Green function Monte Carlo with stochastic reconfiguration: An effective remedy for the sign problem

Sandro Sorella and Luca Capriotti
Istituto Nazionale di Fisica della Materia and International School for Advanced Studies, Via Beirut 4, 34013 Trieste, Italy
(Received 11 February 1999; revised manuscript received 10 June 1999)

A recent technique, proposed to alleviate the ''sign problem disease,'' is discussed in detail. As is well known, the ground state of a given Hamiltonian H can be obtained by applying the propagator $e^{-H\tau}$ to a trial wave function ψ_T and sampling statistically the state $\psi_\tau = e^{-H\tau}\psi_T$ for large imaginary time τ. However, the sign problem may appear in the simulation and such statistical propagation would be practically impossible without employing some approximation such as the ''fixed node'' (FN) one. The present method allows the improvement of the FN dynamics with a systematic correction scheme. This is possible by the simple requirement that, after a short imaginary time propagation via the FN Hamiltonian, a number p of correlation functions can be further constrained to be *exact* by small perturbations of the FN state, which is free from the sign problem. By iterating this procedure, the Monte Carlo average sign, which is almost zero when there is a sign problem, remains stable and finite even for large τ. The proposed algorithm is tested against exact diagonalization data available on finite lattices. It is also shown, in some test cases, that the dependence of the results upon the few parameters entering the stochastic technique can be very easily controlled, unless for exceptional cases.

I. INTRODUCTION

In the last few years, enormous progress in computational techniques has been accompanied by better and better performances of modern computers. All these developments have certainly contributed to the ''feeling'' that the many body problem of solving a strongly correlated Hamiltonian, with many electrons on a reasonably large system size, is becoming possible with some computational effort.

The various numerical methods proposed so far in order to find the ground state of a physically interesting Hamiltonian, can be classified into two main branches developing from two root methods: the exact diagonalization technique (ED) and the variational Monte Carlo method (VMC).

The first technique is a brute force diagonalization of the Hamiltonian matrix, which represents a prohibitive task for large systems as the linear dimension of this matrix grows exponentially with the number of electrons and the size. The use of spatial symmetries and the very efficient Lanczos technique have recently made possible the exact ground-state evaluation of up to ~ 30 electrons for simple lattice Hamiltonians like: the Heisenberg model,[1] the $t-J$ model,[2] the Hubbard model and related ones.[3] However, this is far from being enough for the determination of the physical properties in the thermodynamic limit. Recent progress has been made by using ED within the so called density-matrix renormalization group technique (DMRG),[4,5] which allows to obtain almost exact—at least in one dimension—large size ground-state properties.

The second branch of development starts from the VMC technique.[6] The VMC allows one to sample statistically a variational wave function $\psi_G(x)$, defined on a given basis set whose elements $\{x\}$ are represented by simple *configurations*, defined typically by the electron positions and spins. In the simplest formulation, the VMC sampling can be obtained by accepting a new trial configuration x_{n+1} from a given one x_n if a random number ξ between zero and one satisfies $\xi < |\psi_G(x_{n+1})/\psi_G(x_n)|^2$, otherwise the trial configuration is not accepted and $x_{n+1} = x_n$.

The iterative rule that determines a new configuration x_{n+1} starting from a previous one x_n, and depending also on a random number, defines a Markov chain that allows one to obtain statistical estimates of the variational expectation values. This is possible even if the dimension of the Hilbert space is very large, a property representing the most important advantage of the statistical methods over the ED technique.

From this point of view, the Green-function Monte Carlo (GFMC) technique[7] can be considered as a development of the VMC, because it allows to sample statistically the exact ground state of a many body Hamiltonian H, instead of being restricted to the variational wave function. In the GFMC, the ground state is statistically sampled by a set of M *walkers* (w_i, x_i), $i = 1, \ldots, M$, i.e., at each configuration x_i is associated a weight w_i in order to represent the amplitude of the wave function on the element x_i of the large (or even infinite) Hilbert space. A Markov chain–slightly more complicated than the variational one–can be easily defined for the GFMC as well. As it will be shown later on, the new configurations and weights $(w_i, x_i)_{n+1}$ depend only on the previous weights and configurations $(w_i, x_i)_n$ and M random numbers ξ_i. This iteration is equivalent, statistically, to a matrix-vector product

$$\psi_{n+1}(x') = \sum_x G_{x',x}\psi_n(x), \tag{1}$$

where $G_{x',x}$ is the lattice Green function, which is simply related to the Hamiltonian matrix in the given basis

$$G_{x',x} = \Lambda \delta_{x',x} - H_{x',x}, \tag{2}$$

where Λ is a suitable constant, allowing the convergence of Eq. (1) to the ground state of H for large n. At each Markov iteration n, the state $\psi_n(x)$ is sampled statistically by the walkers, whose number M can be large, but is typically a negligible fraction of the total Hilbert space dimension.

In the statistical iteration process the weights w_i of the walkers increase or decrease exponentially, so that after a few iterations most of the walkers have an irrelevant weight w and some kind of reconfiguration becomes necessary in order to avoid large statistical errors.[7] The process to eliminate the irrelevant walkers from the statistical sampling is called ''branching.'' This consists for instance in duplicating a walker with large w_i in two walkers with half the weights $w_i/2$ acting on the same configuration, or in dropping the walkers with too small weights. For long–and therefore more accurate–Markov chains, it is also necessary to control the number of walkers otherwise the simulation will exceed the maximum available memory or it will terminate for lack of walkers. This reconfiguration of the walker population introduces some amount of bias. Recently, a rigorous and simple way of working at finite number of walkers has been proposed,[8] which simplifies the GFMC technique by controlling and eventually eliminating the bias due to the finite walker population.

With a slight generalization of the previous technique it is also possible to alleviate the infamous ''sign problem,'' which occurs when the matrix elements of the lattice Green function $G_{x',x}$ are not always positive definite. In this case the iteration (1) can still have a statistical meaning at the price of having walkers with weights w_i, which are no longer restricted to be positive. It then happens that the average weight sign $\langle s\rangle_n = \langle \Sigma_{i=1}^M w_i\rangle_n / \langle \Sigma_{i=1}^M |w_i|\rangle_n$ at a Markov iteration n is exponentially decreasing with n, implying a dramatic decrease of the signal-to-noise ratio for all correlation functions. A remarkable improvement of the GFMC on a lattice was realized when the fixed node (FN) approximation, largely employed for fermions in the continuum, was extended to lattice Hamiltonians.[9] In this case, the ''dangerous'' negative off-diagonal elements of the Green function are neglected, and stable simulations with positive walker weights w_i can be performed at the price of having only an approximate solution of the ground-state wave function.

The Green-function Monte Carlo with Stochastic Reconfiguration[10] (GFMCSR) represents a successful attempt of improving over the FN, with a stable simulation without any sign problem instability. In this scheme, better and better approximations of the ground-state correlation functions may be obtained by performing controlled Markov chain simulations with average walker sign $\langle s\rangle_n$ very close to 1 for each iteration n. For the sake of simplicity we will restrict the forthcoming derivation to lattice Hamiltonians but the basic ideas can be straightforwardly extended to the continuum case. This method is based upon the simple requirement that after a few iterations of Eq. (1) via the approximate FN dynamics, a number p of correlation functions can be further constrained to be *exact* by appropriate small perturbations of the propagated FN state ψ_n^{eff}, which is free from the sign problem. By iterating this procedure the average sign remains stable even for large n and, in this limit, the method has the important property of being in principle *exact* if all possible correlation functions are included.

In the first five sections of this paper we will briefly review the basic steps of the GFMC for the general case when the sign problem affects the practical implementation of the algorithm. In Sec. VI, we will introduce the Stochastic Reconfiguration (SR) technique. In the remaining sections we will present the details of the algorithm and some test results, useful to understand the practical implementation, for an efficient and controlled improvement of the FN, even for large system sizes.

II. THE GFMC TECHNIQUE

From a general point of view, the ground state ψ_0 of a lattice Hamiltonian H can be obtained by iterating the well known power method Eqs. (1) and (2) so that $\psi_n \to \psi_0$ for large n, provided the initial state ψ_T at the first iteration of Eq. (1) ($\psi_n = \psi_T$ for $n=1$) is a trial state not orthogonal to the ground state ψ_0.

A stochastic approach is possible if one can sample statistically the matrix-vector iterations (1). This is particularly important since for large systems, only a few power iterations, at most, can be applied exactly.

Following Ref. 8, it is first convenient to define the basic element of the stochastic approach; the so called walker. A walker is determined by an index x corresponding to a given element $|x\rangle$ of the chosen basis and a weight w. Within the stochastic approach the walker ''walks'' in the Hilbert space of the matrix H and assumes a configuration w,x according to a given probability distribution $P(w,x)$. The task of the GFMC approach is to define a Markov chain, yielding a probability distribution $P_n(w,x)$ for the walker, which determines the iterated wave function ψ_n:

$$\psi_n(x) = \langle x|\psi_n\rangle = \int dw\, w P_n(w,x). \tag{3}$$

III. IMPORTANCE SAMPLING

One of the most important advantages of the GFMC technique is the possibility of reducing the variance of the energy by exploiting some information on the ground-state wave function, known a priori on physical grounds. Following Ceperley and Kalos,[11] one can consider in the iteration (1) not the original matrix G, but the slightly more involved nonsymmetric one

$$\bar{G}_{x',x} = \psi_G(x') G_{x',x} / \psi_G(x), \tag{4}$$

where ψ_G is the so called *guiding wave function*, that has to be as simple as possible to be efficiently implemented in the calculation of the matrix elements, and as close as possible to the ground state of H. Here and in the following, we assume that $\psi_G(x)$ is always non-vanishing for all x. It is obvious that $\bar{G}$, though nonsymmetric, has the same spectrum of G as for any eigenvector $\psi_k(x)$ of G with energy $\Lambda - E_k$, $\psi_G(x)\psi_k(x)$ is a right eigenvector of $\bar{G}$ with the same eigenvalue.

As shown later on, by sampling statistically the iteration (1) with $\bar{G}$ instead of G, the walkers (w,x) will be distributed for large n according to $\psi_0(x)\psi_G(x)$, namely $\psi_n(x)$

$\propto \psi_0(x)\psi_G(x)$ in Eq. (3). In order to evaluate the ground-state energy, it is then enough to average the so called local energy,

$$E_x = \frac{\langle \psi_G|H|x\rangle}{\langle \psi_G|x\rangle} = \sum_{x'} \psi_G(x') H_{x',x}/\psi_G(x), \tag{5}$$

over the statistically sampled walkers.

Analogously, after the transformation (4), all mixed average correlation functions

$$\frac{\langle \psi_G|O^k|\psi_0\rangle}{\langle \psi_G|\psi_0\rangle} \tag{6}$$

are easily accessible by GFMC for arbitrary linear operators O^k. The local estimator corresponding to Eq. (6) is, analogously to Eq. (5), given by

$$O^k_x = \sum_{x'} \overline{O}^k_{x',x}, \tag{7}$$

where

$$\overline{O}^k_{x',x} = \psi_G(x') O^k_{x',x}/\psi_G(x), \tag{8}$$

are the operator matrix elements transformed according to the guiding wave function. Summarizing, in order to implement the ''importance sampling'' strategy it is sufficient to replace all the original matrices $O^k_{x',x}$ and G with the transformed nonsymmetric matrices $\overline{O}^k$ (8) and $\overline{G}$ (4). In the following, for simplicity of notations, we put a bar over the symbols corresponding to all the transformed matrices (4) and (8). We finally remark that, since the convergence of the power method (1) is not limited to symmetric matrices, the GFMC method can be more generally considered an efficient tool to find the maximum eigenvalue and eigenvector of a generic matrix $\overline{G}$.

IV. SINGLE WALKER FORMULATION

In general the distribution $P_n(w,x)$ in Eq. (3) is sampled by a finite number M of walkers. Let us first consider the simpler case $M=1$. In order to define a statistical implementation of the matrix multiplication (1), the standard approach is to determine first the nonvanishing Green function matrix elements $\overline{G}_{x',x}$ for all $\{x'\}$. These matrix elements can be generally written in terms of three factors

$$\overline{G}_{x',x} = s_{x',x} p_{x',x} b_x, \tag{9}$$

where b_x is a positive normalization factor, $s_{x',x}$ takes into account the sign of the matrix element and $p_{x',x}$ is a stochastic matrix. All these terms will be defined explicitly below.

The basic step of the GFMC method on a lattice is to define properly the matrix $p_{x',x}$, because it represents the term in the decomposition (9) allowing to select statistically only *one* configuration among all the possible $\{x'\}$ connected to x. Therefore, $p_{x',x}$ has to represent a probability and is restricted to be (i) normalized $\Sigma_{x'} p_{x',x}=1$ and (ii) with all positive matrix elements $p_{x',x}\geqslant 0$. This is just the definition of a *stochastic matrix*.[12] Since the matrix elements of $\overline{G}$ are not restricted to be positive (sign problem) $p_{x',x}$ is more clearly defined in terms of an appropriate Green function $\overline{G}^{eff}$ with all positive matrix elements. Even if the latter restriction may appear rather strong, it is however possible that for large n the approximate propagation of the state ψ_n^{eff} by the Green function $\overline{G}^{eff}$ is not far, in a sense to be specified below, from the true propagation of ψ_n by the exact Green function $\overline{G}$ in Eq. (1). $\overline{G}^{eff}_{x',x}$ needs not to be normalized, as its normalization can be included in the definition of the positive constant

$$b_x = \sum_{x'} \overline{G}^{eff}_{x',x} \tag{10}$$

so that

$$\overline{G}^{eff}_{x',x} = p_{x',x} b_x. \tag{11}$$

Here, we follow a recent development of the FN method on a lattice,[13] and we choose for $\overline{G}^{eff}$ the FN Green function (with importance sampling):

$$\overline{G}^{eff}_{x',x} = \Lambda \delta_{x',x} - \overline{H}^{eff}_{x',x}. \tag{12}$$

The constant shift Λ has to be large enough that all the diagonal elements of $\overline{G}^{eff}$ are strictly positive. This is possible in general for the diagonal elements. If H^{eff} is appropriately defined,[13] one can prove that its ground state *is a variational state of H with an energy better than the guiding wave function one*. Here, we slightly modify this approach which neglects all the matrix elements of H crossing the nodes of the guiding wave function, namely the ones with $\overline{H}_{x',x}>0$, by defining a matrix element $\overline{H}^{eff}_{x',x}<0$ even when $\overline{H}_{x',x}>0$ (see below). The generalization of the above ''FN theorem'' to this case is straightforward and is reported in the Appendix A. Our experience has shown that it is extremely important to cross the nodes on a lattice within the above variational scheme. For instance, the simplest choice for a positive $\overline{G}^{eff}$, i.e., to take the absolute value of the exact Green function, $G^{eff}_{x',x}=|G_{x',x}|$, leads to much larger statistical errors and much less accurate results, especially for fermion systems.

More in detail, the definition of the $\overline{H}^{eff}$ we use is as follows. The off-diagonal matrix elements are given by

$$\overline{H}^{eff}_{x',x} = \begin{cases} \overline{H}_{x',x} & \text{if} \quad \overline{H}_{x',x}\leqslant 0 \\ -\gamma \overline{H}_{x',x} & \text{if} \quad \overline{H}_{x',x}>0, \end{cases} \tag{13}$$

where γ is positive constant, and the diagonal ones by

$$H^{eff}_{x,x} = H_{x,x} + (1+\gamma)\mathcal{V}_{\rm sf}(x), \tag{14}$$

where the *sign-flip* contribution is[13]

$$\mathcal{V}_{\rm sf}(x) = \sum_{\{\overline{H}_{x',x}>0,\ x'\neq x\}} \overline{H}_{x',x}. \tag{15}$$

Notice that there is no difference between the diagonal elements of the Hamiltonian H^{eff} (H) and the ones of the transformed matrix $\overline{H}^{eff}$ ($\overline{H}$), as defined by Eq. (8).

With these definitions, Eq. (9) for $\bar{G}$ holds if the factor $s_{x',x}$ is given by

$$s_{x',x}=\begin{cases} 1 & \text{if} \quad \bar{G}_{x',x}\geq 0 \\ -1/\gamma & \text{if} \quad \bar{G}_{x',x}<0 \\ \dfrac{\Lambda-H_{x,x}}{\Lambda-H^{eff}_{x,x}} & \text{if} \quad x'=x. \end{cases} \tag{16}$$

The appropriate stochastic process (*Markov iteration*) relative to the Hamiltonian H can be defined by the following three steps, where we simply allow the weight w of the walker to become also negative

(1) Given the walker (w,x), change the weight w by rescaling it with b_x as defined in Eq. (10):

$$w'=b_x w.$$

(2) Generate at random a new configuration x' according to the stochastic matrix $p_{x',x}$.

(3) Multiply the new weight w' by the sign factor $s_{x',x}$ as given by Eq. (16):

$$w'\rightarrow w' s_{x',x}. \tag{17}$$

Without the step (3), one is actually sampling the Hamiltonian H^{eff}, which we expect (or assume) to have a ground state close to the one of H, for suitably chosen guiding wave function. During the Markov iteration (17) it is straightforward therefore to update both the weight w associated to the true Hamiltonian, and the one w^{eff} associated to the approximate FN one H^{eff}. From now on therefore the walker will be characterized by the triad

$$(w,w^{eff},x).$$

The previous algorithm (17) allows us to define the evolution of the probability density for having a walker with weights w, and w^{eff} (>0), in the configuration x, namely:

$$P_{n+1}(w',w^{eff\,\prime},x')=\sum_x \frac{p_{x',x}}{b_x^2|s_{x',x}|} P_n\left(\frac{w'}{b_x s_{x',x}},\frac{w^{eff\,\prime}}{b_x},x\right). \tag{18}$$

The first moments of the distribution P over w and w^{eff} give the state $\psi_n(x)$ propagated with the exact Green function $\bar{G}$ and the state $\psi_n^{eff}(x)$ propagated with the FN Green function $\bar{G}^{eff}$, respectively. Indeed, by defining the propagated wave functions as

$$\psi_n(x)=\int dw^{eff}\int dw\, w\, P_n(w,w^{eff},x), \tag{19}$$

$$\psi_n^{eff}(x)=\int dw^{eff}\int dw\, w^{eff}\, P_n(w,w^{eff},x), \tag{20}$$

one can readily verify, using Eq. (18), that ψ_n and ψ_n^{eff} satisfy the iteration condition (1) with $\bar{G}$ and $\bar{G}^{eff}$, respectively.

At this stage the algorithm is exact, and the Markov iteration allows us to sample the ground state of H (with sign problem) and H^{eff} (without sign problem) within statistical errors: unfortunately these errors may be very large, and increasing with the iteration number n, especially when there is sign problem.

The configurations x_n that are generated in the Markov process are distributed, after many iterations, according to the maximum right eigenstate of the matrix $p_{x',x}$ (as only the matrix p is effective in the matrix product (1), if we neglect the weights of the walkers). This state is in general different from the state $\psi_G(x)\psi_0(x)$ we are interested in. So after many iterations the sampled configurations x_n are distributed according to an approximate state, but we can consider this state as a trial state ψ_T for the initial iteration $(n=1)$ in the power method (1). At any Markov iteration n, we can compute the weight of the walker assuming that L iterations before its value was simply $w=1$. In this way, it is simple to compute the resulting weight of the walker after L power Green function $\bar{G}$ applications

$$G_n^L=\prod_{j=1}^{L} b_{x_{n-j}} s_{x_{n-j+1},x_{n-j}}. \tag{21}$$

Therefore, for instance, in order to compute the energy with a single Markov chain of many iterations, the following quantity is usually sampled

$$E_0=\frac{\Sigma_n E_{x_n} G_n^L}{\Sigma_n G_n^L}, \tag{22}$$

with L fixed.[12,8]

This would conclude the GFMC scheme, if averages over the weight variable G_n^L were possible in a stable and controlled manner. However, there are two important drawbacks for the single walker formulation. The first one arises because the weight G_n^L of the walker grows exponentially with L–simply as a result of the L independent products in Eq. (21)–and can assume very large values, implying diverging variances in the above averages. This problem has a very well-established solution by generalizing the GFMC to many walkers and introducing a scheme (branching) that enables to carry out walkers with reasonable values of the weights, by dropping the irrelevant walkers with small weights and splitting the ones with large weights. Recently a simple formulation of this scheme was defined at fixed number of walkers[8] in a way that allows to control efficiently the residual bias related to the finite walker population. The second drawback is the more difficult one, and is due to the sign problem. The average sign $\langle s_L\rangle=\Sigma_n G_n^L/\Sigma_n|G_n^L|$ vanishes exponentially with L. In the formulation of Ref. 8 this problem looks quite similar to the first simple one. As we will see later on, some kind of remedy can be defined by a simple generalization of the SR, which is exact in the case with no sign problem.

V. CARRYING MANY CONFIGURATIONS SIMULTANEOUSLY

Given M walkers we denote the corresponding configurations and weights with two vectors $(\underline{w},\underline{x})$, where each vector component (w_i,w_i^{eff},x_i) $i=1,\ldots,M$, corresponds to the

ith walker. Following Ref. 8, it is easy to generalize Eq. (18) to many walkers by the corresponding probability $P_n(\underline{w},\underline{x})$ of having the M walkers with weights and configurations $(\underline{w},\underline{x})$ at the iteration n. Similarly to the single walker formulation, the propagated wave functions $\psi_n(x)$ and $\psi_n^{eff}(x)$ with the true Green function $\overline{G}$ and the approximate one $\overline{G}^{eff}$ read

$$\psi_n(x)=\int [d\underline{w}]\sum_{\underline{x}} \frac{\Sigma_j w_j \delta_{x,x_j}}{M} P_n(\underline{w},\underline{x})$$

$$\psi_n^{eff}(x)=\int [d\underline{w}]\sum_{\underline{x}} \frac{\Sigma_j w_j^{eff} \delta_{x,x_j}}{M} P_n(\underline{w},\underline{x}), \tag{23}$$

where the symbol $\int[d\underline{w}]$ indicates the $2M$ multidimensional integral over the (w_i, w_i^{eff}) variables $i=1,\ldots,M$ ranging from $-\infty$ to ∞ and from 0 to ∞, respectively. Equation (23) shows that the propagated quantum-mechanical states ψ_n and ψ_n^{eff}, which are sampled statistically, do not uniquely determine the walker probability function $P_n(\underline{w},\underline{x})$. In particular, it is perfectly possible to define a statistical process, the SR, which changes the probability distribution P_n without changing the *exact* information content, i.e., ψ_n and ψ_n^{eff}. In this way a linear transformation of P_n, described by a simple kernel $X(\underline{w}',\underline{x}';\underline{w},\underline{x})$, will be explicitly given

$$P_n'(\underline{w}',\underline{x}')=\int [d\underline{w}]\sum_{\underline{x}} X(\underline{w}',\underline{x}';\underline{w},\underline{x}) P_n(\underline{w},\underline{x}). \tag{24}$$

When there is no sign problem ($w^{eff}=w$) it is possible to define the kernel X (Ref. 8) by requiring that the weights w_j' are all equal to $\Sigma_j w_j/M$ after the SR. In this case, the algorithm is exact, and allows to perform stable simulations by applying the SR each few k_p iterations. Furthermore, by increasing the number of walkers M, the exponential growth in the variance of the weights w_j can always be reduced and systematically controlled. In fact, for large enough M, it is possible to work with L sufficiently large ($L \propto M$) obtaining results already converged in the power method iteration (1), and with small error bars.

VI. STOCHASTIC RECONFIGURATION, STABILIZATION OF THE SIGN PROBLEM

In order to avoid the sign problem instability, at least in an approximate way, we can follow the previous scheme by using the following kernel X that defines the SR (24)

$$X(\underline{w}',\underline{x}';\underline{w},\underline{x})=\prod_{i=1}^{M}\left(\frac{\Sigma_j |p_{x_j}|\,\delta_{x_i',x_j}}{\sum_j |p_{x_j}|}\right) \times \delta\left(w_i'-\beta^{-1}\frac{\Sigma_j w_j}{M}\operatorname{sgn} p_{x_i'}\right) \times \delta(w_i^{eff\prime}-|w_i'|), \tag{25}$$

where the coefficients p_{x_j} will be defined in the following, and $\beta=\Sigma_j p_{x_j}/\Sigma_j |p_{x_j}|$. The kernel (25) has a particularly simple form since the outcoming variables x_j' and w_j' are completely independent for different j values. In particular, it is possible to integrate easily each of the M factors of the kernel in the variables w_j', $w_j^{eff\,\prime}$ and to sum over the configurations x_j', the result being simply one, as it is required by the normalization condition of the probability density P' in Eq. (24). In general, the SR defines new states $\psi_n'(x)$ and $\psi_n'^{\,eff}(x)$ from the given states $\psi_n(x)$ and $\psi_n^{eff}(x)$ at the given Markov iteration n. The new states $\psi_n'(x)$ and $\psi_n'^{\,eff}(x)$ are simply obtained by replacing P with P' in Eq. (23). The SR will be exact if it does not affect the evolution of the state $\psi_n(x)$ during the Markov chain, namely whenever

$$\psi_n'(x)=\psi_n(x). \tag{26}$$

In the SR, the new configurations x_i' are taken randomly among the old ones $\{x_j\}$, according to the probability $|p_{x_i}|/\Sigma_j|p_{x_j}|$, defined below in terms of the given weights $\{w_j\}$, $\{w_j^{eff}\}$ and configurations $\{x_j\}$. Moreover, the weights w_i' are changed consistently to Eq. (25) by $w_i'=\beta^{-1}\Sigma_j w_j/M \operatorname{sgn} p_{x_i'}$ and the FN weights, restricted to be positive, are defined by taking their absolute values $w_i^{eff\prime}=|w_i'|$.

The coefficient $\beta=\Sigma_j p_{x_j}/\Sigma_j|p_{x_j}|$ guarantees that the normalization is preserved by the SR, namely $\Sigma_x \psi_n'(x)=\Sigma_x \psi_n(x)$. This coefficient β represents also the expected average walker sign $\langle s\rangle'=\Sigma_j w_j'/\Sigma_j|w_j'|$ after the reconfiguration. It is supposed to be much higher than the average sign before the reconfiguration $\langle s\rangle=\Sigma_j w_j/\Sigma_j|w_j|$, so that a stable simulation with approximately constant average sign $\langle s\rangle'$ can be obtained by iteratively applying the SR every few k_p steps of the power method iteration (1).

In the actual implementation of this algorithm (see Sec. VII for the details) the weights are reset to unit values after the SR: $w_i'=\operatorname{sgn} p_{x_i'}$ and $w_i^{eff\prime}=1$, whereas only the overall constant $\beta^{-1}\Sigma_j w_j/M$, common to all the different walkers, is stored in a sequential file. As in the single walker formulation we can assume that, at any given iteration n, L iterations before the trial state ψ_T is given by the equilibrium distribution of walkers with unit weights $w_j=\operatorname{sgn} p_{x_j}$. Therefore, in order to obtain the weights predicted by the Eq. (25) for L power method iterations starting from ψ_T it is enough to multiply the previous L/k_p saved factors $f_n=\beta^{-1}\Sigma_j w_j/M$. This yields a natural extension of the factors G_n^L in Eq. (21) to the many walker case

$$G_n^L=\prod_{k=1}^{L/k_p} f_{n-k\times k_p} \tag{27}$$

and the corresponding mixed average correlation functions are obtained by averaging the local estimators over all the iterations n just before the SR (i.e., n is a multiple of k_p)

$$\langle O^k\rangle=\frac{\Sigma_n G_n^L \Sigma_j w_j O_{x_j}^k}{\Sigma_n G_n^L \Sigma_j w_j}, \tag{28}$$

where, in the above equation, the weights w_j and the local estimators $O_{x_j}^k$ are evaluated only before the SR.

A. Choice of the coefficients p_{x_j}

The only quantity which we still need to define properly the whole algorithm, in the kernel of Eq. (25) are the coefficients p_{x_j} which *have not* to be assumed positive. These coefficients may depend on all the weights w_j, the configurations x_j and the FN weights w_j^{eff}.

The choice $p_{x_j}=w_j$ is exact in the sense that $\psi_n'(x)=\psi_n(x)$, and coincides with the one for the case without sign problem.[8] However, this choice is obviously not convenient, because this reconfiguration will not improve the sign, which will decay exponentially in the same way.

Instead, in the case with sign problem, we can parameterize the coefficients p_{x_j} by assuming they are close enough to the positive definite weights $\{w_j^{eff}\}$, the ones obtained with the FN Green function G^{eff}. The reason for this choice is that, though the weights w_j^{eff} may be occasionally very different from the exact weights w_j–namely their sign can be wrong–they sample a state $\psi_n^{eff}(x)$, which is supposed to be quite close to the exact propagated state $\psi_n(x)$. This condition is clearly verified for an appropriate choice of the guiding wave function ψ_G, making the FN accurate. Then, we assume that small perturbations over the state $\psi_n^{eff}(x)$ may lead to fulfill the equality (26) with an arbitrarily small error. This error will affect the equilibrium walker distribution P_n for large n, but there will be no problem if this error (i) is small and (ii) can be reduced within the desired accuracy.

In the simplest and most practical formulation we require that only the average energy before and after the SR coincide

$$\sum_{x',x} \overline{H}_{x',x}\psi_n(x)=\sum_{x',x} \overline{H}_{x',x}\psi_n'(x) \tag{29}$$

[the denominators in the mixed averages (6) are already equal by definition, as $\Sigma_x\psi_n(x)=\Sigma_x\psi_n'(x)$ for the chosen β in Eq. (25)]. Then, we define

$$p_{x_j}=w_j^{eff}[1+\alpha(E_{x_j}-\overline{E}_{eff})]$$

and

$$\overline{E}_{eff}=\frac{\Sigma_j w_j^{eff}E_{x_j}}{\Sigma_j w_j^{eff}}$$

$$\overline{E}=\frac{\Sigma_j w_j E_{x_j}}{\Sigma_j w_j}, \tag{30}$$

where E_{x_j} is the local energy (5) associated to the configuration x_j. Thus, $\overline{E}$ represents the estimate of the average energy correctly sampled with the sign, whereas $\overline{E}_{eff}$ is the corresponding FN one. In order to satisfy the requirement (29) we just determine α by

$$\alpha=\frac{\overline{E}-\overline{E}_{eff}}{\overline{E}^2_{eff}-(\overline{E}_{eff})^2}, \tag{31}$$

where $\overline{E}^2_{eff}=\Sigma_j w_j^{eff}E^2_{x_j}/\Sigma_j w_j^{eff}$ is the average square energy over the positive weights w_j^{eff}.

A simple calculation shows that with this reconfiguration, that clearly improves the sign, the value of the energy (the mixed average energy) remains statistically the same before and after the SR (see Appendices B and C). It is clear, however, that this is not enough to guarantee convergence to the exact ground state, because fulfillment of Eq. (29) does not imply the exact equality (26). We can improve the definition of the constants p_{x_j} by including an arbitrary number p of parameters with, $p \ll M$,

$$p_{x_j}=w_j^{eff}[1+\alpha_1(O^1_{x_j}-\overline{O}^1_{eff})+\cdots+\alpha_p(O^p_{x_j}-\overline{O}^p_{eff})] \tag{32}$$

proportional to the fluctuations $O^k_{x_j}-\overline{O}^k_{eff}$ of p different operators O^k with corresponding local estimators $O^k_{x_j}=\langle\psi_G|O^k|x_j\rangle/\langle\psi_G|x_j\rangle$ $(k=1,\cdots,p)$, and average value over the positive weights $\overline{O}^k_{eff}=\Sigma_j w_j^{eff}O^k_{x_j}/\Sigma w_j^{eff}$. With the more general form (32) for the coefficients p_{x_j} it is possible to fulfill that all the mixed averages for the chosen p operators have the same value before and after the SR

$$\sum_{x',x} \overline{O}^k_{x',x}\psi_n(x)=\sum_{x',x} \overline{O}^k_{x',x}\psi_n'(x). \tag{33}$$

In general, the reference weights w_j^{eff} in Eq. (32) may be also different from the ones generated by the FN Green function, the only restriction is that $w_j^{eff}>0$ for each walker j (see Appendix C 2).

It can be proven that, in order to fulfill exactly the SR conditions (33), it is *sufficient* that the coefficients p_{x_j} are chosen in a way that

$$\frac{\Sigma_j p_{x_j}O^k_{x_j}}{\Sigma_j p_{x_j}}=\frac{\Sigma_j w_j O^k_{x_j}}{\Sigma_j w_j}, \tag{34}$$

which can be fulfilled with a solution of a simple linear system for the unknown variables α_k, for $k=1,\ldots,p$, as described in the Appendix C. The conditions (34) are much simpler to handle, because they can be satisfied at a given iteration of the Markov process. A theorem, proven in Appendix B, guarantees indeed that the exact conditions (33) are implied by the constraints (34) after the complete statistical average over the walker probability distribution P_n.

B. Proof of the asymptotical convergence of the GFMCSR to the exact result

Asymptotically, by adding more and more parameters $\{\alpha_j\}$, we can achieve $\psi'_n(x)=\psi_n(x)$ strictly, since the distribution $\psi_n(x)$ is completely determined by its correlation functions. The proof of this important statement is very simple. Consider first the diagonal operators. All these operators may be written as linear combinations of the ''elementary'' ones $O^{x_0}_{x',x}=\delta_{x',x}\delta_{x,x_0}$ acting on a single configuration x_0, plus at most some constants. If conditions (33) are satisfied for *all* the elementary operators O^{x_0} it immediately follows that $\psi_n'(x_0)=\psi_n(x_0)$ for all x_0, which is the exact SR condition (26). Then it is simple to show that the coefficients p_{x_j}, determining P_n' and ψ_n', are invariant for any

constant shift of the operators O^k. Furthermore with a little algebra it turns out that these coefficients p_{x_j} do not change for any arbitrary linear transformation of the chosen operator set: $O^{k\prime}=\Sigma_k L_{k',k} O^k$ (with real L and $\det L \neq 0$) (see Appendix C 1). Thus the proven convergence of the GFMCSR is obtained for any sequence of diagonal operators, that, with increasing p, becomes complete. For nondiagonal operators $O_{x',x}$ we simply note that they assume the same mixed average values of the equivalent diagonal ones $O^{diag}_{x',x} = \delta_{x',x}\Sigma_{x'} O_{x',x}$. Thus, the proof that GFMCSR converges in principle to the exact solution is valid in general even when nondiagonal operators, such as the Hamiltonian itself, are included in the conditions (33).

VII. DETAILS OF THE ALGORITHM

In this section the flow chart of the GFMCSR algorithm is briefly sketched. As described in Appendix D, it is possible to work without the extra constant shift Λ and apply directly $e^{-H\tau}$, the usual imaginary time propagator, to filter out the ground state from the chosen trial wave function ψ_T.

For practical purposes, the algorithm can be divided into three steps: (1) the Green function (GF) evolution, (2) the SR, and (3) the measurements of physical mixed average correlation functions. These three steps are iterated until a satisfactory statistical accuracy is obtained for the latter quantities.

The algorithm works with a finite number M of walkers, which is kept fixed. Starting from the first walker ($j=1$), the basic steps of the algorithm are described below:

1. In the GF evolution, the exact propagator $e^{-H\Delta\tau}$ and the FN one $e^{-H^{eff}\Delta\tau}$ are applied statistically for a given imaginary time interval $\Delta\tau$. In practice this can be done by setting initially $\Delta\tau_l=\Delta\tau$ and repeating the following steps until $\Delta\tau_l>0$:

(a) Given the configuration of the walker, x_j, the quantities E_{x_j}, $\mathcal{V}_{sf}(x_j)$ and $H^{eff}_{x_j,x_j}$ Eqs. (5), (15), and (14) are evaluated. Then the interval $\Delta\tau_d$ during which the walker is expected to perform only diagonal moves (see Appendix D) is computed using the relation $\Delta\tau_d=\min(\Delta\tau_l, \ln\xi/\pi_d)$, where ξ is a random number between 0 and 1 and $\pi_d = \lim_{\Lambda\to\infty}\Lambda \ln p_d = E_{x_j} - H^{eff}_{x_j,x_j}$ according to Eq. (D1).

(b) $\Delta\tau_l$ is updated $\Delta\tau_l \to \Delta\tau_l - \Delta\tau_d$ and the walker weights (w_j, w_j^{eff}) are multiplied respectively by $e^{(-E_{x_j}+(1+\gamma)V_{sf}(x_j))\Delta\tau_d}$ and $e^{-E_{x_j}\Delta\tau_d}$. Then if $\Delta\tau_l>0$ a new configuration $x_j' \neq x_j$ is chosen according to the probability table defined only by the normalized off-diagonal matrix elements of p_{x',x_j},

$$\frac{p_{x',x_j}}{\sum_{x'\neq x_j} p_{x',x_j}},$$

and the weight w_j is multiplied by s_{x_j',x_j} (16). The GF evolution then restarts from (a). Otherwise, if $\Delta\tau_l=0$ the GF evolution for the walker j terminates and the algorithm proceeds for the next walker starting from step (1).

2. After all the walkers (w_j, w_j^{eff}, x_j) have been propagated for the total imaginary time interval $\Delta\tau$, the SR can be applied. The mixed averages $O^k_{x_j} = \langle\psi_G|O|x_j\rangle / \langle\psi_G|x_j\rangle$ are computed at the end of such propagation for the chosen set of operators O^k. With these quantities both $\overline{O}^k_{eff} = \Sigma_j w_j^{eff} O^k_{x_j} / \Sigma_j w_j^{eff}$ and the covariance matrix $s_{k,k'}$ in Eq. (C4) are evaluated. By using the latter quantities in the linear system (C3), the coefficients α_k are computed and the table p_{x_j} is determined according to Eq. (C2). At this stage the reconfiguration procedure for the walkers can finally be performed, i.e., the new M configurations of the walkers are chosen among the old ones according to the probability $|p_{x_j}|/\Sigma_k|p_{x_k}|$.

(3) The mixed averages of the physical observables O^k_j and the quantity

$$\frac{\Sigma_k w_k}{M}\frac{\Sigma_k |p_{x_k}|}{\Sigma_k p_{x_k}},$$

needed for the calculation of the statistical averages, are stored. The walker weights are set to $w_j = \mathrm{sgn}\, p_{x_j}$ and $w_j^{eff}=1$, and the GF evolution can continue from step (1), starting again from the first walker.

In the practical implementation of the algorithm the FN dynamic can be worked out at fixed γ, where γ has to be a non-zero number otherwise the exact GF would not be sampled [see Eqs. (13,14)]. On the other hand the FN is more accurate for $\gamma=0$. A good compromise is to work with $\gamma=0.5$ fixed. An alternative choice is to implement a few runs with different nonzero γ, and to extrapolate the results for $\gamma=0$, which should represent the most accurate calculation. Typically, this extra effort is not necessary because there is a very weak dependence of the results upon γ. However, the extrapolation to $\gamma\to 0$ is an interesting possibility for the extension of the method to continuum models, since, in this case, there is no practical way of crossing the nodes with a variational FN approach (see Appendix A).

VIII. THE LIMIT OF SMALL $\Delta\tau$ AND LARGE NUMBER OF WALKERS

In this section some general properties of the GFMCSR technique are discussed and explicitly tested on the J_1-J_2 Heisenberg model

$$H = J_1 \sum_{\langle i,j\rangle} \mathbf{S}_i\cdot\mathbf{S}_j + J_2 \sum_{\langle\langle i,j\rangle\rangle} \mathbf{S}_i\cdot\mathbf{S}_j, \tag{35}$$

where $\mathbf{S}_i$ are the s-1/2 operators sitting on the sites of a square lattice. $J_1=1$ and $J_2=0.5$ are the antiferromagnetic super-exchange couplings between nearest- and next-nearest-neighbor pairs of spins, respectively. For the chosen values of the parameters of the Hamiltonian the GS of the model is likely to have no magnetic long range order.[10] In the following we will consider finite square clusters of N sites with periodic boundary conditions. We use the same guiding wave function of Ref. 10 and report here some test results useful to understand the crucial dependence of GFMCSR on the number of walkers M and the distance in imaginary time $\Delta\tau$ between two consecutive SR. In fact, after the selection

of a given number p of correlation functions in Eq. (33), the results depend only on the number of walkers M and the frequency of reconfiguration $\Delta\tau$. In the limit of large number of walkers, at fixed p, the algorithm has the important property that the fluctuations of the coefficients α_k and $\overline{O^k}$ in Eq. (32) are obviously vanishing, because they depend on ''averages'' of a very large number of samples of many different walkers, implying that these fluctuations are decreasing with $1/\sqrt{M}$. In this limit it is possible to recover an important property of the FN: *if the guiding wave function is exact, the FN averages are also exact*. In fact suppose we begin to apply the propagator $e^{-H\tau}$ starting at $\tau=0$ from the exact sampling of the ground state ψ_0 determined by FN with the exact guiding wave function $\psi_G=\psi_0$. Then at any Markov iteration n, before the SR is applied, both the mixed average correlation functions calculated with the FN weights w^{eff} ($\langle O^k\rangle=\Sigma_j w_j^{eff} O^k_{x_j}/\Sigma_j w_j^{eff}$) and the weights with arbitrary signs w ($\langle O^k\rangle=\Sigma_j w_j O^k_{x_j}/\Sigma_j w_j$) sample statistically the true quantum average $\langle\psi_0|O^k|\psi_0\rangle$. If, for large M, we can neglect statistical fluctuations of these averages, then by Eq. (34) $\alpha_k=0$ and the SR algorithm just replace the weights w_j (with sign problem) with the FN weights w_j^{eff}, which also sample ψ_0 exactly if $\psi_G=\psi_0$. This means that the SR approach does not affect this important property of the FN approach, at least in the limit $M\to\infty$.

Another reason to work in the limit $M\to\infty$ is the following. In this limit it is not necessary to include in the SR conditions (34) operators O^k that vanish for some symmetry that is satisfied both by the true Hamiltonian $\overline{H}$ and the FN one $\overline{H}^{eff}$. In fact, if the coefficients p_{x_j} are defined in terms of operators O^k that conserve the above mentioned symmetries (e.g., translation invariance, rotation by 90° degree of the lattice, etc.) by definition Eq. (33) is satisfied for all the remaining nonsymmetric operators, which have vanishing expectation value due to symmetry constraints (such as, e.g., an operator that changes sign for a rotation operation which is a symmetry of $\overline{H}$ and $\overline{H}^{eff}$). In this case, both sides of Eq. (33) are zero by such symmetry constraints. Moreover, for $M\to\infty$ the statistical fluctuations are negligible and for the same reason Eq. (34) is also automatically satisfied with vanishing α_k for the above mentioned nonsymmetric operators. In this limit, it is therefore useless to include nonsymmetric operators in the SR (34).

Finally, it is interesting that in this important limit $M\to\infty$, within the assumption that we can neglect the fluctuations of α_k and $\overline{O}^k_{eff}$, the SR depends only on the propagated states $\psi_n^{eff}(x)$ and $\psi_n(x)$. In fact given the state $\psi_n(x)$ and the FN one $\psi_n^{eff}(x)$, then the state $\psi_n'(x)$ after the SR will be

$$\psi_n'(x)=C\left[1+\sum_k \alpha_k(O^k_x-\overline{O}^k_{eff})\right]\psi_n^{eff}(x)$$
$$\psi_n^{eff\,\prime}(x)=|\psi_n'(x)| \tag{36}$$

where now the α_k are uniquely determined by the conditions (33), whereas the normalization constant $C=\Sigma_x\psi_n(x)/\Sigma_x\psi_n^{eff}(x)$, and, finally, $\psi_n^{eff\,\prime}$ replace the FN propagated state ψ_n^{eff} after the SR (due to the condition $w_j^{eff\,\prime}=|w_j'|$). In this limit the dynamics described by the SR constraints is therefore perfectly defined and meaningful even in an exact calculation without the Monte Carlo sampling.

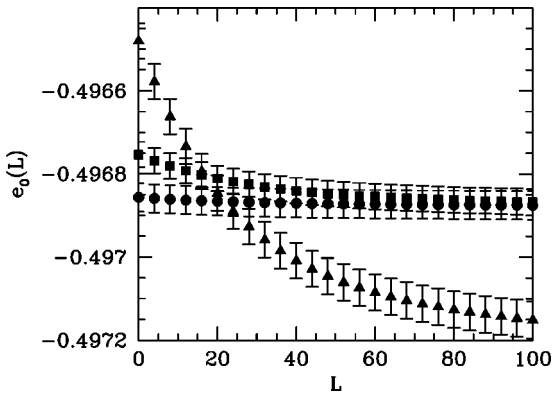

FIG. 1. Dependence on the number L of correcting factors of the estimated ground state energy per site for $N=64$ and $J_2=0.5$ obtained with the GFMCSR technique ($\Delta\tau=0.01$) with $M=200$ (triangles), 1500 (squares), and 10000 (circles). The GFMCSR technique is applied using in the SR the energy, all $S^z(q)$, the spin square, and the order parameter $m^{\dagger 2}$.

The way the computed results depend on the number of walkers is shown in Fig. 1, as a function of the number of correcting factors. As shown in Ref. 8 these correcting factors allow to eliminate the bias due to the finite population of walkers in the case there is no sign problem. In this case, instead the finite population bias cannot be eliminated even by an infinite number of factors and a properly large number M of walkers has to be taken for unbiased simulations. In fact for $M\to\infty$ the fluctuations of the G_n^L factors are bounded by the central limit theorem by $O(1/\sqrt{M})$. Therefore, for given L and large enough M, they do not play any role in the average quantities (28).

As it is evident for large number of walkers ($M\to\infty$) the correcting factors do not play any role and the estimate with minimum statistical error is obtained by simply ignoring the correcting factors. This is actually a common approach in GFMC, to consider a large number of walkers so that the bias of the finite walker population becomes negligible, and typically decreasing as $1/M$ (see, e.g., Fig. 2). However from the picture it is also evident that, for large enough M, the predicted results obtained by including or by neglecting the correcting factors are both consistent. The convergence to the $M\to\infty$ limit is however faster for the first method. Thus the inclusion of the correcting factors G_n^L in Eq. (28), though increasing the error bars, may be useful to reach the $M\to\infty$ limit with a smaller number of walkers. The fact that the two types of extrapolation to infinite M–the one including the correcting factors and the one neglecting them–converge to the same value (see Fig. 2) shows that the theoretical limit when Eq. (36) holds can be reached with a reasonable number of walkers, much smaller than the dimension of the Hilbert space.

The other parameter affecting the accuracy of the SR approach is the imaginary time distance $\Delta\tau$ between two consecutive SR. It is then natural to ask whether by increasing the frequency of the reconfigurations, one reaches a well defined dynamical limit for $\Delta\tau\to 0$. This is important since,

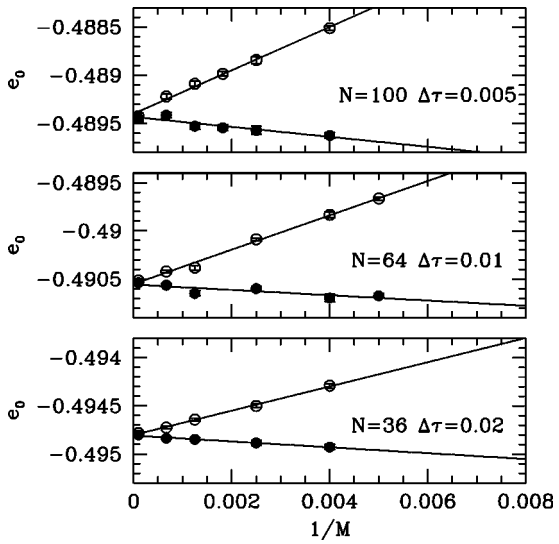

FIG. 2. Ground state energy per site for $J_2=0.5$ obtained for different clusters and different number of walkers. Empty dots are data obtained with zero correcting factors while full dots refer to converged values in L. The GFMCSR technique is applied using in the SR only the Hamiltonian ($p=1$).

due to the sign problem, for large system size N the time interval $\Delta\tau$ *has* to be decreased at least by a factor inversely proportional to N, because the average walker sign vanishes exponentially $\sim e^{-\Delta_s \tau}$ with an exponent Δ_s, which diverges with N. Different calculations, performed for different sizes can be compared only when the finite $\Delta\tau$ error (the difference between $\Delta\tau \to 0$ and finite $\Delta\tau$) is negligible.

As shown in Fig. 3, whenever the simulation is stable for $\Delta\tau \to 0$, the limit $\Delta\tau \to 0$ can be reached with a linear extrapolation. This property can be easily understood since in the limit of a large number of walkers, the variation of the average correlation functions Eq. (28) both for the FN dynamics and the exact dynamics in a time interval between two consecutive SR differ clearly by $O(\Delta\tau)$.

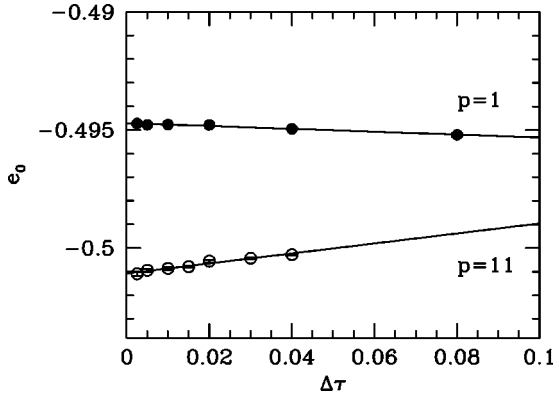

FIG. 3. Dependence of the ground state energy per site on the imaginary time step $\Delta\tau$ obtained for $J_2=0.5$ and $N=36$ with the GFMCSR technique by using in the SR the energy ($p=1$, full dots), all $S^z(q)$, the spin square and the order parameter $m^{\dagger 2}$ ($p=11$, empty dots). The number of walkers was fixed to $M=10000$, so that the finite-M bias can be neglected on this scale. The lower horizontal axis coincides with the exact diagonalization result.

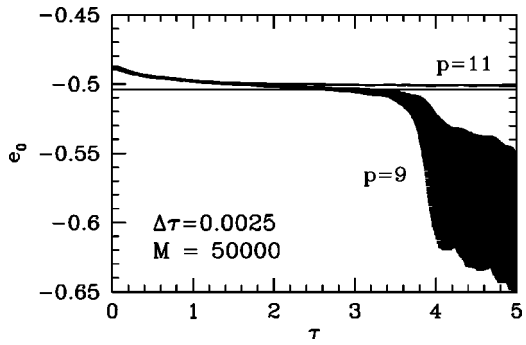

FIG. 4. Stable ($p=9$, upper curve) and unstable ($p=11$, lower curve) imaginary time evolution of the GFMCSR estimates of the ground-energy per site for $J_2=0.5$ and the $N=36$ cluster. The horizontal line indicates the exact result.

In order to show more clearly how the method is working and systematically correcting the FN we have implemented a slightly different but more straightforward ''release node'' technique.[14] We first apply the standard FN [with $\gamma=0$, see Eq. (13)] for a given number of walkers M and for long simulation time. We store the M-walkers configurations, after some equilibration at time interval large enough to allow uncorrelated and independent samples of the FN ground state. In a second step we recover each of these M-walker configurations and apply GFMCSR for a fixed imaginary time τ, so that we can see how the energy expectation value evolves from the FN to a more accurate determination. Typically one obtains a reasonable behavior for these curves that always coincides with the exact dynamics in the initial part where an exact sampling of the sign is possible. However, for large imaginary time, exceedingly small $\Delta\tau$ and large number of walkers, some instability may occur leading to results clearly off, as shown in Fig. 4. In this case, the instability is due to the fact that the correlation functions $S^z(q)=1/N^2\Sigma_{i,j}S_i^zS_j^z e^{iq(i-j)}$ which we have used in the SR ($p=9$),[10] introduce some uncontrolled fluctuations for the momentum $Q=(\pi,\pi)$ relevant for the antiferromagnetic order parameter. If we include in the SR technique also the spin isotropic operator corresponding to the order parameter $m^{\dagger 2}=1/N^2\Sigma_{i,j}\vec{S}_i\cdot\vec{S}_j e^{iQ(i-j)}$ and the total spin square ($p=11$) this instability disappears (see, Fig. 4, stable results, not shown in the picture, are obtained even without the total spin square, i.e., with $p=10$). This is a reasonable effect, since the order parameter has important fluctuations in all spin directions.

IX. CONCLUSIONS

In this paper, we have tried to describe in detail a recently proposed technique, the Green function Monte Carlo with Stochastic Reconfiguration (GFMCSR), that allows to work out the ground state energy and related mixed average correlation functions within a controlled accuracy even for models where the conventional Quantum Monte Carlo technique cannot be used because of the well known sign problem.

This method is rather general: in principle, convergence is achieved within an arbitrary accuracy if a sufficiently large number p of correlation functions are constrained to be equal before and after the SR, the basic statistical step used to

TABLE I. Variational estimate (VMC) and mixed averages (FN, SR, and Exact) of the ground energy per site $e_0=\langle H\rangle/N$, the total spin square and the order parameter (defined as in Ref. 15) for the triangular Heisenberg antiferromagnet for various system sizes. SR data are obtained using the short-range correlation functions generated by H $(p=2)$ and H^2 $(p=7)$ reported in Ref. 15. All the values reported in this table are obtained with large enough M and $1/\Delta\tau$, practically converged in the limit of $\Delta\tau\to 0$ and $M\to\infty$. Exact results are obtained using Lanczos technique.

	N	VMC	FN	SR($p=2$)	SR($p=7$)	Exact
e_0	12	-0.5981	$-0.6083(1)$	$-0.6085(1)$	$-0.6105(1)$	-0.6103
	36	-0.5396	$-0.5469(1)$	$-0.5534(1)$	$-0.5581(1)$	-0.5604
	48	$-0.5366(1)$	$-0.5426(1)$	$-0.5495(1)$	$-0.5541(1)$	
	108	$-0.5333(1)$	$-0.5387(1)$	$-0.5453(1)$	$-0.5482(1)$	
S^2_{tot}	12	0.235	0.0111(2)	0.006(4)	$-0.002(4)$	0.00
	36	1.71	1.20(1)	0.65(1)	0.02(1)	0.00
	48	2.55(1)	2.12(2)	1.44(1)	0.23(3)	0.00
	108	6.36(4)	5.66(3)	4.35(4)	2.7(1)	0.00
$m^{\dagger 2}$	12	0.9241	0.9286(1)	0.9210(2)	0.9132(6)	0.9109
	36	0.7791	0.7701(4)	0.7659(2)	0.7512(3)	0.7394
	48	0.7496(3)	0.7243(5)	0.7177(2)	0.7080(5)	
	108	0.6338(7)	0.6182(4)	0.6040(3)	0.5836(5)	

stabilize the sign problem instability. However, this is only a theoretical limit because the number of correlation functions p required to obtain the exact result scales exponentially with the system size, yielding a computational effort similar to the exact diagonalization methods.

In order to minimize the number p of correlation functions used in the SR, one is limited to use an empirical approach, based on physical intuition, and/or by comparison with exact results obtained for small sizes with the exact diagonalization technique. Typically, the fundamental ingredient that we have found to be important for strongly correlated Hamiltonians is the ''locality.'' The most useful correlation functions are the short-ranged ones appearing in the Hamiltonian H. A successful example is the application of the method to the Heisenberg model on the triangular lattice[15] where a remarkable accuracy is obtained by including also the short-range correlation functions generated by the application of the square of the Hamiltonian. Table I reports all the values of the ground state energy per site, the total spin square and the antiferromagnetic order parameter $m^{\dagger 2}$ obtained with VMC, FN and GFMCSR (for two different p's), up to $N=108$. However, the method of increasing systematically p, by including in the SR the short-range correlation functions generated by $H, H^2 \cdots$, does not seem general enough. For instance, it does not work for the J_1-J_2 Heisenberg model where the inclusion in the SR Eq. (33) of long-range operators, such as the spin-spin correlation function $S^z_i S^z_j$ at large distance $|i-j|$, is crucial to improve the accuracy of the method, whereas the short range ones generated by H^2 do not give any significant improvement.

Similarly to FN, the GFMCSR turns out to be size-consistent, in the sense that at fixed p the average correlation functions can be sizeably improved with respect to the variational guess, *even in the thermodynamic limit* (see Fig. 5). This is a non trivial property because, whenever there is sign problem, it is basically impossible to improve the best variational FN guess by the conventional ''release node''[14] since, for large sizes the variances become untractable even for a very short imaginary time propagation.

This kind of size consistency is a very important property of the present algorithm because the stability of the average sign at fixed p allows a *polynomial* complexity of the algorithm as a function of the system size. The algorithm, however, is typically a large factor ($\simeq 100$) more expensive than the standard FN as far as the computational time is concerned, for a given statistical error on correlation functions.

Until now the method has been extended rather successfully to several models: the mentioned J_1-J_2 and triangular lattice Heisenberg models, the $t-J$ model,[16] and the Hubbard model, where preliminary results[17] show that a similar improvement of the standard FN can also be obtained. In the latter case it is worth mentioning that a different approach, the Constrained Path Monte Carlo[18] (CPMC) also represents a very good remedy for the sign problem disease at least for intermediate coupling ($U/t\leqslant 8$). On the other hand, different schemes to get rid of the sign problem for continuum systems were previously proposed and successfully applied to small electron systems.[19]

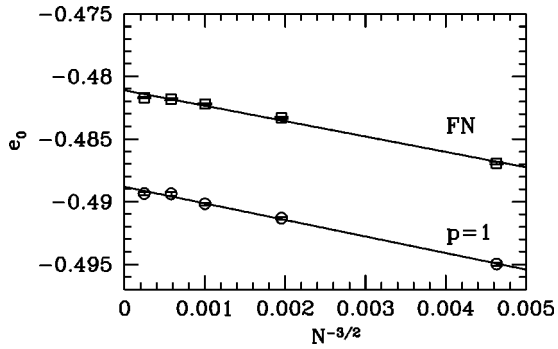

FIG. 5. Finite size scaling of the GS energy per site for $J_2=0.5$ obtained with the FN and GFMCSR technique applied reconfiguring the Hamiltonian ($p=1$).

Although the GFMCSR is far from being the definite solution to the sign problem in the Monte Carlo simulation, it certainly represents an interesting tool to alleviate this instability even for large system sizes. Its extension to continuum systems, and also to CPMC, is indeed straightforward, even though, in these cases, crossing the nodal surface in a variational way (see Appendix A) is not possible at present.

ACKNOWLEDGMENTS

We acknowledge useful discussions with M. Calandra, F. Becca, M. Capone, G. Bachelet, J. Carlson and G. Santoro, to whom we are also grateful for critical and careful reading of the manuscript. This work was partially supported by INFM (PRA HTSC and LOTUS) and by ''MURST-Progetti di ricerca di rilevante interesse nazionale.''

APPENDIX A: PROOF OF THE UPPER BOUND

Here we follow the paper[13] to prove rigorously the upper bound property of the ground state energy for H^{eff}. We want to show that the prescription given in Eqs. (13) and (14) for H^{eff} leads to an upper bound for the ground state energy of H. When importance sampling is used it is important to change slightly the definition of the sign-flip term as in Eq. (15):

$$\mathcal{V}_{\rm sf}(x)=\sum_{\{\psi_G(x')H_{x',x}/\psi_G(x)>0,\ x'\neq x\}}\psi_G(x')H_{x',x}/\psi_G(x). \tag{A1}$$

We now take *any* state

$$|\psi\rangle=\sum_x \psi(x)|x\rangle, \tag{A2}$$

and we compare its energy with respect to H and to H^{eff}:

$$\Delta E=\langle\psi|(H^{eff}-H)|\psi\rangle. \tag{A3}$$

ΔE can be written explicitly in terms of the matrix elements of H, using the definitions given in Eqs. (13), (14), and (A1)

$$\Delta E=(1+\gamma)\sum_x \psi(x)^*\left[\sum_{x'}^{\rm sf} H_{x,x'}\frac{\psi_G(x')}{\psi_G(x)}\psi(x)-\sum_{x'}^{\rm sf} H_{x,x'}\psi(x')\right], \tag{A4}$$

where the notation sf indicates conventionally the summation over the off-diagonal elements such that $\psi_G(x)H_{x,x'}/\psi_G(x')>0$. In this double summation each pair of configurations x and x' occurs twice. We combine these terms and rewrite Eq. (A4) as a summation over pairs:

$$\Delta E=(1+\gamma)\sum_{(x,x')}^{\rm sf} H_{x,x'}\left[|\psi(x)|^2\frac{\psi_G(x')}{\psi_G(x)}+|\psi(x')|^2\frac{\psi_G(x)}{\psi_G(x')}-\psi(x)^*\psi(x')-\psi(x')^*\psi(x)\right]. \tag{A5}$$

Denoting by $sH(x,x')$ the sign of the matrix element $H_{x,x'}$, and using the fact that for all terms in this summation the condition $\psi_G(x')H_{x',x}\psi_G(x)>0$ is satisfied, we can finally write ΔE as

$$\Delta E=(1+\gamma)\sum_{(x,x')}^{\rm sf}|H_{x,x'}|\left|\psi(x)\sqrt{\left|\frac{\psi_G(x')}{\psi_G(x)}\right|}-sH(x,x')\psi(x')\sqrt{\left|\frac{\psi_G(x)}{\psi_G(x')}\right|}\right|^2. \tag{A6}$$

Obviously, ΔE is positive for any wave function ψ. Thus the ground-state energy of H^{eff} is an upper bound for the ground-state energy of the original Hamiltonian H.

Now the GFMC method can calculate the exact ground-state energy E_0^{eff} and wave function ψ^{eff} of H^{eff}, without any sign problem. Hence, $E_0^{eff}\geqslant\langle\psi^{eff}|H|\psi^{eff}\rangle\geqslant E_0$, where the second inequality follows from the usual variational principle. We conclude therefore that the FN energy is an upper bound to the true ground state energy. One can easily verify that $\langle\psi_G|H|\psi_G\rangle=\langle\psi_G|H^{eff}|\psi_G\rangle$, and thus one can be sure that the GFMC procedure improves on the energy of the guiding wave function: $E_0^{eff}\leqslant\langle\psi_G|H^{eff}|\psi_G\rangle=\langle\psi_G|H|\psi_G\rangle$.

Note that the standard lattice FN approach[13] is obtained for the particular parameter $\gamma=0$.

APPENDIX B: FORMAL PROOF OF THE GFMCSR CONDITIONS

As stated in Sec. VI the SR conditions (33) read

$$\sum_{x',x}\overline{O}^k_{x',x}\psi'_n(x)=\sum_{x',x}\overline{O}^k_{x',x}\psi_n(x), \tag{B1}$$

for $k=1,\ldots,p$, with the normalization one $\Sigma_x\psi'_n(x)=\Sigma_x\psi_n(x)$.

The wave function $\psi'_n(x)$ after the SR conditions defined by Eq. (25) can be explicitly written in terms of the original walker probability distribution. To this purpose, we single out in the definition of $\psi'_n(x)$

$$\psi'_n(x)=\int[d\underline{w}']\sum_{\underline{x}'}P'_n(\underline{w}',\underline{x}')\frac{\Sigma_j\delta_{x,x'_j}w'_j}{M}, \tag{B2}$$

a term k in the above summation over j which gives an additive contribution to ψ'_n, namely $\psi'_n=1/M\Sigma_k\{\psi'_n\}_k$ with

$$\{\psi'_n(x)\}_k=\int[d\underline{w}']\sum_{\underline{x}'}\int[d\underline{w}]\times\sum_{\underline{x}}X(\underline{w}',\underline{x}';\underline{w},\underline{x})P_n(\underline{w},\underline{x})\delta_{x,x'_k}w'_k, \tag{B3}$$

where in the above equation we have substituted the definition of P' in terms of P given by Eqs. (24) and (25). In the latter equation it is easy to integrate over all variables $w'_j, w_j^{eff'}, x'_j$ for $j\neq k$ using that the kernel X is particularly simple as discussed in Sec. VI. Then, the remaining three integrals and summations over $w'_k, w_k^{eff'}, x'_k$ can be easily

performed using the simple δ functions that appear in the kernel X and the definition of $\beta=\Sigma_j p_{x_j}/\Sigma_j|p_{x_j}|$, so that one easily obtains

$$\{\psi_n'(x)\}_k=\int [d\underline{w}]\sum_{\underline{x}} P_n(\underline{w},\underline{x})\frac{\Sigma_j w_j}{M}\operatorname{sgn} p_x\frac{\Sigma_j|p_{x_j}|\delta_{x,x_j}}{\Sigma_j p_{x_j}}. \tag{B4}$$

We can replace in general $\operatorname{sgn} p_x\Sigma_j|p_{x_j}|\delta_{x,x_j}/\Sigma_j p_{x_j}=\Sigma_j p_{x_j}\delta_{x,x_j}/\Sigma_j p_{x_j}$ even when, occasionally, more configurations satisfy $x_j=x$.[20] Thus, we obtain a closed expression for $\psi_n'(x)$ after the simple summation on the index k:

$$\psi_n'(x)=\int [d\underline{w}]\sum_{\underline{x}} P_n(\underline{w},\underline{x})\left(\frac{\Sigma_j w_j}{M}\right)\frac{\Sigma_j p_{x_j}\delta_{x,x_j}}{\Sigma_j p_{x_j}}. \tag{B5}$$

Then the normalization condition $\Sigma_x\psi_n'(x)=\int[d\underline{w}]\Sigma_{\underline{x}}P_n(\underline{w},\underline{x})(\Sigma_j w_j/M)=\Sigma_x\psi_n(x)$ easily follows. On the other hand, the left-hand side of Eq. (33) can be also computed easily, yielding

$$\sum_{x',x}\overline{O}^k_{x',x}\psi_n'(x)=\int [d\underline{w}]\sum_{\underline{x}} P_n(\underline{w},\underline{x})\left(\frac{\Sigma_j w_j}{M}\right)\frac{\Sigma_j p_{x_j}O^k_{x_j}}{\Sigma_j p_{x_j}}, \tag{B6}$$

where $O^k_{x_j}=\Sigma_{x'}\overline{O}_{x',x_j}$ is the mixed estimator of the operator O^k.

Finally, by substituting the condition (34) into the previous equation, one obtains

$$\sum_{x',x}\overline{O}^k_{x',x}\psi_n'(x)=\int [d\underline{w}]\sum_{\underline{x}} P_n(\underline{w},\underline{x})\frac{\Sigma_j w_j O^k_{x_j}}{M}$$

$$=\sum_{x',x}\overline{O}^k_{x',x}\psi_n(x), \tag{B7}$$

which proves the statement at the beginning of this section.

APPENDIX C: PROOF OF EXISTENCE AND UNIQUENESS OF SOLUTION FOR THE RECONFIGURATION

In this appendix, we prove that given the $p+1$ SR conditions (34) the elements of the table p_{x_j} are uniquely determined for each walker configuration $(\underline{w},\underline{x})$.

We define here the quantity

$$v_j^k=(O^k_{x_j}-\overline{O}^k_f), \tag{C1}$$

for each configuration j, where $\overline{O}^k_f=\Sigma_j w_j^f O^k_{x_j}/\Sigma_j w_j^f$ is the average value over the reference weights, w_j^f, of the operator considered, labeled by the number k. The reference weights w_j^f are restricted to be strictly positive but can be in general arbitrary functions of all the FN weights $\{w_j^{eff}\}$ the exact weights $\{w_j\}$ and the configurations $\{x_j\}$. It is easy to show that, in order that

$$p_{x_j}=w_j^f\left(1+\sum_k \alpha_k v_j^k\right) \tag{C2}$$

allows to satisfy the SR conditions (33), it is sufficient that α_k are determined by the simple linear equation

$$\sum_{k'} s_{k,k'}\alpha_{k'}=\frac{\Sigma_j w_j v_j^k}{\Sigma_j w_j}, \tag{C3}$$

where

$$s_{k,k'}=\frac{\Sigma_j w_j^f v_j^k v_j^{k'}}{\Sigma_j w_j^f} \tag{C4}$$

is the covariance matrix of the operators O^k over the reference weights w_j^f. The solution to Eq. (C3) is possible if the determinant of $s_{k,k'}$ is non-vanishing. Since s represents an overlap matrix defined with a nonsingular scalar product $\langle v^k|v^{k'}\rangle=\Sigma_j w_j^f v_j^k v_j^{k'}/\Sigma_j w_j^f$ as w_j^f are positive, its determinant is always nonzero provided the vectors v^k are linearly independent. Thus, in the latter case, the solution to Eq. (C3) exists and is unique.

On the other hand suppose that among the p vectors v^k only $p'<p$ are linearly independent. Thus, the remaining $p-p'$ vectors can be written as linear combination of p' linearly independent ones (henceforth we assume that these linearly independent vectors are labeled by the consecutive indices $k=1,\ldots,p'$)

$$v_j^{k'}=\sum_{k=1}^{p'} x_k^{k'} v_j^k, \tag{C5}$$

for $k'>p'$, where $x_k^{k'}$ are suitable coefficients. The same previous considerations allow to satisfy the first p' SR conditions as for Eq. (C3) a unique solution exists if we restrict all the sums for k, $k\leqslant p'$, and p_{x_j} is determined only by the first p' linearly independent vectors in Eq. (C2). With the determined p_{x_j} it is obvious that

$$\frac{\Sigma_j p_{x_j} v_j^k}{\Sigma_j p_{x_j}}=\frac{\Sigma_j w_j v_j^k}{\Sigma_j w_j} \tag{C6}$$

is verified for $k=1,\ldots,p'$.

On the other hand we can easily show that all the remaining SR conditions (C6) for $k'>p'$ are identically satisfied. In fact, in this case the LHS of Eq. (C6) can be manipulated as follows, using definition (C5)

$$\frac{\Sigma_j p_{x_j} v_j^{k'}}{\Sigma_j p_{x_j}}=\sum_{k=1}^{p'} x_k^{k'}\left(\frac{\Sigma_j v_j^k p_{x_j}}{\Sigma_j p_{x_j}}\right)$$

$$=\sum_{k=1}^{p'} x_k^{k'}\left(\frac{\Sigma_j v_j^k w_j}{\Sigma_j w_j}\right)$$

$$=\frac{\Sigma_j v_j^{k'} w_j}{\Sigma_j w_j}, \tag{C7}$$

where in the intermediate steps we have used Eq. C6 for $k\leqslant p'$. Thus, the SR conditions determine uniquely p_{x_j} in any case and this conclude the important statement of this Appendix.

1. Remark

With the above definitions it is also possible to show that p_{x_j} remains unchanged for any linear transformation of the operator set. Namely, suppose we consider the new operators

$$\tilde{O}^{k'} = \sum_k L_{k',k} O^k + \beta_{k'} \tag{C8}$$

in the SR conditions, where the real matrix L is assumed to have nonvanishing determinant. Within this assumption it is simple to show that p_{x_j} will remain unchanged.

In fact, the new set of operators will define a new covariance matrix between the new vectors

$$\tilde{v}_j^{k'} = \sum_k L_{k',k} v_j^k, \tag{C9}$$

i.e., $\tilde{v} = Lv$, $\tilde{s} = LsL^T$, where L^T is the transposed of L and the set of new equations

$$\sum_{k'} \tilde{s}_{k,k'} \tilde{\alpha}_{k'} = \frac{\Sigma_j w_j \tilde{v}_j^k}{\Sigma_j w_j}$$

is obviously satisfied by

$$\tilde{\alpha} = (L^{-1})^T \alpha, \tag{C10}$$

where α is the solution of the SR conditions before the transformation (C8). Whenever the number p' of linearly independent v^k is less than p, also the number of linearly independent $\tilde{v}^k$ will be p' as L is nonsingular. The solutions α and $\tilde{\alpha}$, as described previously, refer therefore to the first p' components, and all the matrix involved, such as $\tilde{L}$ and $\tilde{s}$ are in this case restricted to this subspace.

Then, by Eq. (C10) and Eq. (C9), it easily follows that the new coefficients $\tilde{p}_{x_j} = w_j^f (1 + \Sigma_k \tilde{\alpha}_k \tilde{v}_j^k) = w_j^f (1 + \Sigma_k \alpha_k v_j^k) = p_{x_j}$, which finally proves the statement of this remark.

2. Optimization of the weights

The definition of the weights p_{x_j} that satisfy the SR condition (33) is highly arbitrary because as we have mentioned before the probabilities P_n and P'_n do not uniquely determine the quantum states ψ_n and ψ'_n that are subject to the conditions (33). In this sense there may be different definitions of the weights p_{x_j} that may behave differently at finite p with less or more accuracy. Though Eqs. (33) are equally satisfied for different choices of the coefficients p_{x_j} the two states ψ_n and ψ'_n may be much closer (less bias) for an optimal choice. The optimal choice that minimizes the distance $|\psi_n - \psi'_n|$, at fixed number p of correlation functions included in the SR, has not probably been found yet. We have attempted several choices for the reference weights w_j^f of Eq. (C2) but until now no significant improvement of the simplest FN ones[10] has been obtained.

APPENDIX D: THE LIMIT $\Lambda \to \infty$ FOR THE POWER METHOD

The constant Λ, which defines the Green function $G_{x',x} = \Lambda \delta_{x',x} - H_{x',x}$ and the FN one G^{eff} (12) has to be taken large enough to determine that all the diagonal elements of G^{eff} are non-negative (by definition the off-diagonal ones of G^{eff} are always non-negative). This requirement often determines a very large constant shift, which increases with larger size and is not known *a priori*. The trouble in the simulation may be quite tedious, as if for the chosen Λ a negative diagonal element is found for G^{eff}, one needs to increase Λ and start again with a completely new simulation. The way out is to work with exceedingly large Λ, but this may slow down the efficiency of the algorithm as in the stochastic matrix $p_{x',x}$ the probability to remain in the same configuration p_d may become very close to one

$$p_d = \frac{\Lambda - H_{x,x} - (1+\gamma)\mathcal{V}_{\rm sf}(x)}{\Lambda - E_x}, \tag{D1}$$

where $\mathcal{V}_{\rm sf}(x)$ is given in Eq. (15) and E_x is the local energy Eq. (5) that do not depend on Λ given the configuration x.

Following Ref. 7 the problem of working with large Λ can be easily solved with no loss of efficiency. We report this simple idea applied to our particular algorithm at fixed number of walkers. If Λ is large it is possible to take a large value of k_p (of order Λ) iterations between two consecutive reconfigurations, because in most iterations the configuration x is not changed. The idea is that one can determine *a priori*, given p_d, what is the probability $t(k)$ to make k diagonal moves before the first acceptance of a new configuration with $x' \neq x$. This is given by $t(k) = p_d^k (1 - p_d)$ for $k = 0, \ldots, n_l - 1$ and $t(n_l) = p_d^{n_l}$ if no off-diagonal moves are accepted during the n_l trials that are left to complete the loop without reconfigurations.

It is a simple exercise to show that, in order to sample $t(k)$ one needs one random number $0 < \xi < 1$, so that the stochastic integer number k can be computed by the simple formula

$$k = \min\left(n_l, \left[\frac{\ln \xi}{\ln p_d} \right] \right), \tag{D2}$$

where the brackets indicate the integer part. During the k_p iterations one can iteratively apply this formula by bookkeeping the number of iterations n_l that are left to complete the loop without reconfigurations. At the first iteration $n_l = k_p$, then k is extracted using Eq. (D2), and the weights (w, w^{eff}) of the walker are updated according to k diagonal moves and if $k < n_l$ a new configuration is extracted at random according to the off-diagonal matrix elements of $p_{x',x}$. The weights are correspondingly updated for this off-diagonal move, and finally, if $k < n_l$, n_l is changed to $n_l - k - 1$, so that one can continue to use Eq. (D2) until all the k_p steps are executed for each walker.

The interesting thing of this method is that it can be readily generalized for $\Lambda \to \infty$ by increasing k_p with Λ, namely $k_p = [\Lambda \Delta\tau]$, where $\Delta\tau$ represents now exactly the imaginary time difference between two consecutive reconfigurations when the exact propagator $e^{-H\Delta\tau}$ or $e^{-H^{eff}\Delta\tau}$ is applied statistically.

[1] H. J. Schulz, T. A. L. Ziman, and D. Poilblanc, J. Phys. I **6**, 675 (1996); B. Bernu, P. Lecheminant, C. Lhuillier, and L. Pierre, Phys. Rev. B **50**, 10 048 (1994).
[2] D. Poilblanc, Phys. Rev. B **52**, 9201 (1995); K. J. Szczepanski, P. Horsch, W. Stephan, and M. Ziegler, *ibid.* **41**, 2017 (1990); P. Prelovsek, I. Sega, and J. Bonca, *ibid.* **42**, 10 706 (1990); B. Eder and Y. Ohta, *ibid.* **50**, 10 043 (1994).
[3] Overviews of exact diagonalization studies can be found in E. Dagotto, Rev. Mod. Phys. **66**, 763 (1994); E. Dagotto, A. Moreo, F. Ortolani, D. Poilblanc, and J. Riera, Phys. Rev. B **45**, 10 741 (1992).
[4] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992); Phys. Rev. B **48**, 10 345 (1993).
[5] S. R. White and R. L. Martin, cond-mat/9808118 (unpublished).
[6] C. Gros, Phys. Rev. B **38**, 931 (1988); H. Yokoyama and H. Shiba, J. Phys. Soc. Jpn. **57**, 2482 (1988); T. Giamarchi and C. Lhuillier, Phys. Rev. B **42**, 10 641 (1990); **43**, 12 943 (1991).
[7] N. Trivedi and D. M. Ceperley, Phys. Rev. B **41**, 4552 (1990).
[8] M. Calandra and S. Sorella, Phys. Rev. B **57**, 11 446 (1998).
[9] H. J. M. van Bemmel, D. F. B. ten Haaf, W. van Saarlos, J. M. J. van Leeuwen, and G. An, Phys. Rev. Lett. **72**, 2442 (1994).
[10] S. Sorella, Phys. Rev. Lett. **80**, 4558 (1998).
[11] D. M. Ceperley and M. H. Kalos, in *Monte Carlo Method in Statistical Physics*, edited by K. Binder (Springer-Verlag, Heidelberg, 1992).
[12] J. H. Hetherington, Phys. Rev. A **30**, 2713 (1984).
[13] D. F. B. ten Haaf, H. J. M. van Bemmel, J. M. J. van Leeuwen, W. van Saarloos, and D. M. Ceperley, Phys. Rev. B **51**, 13 039 (1995).
[14] D. M. Ceperley and B. J. Alder, J. Chem. Phys. **81**, 5833 (1984).
[15] L. Capriotti, A. E. Trumper, and S. Sorella, Phys. Rev. Lett. **82**, 3899 (1999).
[16] M. Calandra, F. Becca, and S. Sorella, Phys. Rev. Lett. **81**, 5185 (1998).
[17] F. Becca, M. Capone, and S. Sorella (unpublished).
[18] S. Zhang, J. Carlson, and J. E. Gubernatis, Phys. Rev. Lett. **74**, 3652 (1995); Phys. Rev. B **55**, 7464 (1997).
[19] S. Zhang and M. H. Kalos, Phys. Rev. Lett. **67**, 3074 (1991).
[20] In fact, $\mathrm{sgn}\, p_x$ is the same for all the corresponding indices matching $x_j = x$, as implied by the definition (32) of p_{x_j} and the condition $w_j^{eff} > 0$ valid for all j.