

## Green Function Monte Carlo with Stochastic Reconfiguration

Sandro Sorella

*Istituto Nazionale di Fisica della Materia and International School for Advanced Studies, Via Beirut 4,  
34013 Trieste, Italy*

(Received 17 November 1997)

A new method for the stabilization of the sign problem in the Green function Monte Carlo technique is proposed. The method is devised for real lattice Hamiltonians and is based on an iterative “stochastic reconfiguration” scheme which introduces some bias but allows a stable simulation with constant sign. The systematic reduction of this bias is possible in principle. The method is applied to the frustrated  $J_1 - J_2$  Heisenberg model, and tested against exact diagonalization data. Evidence of a finite spin gap for  $J_2/J_1 > \sim 0.4$  is found in the thermodynamic limit. [S0031-9007(98)06070-0]

PACS numbers: 75.10.Jm, 02.70.Lq, 75.40.Mg

As is well known the Green function Monte Carlo method (GFMC) allows one to obtain the exact ground state properties of a many body Hamiltonian with a statistical method. One of the most severe restrictions is that only positive definite Green function (GF) can be sampled; otherwise, the method is facing the well known “sign problem.” Approximate techniques like the fixed node approximation (FN) have been developed to circumvent the sign problem but at the very least they cannot be systematically improved to achieve the exact answer within statistical errors. This property has severely limited the applications of GFMC to fermions and frustrated boson models. In this Letter I propose a new approach to stabilize the sign problem, the GFMC with stochastic reconfiguration (GFMC SR), which will be shortly described below, revisiting also the basic steps of the standard GFMC on a lattice [1,2].

In order to filter out the ground state of a given lattice Hamiltonian  $H$  the standard power method may be applied iteratively:

$$\psi_{n+1}(x') = \sum_x (\Lambda \delta_{x',x} - H_{x',x}) \psi_n(x), \quad (1)$$

where  $x$  represents conventionally the index of a complete basis  $|x\rangle$ ,  $H_{x',x}$  being the corresponding matrix elements of the Hamiltonian which in the following are assumed real, and  $\Lambda$  is a positive constant that allows the convergence of  $\psi_n$  to the ground state  $\psi_0(x)$ , for large  $n$ . In numerical calculations of interesting lattice Hamiltonians the dimension of the basis grows exponentially with the size and the particle number, though the matrix itself is very sparse and all its elements  $H_{x',x}$ , for given  $x$ , can be generally computed even for a large system size. In this case an exact application of (1) is impossible unless for few steps. A way out is to use a stochastic approach, like GFMC, which is particularly simple on a lattice.

In order to implement stochastically the iteration (1) the corresponding lattice GF

$$G_{x',x} = \Lambda \delta_{x',x} - H_{x',x} \quad (2)$$

may be decomposed in the following way:

$$G_{x',x} = s_{x',x} p_{x',x} b_x, \quad (3)$$

where  $p_{x',x}$  is a normalized stochastic matrix,  $b_x \geq 0$  is a normalization constant, and the matrix  $s$  takes into account the sign of the GF. The typical choice is to take  $p_{x',x} = |G_{x',x}|/b_x$ ,  $b_x = \sum_{x'} |G_{x',x}|$ , and  $s_{x',x} = \text{sgn } G_{x',x}$ , which is identically one if there is no sign problem.

In the GFMC method the so called “walker” is defined by a weight  $w$  and a configuration  $x$ . At a given iteration  $n$  the walker is assumed to sample statistically the state  $\psi_n(x)$  in Eq. (1), in the sense that the probability  $P_n(w, x)$  to have the walker with weight  $w$  (not restricted to be positive) in a given configuration  $x$  satisfies  $\int dw P_n(w, x) w = \psi_n(x)$ . Then the matrix multiplication (1) can be implemented statistically, in the precise sense that  $\int dw P_{n+1}(w, x) w = \psi_{n+1}(x)$ , by the following three steps: (i) scale the walker weight by  $b_x$ :  $w' = b_x w$ . (ii) Select randomly a new configuration  $x'$  according to the stochastic matrix  $p_{x',x}$ . (iii) Finally multiply the weight of the walker by the sign factor  $s_{x',x}$ :  $w' \rightarrow w' s_{x',x}$  (MI). In principle the previous Markov process determines, for large  $n$ , the ground state of  $H$  even with a single walker. In practice it is convenient to use a large number  $M$  of walkers, which I indicate by  $(w_j, x_j) j = 1, \dots, M$  shorthand in the following also by vector notations  $\underline{w}, \underline{x}$ .

If there is a sign problem the average walker sign  $\langle s \rangle_n = \langle \sum_j w_j \rangle_n / \langle \sum_j |w_j| \rangle_n$  decreases exponentially to zero as the Markov iteration (MI) is repeatedly applied and it is basically impossible to reach a reasonably large value of  $n$ .

Recently remarkable progress in GFMC on a lattice was the extension of the FN to this case. The method is based on a definition of an effective GF  $G_{x',x}^f$  which is always positive definite but yields a good *variational* estimate of the energy. For later purposes we define this effective GF in a slightly different way, by introducing a parameter  $\gamma$ , which allows one to sample also the negative elements of the GF:

$$G_{x',x}^f = \begin{cases} -H_{x',x} & \text{if } H_{x',x} \leq 0, \\ \gamma H_{x',x} & \text{if } H_{x',x} > 0, \\ \Lambda - H_{x,x} - (1 + \gamma) \mathcal{V}_{\text{sf}}(x) & \text{if } x = x', \end{cases} \quad (4)$$

where the diagonal *sign-flip* contribution is given by [3,4]

$$\mathcal{V}_{\text{sf}}(x) = \sum_{H_{x',x} > 0, x' \neq x} H_{x',x}. \quad (5)$$

For  $\gamma = 0$  the usual formulation [4] is recovered, whereas for  $\gamma > 0$  [5] the crossing to the negative sign region is allowed so that the exact GF can be written as  $G_{x',x} = s_{x',x} G_{x',x}^f$ , where  $s_{x',x}$  is finite and nonzero and is determined by the ratio  $G_{x',x}/G_{x',x}^f$  with  $G$  and  $G^f$  given by Eqs. (2) and (4), respectively. The value of the constant  $\gamma$  necessary to cross the “nodal surface” was chosen to be  $\frac{1}{2}$  in all forthcoming applications.

In the basic decomposition (3) the stochastic matrix  $p_{x',x} = G_{x',x}^f/b_x$  and the normalization coefficient  $b_x = \sum_{x'} G_{x',x}^f$  are instead determined only by  $G^f$ .

By omitting the last step  $w' \rightarrow ws_{x',x}$  in the Markov iteration process MI, the state  $\psi_n$  is indeed propagated through the positive GF  $G^f$ . The main property used in the following is that at any Markov iteration  $n$  we can have a statistic knowledge of both the state  $\psi_n(x)$  obtained with the exact GF and of  $\psi_n^f(x)$  obtained instead with the approximate but positive definite one  $G^f$ . To this purpose the  $j$ th walker is defined by two weights  $w_j^f$  and  $w_j$  corresponding to the propagation of the walker by  $G^f$  and  $G$ , respectively. These weights act on *the same* configuration  $x_j$ . Hereafter the vector  $\underline{w}$  represents therefore a shorthand notation for the  $2M$  components  $w_j, w_j^f$  for  $j = 1, \dots, M$ .

The walker vector  $\underline{w}, \underline{x}$  allows one to determine statistically the state

$$\psi_n(x) = \int d[\underline{x}] \sum_{\underline{x}} P_n(\underline{w}, \underline{x}) \sum_j \delta_{x,x_j} w_j / M \quad (6)$$

and analogously  $\psi_n^f(x)$  by replacing the weights  $w_j$  with the positive ones  $w_j^f$  in the previous equation. In this way the configurations generated by the described Markov process MI, if weighted with the constants  $w_j^f$ , are distributed for large  $n$ , according to the variational state corresponding to  $G^f$ . This is a reasonable variational wave function (WF), which will be the initial approximation to which systematic corrections will be applied, as described later on.

Apart for the previous technical definitions, we can explain in a few words the basic idea used for the stabilization of the sign problem. The iteration MI converges to the ground state, but due to the sign problem, only a few iterations can be performed with a reasonable statistical accuracy. However, the representation of the state  $\psi_n(x)$  in terms of the walker population  $x_j, w_j$  is not unique. In fact, it is perfectly possible to represent the same state  $\psi_n(x)$  either with a walker population with very small average sign or with a one with a very large average sign. If such reconfigurations are possible each few  $k_p$  steps, the average sign may be stabilized to a large value during the iteration (1) and there will be no difficulty to sample the ground state for  $n \rightarrow \infty$ , *with no sign problem*.

I will show that this reconfiguration is well defined and indeed possible. The set of  $M$  walkers  $(\underline{w}, \underline{x})$  is defined via their probability function  $P_n(\underline{w}, \underline{x})$  which in turn defines the state  $\psi_n(x)$  by Eq. (6). The task is to change  $P_n$  onto a new probability distribution  $P'_n$  corresponding to a steadily high sign for the walker population, and this *without changing the information content*, the state  $\psi_n(x)$ .

Let us define the new state  $\psi_n'(x)$  as the one obtained by averaging over  $P'_n$  in Eq. (6), then the reconfiguration is exact if  $P'_n$  is such that

$$\psi_n'(x) = \psi_n(x) \quad \text{for all } x. \quad (7)$$

In general it is difficult or impractical to realize all these conditions (7) as their number equals the dimension of the Hilbert space. I consider therefore a set of operators  $O^k, k = 1, \dots, p \ll M$  and require only  $p + 1$  stochastic reconfiguration conditions:

$$\sum_{x',x} O_{x',x}^k \psi_n'(x) = \sum_{x',x} O_{x',x}^k \psi_n(x) \quad (8)$$

for  $k = 1, \dots, p$ , beyond the normalization one  $\sum_x \psi_n'(x) = \sum_x \psi_n(x)$ .

The previous equations (8) mean that the so called “mixed averages” of the operators  $O^k$  coincide before and after the reconfiguration [6].

The main idea of this Letter is that these  $p + 1$  conditions can be fulfilled *exactly* (for chosen operators) by defining the reconfiguration in the following form:

$$P'_n(\underline{w}', \underline{x}') = \int d[\underline{w}] \sum_{\underline{x}} \prod_{i=1}^M \times \left\{ \frac{\sum_j |p_{x_j}| \delta_{x'_i, x_j}}{\sum_j |p_{x_j}|} \delta \left( w'_i - \frac{\sum_j w_j}{\beta M} \text{sgn } p_{x'_i} \right) \times \delta(w'_i - |w'_i|) \right\} P_n(\underline{w}, \underline{x}), \quad (9)$$

where  $\beta = \sum_j p_{x_j} / \sum_j |p_{x_j}|$  is the average sign after the reconfiguration, which is supposed to be much higher to stabilize the process. The new configurations  $x'_i$  are taken randomly among the old ones  $\{x_j\}$ , according to the table  $p_{x_j}$ , defined below. The positive weights  $w_j^f$  represent a good starting point for the definition of a reconfiguration with large  $\beta$ . Though there is some arbitrariness in the definition of the coefficients  $p_{x_j}$ , a simple and convenient choice is

$$p_{x_j} = w_j^f \left[ 1 + \sum_k \alpha_k (O_j^k - \bar{O}_f^k) \right],$$

where  $\bar{O}_f^k = \sum_j w_j^f O_j^k / \sum_j w_j^f$  are the averages over the positive weights  $w_j^f$  of the mixed estimates  $O_j^k = \sum_{x'} O_{x',x_j}^k$  corresponding to the operator  $O^k$  and the configuration  $x_j$ .

Then, in order to satisfy the WF conditions (8), by using the definition (9), it is *sufficient* that the coefficients  $p_{x_j}$  satisfy the following Markovian conditions:

$$\frac{\sum_j p_{x_j} O_j^k}{\sum_j p_{x_j}} = \frac{\sum_j w_j O_j^k}{\sum_j w_j}, \quad (10)$$

which in turn determine the unknown variables  $\alpha_k$ , for  $k = 1, \dots, p$ , for given  $\underline{w}, \underline{x}$ .

For Hamiltonian not affected by the sign problem ( $G^f = G$ ,  $\alpha_k = 0$  and  $\beta = 1$ ) this reconfiguration was already used to control the walker population size without introducing any source of systematic error [7]. The present more general reconfiguration (9) can be easily and efficiently implemented in a similar way.

Obviously the reconfiguration conditions (8) are equivalent to the *exact* ones (7), when the number  $p$  of linearly independent operators considered in (8) is equal to the large dimension of the Hilbert space. An important applicative issue is whether GFMCSR converges, within a reasonable accuracy, even with a small number  $p$  of meaningful operators  $O^k$ .

We consider the frustrated  $J_1 - J_2$  Heisenberg spin  $\frac{1}{2}$  model on a finite square lattice with  $L$  sites and periodic boundary conditions (tilted by  $45^\circ$  for the  $L = 32$  size only). The model Hamiltonian is determined by an antiferromagnetic coupling  $J_1 > 0$  between nearest neighbor spins and a frustrating coupling  $J_2 > 0$  between next neighbor ones [8–10]. In all forthcoming examples the stochastic reconfigurations were applied frequently enough to maintain the average sign before reconfiguration  $\sim 0.8$ , a condition that minimizes the statistical fluctuations. Moreover in each simulation it is important to work with a fairly large number of walkers, since in the  $M \rightarrow \infty$  limit, the GFMCSR results are practically independent of the frequency of reconfigurations, as well as the overall constant energy shift  $\Lambda$ .

The accuracy of GFMCSR for the ground state is displayed in Table I, and compared with other methods. The variational WF (used also for GFMC importance sampling [6]) contains a Jastrow-like factor

$$\exp\left(\frac{\eta}{2} \sum_{R,R'} v(R - R') S_R^z S_{R'}^z\right)$$

to mimic the interaction between the spins  $S_R^z = \pm 1/2$  at sites  $R, R'$ , where  $\eta$  is a variational parameter and the two-spin interaction  $v$  can be derived by using the method described in [11], yielding an explicit Fourier transform for  $v$ :

$$2 - 2\sqrt{\frac{2 - \sigma(1 - \cos q_x \cos q_y) + \cos q_x + \cos q_y}{2 - \sigma(1 - \cos q_x \cos q_y) - \cos q_x - \cos q_y}}$$

with  $\sigma = 2J_2/J_1$ . This potential is not defined for  $J_2/J_1 = 1/2$ , and in such a case I have chosen to work with  $\sigma = 0.8$ . Restriction to any subspace of total spin projection  $S_{\text{tot}}^z = \sum_R S_R^z$  allows one to evaluate the spin gap by performing two simulations for  $S_{\text{tot}}^z = 0$  and  $S_{\text{tot}}^z = 1$ . Henceforth I will use the same potential  $v$  in both subspaces, by optimizing  $\eta$  for the  $S_{\text{tot}}^z = 0$  energy.

As shown in Table I the accuracy of the variational WF is rather poor, and is considerably improved by the FN, at least for small  $J_2$ . This kind of accuracy is however not enough to determine the rapid increase of the spin gap as  $J_2/J_1$  approaches the value  $\frac{1}{2}$  of the classical transition. Instead, as shown in Fig. 1 the GFMCSR allows one to achieve a good accuracy also on this delicate quantity by considering in the reconfigurations only the energy and the spin structure factor  $S_q^z = \sum_{R,R'} e^{iq(R-R')} S_R^z S_{R'}^z$ , symmetrized over all directions and for all nonequivalent wave vectors  $q$ . Remarkably also mixed averages of correlation functions that are not included in such reconfiguration conditions (8) are also significantly improved (see Table I).

TABLE I. Percentage error of the energy (square antiferromagnetic order parameter  $\vec{m}^2$  as in [7]) for the various methods: variational (VMC), fixed node (FN),  $p = 1$  GFMCSR (SRe) with the energy alone and  $p = 5, 8, 9$  GFMCSR estimate (SR) with the energy, and  $S_q^z$  for  $L = 16, 32$ , and  $36$ . The statistical errors are about one place in the last digit.

$J_2/J_1$	$L$	$\eta$	% VMC	% FN	% SRe	% SR
0.1	16	1.2	2.84 (2.2)	0.17 (0.1)	-0.03 (0.0)	0.02 (0.0)
0.2	16	1.15	2.80 (2.5)	0.41 (0.4)	0.00 (0.2)	0.03 (0.0)
0.3	16	1.1	3.25 (2.5)	0.87 (0.7)	0.12 (0.8)	0.05 (0.1)
0.4	16	0.8	3.38 (2.4)	1.76 (3.2)	0.56 (4.5)	0.26 (0.2)
0.5	16	0.85	5.65 (10.9)	3.84 (8.9)	2.08 (8.9)	0.66 (1.1)
0.1	32	1	1.55 (2.5)	0.22 (0.3)	0.05 (0.1)	0.02 (0.0)
0.2	32	1	1.78 (2.5)	0.48 (0.6)	0.15 (0.6)	0.05 (0.1)
0.3	32	1	2.23 (2.1)	0.85 (0.9)	0.30 (1.4)	0.10 (0.0)
0.4	32	0.8	3.07 (4.0)	1.61 (3.1)	0.26 (5.6)	0.21 (0.1)
0.5	32	0.9	4.51 (10.0)	2.92 (7.2)	1.52 (7.7)	0.46 (0.9)
0.1	36	1.1	1.86 (2.8)	0.21 (0.2)	0.10 (0.1)	0.02 (0.1)
0.2	36	1.1	2.22 (2.8)	0.47 (0.5)	0.16 (0.5)	0.07 (0.1)
0.3	36	1	2.31 (2.8)	0.91 (1.4)	0.35 (2.0)	0.11 (0.1)
0.4	36	0.8	3.34 (5.5)	1.74 (4.5)	0.51 (6.8)	0.26 (0.3)
0.5	36	0.9	5.09 (14.4)	3.34 (11.1)	1.83 (11.8)	0.62 (2.1)

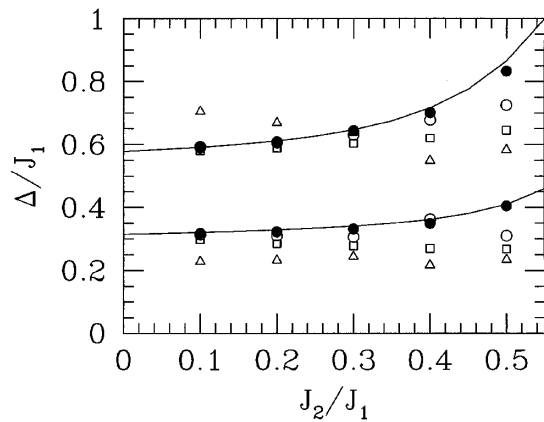


FIG. 1. Estimate of the spin gap for several methods: variational (empty triangles), FN (empty squares), GFMCSR  $p = 1$  (empty dots), GFMCSR (full dots) as in Table I for  $L = 16$  (upper points) and  $L = 32$  (lower points). The exact results are connected by continuous lines.

The way GFMCSR reaches the large  $n$  limit (at fixed number of operators  $p$ ) is displayed in Fig. 2 where the initial  $n = 0$  distribution was obtained by the FN for  $\gamma = 0$ . For fixed  $p$  the algorithm is Markovian and reaches an equilibrium distribution for  $n \rightarrow \infty$ , independent of the initial one (see example in Fig. 2 where  $p$  was changed at the iteration indicated by the arrow); this in turn will converge to the ground state distribution for large  $p$ . A comparison with the standard “release nodes” estimate is also shown in the picture. It is clear that there is no hope to obtain meaningful results in this case by the direct sampling of the sign.

On the contrary this method looks very stable and, though approximate, a convergence to a reasonable accuracy is obtained even with a very small number of operators, compared to the dimension of the Hilbert space.

The data shown in Table I and Fig. 2 indicate that the accuracy of GFMCSR may become rather size independent with a relatively small increase of the operator num-

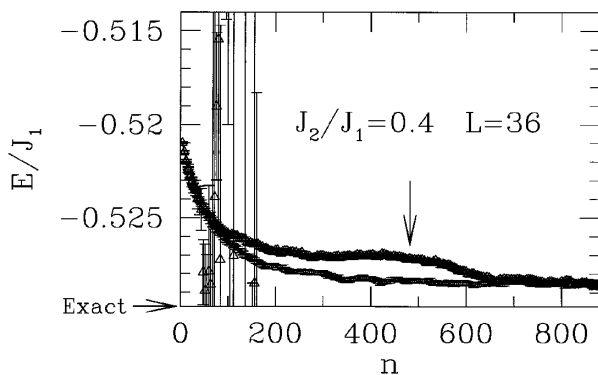


FIG. 2. Energy per site vs  $n$  for GFMCSR with  $p = 1$  (upper curve to the left of the arrow) and  $p = 9$  (remaining curves). The triangles represent the standard method with sign problem, i.e., with large error bars already for  $n > 15$ .

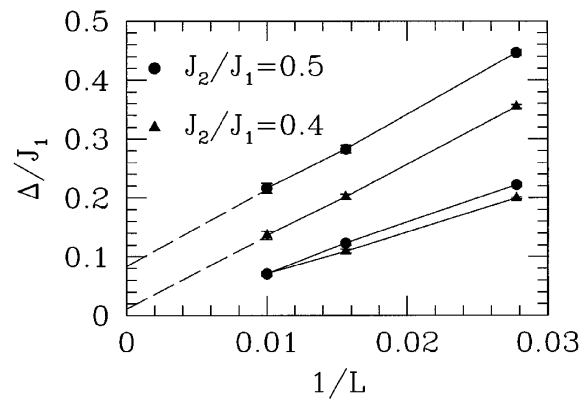


FIG. 3. Size scaling of the spin gap. The dashed lines are linear fits of the GFMCSR data with  $p = 9, 14,$  and  $20$  for  $L = 36, 64,$  and  $100$ , respectively. Lower curves are the variational estimates and the continuous lines are guides to the eye.

ber  $p$ . The error to work at finite small  $p$  is systematic. Thus there is a considerable cancellation of this error for the determination of the spin gap displayed in Fig. 1.

The calculation was therefore extended to the large size system up to  $L = 100$  where exact diagonalization is not possible. The spin gap as a function of the system size is displayed in Fig. 3. This figure is consistent with the opening of a finite spin gap for  $J_2/J_1 \geq \sim 0.4$ . This gap is certainly not an artifact of the variational WF, which is obviously gapless, as also confirmed numerically in the same figure. The present numerical results confirm that the transition to a spin liquid state with a finite spin gap but no classical order parameter should be close to  $J_2/J_1 = 0.4$  [10].

This work was supported in part by INFM (PRA HTSC) and CINECA grant. Thanks to C. Cavazzoni for technical help on SGI-origin.

- [1] N. Trivedi and D. Ceperley, Phys. Rev. B **41**, 4552 (1990).
- [2] K. Runge, Phys. Rev. B **45**, 7223 (1992).
- [3] H. van Bemmelen *et al.*, Phys. Rev. Lett. **72**, 2442 (1994).
- [4] D. ten Haaf *et al.*, Phys. Rev. B **51**, 13039 (1995).
- [5] It is possible to prove that the method is variational also for  $\gamma > 0$ , with a simple extension of the proof in [4].
- [6] All the analysis remains unchanged if a guiding function  $\psi_G(x)$  is used for importance sampling. The matrix elements of all the operators  $O_{x',x}$  (including the GF) have to be changed accordingly:  $O_{x',x} \rightarrow \psi_G(x') O_{x',x} / \psi_G(x)$ .
- [7] M. Calandra and S. Sorella (to be published).
- [8] E. Dagotto and A. Moreo, Phys. Rev. Lett. **63**, 2148 (1989).
- [9] T. Nakamura *et al.*, J. Phys. Soc. Jpn. **61**, 3494 (1992).
- [10] J. Schulz *et al.*, J. Phys. I (France) **6**, 675 (1996).
- [11] F. Franjic and S. Sorella, Prog. Theor. Phys. **97**, 399 (1997).