{R}\Pi_t\rangle}{\langle\psi_T|\mathcal{R}\Pi_t\rangle}. \tag{107}$$

Let us now turn back to our discussion of the stability of FMC. For that, we need to compare the lowest eigenvalue $\tilde{E}_0^B$ of D_{FMC} and, E_0^B, the lowest eigenvalue of H. Now, it is clear from the definition of D_c, equation (103), that the following relation holds:

$$\tilde{E}_0^B > E_0^B. \tag{108}$$

Indeed, the action of P_c present in the definition of D_c, equation (103), consists in removing positive coefficients within the extradiagonal part of D. As well known, a consequence of such a matrix (or operator) manipulation is to increase the energy of the lowest eigenvalue of the matrix. Expressed in a more physical way, the cancellation process reduces the growth of the population of pairs: $e^{-(\tilde{E}_0^B - E_T)t} < e^{-(E_0^B - E_T)t}$.

To summarize, we have shown that FMC reduces the instability of fermion simulations. The signal-over-noise ratio decreases as $e^{(\tilde{E}_0^B - E_0^F)t}$, where $\tilde{E}_0^B$ is the lowest eigenvalue of the FMC operator, D_{FMC}. Because of inequality (108), this ratio decreases slower in FMC than in any standard transient DMC or nodal release methods. We have shown that the cancellation process is at the origin of this improvement; however, as we shall see in the following section, the cancellation process is efficient (i.e., we have a small difference $\tilde{E}_0^B - E_0^F$) only if the correlation between walkers described by the coupling terms $c_{\mu\nu}$ is introduced. This feature is important, particularly in high-dimensional spaces where the probability of meeting and cancelling becomes extremely small for independent walkers. As a result, the correlation of positive and negative walkers is a fundamental feature of FMC. The quantitative effect of the correlation on the stabilization of the algorithm is not easy to study theoretically and to optimize in the general case. In the following section we will give a numerical illustration, for a simple system, of the interplay between cancellation and correlation (via the $c_{\mu\nu}$ parameters), and also of the role of the choice of the guiding functions, ψ_G^+ and ψ_G^-.

6. Numerical study

6.1. The model: 2D-harmonic oscillator on a finite grid

In this section, we study the FMC method on a very simple model on a lattice. For this model it is possible to calculate E_0^F (fermionic ground-state energy), E_0^B (bosonic ground-state energy) and $\tilde{E}_0^B$ the lowest eigenvalue of the FMC operator by a standard deterministic method (exact diagonalization). The results obtained for this simple model will provide us a well-grounded framework to interpret the fermion Monte Carlo simulations. The second motivation is that, using such a simple model, it is possible to study the limit of a large number of walkers, large with respect to the dimension of the Hilbert space considered. This possibility turns out to be essential to better understand the FMC algorithm.

Our model is based on the discretization of a system describing two-coupled harmonic oscillators

$$H = -\frac{1}{2}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) + V(x, y), \tag{109}$$

with

$$V(x, y) = \tfrac{1}{2}x^2 + \tfrac{1}{2}\lambda y^2 + xy. \tag{110}$$

In the following we shall take $\lambda = 2$. Now, we define the discretization of this model on a $N \times N$ regular grid (N odd). A grid point $\mathbf{R}_i$ $i \in (1, \ldots, N^2)$ has the following coordinates:

$$\mathbf{R}_i \equiv \left(\left(-\frac{N}{2} + k - 1\right)\delta_x, \left(-\frac{N}{2} + l - 1\right)\delta_y\right) \quad k \in [1, \ldots, N], \quad l \in [1, \ldots, N], \tag{111}$$

where

$$\delta_x = \delta_y = x_{\max}/N. \tag{112}$$

On this lattice the Hamiltonian has a corresponding discrete representation given by a finite matrix. The diagonal part of the matrix reads

$$H_{ii} = \frac{1}{\delta_x{}^2} + \frac{1}{\delta_y{}^2} + V(R_i) \qquad i \in (1, \dots, N^2) \tag{113}$$

and the off-diagonal part reads

$$H_{ij} = -\frac{1}{2\delta_x{}^2} \quad \text{or} \quad -\frac{1}{2\delta_y{}^2} \quad \text{when } R_i \text{ and } R_j \text{ are nearest-neighbours on the lattice} \tag{114}$$

$$H_{ij} = 0 \quad \text{otherwise}.$$

This Hamiltonian is symmetric with respect to the inversion P of centre $O = (0, 0)$:

$$P(x, y) \equiv (-x, -y). \tag{115}$$

As a consequence, the eigenfunctions are either symmetric or antisymmetric under P. We are interested in the energy E_0^F of the lowest antisymmetric eigenstate, ϕ_0^F. Even for this simple system, we are confronted with a sign instability and a genuine 'sign problem'. Indeed, the sign of ϕ_0^F for each grid point cannot be entirely determined by symmetry. Symmetry implies only that ϕ_0^F vanishes at the inversion centre and that the two-dimensional pattern of positive and negative values for ϕ_0^F is symmetric by inversion. The precise delimitation between a positive and a negative zones of the wavefunction (analogous to 'nodal surfaces' for continuous systems) is not known.

Let us now introduce the trial functions ψ_T and ψ_S. ψ_T has to be an (antisymmetric) approximation of ϕ_0^F and ψ_S some symmetric and positive approximation of the lowest eigenstate, ϕ_0^B. We have chosen them as discretizations of the exact solutions of the initial continuous model. To find these solutions we perform a diagonalization of the quadratic form

$$V(x, y) = \tfrac{1}{2}x^2 + \tfrac{1}{2}\lambda y^2 + xy = \tfrac{1}{2}k_1 \tilde{x}^2 + \tfrac{1}{2}k_2 \tilde{y}^2. \tag{116}$$

It is trivial to verify that

$$k_1 = \frac{\cos^2\theta - \lambda\sin^2\theta}{\cos 2\theta} \qquad k_2 = \frac{\lambda\cos^2\theta - \sin^2\theta}{\cos 2\theta} \qquad \tan 2\theta = \frac{2}{\lambda - 1},$$

with

$$\tilde{x} = x\cos\theta - y\sin\theta \tag{117}$$

and

$$\tilde{y} = x\sin\theta + y\cos\theta. \tag{118}$$

If $k_1 < k_2$, we choose as trial wavefunction

$$\psi_T = \tilde{x}\exp\left(-\frac{\sqrt{k_1}}{2}\tilde{x}^2 - \frac{\sqrt{k_2}}{2}\tilde{y}^2\right), \tag{119}$$

while, in the other case, we take

$$\psi_T = \tilde{y}\exp\left(-\frac{\sqrt{k_1}}{2}\tilde{x}^2 - \frac{\sqrt{k_2}}{2}\tilde{y}^2\right). \tag{120}$$

The lowest (symmetric) eigenstate is chosen to be

$$\psi_S = \exp\left(-\frac{\sqrt{k_1}}{2}\tilde{x}^2 - \frac{\sqrt{k_2}}{2}\tilde{y}^2\right). \tag{121}$$

Note that, in the limit of a very large system the trial functions, ψ_T and ψ_S, reduce to two exact eigenstates of H; however, this is not the case for finite systems.

6.2. FMC on the lattice

Before presenting our results, let us say a few words about the implementation of the FMC on the lattice. The same ingredients as in the continuum case hold, except that in the lattice case the Langevin process is realized through a discrete transition probability matrix. The probability for a (positive or negative) walker i to go to j, after a time step τ is

$$P^{\pm}(i \rightarrow j) \equiv \frac{\psi_G^{\pm}(\mathbf{R}_j)}{\psi_G^{\pm}(\mathbf{R}_i)} \langle \mathbf{R}_j | 1 - \tau (H - E_L^{\pm}) | \mathbf{R}_i \rangle, \tag{122}$$

where τ is small enough to have a positive density, namely

$$\tau < \frac{1}{\text{Max}[H_{ii} - E_L^{\pm}(i)]}. \tag{123}$$

The local energies are defined as in the continuum, equation (43),

$$E_L^{\pm}(\mathbf{R}) = \frac{H \psi_G^{\pm}}{\psi_G^{\pm}}(\mathbf{R}), \tag{124}$$

with the same expression for the guiding functions $\psi_G^{\pm}$, equation (47),

$$\psi_G^{\pm}(\mathbf{R}) \equiv \sqrt{\psi_S^2 + c^2 \psi_T^2} \pm c \psi_T. \tag{125}$$

Let (i_1, i_2) represent a given pair of positive and negative walkers $(\mathbf{R}_{i_1}^+, \mathbf{R}_{i_2}^-)$. In a standard diffusion Monte Carlo (no correlation and no cancellation of positive and negative walkers), the density of pairs of positive and negative walkers $\Pi_{i_1 i_2}^{(k)}$ evolves as follows in one time-step:

$$\Pi_{i_1 i_2}^{(k+1)} = \sum_{j_1 j_2} \Pi_{j_1 j_2}^{(k)} P^+(j_1 \rightarrow i_1) W_{j_1}^+ P^-(j_2 \rightarrow i_2) W_{j_2}^-, \tag{126}$$

where $W^{\pm}$ is the Feynman–Kac weight, equation (45). To build the FMC algorithm, one first correlate the two stochastic processes P^+ and P^-, equation (122). The way it is performed here is the counterpart of the correlation term introduced by Liu *et al* [20] in the continuum case, equation (63). With such a choice, the positive and negative walkers of a pair tend to get closer or to move away in a concerted way. For the lattice case it is done as follows. The positive walker j_1 is connected by P^+ to a finite number of states j_1^c (here, maximum five) with probability $P^+(j_1 \rightarrow j_1^c)$. The negative walker j_2 is connected to a finite number of states j_2^c with the probability $P^-(j_2 \rightarrow j_2^c)$. The states j_1^c are ordered taking as criterion the distance to the negative walker j_2, $|\mathbf{R}_{j_1^c} - \mathbf{R}_{j_2}|$. We do the same for the states i_2^c ordered by their distance with respect to the positive walker. A *unique* random number uniformly distributed between 0 and 1 is then drawn and the repartition functions of the two probability measures, $p(j_1^c) \equiv P^+(j_1 \rightarrow j_1^c)$ and $p(j_2^c) \equiv P^-(j_2 \rightarrow j_2^c)$ are then sampled using this common random number. The new pair (j_1^c, j_2^c) is drawn accordingly. Such a procedure defines a *correlated* transition probability in the space of pairs

$$P_c(j_1 j_2 \rightarrow i_1 i_2) \neq P(j_1 \rightarrow i_2) P(j_1 \rightarrow i_2), \tag{127}$$

whose role is to enhance the probability of having positive and negative walkers meeting at the same site. Note that, by construction, the correlation introduced via P_c does not change the individual (reduced) densities associated with each type of walker (positive/negative). Now, let us write explicitly the FMC rules in our lattice case, that is, the one time-step $(k \rightarrow k + 1)$ evolution of the density of pairs $\Pi_{j_1 j_2}^{(k)}$.

(i) *Correlation and branching.* The diffusion and branching of an individual pair (j_1, j_2), equations (90) and (91), correspond to the following evolution of the density:

$$\Pi_{i_1 i_2}^{(k+1)} = \Pi_{i_1 i_2}^{(k)} + \sum_{j_1 j_2} \Pi_{j_1 j_2}^{(k)} P_c(j_1 j_2 \rightarrow i_1 i_2) \text{Min}\left[W_{i_1}^+, W_{i_2}^-\right]. \tag{128}$$

The creation of pairs, equations (92) and (93), can be written as follows:

$$\Pi_{i_1 P(i_1)}^{(k+1)} = \Pi_{i_1 P(i_1)}^{(k)} + \sum_{j_1 j_2} \Pi_{j_1 j_2}^{(k)} P_c(j_1 j_2 \rightarrow i_1 i_2) \theta\left[W_{i_1}^+, W_{i_2}^-\right] \left|W_{i_1}^+ - W_{i_2}^-\right|/2 \tag{129}$$

$$\Pi_{P(i_2) i_2}^{(k+1)} = \Pi_{P(i_2) i_2}^{(k)} + \sum_{j_1 j_2} \Pi_{j_1 j_2}^{(k)} P_c(j_1 j_2 \rightarrow i_1 i_2) \left(1 - \theta\left[W_{i_1}^+, W_{i_2}^-\right]\right) \left|W_{i_1}^+ - W_{i_2}^-\right|/2, \tag{130}$$

where $\theta(x, y) = 1$ if $x > y$, $\theta(x, y) = 0$, otherwise.

(ii) *Cancellation.* The cancellation process is done when a positive walker and a negative walker meet $i_1 = i_2 = i$:

$$\Pi_{ii}^{(k+1)} = \left[1 - \min\left(\frac{\psi_G^+}{\psi_G^-}(i), \frac{\psi_G^-}{\psi_G^+}(i)\right)\right] \Pi_{ii}^{(k)}. \tag{131}$$

If $\psi_G^+(i) > \psi_G^-(i)$ we have

$$\Pi_{P(i)i}^{(k+1)} = \Pi_{P(i)i}^{(k)} + \frac{\left[1 - \psi_G^-(i)/\psi_G^+(i)\right]}{2} \Pi_{ii}^{(k)}. \tag{132}$$

If $\psi_G^+(i) < \psi_G^-(i)$ we have

$$\Pi_{i P(i)}^{(k+1)} = \Pi_{i P(i)}^{(k)} + \frac{\left[1 - \psi_G^+(i)/\psi_G^-(i)\right]}{2} \Pi_{ii}^{(k)}. \tag{133}$$

Operations (128)–(133) describe the one time-step dynamics of the simulation. At iteration k, the distribution of pairs $\Pi^{(k)}(i_1, i_2)$ is obtained and the transient estimator of the energy (58) can be computed from

$$E(k) = \frac{\sum_{i_1, i_2} \Pi_{i_1, i_2}^{(k)} \left[\frac{H\psi_T(i_1)}{\psi_G^+(i_1)} - \frac{H\psi_T(i_2)}{\psi_G^-(i_2)}\right]}{\sum_{i_1, i_2} \Pi_{i_1, i_2}^{(k)} \left[\frac{\psi_T(i_1)}{\psi_G^+(i_1)} - \frac{\psi_T(i_2)}{\psi_G^-(i_2)}\right]}. \tag{134}$$

This estimator converges to E_0^F when $k \rightarrow \infty$.

Now, it is important to realize that the FMC rules just presented have, in principle, no stochastic character at all. For a finite system the FMC rules can be viewed as simple deterministic matrix manipulations between finite vectors of the Hilbert space (here, the multiplications to be performed have been explicitly written). At the beginning of the simulation (iteration $k = 0$) some arbitrary starting vector $\Pi_{i_1, i_2}^{(0)}$ is chosen and, then, iterations are performed up to convergence. This important remark will allow us to organize our discussion of the FMC results into two parts. In the first part (section 6.3), we perform explicitly the matrix multiplications involved, and any stochastic aspect is removed from the results. Stated differently, this procedure can be viewed as performing a standard FMC simulation with an infinite number of walkers (the distribution at each point of the configuration space is exactly obtained, no statistical fluctuations are present). In the second part (section 6.4), we implement the usual stochastic interpretation of the FMC rules using a finite number of walkers. This second part will allow us to discuss the important consequences of the finiteness of the number of walkers and, in particular, the role played by the use of population control techniques.

Table 1. Average meeting times $\frac{\langle T \rangle}{N}$ for the correlated and uncorrelated cases in the non-symmetric case $c = 4$, equation (125). In this example, $x_{max} = 3$, equation (112), and $\tau = 0.9\,\tau_{max}$, where τ_{max} is the maximal time-step defined in equation (123).

Linear size N	$\frac{\langle T \rangle}{N}$ Uncorrelated cases	$\frac{\langle T \rangle}{N}$ Correlated cases
$N = 3$	2152 (23)	134 (1)
$N = 5$	2162 (20)	92 (1)
$N = 7$	2234 (26)	75 (1)
$N = 9$	2834 (27)	76 (1)
$N = 11$	3546 (38)	82 (1)
$N = 13$	4214 (41)	88 (1)
$N = 15$		94 (1)
$N = 17$		102 (1)

6.3. FMC using an infinite number of walkers: the deterministic approach

6.3.1. No systematic error: FMC is an exact method. In this section, we verify on our simple example that the FMC rules do not introduce any systematic error (bias). The energy expression (134) has been computed by iterating the applications of the elementary operators defined by (128)–(133). In practice, this corresponds to iterate a matrix $G_{FMC}(\tau)$. The distribution $\Pi^{(k+1)}$ at iteration $k + 1$ is obtained from $\Pi^{(k)}$ as follows:

$$\Pi^{(k+1)} = G_{FMC}(\tau)\Pi^{(k)}. \tag{135}$$

The operator $G_{FMC}(\tau)$ has been applied, a large number of times, to some initial density $\Pi^{(0)}_{j_1 j_2}$ (an N^4 component vector, N being the linear size of our lattice), and the energy (134) has been computed at each iteration k. We have checked that the energy converges to the exact value, E_0^F, corresponding the lowest antisymmetric state, with all decimal places. We have verified that this is true for several cases corresponding to N ranging from 4 to 17. For this specific problem these results confirm numerically that FMC is an exact method.

6.3.2. Meeting time between a positive and a negative walkers. Here, we want to illustrate quantitatively the fact that the correlation introduced via the probability transition helps greatly to lower the meeting time between a positive and a negative walkers. The influence of the choice of the guiding functions (here, parameter c in equation (125)) on the meeting time is also examined. The meeting time is defined and evaluated as follows. We start with a configuration consisting of a positive walker located at a corner of the lattice and a negative walker located at the opposite corner. The positive and negative walkers are moving stochastically with the transition probability defined in (122). We test the two cases corresponding to uncorrelated and correlated moves. The average time $\langle T \rangle$ (number of Monte Carlo steps times τ) before the walkers meet is computed. Our results are presented in table 1 and are given for different linear sizes of the grid. In this first case the guiding functions are chosen with a large antisymmetric component, $c = 4$. The results indicate clearly that more than one order of magnitude is gained by correlating the moves of the two stochastic processes. In table 2 the same calculations are done, except that a symmetric guiding function $c = 0$ $\left(\psi_G^+ = \psi_G^- = \psi_S\right)$ is employed. The average meeting time is found to be much lower than in the non-symmetric case, $c = 4$, by nearly two orders of magnitude. This is true whether or not the stochastic processes are correlated. This behaviour of the meeting time as a function of c is not surprising. When c is large the two functions ψ_G^+ and ψ_G^- are localized in the nodal pockets of ψ_T. In the large-c limit ψ_G^+ is zero whenever ψ_T is negative, and ψ_G^- is zero whenever ψ_T is positive. In this limit the overlap between the two distributions ψ_G^+ and ψ_G^- is zero, and we have a similar result

Table 2. Average meeting times $\frac{\langle T \rangle}{N}$ for the correlated and uncorrelated cases in the symmetric case ($c = 0$, symmetric guiding function), equation (125). In this example, $x_{max} = 3$, equation (112), and $\tau = 0.9\,\tau_{max}$, where τ_{max} is the maximal time-step defined in equation (123).

Linear size N	$\frac{\langle T \rangle}{N}$ Uncorrelated cases	$\frac{\langle T \rangle}{N}$ Correlated cases
$N = 3$	3.32 (2)	1.634 (9)
$N = 5$	5.64 (6)	2.27 (1)
$N = 7$	7.6 (1)	2.65 (1.6)
$N = 9$	10.3 (1)	3.32 (2.5)
$N = 11$	12.95 (8)	4.18 (4)
$N = 13$	15.7 (1)	4.78 (4)
$N = 15$	18.5 (2)	5.52 (6)
$N = 17$	21.6 (2)	6.26 (7)

Table 3. $N = 3$, reduced Bose–Fermi gap $\tilde{\Delta}_{B-F}$, equation (136), with or without correlation for different values of c. The average meeting times are also indicated in parentheses. The bare Bose–Fermi gap, Δ_{B-F}, is ~ 0.7695 (here, $x_{max} = 3$. and $\tau = 0.09\,\tau_{max}$).

Value of c	Correlated process	Uncorrelated process
$c = 0$	$\tilde{\Delta}_{B-F} = 0.0366 \left[\frac{\langle T \rangle}{N} = 1.634\,(9) \right]$	$\tilde{\Delta}_{B-F} = 0.1629; \left[\frac{\langle T \rangle}{N} = 3.32\,(2) \right]$
$c = 1$	$\tilde{\Delta}_{B-F} = 0.0917 \left[\frac{\langle T \rangle}{N} = 6.37\,(7) \right]$	$\tilde{\Delta}_{B-F} = 0.2540; \left[\frac{\langle T \rangle}{N} = 16.5\,(2) \right]$
$c = 2$	$\tilde{\Delta}_{B-F} = 0.1336 \left[\frac{\langle T \rangle}{N} = 24.6\,(2) \right]$	$\tilde{\Delta}_{B-F} = 0.2277; \left[\frac{\langle T \rangle}{N} = 130.5\,(5) \right]$
$c = 3$	$\tilde{\Delta}_{B-F} = 0.1026 \left[\frac{\langle T \rangle}{N} = 62.9\,(3) \right]$	$\tilde{\Delta}_{B-F} = 0.1981; \left[\frac{\langle T \rangle}{N} = 627\,(3) \right]$
$c = 4$	$\tilde{\Delta}_{B-F} = 0.1092 \left[\frac{\langle T \rangle}{N} = 134\,(1) \right]$	$\tilde{\Delta}_{B-F} = 0.1787; \left[\frac{\langle T \rangle}{N} = 2152\,(23) \right]$

for the probability that walkers meet. From these preliminary results the introduction of non-symmetric wavefunctions seems to deteriorate the stability; this property will be confirmed in the following section.

6.3.3. Stability in time of FMC. We know from section 5 that the stability in time is directly related to the magnitude of the reduced Bose–Fermi energy gap:

$$\tilde{\Delta}_{B-F} \equiv E_0^F - \tilde{E}_0^B, \tag{136}$$

where $\tilde{E}_0^B$ is the lowest eigenvalue of the FMC diffusion operator. The greater this gap is, the faster the signal-over-noise ratio of the Monte Carlo simulation deteriorates; the full stability being obtained only when this gap vanishes. The ultimate goal of an efficient FMC algorithm is to reduce the Bose–Fermi gap from its bare value $\Delta_{B-F} = E_0^F - E_0^B$ to a value very close to zero (ideally, zero). Energies $\tilde{E}_0^B$ and E_0^F can be calculated by exact diagonalization of the fermion Monte Carlo operator, G_{FMC}. In practice, we have chosen here to extract this information from the large-time behaviour of the denominator of the energy, equation (134). In this regime the denominator behaves as in equation (78) where the reference energy is adjusted to keep the number of pairs constant, $E_T = \tilde{E}_0^B$:

$$\langle \mathcal{D}(t = k\tau) \rangle \propto e^{-(E_0^F - \tilde{E}_0^B)t}. \tag{137}$$

The gap $E_0^F - \tilde{E}_0^B$ has been extracted from the large-k values of $\mathcal{D}\,(t = k\tau)$, a quantity calculated deterministically by iterating the matrix G_{FMC}. The results for different values of c are reported in table 3. For both the correlated and uncorrelated processes it is found that the

Table 4. Comparison between the reduced Bose–Fermi gap, $\tilde{\Delta}_{\text{B–F}}$, and the bare Bose–Fermi gap, $\Delta_{\text{B–F}}$, in the symmetric case ($c = 0$, symmetric guiding function) as a function of N.

Value of N	$\tilde{\Delta}_{\text{B–F}}$	$\Delta_{\text{B–F}}$	Gap ratio
$N = 3$	$\tilde{\Delta}_{\text{B–F}} = 0.0366$	$\Delta_{\text{B–F}} = 0.7695$	$\frac{\tilde{\Delta}_{\text{B–F}}}{\Delta_{\text{B–F}}} = 0.0476$
$N = 5$	$\tilde{\Delta}_{\text{B–F}} = 0.0516$	$\Delta_{\text{B–F}} = 1.0195$	$\frac{\tilde{\Delta}_{\text{B–F}}}{\Delta_{\text{B–F}}} = 0.0506$
$N = 7$	$\tilde{\Delta}_{\text{B–F}} = 0.0577$	$\Delta_{\text{B–F}} = 1.1782$	$\frac{\tilde{\Delta}_{\text{B–F}}}{\Delta_{\text{B–F}}} = 0.0490$

gap increases with c. The minimal gap is obtained for $c = 0$, that is when the guiding functions are symmetric $\psi_G^+ = \psi_G^- = \psi_S$. This result is easily explained from the fact that there are two factors which favour the cancellation of walkers. First, as we have already seen, the average meeting time is minimal when $c = 0$ since, in this case, the overlap between the functions ψ_G^+ and ψ_G^- is maximal. Furthermore, a full cancellation between the walkers is precisely obtained when $c = 0$. In conclusion, the greatest stability is obtained for a symmetric guiding function.

In table 4 the gaps obtained at $c = 0$ for different linear sizes N are reported. This table shows that, in the limit of large grids, that is for a system close to the continuous model, the FMC algorithm reduces the gap by a factor $\sim$20. Such a result corresponds to a huge gain in the stability since projection times about twenty times larger than in a standard nodal release method can be used.

6.4. FMC using a finite number of walkers: the stochastic approach

In the previous section the fermion Monte Carlo method has been discussed and implemented by manipulating the exact fermion Monte Carlo diffusion operator without making reference to any stochastic aspect (as already mentioned it is formally equivalent to use an infinite number of walkers). Of course, for non-trivial systems it is not possible to propagate exactly the dynamics of the FMC operator. Accordingly, a finite population of walkers is introduced and specific stochastic rules allowing to simulate in average the action of the FMC operator are defined. Now, the important point is that in practice—like in any DMC method—one does not sample exactly the dynamics of the FMC operator because of the population control step needed to keep the finite number of walkers roughly constant [22–24]. This step introduces a small modification of the sampled diffusion operator which is at the origin of a systematic error known as the population control error. For a bosonic system, the error on the ground-state energy behaves as $\frac{1}{M}$ (M is the average size of the population) and an extrapolation in $\frac{1}{M}$ can be done to obtain the exact energy. For a fermionic system, as we shall see below, this behaviour is qualitatively different and, furthermore, depends on the guiding function used. To have a precise estimate of the mathematical behaviour of the population control error is fundamental since, in practice, it is essential to be able to reach the exact Fermi result using a reasonable number of walkers. As we shall see later, this will not be in general possible with FMC.

In this section, the fermion Monte Carlo simulations are performed using equations (128)–(133) which allow us to propagate stochastically a population of M walkers. The population is kept constant during the simulation by using the stochastic reconfiguration Monte Carlo (SRMC) method [23, 24]. In short, the SRMC method is a DMC method in which a reconfiguration step replaces the branching step. A reconfiguration step consists in drawing M new walkers among the M previous ones according to their respective Feynman–Kac weight (for the details, see the references given above).
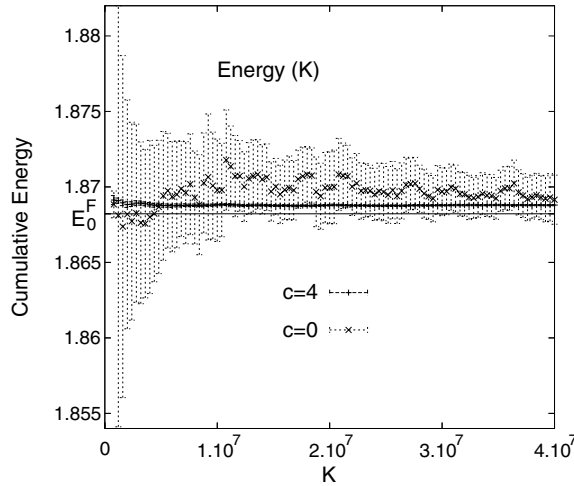
Figure 1. $N = 3$, energy estimator as a function of the projection time (number of iterations K). Comparison between $c = 0$ (large fluctuations) and $c = 4$ (small fluctuations). The exact energy is $E_0^F = 1.868\,22\ldots$. Number of walkers $M = 100$. Number of Monte Carlo steps: 4×10^7.

In figure 1 the time-averaged energy defined as

$$E(K) \equiv \frac{\sum_{k=1}^K \mathcal{N}(k)}{\sum_{k=1}^K \mathcal{D}(k)} \tag{138}$$

is plotted as a function of K. In this formula $\mathcal{N}(k)$ and $\mathcal{D}(k)$ represent the numerator and denominator at iteration k of the estimator (75) evaluated as an average over the population of pairs. In figure 2 we plot the time-averaged denominator given by

$$\mathcal{D}_K = \frac{1}{K} \sum_{k=1}^K \mathcal{D}(k). \tag{139}$$

The time dependence of this quantity is interesting since it can be used as a measure of the stability of the algorithm [18]. As we have shown above, the algorithm is stable only when the reduced Bose–Fermi energy gap, $\tilde{\Delta}_{B-F} = E_0^F - \tilde{E}_0^B$, is equal to zero. Equivalently, the denominator (139) must converge to a constant different from zero. In our simulations the number of walkers was chosen to be $M = 100$, a value which is much larger than the total number of states of the system (here, nine states). Of course, such a study is possible only for very simple systems. As seen in the figures 1, 2 and 3, the results obtained in the case $c = 0$ (symmetric guiding function) and $c \neq 0$ are qualitatively very different. Figure 1 shows that, within statistical error bars, there is no systematic error on the energy when a symmetric guiding function is used, $c = 0$. However, the price to pay is that the statistical fluctuations are very large. This point can be easily understood by looking at the behaviour of the denominator, figure 2. Indeed, this denominator vanishes at large times, thus indicating that the simulation is not stable. In sharp contrast, for $c = 4$ (non-symmetric guiding functions), the statistical fluctuations are much more smaller (by a factor of about 40) but a systematic error appears for the energy. Furthermore, the denominator plotted in figure 2 is seen to converge to a finite value. The stability observed in the case $c = 4$ seems to confirm the results of Kalos *et al* [18] for non-symmetric guiding functions ($c \neq 0$). However, the situation deserves a closer look. Indeed, the existence of this finite asymptotic value seems to be in contradiction with
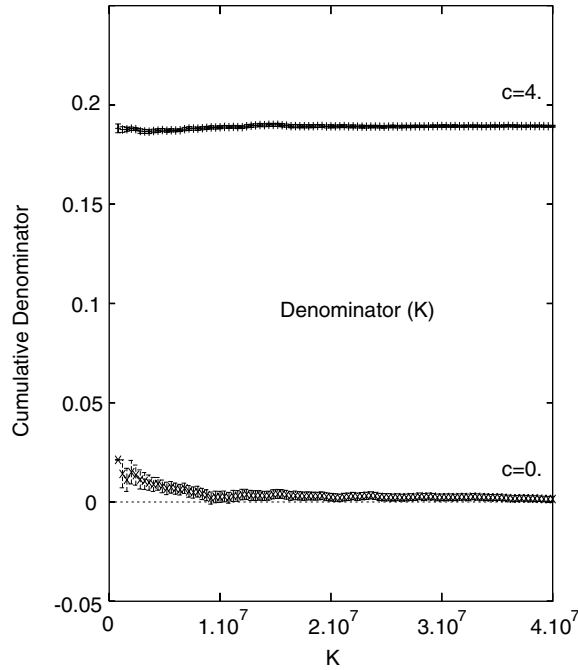
Figure 2. $N = 3$, denominator as a function of the projection time (number of iterations K). Comparison between $c = 4$ (upper curve) and $c = 0$ (lower curve). Number of walkers $M = 100$. Number of Monte Carlo steps: 4×10^7.

our theoretical analysis: the denominator should converge exponentially fast to zero, and the algorithm should not be stable ($\tilde{\Delta}_{B-F} > 0$). In fact, as we shall show now, the asymptotic value obtained for $c = 4$ and the corresponding stability results from a population control error. To illustrate this point, we have computed the average denominator as a function of the population size M. Results are reported in figure 3. On this plot, we compare the population dependence of the denominator (139) for $c = 4$ and for a much smaller value of $c = 0.5$. The values of M range from $M = 100$ to $M = 6400$. A first remark is that the population control error can be quite large and is much larger for $c = 4$ than for $c = 0.5$. In the appendix it is shown that the theoretical asymptotic behaviour of the error as a function of M is expected to be in $1/M$. In the $c = 0.5$ case, the denominator is clearly seen to extrapolate to zero like $\frac{1}{M}$. In the $c = 4$ case, we can just say that the data are compatible with such a behaviour; but even for the largest M reported in figure 3 ($M = 6400$) this asymptotic regime is not yet reached. Much larger populations would be necessary. This result illustrates the great difficulty in reaching the asymptotic regime, even for such a simple system having only nine states. Stated differently, the stability observed when using non-symmetric guiding functions disappear for a large number of walkers, thus confirming that the stability obtained at finite M is a control population artefact. Note that a large population control error on the denominator is not surprising. Indeed, when $c \neq 0$, the local energies of the guiding functions, equation (43), have strong fluctuations because $\psi_G^\pm$ is far from any eigenstate of H ($\psi_G^\pm$ contains a symmetric and an antisymmetric components). In the case of a symmetric guiding function ($c = 0$), the distribution of walkers is also symmetric at large times and, thus, the average of this distribution on the antisymmetric wavefunction ψ_T must necessarily be zero. Consequently, in the $c = 0$ case there is no control population error on the denominator, figure 2.
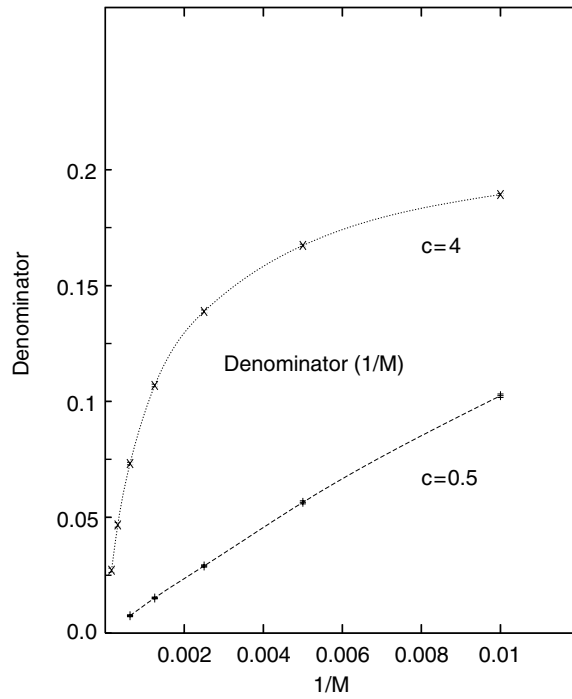
Figure 3. $N = 3$, time-averaged denominator, equation (139), as a function of $1/M$.
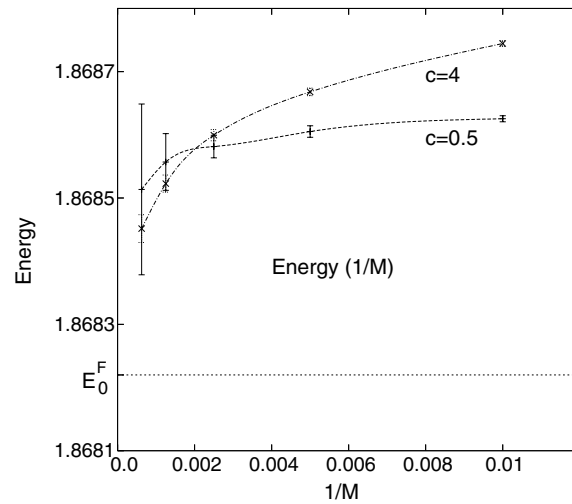
Figure 4. $N = 3$, energy, equation (134), as a function of $1/M$ for the $c = 0$ and $c = 4$ cases. Exact energy: $E_0^F = 1.868\,22\ldots$ (horizontal solid line).

An example of the behaviour of the energy as a function of the population size M is presented in figure 4. Some theoretical estimates of the energy bias dependence on M are derived in the appendix. Let us summarize the results obtained. When $c = 0$ (use of a

symmetric guiding function), the systematic error behaves as in a standard DMC calculation for a bosonic system:

$$\delta E \propto \frac{1}{M}. \tag{140}$$

However, the statistical fluctuations are exponentially large since the calculation is no longer stable. Now, when $c > 0$ the systematic error has a radically different behaviour. For a population size M the control population error grows exponentially as a function of the projection time t:

$$\delta E \propto \frac{1}{M} e^{t \tilde{\Delta}_{B-F}}, \tag{141}$$

where $\tilde{\Delta}_{B-F}$ is the reduced Bose–Fermi energy gap. This dependence of the control population error as a function of the projection time is of course pathological and is a direct consequence of the use of non-symmetric guiding functions. Now, because of the form (141) it is clear that a population size exponentially larger than the projection time is necessary to remove the systematic population control error. In practical calculations, for a given population of walkers M, one has to choose a finite projection time, t. This time has to be small enough to have a small finite population control error but, at the same time, large enough to extract the exact fermionic ground state from the initial distribution of walkers. The best compromise is easily calculated and leads to the following expression of the systematic error as a function of the number of walkers (for more details, see the appendix):

$$\delta E \propto \frac{1}{M^\gamma}, \tag{142}$$

where

$$\gamma \equiv \frac{\Delta_F}{\tilde{\Delta}_{B-F} + \Delta_F} < 1, \tag{143}$$

and Δ_F is the usual Fermi gap (energy difference between the two lowest fermionic states).

In figure 4 some numerical results for the $c = 4$ and $c = 0.5$ cases are presented. The numbers of walkers considered are $M = 100, 200, 400, 800$ and 1600. No data are shown for the symmetric case, $c = 0$, because of the very large error bars, see figure 1. The calculations have been done for the smallest system, $N = 3$ (recall that the finite configuration space consists of only nine states) and for very large numbers of Monte Carlo steps (more than 10^8). As it should be, the systematic errors are found to be larger for the $c = 4$ case than for the $c = 0.5$ case (note that the data corresponding to $M = 800$ and $M = 1600$ must not be considered because of their large statistical noise). The concavity of both curves confirms our theoretical result, $\gamma < 1$, equations (142) and (143). However, it is clear that getting a quantitative estimate of this exponent is hopeless because of the rapid increase of error bars as a function of M. The only qualitative conclusion which can be drawn by looking at the curves is that $\gamma_{c=4} < \gamma_{c=0.5}$, in agreement with our formula (143). Finally, let us insist on the fact that, despite these very intensive calculations for a nine-state configuration space, no controlled extrapolation to the exact energy is possible.

To summarize, when ψ_G has an antisymmetric component, the error is expected to decrease—for M large enough—very slowly as a function of the population size (algebraically with a (very) small exponent), while in the symmetric case the bias has a much more interesting $\frac{1}{M}$ behaviour. However, in this latter case the price to pay is the presence of an exponential growth of the statistical error. In both cases, and this is the fundamental point, the number of walkers needed to get a given accuracy grows pathologically. In addition, as illustrated by our

data for the very simple model problem treated here, the asymptotic regimes corresponding to the $1/M^\gamma$ behaviour appear to be very difficult to reach in practice (very large values of M are needed).

7. Conclusion and perspectives

The FMC method differs from the DMC method by correlating the diffusion of the walkers and introducing a cancellation procedure between the positive and the negative walkers whenever they meet. In this work we have shown that the fermion Monte Carlo approach is exact but, in general, not stable. FMC can be viewed as belonging to the class of transient DMC methods, the most famous one being probably the nodal release approach [14, 34]. However, in contrast with the standard transient methods, FMC allows us to reduce in a systematic way the Fermi instability. The importance of this instability is directly related to the magnitude of some 'effective' Bose–Fermi energy gap, $\tilde{\Delta}_{B-F} = E_0^F - \tilde{E}_0^B$, where E_0^F is the exact Fermi energy and $\tilde{E}_0^B$ some effective Bose energy. We have seen that this gap is intimately connected to the cancellation rate, that is to say, to the speed at which the positive and negative walkers cancel. We have shown that $E_0^B < \tilde{E}_0^B < E_0^F$, where E_0^B is the standard bosonic ground-state energy. As an important consequence, the closest the $\tilde{E}_0^B$ is from the exact fermionic energy, the smoother the sign problem is. For the toy model considered, the reduction obtained for the instability is very large (orders of magnitude). For large-dimensional systems, there are also strong indications in favour of an important reduction. A first argument is purely theoretical. In FMC, the walkers within a pair $\left(\mathbf{R}_i^+, \mathbf{R}_i^-\right)$ are correlated in such a way that $\mathbf{R}_i^+ - \mathbf{R}_i^-$ makes a random move only in one dimension. As a result there is a high probability that the walkers meet in a finite time even if they move in a high-dimensional space. The second argument is numerical. As shown by previous authors, the impact of correlating walkers on the average meeting time is important even for much larger systems [18, 21].

We have shown that the recent introduction of nonsymmetric guiding functions in FMC introduces a large systematic error which goes to zero very slowly as a function of the population size ($\sim 1/M^\gamma$, $\gamma = \Delta_F/(\tilde{\Delta}_{B-F} + \Delta_F)$ and $\Delta_F = E_1^F - E_0^F$ is the usual fermionic gap). For an infinite number of walkers, this systematic error is removed but the algorithm recovers the Fermi instability. Moreover, we have shown that using such guiding functions does not, in general, improve the stability. For a large enough number of walkers, the simulation can be less stable than the simulation using a symmetric guiding function. Finally, it is important to emphasize that the conclusion of this work is that the FMC algorithm is not a solution to the sign problem. However, it is a promising way towards 'improved' transient methods. As a transient method, FMC is expected to converge much better than a standard nodal release method. We are presently working in this direction.

Appendix. Population control error in FMC

FMC, like any Monte Carlo method using a branching process, suffers from a so-called population control bias. This systematic error appears because the branching rules (creation/annihilation of walkers) are implemented using a population consisting of a large but *finite* number of walkers. Nothing preventing the population size from imploding or exploding, a population control step is required to keep the average number of walkers finite. A standard strategy to cope with this difficulty consists in introducing a time-dependent reference energy whose effect is to slightly modify the elementary weights of each walker by a common multiplicative factor (close to 1) so that the total weight of the population remains nearly constant during the simulation. This step, which introduces some correlation between walkers and, therefore, slightly modifies the stationary density, must be performed very smoothly to keep the population control error as small as possible. In practice, for standard DMC calculations done with accurate trial wavefunctions and population sizes large enough, the error is found to be very small, in general, much smaller than the statistical fluctuations. As a consequence, the presence of a population control bias is usually not considered as critical. Here, the situation is rather different. In FMC, the use of non-symmetric guiding functions introduces very large fluctuations of the local energy and the cancellation rules a very small signal-over-noise ratio for fermionic properties. In this case, it is not clear whether the bias can be kept small with a reasonable number of walkers.

In this section, we present an estimate of the population control bias in FMC. As we shall see our estimate shows that the sign problem is actually not solved but attenuated in FMC (an *exponentially* large number of walkers is needed to maintain a constant bias as the number of electrons is increased). The derivation presented in this section is very general: it is valid for any exact fermion QMC method based on the use of a nodeless bosonic-type reference process and some projection to extract the Fermi ground state. Accordingly, we have chosen not to use the specific framework and notations of FMC but, instead, notations of a general DMC algorithm (transient method). Of course, we do not need FMC to introduce nonsymmetric guiding functions. The adaptation of what follows to FMC is straightforward.

In quantum Monte Carlo, we evaluate stochastically the following expression for the lowest eigenstate of energy E_0^{F}:

$$E_{\mathrm{F}}^t = \frac{\langle \psi_T | H \, \mathrm{e}^{-t(H-E_T)} f_0 \rangle}{\langle \psi_T | \mathrm{e}^{-t(H-E_T)} f_0 \rangle}, \tag{A.1}$$

where f_0 is some positive initial distribution and ψ_T an approximation of the eigenstate with energy E_0^{F}. Here, we deal with a fermionic problem so ψ_T must be antisymmetric. Expression (A.1) gives the exact energy E_0^{F} only when taking the limit $t \to \infty$. In practice for a finite t, there is a systematic error

$$\Delta E_{\mathrm{F}}^t \equiv E_{\mathrm{F}}^t - E_0^{\mathrm{F}} \propto \mathrm{e}^{-\Delta_{\mathrm{F}} t} \equiv \mathrm{e}^{-(E_1^{\mathrm{F}} - E_0^{\mathrm{F}})t}, \tag{A.2}$$

where E_1^{F} is the energy of the first excited state in the antisymmetric sector and Δ_{F} is the fermionic energy gap. For an exact algorithm with one walker (e.g. pure diffusion Monte Carlo [36, 37]) one computes the RHS of (A.1) by evaluating the following expression:

$$E_{\mathrm{F}}^t = \frac{\left\langle \frac{H\psi_T}{\psi_G}[R(t)] \, \mathrm{e}^{-\int_0^t ds(E_L[R(s)]-E_T)} \right\rangle}{\left\langle \frac{\psi_T}{\psi_G}[R(t)] \, \mathrm{e}^{-\int_0^t ds(E_L[R(s)]-E_T)} \right\rangle}, \tag{A.3}$$

where ψ_G is the guiding function, strictly positive, with eventually an antisymmetric component. We have also introduced in this expression, the local energy of the guiding function:

$$E_L = \frac{H\psi_G}{\psi_G}. \tag{A.4}$$

The integral in (A.3) is done over the drifted random walks going from 0 to t. To simplify the notations we note (A.3) as follows:

$$E_F^t = \frac{\langle h^t W^t \rangle}{\langle p^t W^t \rangle}, \tag{A.5}$$

where

$$h^t \equiv \frac{H\psi_T}{\psi_G}[R(t)] \tag{A.6}$$

$$W^t \equiv e^{-\int_0^t ds\, E_L[R(s)]} \tag{A.7}$$

$$p^t \equiv \frac{\psi_T}{\psi_G}[R(t)]. \tag{A.8}$$

For M-independent walkers R_i one has

$$E_F^t = \frac{\left\langle \frac{1}{M}\sum_i h_i^t W_i^t \right\rangle}{\left\langle \frac{1}{M}\sum_i p_i^t W_i^t \right\rangle}. \tag{A.9}$$

In the analysis presented here based on a population of M, the walkers branched according to their relative multiplicities (the M walkers are therefore no longer independent); one replaces the individual weight W_i by a global weight [24]:

$$\bar{W}^t = \frac{1}{M}\sum_i W_i^t. \tag{A.10}$$

As a result the energy may be written as

$$E_F^t = \frac{\langle \bar{h}^t \bar{W}^t \rangle}{\langle \bar{p}^t \bar{W}^t \rangle}, \tag{A.11}$$

where

$$\bar{h}^t \equiv \frac{1}{M}\sum_i h_i^t \tag{A.12}$$

$$\bar{p}^t \equiv \frac{1}{M}\sum_i p_i^t. \tag{A.13}$$

Expression (A.11) is exact when the weights $\bar{W}^t$ are included. A control population error arises when one does not take into account the weights in expression (A.11). This population control error is thus given by

$$\Delta E_F^M = \frac{\langle \bar{h}^t \rangle}{\langle \bar{p}^t \rangle} - \frac{\langle \bar{h}^t \bar{W}^t \rangle}{\langle \bar{p}^t \bar{W}^t \rangle} \tag{A.14}$$

$$= \frac{\langle \bar{h}^t \rangle}{\langle \bar{p}^t \rangle} - \frac{\mathrm{cov}(\bar{h}^t, \bar{W}^t) + \langle \bar{h}^t \rangle \langle \bar{W}^t \rangle}{\mathrm{cov}(\bar{p}^t, \bar{W}^t) + \langle \bar{p}^t \rangle \langle \bar{W}^t \rangle}, \tag{A.15}$$

or, after normalizing $\bar{W}^t$ in such a way that $\langle \bar{W}^t \rangle = 1$ (for example, by suitably adjusting the reference energy E_T),

$$\Delta E_{\mathrm{F}}^{\mathrm{M}} = \frac{\langle \bar{h}^t \rangle}{\langle \bar{p}^t \rangle} - \frac{\mathrm{cov}(\bar{h}^t, \bar{W}^t) + \langle \bar{h}^t \rangle}{\mathrm{cov}(\bar{p}^t, \bar{W}^t) + \langle \bar{p}^t \rangle}. \tag{A.16}$$

This is our basic formula expressing the systematic error at time t resulting from the use of a finite population. Now, let us evaluate this expression in the large time t and large M regimes. First, we consider the two denominators appearing in the RHS of equation (A.16). Let us begin with the denominator of the second ratio:

$$D_e \equiv \mathrm{cov}(\bar{p}^t, \bar{W}^t) + \langle \bar{p}^t \rangle. \tag{A.17}$$

Because this denominator is nothing but the denominator of the RHS of equation (A.3), we can conclude that D_e does not depend on the population size M and that it vanishes exponentially fast:

$$D_e = K_e \, \mathrm{e}^{-(E_0^{\mathrm{F}} - E_T)t}, \tag{A.18}$$

where E_T is the reference energy, $E_T = E_0^{\mathrm{B}}$ for a nodal release-type method, and $E_T = \tilde{E}_0^{\mathrm{B}} > E_0^{\mathrm{B}}$ for the FMC method. Let us now look at the other denominator of equation (A.16)

$$D_a \equiv \langle \bar{p}^t \rangle. \tag{A.19}$$

This denominator is the usual quantity evaluated during the simulation. It is an approximate quantity since it does not include the corrective weights. The asymptotic behaviour of D_a depends on ψ_G. We distinguish two cases as follows.

(i) If ψ_G is symmetric ($c = 0$), the stationary density (t large enough) is symmetric. Consequently, D_a, which is the average of an antisymmetric function, converges to zero exponentially fast at large times. For M large enough, in a regime where the dynamics is close to the exact dynamics of the Hamiltonian, we know that the convergence is given by

$$D_a = K_a(M) \, \mathrm{e}^{-\tilde{\Delta}_{\mathrm{B-F}}t}, \tag{A.20}$$

where the coefficient K_a depends on M in general. This coefficient will be determined later. From equations (A.20) and (A.18), one can evaluate the error on the denominator

$$D_a - D_e = (K_a(M) - K_e) \, \mathrm{e}^{-\tilde{\Delta}_{\mathrm{B-F}}t}. \tag{A.21}$$

We also know from the definitions of D_a and D_e ((A.19) and (A.17)) that the difference $D_a - D_e$ is a covariance of two averages:

$$D_a - D_e = \mathrm{cov}(\bar{p}^t, \bar{W}^t), \tag{A.22}$$

which, due to the central-limit theorem, behaves as

$$D_a - D_e = \frac{1}{M} \mathcal{C}(t), \tag{A.23}$$

where $\mathcal{C}(t)$ is some function of t. Identifying (A.21) and (A.23), one finally obtains a determination of K_a. Finally, we obtain the following behaviour for the systematic error on the denominator:

$$D_a - D_e = \mathrm{cov}(\bar{p}^t, \bar{W}^t) \propto \frac{1}{M} \, \mathrm{e}^{-\tilde{\Delta}_{\mathrm{B-F}}t}. \tag{A.24}$$

(ii) If ψ_G is not symmetric ($c \neq 0$), the stationary density has an antisymmetric component and D_a converges to a constant different from zero at large times. Of course, this constant depends on the number of walkers M. This dependence can be easily found by using the central limit theorem as before:

$$D_a - D_e = \text{cov}(\bar{p}^t, \bar{W}^t) = K \frac{1}{M}. \tag{A.25}$$

Finally, we have just proved that, when the guiding function is not symmetric, the asymptotic behaviour of the denominator is $D_a \propto \frac{1}{M}$ independent of t. This important result is in agreement with the numerical data shown in figure 3. Using exactly the same arguments, the asymptotic behaviour (large M, large t) of the difference of the two numerators in the RHS of expression (A.16) is found to be the same as $D_a - D_e$:

$$N_a - N_e = \text{cov}(\bar{h}^t, \bar{W}^t) \propto D_a - D_e. \tag{A.26}$$

Now, we are ready to write down the expression of the systematic population control error, $\Delta E_{\text{F}}^{\text{M}}$. For M large enough, $\Delta E_{\text{F}}^{\text{M}}$ is well approximated by its first-order contribution in the $\frac{1}{M}$ expansion. Here also, we need to distinguish between the nature of the guiding function.

(i) If ψ_G is symmetric one easily obtains

$$\Delta E_{\text{F}}^{\text{M}} \propto \frac{1}{M}. \tag{A.27}$$

(ii) If ψ_G has an antisymmetric component

$$\Delta E_{\text{F}}^{\text{M}} \propto \frac{1}{M} e^{\tilde{\Delta}_{\text{B–F}} t}. \tag{A.28}$$

Let us now evaluate the total systematic error resulting from using a finite time t and a finite population size M:

$$\Delta E_0^{\text{F}} = \Delta E_{\text{F}}^{\text{M}} + \Delta E_{\text{F}}^t, \tag{A.29}$$

where ΔE_{F}^t, the systematic error coming from a finite simulation time, is given by equation (A.2) and $\Delta E_{\text{F}}^{\text{M}}$ is the error just discussed. The strategy consists in determining, for a given systematic error $\Delta E_0^{\text{F}} \sim \epsilon$, what is the time t and the number of walkers M one should consider. The condition for the total systematic error to be of order ϵ is that both terms in (A.29) are also of order ϵ:

$$\Delta E_{\text{F}}^t \sim \epsilon \tag{A.30}$$

$$\Delta E_{\text{F}}^{\text{M}} \sim \epsilon. \tag{A.31}$$

This is true because no error compensation is present, ΔE_{F}^t and $\Delta E_{\text{F}}^{\text{M}}$ being generally of the same sign (both positive). Our numerical results on the toy model give an illustration of this property. From both equations (A.30) and (A.2) one can deduce the simulation time corresponding to such a systematic error:

$$t \sim -\frac{\ln \epsilon}{\Delta_{\text{F}}}. \tag{A.32}$$

In other words, to obtain an error of order ϵ it is sufficient to stop the simulation at a time t of order (A.32). Now, let us come to the number of walkers M needed. If ψ_G is symmetric, we already know from expression (A.27) that the systematic error does not depend on the projection time and that the number of walkers M and the systematic error ϵ are related as follows:

$$M \propto \frac{1}{\epsilon}. \tag{A.33}$$

If ψ_G is not symmetric, the equation (A.31) can be easily solved. Replacing in equation (A.31), ΔE_F^M by its expression (A.28) and using relation (A.32) one finally finds the number of walkers required to obtain a systematic error ϵ.

$$M \propto \epsilon^{-\frac{\tilde{\Delta}_{B-F}}{\Delta_F}-1}. \tag{A.34}$$

Let us make some important comments. First, note that in this formula the dependence on the guiding function is not included in the exponent, only in the prefactor. Second, this formula shows that the FMC algorithm reduces the systematic error by lowering the exponent. As already mentioned, the gap is equal to $E_0^F - E_0^B$ in a standard release node method and equal to $E_0^F - \tilde{E}_0^B < E_0^F - E_0^B$ in FMC. Third, in the zero-limit gap, the $\frac{1}{M}$ behaviour of the systematic error is recovered. This formula shows that the number of walkers needed for a given accuracy, ϵ, grows exponentially with respect to the number of electrons. Indeed, although the gap is indeed reduced by FMC, there is no reason not to believe that it will still be proportional to the number of electrons. In consequence, the 'sign problem' fully remains in FMC.

Finally, let us write the systematic error as a function of the finite population M by inverting the preceding equation (A.34):

$$\epsilon \propto M^{-\gamma}, \tag{A.35}$$

where

$$\gamma \equiv \frac{\Delta_F}{\tilde{\Delta}_{B-F} + \Delta_F}. \tag{A.36}$$

This latter equation illustrates the respective role played by the Fermi gap, Δ_F, and the reduced Bose–Fermi gap, $\tilde{\Delta}_{B-F}$.

References

[1] Kalos M H and Pederiva F 1999 *NATO-ASI Conference Quantum Monte Carlo Methods in Physics and Chemistry, Cornell* ed M P Nightingale and C J Umrigar NATO-ASI (Dordrecht: Kluwer)
[2] Ceperley D M and Kalos M 1979 *Monte Carlo Methods in Statistical Physics* (Berlin: Springer)
[3] Schmidt K E and Kalos M H 1984 *Applications of the Monte Carlo Method in Statistical Physics* vol 2 (Berlin: Springer)
[4] Hammond B L and Lester W A 1994 *Monte Carlo Methods in Ab Initio Quantum Chemisty* (Singapore: World Scientific)
[5] Alder P J R D C B J and Lester W A 1982 *J. Chem. Phys.* **77** 5593
[6] Anderson J B 1975 *J. Chem. Phys.* **63** 1499
[7] Anderson J B 1976 *J. Chem. Phys.* **65** 4121
[8] Filippi C and Umrigar C J 1996 *J. Chem. Phys.* **105** 213
[9] Healy S B, Filippi C, Kratzer P, Penev E and Scheffler M 2001 *Phys. Rev. Lett.* **87** 016105
[10] Filippi C, Healy S B, Kratzer P, Pehlke E and Scheffler M 2002 *Phys. Rev. Lett.* **89** 166102
[11] Kent P R C, Towler M D, Needs R J and Rajagopal G 2000 *Phys. Rev.* B **62** 15394
[12] Porter A R, Towler M D and Needs R J 2001 *Phys. Rev.* B **64** 035320
[13] Mitas L, Grossman J C, Stich I and Tobik J 2000 *Phys. Rev. Lett.* **84** 1479
[14] Ceperley D M and Alder B J 1984 *J. Chem. Phys.* **81** 5833
[15] Bernu B, Ceperley D and Lester W 1990 *J. Chem. Phys.* **93** 552
[16] Bernu B, Ceperley D and Lester W 1991 *J. Chem. Phys.* **95** 7782
[17] Caffarel M and Ceperley D 1992 *J. Chem. Phys.* **97** 8415
[18] Kalos M H and Pederiva F 2000 *Phys. Rev. Lett.* **85** 3547
[19] Arnow D, Kalos M H, Lee M A and Schmidt K E 1982 *J. Chem. Phys.* **77** 5562
[20] Liu Z, Zhang S and Kalos M H 1994 *Phys. Rev.* E **50** 3220
[21] Colletti L, Pederiva Y and Kalos M H 2005 *Fermion Monte Carlo Calculation of Liquid-3He*
[22] Umrigar C J, Nightingale M P and Runge K J 1993 *J. Chem. Phys.* **99** 2865
[23] Buonaura M C and Sorella S 1998 *Phys. Rev.* B **57** 11446
[24] Assaraf R, Caffarel M and Khelif A 2000 *Phys. Rev.* E **61** 4566

[25] Kalos M H, Levesque D and Verlet L 1974 *Phys. Rev.* A **9** 2178
[26] Langfelder P, Rothstein S M and Vrbik J 1997 *J. Chem. Phys.* **107** 8525
[27] Bos I and Rothstein S M 2004 *J. Chem. Phys.* **121** 4486
[28] Assaraf R and Caffarel M 2000 *J. Chem. Phys.* **113** 4028
[29] Filippi C and Umrigar C J 2000 *Phys. Rev.* B **61** R16291
[30] Caffarel M, Rérat M and Pouchan C 1993 *Phys. Rev.* A **47** 3704
[31] Assaraf R and Caffarel M 2003 *J. Chem. Phys.* **119** 10536
[32] Casalegno M, Mella M and Rappe A M 2003 *J. Chem. Phys.* **118** 7193
[33] Ceperley D M and Alder B J 1986 *Science* **231** 555
[34] Ceperley D M and Alder B J 1980 *Phys. Rev. Lett.* **45** 566
[35] Grossman J 2002 *J. Chem. Phys.* **117** 1434
[36] Caffarel M and Claverie P 1988 *J. Chem. Phys.* **88** 1088
[37] Caffarel M and Claverie P 1988 *J. Chem. Phys.* **88** 1100