Quantum Monte Carlo calculation of the long-range order in the Heisenberg antiferromagnet

Karl J. Runge
Lawrence Livermore National Laboratory, University of California, Livermore, California 94550
(Received 24 June 1991)

A Green's-function Monte Carlo (GFMC) method is used to compute the staggered magnetization $m^\dagger$ in the two-dimensional, spin-$\frac{1}{2}$ Heisenberg antiferromagnet on $L\times L$ square lattices, up to $L=12$. Unlike previous GFMC calculations, the present method, which uses the forward-walking algorithm is unbiased and projects out the exact, rotationally invariant ground state. These calculations provide confirmation of the existence of long-range antiferromagnetic order in the ground state. A known relationship between $m^\dagger$ and the leading finite-size correction, coupled with high-precision ground-state-energy calculations, is used to reduce the error in extrapolating to the thermodynamic limit. The data extrapolate to $m^\dagger=0.3075\pm0.0025$, only slightly different from the spin-wave-theory result, 0.3034. Several perfect singlet trial wave functions used to reduce the statistical error are discussed. A possible explanation as to why exact-diagonalization extrapolations tend to yield low values of $m^\dagger$ is presented.

I. INTRODUCTION

In response to the discovery of the high-T_c superconducting oxides, the recent flurry of theoretical work on the two-dimensional Heisenberg antiferromagnet (HAF) has shown convincingly that its ground state possesses long-range order (LRO).[1–7] In this paper I discuss another method to compute the long-range order parameter (the staggered magnetization $m^\dagger$), via quantum Monte Carlo simulation. The results provide confirmation that the LRO does indeed exist, along with a fairly accurate estimate of the value for $m^\dagger$.

The forward-walking method removes some of the drawbacks present in the previous methods of finite-temperature path-integral Monte Carlo[1,2,8] (PIMC) and of extrapolated Green's-function Monte Carlo (GFMC).[3,4] In the former the calculations are done at nonzero temperature and so one must take care that T is low enough (for a given sized system). It is rigorously known through the Mermin-Wagner theorem[9] that no LRO exists at any $T>0$ for the infinite lattice HAF in two dimensions. Some of the problems that may arise are discussed by Gross, Sanchez-Velasco, and Siggia.[2] One has $T=0$ quite accurately in the GFMC method, however, there exists a technical problem in that the spin configurations are not distributed by the ground-state wave function squared but rather by the ground-state wave function times a known trial wave function. Thus, the average of observables (such as spin-spin correlation functions) are biased unless the trial wave function is exceedingly close to the exact ground state. The bias is usually corrected for by extrapolating away to linear order the (presumed small) difference between the exact and trial wave functions. This procedure is potentially dangerous since from variational calculations it is known that one can obtain very low energies from wave functions both with and without LRO.[10] The extrapolated GFMC method also produces nonrotationally invariant averages in that the LRO order points primarily in the xy plane,[3,4] while it is known that the exact ground state has long-range correlations independent of the spin direction. One can test that the extrapolated predictions do not change as one "deoptimizes" the trial wave function but this method is not completely foolproof. The forward-walking algorithm[11] used here avoids these difficulties by the brute force production of spin configurations distributed by the exact, $T=0$, rotationally invariant singlet ground state.

The forward-walking calculations were begun in an attempt to provide additional evidence on the existence of LRO without having the potential pitfalls mentioned in the previous paragraph. By no means do I imply that the numerical predictions from the finite-temperature PIMC or the extrapolated GFMC are incorrect; in fact, my results agree with those calculations. The point is that since until recently the very *existence* of LRO in this system was uncertain,[12] it seems useful to provide verification by a more "robust" technique.

A somewhat surprising result that has emerged from all of the work on the two-dimensional HAF is that the value of $m^\dagger$ appears to be very close to the spin-wave-theory result $m^\dagger_{\rm SWT}=0.3034$ (the units are those in which the perfectly ordered Néel state has $m^\dagger=\frac{1}{2}$). The close agreement was unexpected because spin-wave theory should be least accurate for the low spin of $\frac{1}{2}$ and also because of the "precarious" nature of the LRO in two dimensions.[13] The value I find for the staggered order is slightly ($\approx 1.5\%$) above $m^\dagger_{\rm SWT}$. This result is consistent with the prediction of Hirsch and Tang[14] based on the difference between their sublattice-symmetric spin-wave-theory and exact results for small lattices with $N\leq 26$, and also consistent with the results of long perturbative series expansions about the Ising limit.[5,7]

The paper is organized as follows. In Sec. II the Green's-function Monte Carlo (or "projection Monte Carlo") algorithm is described including the error reduction technique of importance sampling. The exact

method of calculating expectation values of operators via forward-walking is discussed in detail in Sec. III. Trial wave functions used to accelerate the convergence are outlined in Sec. IV, then followed by the presentation and interpretation of the GFMC results in Sec. V.

II. COMPUTATIONAL METHOD

The antiferromagnetic Heisenberg Hamiltonian is

$$H=J\sum_{\langle ij\rangle}\mathbf{S}_i\cdot\mathbf{S}_j$$
$$=J\sum_{\langle ij\rangle}\frac{1}{2}(S_i^+S_j^-+S_i^-S_j^+)+J\sum_{\langle ij\rangle}S_i^zS_j^z\ . \tag{1}$$

S_i^x, S_i^y, S_i^z are the components of the vector spin-$\frac{1}{2}$ operator at site i. $S_i^{\pm}=S_i^x\pm iS_i^y$ are the raising and lowering operators. J is positive. The sum $\langle ij\rangle$ is over nearest-neighbor pairs on a square lattice consisting of $N=L^2$ sites. The GFMC algorithm described below may be trivially extended to anisotropic coupling in the z direction ($J_z\neq J$),[15] arbitrary spin, and d-dimensional bipartite lattices.

By performing the unitary transformation,

$$S_i^x\rightarrow -S_i^x,\quad S_i^y\rightarrow -S_i^y,\quad S_i^z\rightarrow S_i^z\ , \tag{2}$$

for all i residing on *one* of the two sublattices, the sign of the raising and lowering term in H is reversed. From this one can show that all matrix elements of the operator $K=C-H$ in the S^z basis are positive, where $C=(J/4)N_{\text{bond}}$ with $N_{\text{bond}}=2N=$ the total number of bonds on the periodic lattice. Furthermore, the operator K projects out the ground state $|\psi_0\rangle$ (in a given total magnetization sector), that is,

$$|\psi_0\rangle\cong K^M|\psi_T\rangle\ , \tag{3}$$

where $|\psi_T\rangle$ is any trial state with nonzero overlap with the true ground state. If M in Eq. (3) is taken large enough all excited states are projected out of ψ_T. For example, the contribution of the first excited state, with energy E_1 is negligible if

$$\langle\psi_1|\psi_T\rangle\left[\frac{C-E_1}{C-E_0}\right]^M\ll 1\ , \tag{4}$$

where E_0 is the ground-state energy.

The symbol $\mathcal{S}$ will be used to denote a spin configuration of the entire system: $\mathcal{S}=(S_1^z,\ldots,S_N^z)$. The positivity of K and of the ground-state wave function[16] in the $\mathcal{S}$ basis allows the construction of a Monte Carlo algorithm based on the interpretation of K as a transition probability and of the ground-state wave function[17] as the corresponding steady-state distribution.

Importance sampling[18] is achieved by the following similarity transformation of the matrix K in the $\mathcal{S}$ basis:

$$\tilde{K}(\mathcal{S}',\mathcal{S})=\psi_G(\mathcal{S}')K(\mathcal{S}',\mathcal{S})/\psi_G(\mathcal{S})\ , \tag{5}$$

where $K(\mathcal{S}',\mathcal{S})\equiv\langle\mathcal{S}'|K|\mathcal{S}\rangle$ and the "guiding function" $\psi_G(\mathcal{S})\equiv\langle\mathcal{S}|\psi_G\rangle$ is preferably a good approximate solution to the Schrödinger equation.

Equation (3) rewritten in the $\mathcal{S}$ basis and multiplication by $\psi_G(\mathcal{S})$ yields

$$f(\mathcal{S})\equiv\psi_G(\mathcal{S})\psi_0(\mathcal{S})\cong\sum\tilde{K}(\mathcal{S},\mathcal{S}_M)\tilde{K}(\mathcal{S}_M,\mathcal{S}_{M-1})\cdots\tilde{K}(\mathcal{S}_2,\mathcal{S}_1)\psi_G(\mathcal{S}_1)\psi_T(\mathcal{S}_1)\ , \tag{6}$$

where the sum is over all possible $\{\mathcal{S}_1,\ldots,\mathcal{S}_M\}$. The ground-state energy is given by

$$E_0=\frac{\sum E_{\text{loc}}(\mathcal{S})f(\mathcal{S})}{\sum f(\mathcal{S})}\equiv\langle E_{\text{loc}}(\mathcal{S})\rangle_f\ , \tag{7}$$

where

$$E_{\text{loc}}(\mathcal{S})\equiv\frac{\langle\mathcal{S}|H|\psi_G\rangle}{\langle\mathcal{S}|\psi_G\rangle}\ . \tag{8}$$

The function $E_{\text{loc}}(\mathcal{S})$ is called the "local energy:" it depends on the instantaneous configuration $\mathcal{S}$ and when averaged with respect of $f(\mathcal{S})$ yields the exact ground-state energy.

The GFMC algorithm may be described as follows.[19] A random "walker" is defined to be a spin configuration and weight pair: $(\mathcal{S}_\alpha,w_\alpha)$, where w_α is a positive real number. One starts with an initial population $\{\mathcal{S}_\alpha\}$ ($\alpha=1,\ldots,N_{\text{pop}}$) of walkers drawn (via, say, the Metropolis algorithm) from the known function $\psi_G(\mathcal{S})\psi_T(\mathcal{S})$ and all of the weights are set to unity. One then stochastically evolves each walker from $(\mathcal{S}_\alpha,w_\alpha)$ to $(\mathcal{S}'_\alpha,w'_\alpha)$, where $w'_\alpha=\beta w_\alpha$,

$$\beta=\sum_{\mathcal{S}}\tilde{K}(\mathcal{S},\mathcal{S}_\alpha)\ , \tag{9}$$

and $\mathcal{S}'_\alpha$ is sampled from the known, normalized transition probability

$$P(\mathcal{S}'_\alpha)=\tilde{K}(S'_\alpha,\mathcal{S}_\alpha)/\beta\ . \tag{10}$$

Note that for the present application $\tilde{K}(\mathcal{S}'_\alpha,\mathcal{S}_\alpha)$ is nonzero only when $\mathcal{S}'_\alpha=\mathcal{S}_\alpha$ or when $\mathcal{S}'_\alpha$ and $\mathcal{S}_\alpha$ differ by at most a single interchange of nearest neighbor $+$ and $-$ spins. $\tilde{K}$ is therefore very sparse. Updating all members of the population will be referred to as advancing one generation. The steady-state limit of the random process is a set of walkers distributed by the function $f(\mathcal{S})$, and so the average of $E_{\text{loc}}(\mathcal{S})$ over this set provides an estimate of the ground-state energy.

In the implementation of the above method one finds the weights w_α quickly get out of hand. Each weight has a tendency to grow or decline exponentially with time, and so after a certain number of generations only a few walkers possess most of the total weight. One wastes computational effort on the remaining walkers, and so some sort of reconfiguration should be done.[18,20] What I have done is to split a walker into two copies (each with

weight $w_\alpha/2$) whenever w_α exceeds 2. Walkers are combined when their weights get below $\frac{1}{2}$.[21] This form of reconfiguration leads to a fluctuating number of random walkers (branching). There is a chance, therefore, that either all of the walkers will die off or the number will exceed the maximum value one desires to keep in computer memory. This drawback may be removed by means of a second reconfiguration. One scheme is to use at the nth generation the kernel

$$\tilde{K}^{(n)}=g(n)\tilde{K} \tag{11}$$

rather than $\tilde{K}$. The factor $g(n)$ drives the population to the desired base level. This latter reconfiguration introduces a (usually small) bias to all estimated quantities because the factor $g(n)$ is correlated with the recent history of the walk. The bias may be systematically removed by taking larger base populations or by dividing out a sufficiently large number k of factors (i.e., $g(n)\cdots g(n-k)$ for a measurement at generation n) from all estimates.[22] In my calculations I have found there is often a small, but statistically significant, bias for $N_{\rm pop}$ as large as 6000. The bias increases with the system size L, evidently because of the reduced effective number of walkers resulting from longer autocorrelation times. The Appendix contains additional details.

The efficiency of the algorithm can be greatly improved by using better guiding functions. From Eqs. (7) and (8) it is evident that the closer ψ_G is to the exact ground state the smaller will be the fluctuations in the local energy. From Eqs. (5) and (9) it can be seen that no branching occurs in the limit $\psi_G\rightarrow\psi_0$.

I have taken the guiding function of the form

$$\psi_G(\mathcal{S})=\prod_{\langle ij\rangle}\exp(-bS_i^zS_j^z)\ , \tag{12}$$

that is, a nearest-neighbor Gutzwiller wave function.[23] The variational parameter b is chosen to give the lowest expectation value of energy $\langle\psi_G|H|\psi_G\rangle$.[24] Positive b induces antiferromagnetic correlations in the z components of spin. Given the simple form for ψ_G it is possible to sample the (importance sampled) kernel $\tilde{K}(\mathcal{S}',\mathcal{S})$ *exactly* by rejection. This is fortuante because it cuts down on unnecessary branching induced by including a factor such as $\psi_G(\mathcal{S}')/\psi_G(\mathcal{S})$ as an additional weight in β. The form in Eq. (12) also allows a straight-forward bookkeeping scheme that enables the sampling of $P(\mathcal{S}')$ and the updating of β to be independent of the system size N. Thus, the time to evolve a given number of walkers through a fixed number of generations is independent of N.[25]

Much better wave functions than Eq. (12) exist. The wave function in Eq. (12) does not yield a particularly low variational energy, is not rotationally invariant, and possesses no long-range order. The best wave functions involve long-range spin correlations.[26,27,10] For example, the one due to Manousakis[26] is given by

$$\psi_{\rm LR}=\prod_{1\le i<j\le N}\exp(-u_{ij}S_i^zS_j^z)\ , \tag{13}$$

where the quantities u_{ij} may be obtained numerically from finite lattice sums. The slow decay of u_{ij} ($\sim 1/r_{ij}$) leads to long-range order in the xy plane. The function $\psi_{\rm LR}$ is not rotationally invariant, however. Unfortunately, the extra computational cost (a factor of N) from using $\psi_{\rm LR}$ as a guiding function does not appear to be offset by the gain in efficiency due to diminished branching and shorter projection times.[28]

III. FORWARD WALKING

The quantity we are interested in computing is the staggered magnetization given by

$$m^\dagger\equiv\sqrt{\langle\psi_0|(\mathbf{M}^\dagger)^2|\psi_0\rangle}\ , \tag{14}$$

$$\mathbf{M}^\dagger=\frac{1}{N}\sum_{i=1}^{N}\epsilon_i\mathbf{S}_i\ . \tag{15}$$

The quantity ϵ_i is $+1$ if the site i is on one sublattice and is -1 if i is on the other sublattice. For the classical Néel state $m^\dagger$ is $\frac{1}{2}$. For rotationally invariant states, such as the true ground state, one may use $m^\dagger=\sqrt{3\langle\psi_0|(M_z^\dagger)^2|\psi_0\rangle}$, which is computationally convenient because $M_z^\dagger$ is diagonal in the random walk basis.

For operators like $M_z^\dagger$ that do not commute with H something must be done to project the trial state in $f(\mathcal{S})$ into ψ_0 [Eq. (6)]. This may be achieved by the forward-walking algorithm,[11] based on the equation

$$\langle\mathcal{O}\rangle\equiv\frac{\langle\psi_0|\mathcal{O}|\psi_0\rangle}{\langle\psi_0|\psi_0\rangle}\cong\frac{\langle\psi_T|K^M\mathcal{O}|\psi_0\rangle}{\langle\psi_T|K^M|\psi_0\rangle} \tag{16a}$$

or in the language of the GFMC random walk,

$$\langle\mathcal{O}\rangle=\frac{\sum W(\mathcal{S}_M)\tilde{K}(\mathcal{S}_M,\mathcal{S}_{M-1})\cdots\tilde{K}(\mathcal{S}_2,\mathcal{S}_1)\mathcal{O}(\mathcal{S}_1)f(\mathcal{S}_1)}{\sum W(\mathcal{S}_M)\tilde{K}(\mathcal{S}_M,\mathcal{S}_{M-1})\cdots\tilde{K}(\mathcal{S}_2,\mathcal{S}_1)f(\mathcal{S}_1)} \tag{16b}$$

where I have assumed $\mathcal{O}$ is an operator diagonal in the $\mathcal{S}$ basis and have set $\mathcal{O}(\mathcal{S})\equiv\langle\mathcal{S}|\mathcal{O}|\mathcal{S}\rangle$. $W(\mathcal{S})$ is a factor that accounts for projecting the trial wave function, rather than the guiding wave function, into the ground state. The factor is

$$W(\mathcal{S})=\frac{\psi_T(\mathcal{S})}{\psi_G(\mathcal{S})}\ . \tag{17}$$

Equation (16) may be interpreted as follows. For each walker sampled from $f(\mathcal{S})$ evaluate and store $\mathcal{O}(\mathcal{S})$, let the random walk proceed, and keep track of the lineage of each walker, and then take the *weighted* average of the $\mathcal{O}$'s with respect to how much each walker and its successors contribute to the population M generations after the measurement[29] [and reweight with $W(\mathcal{S})$ also]. One sees that if a walker and its successors all die off before generation M is reached then their measurement $\mathcal{O}(\mathcal{S})$ is excluded from the estimate. This fact reveals the primary drawback of the forward-walking algorithm, namely, since the population is held approximately constant, there will be some amount of time after which all members of the population have sprung from a *single* predecessor. Thus, all of the measurements of $\mathcal{O}$, save

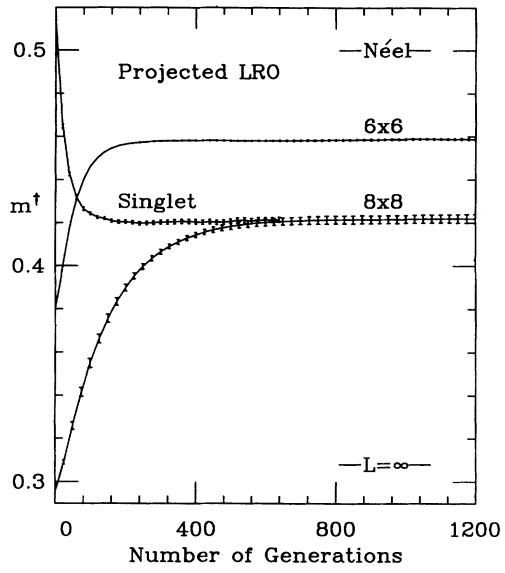

FIG. 1. Staggered magnetization (long-range order parameter) as a function of generation time in the forward-walking projection. The value at zero generations is the "mixed" ($\psi_T\psi_0$) estimate, while the asymptote at long time is the true singlet ground-state expectation value. For the 8×8 lattice results for both the nearest-neighbor Gutzwiller and Néel singlet projection functions ψ_T are shown. "Néel" "$L=\infty$" indicate the values for the classical Néel state and the infinite Heisenberg antiferromagnet, respectively.

one, are lost from consideration, which leads to large statistical (and often systematic) errors. It is therefore crucial to use a good guiding wave function ψ_G to minimize branching (and hence the rate at which families may die off) and to choose a good projecting wave function ψ_T to minimize the number of iterations M required to reach the ground state. The system size dependence of the statistical error is discussed in the Appendix. In a single simulation one may investigate a range of projection times M, spread the "origin of the forward walk" (where the measurements are made) over the entire run, and investigate a variety of projection trial functions ψ_T. In these ways correlated sampling is taken advantage of.[30] The program and algorithm were extensively checked against the known $m^\dagger=0.525\,858$ result for the $L=4$ case. The convergence of the forward-walking projection for $L=6$ and 8 and two different projection wave functions ψ_T is shown in Fig. 1.

IV. CHOICES OF ψ_T

In this work three different trial wave functions have been explored as projection functions. The first is simply the nearest-neighbor Gutzwiller function [Eq. (12)]. As it possesses no LRO, the mixed estimate of $(M_z^\dagger)^2$ vanishes in the thermodynamic limit. This fact, coupled with the knowledge[31] that the relative gap in the energy spectrum vanishes as $1/L^4$, implies longer forward-walking projection times are requred for larger lattices. For the lattice sizes examined here the projection time is dominated by the closing of the gap and so goes as L^4.

I have also used the long-range wave function $\psi_{\rm LR}$ of Manousakis [Eq. (13)]. Unfortunately, it proved to be only slightly better than the nearest-neighbor Gutzwiller function. This must be due to the fact that although $\psi_{\rm LR}$ has LRO, it lies primarily in the xy plane. Presumably the quantity $\langle\psi_{\rm LR}|K^M(\mathbf{M}^\dagger)^2|\psi_0\rangle$ would converge much faster than the z component alone, however, it would be quite difficult to implement this in a forward walk because the xy piece of the observable is not diagonal in the S^z basis. Perhaps an algorithm in which occasional steps are made with the observable rather than K could be implemented,[32] however, such a method was not attempted in this work.

The final wave function used was a perfect singlet Néel state. Let $|\text{Néel}\rangle$ denote either of the two Néel states in the S^z basis. A perfect singlet (i.e., rotationally invariant) version of this state may be obtained by "spherically averaging:"

$$|\psi_{\rm NS}\rangle\equiv\mathcal{P}|\text{Néel}\rangle\ , \tag{18}$$

"NS" denotes Néel singlet. The projection operator $\mathcal{P}$ is given by an equal amplitude summation over all possible rotations:

$$\mathcal{P}=\int d\Omega\int_0^\pi d\alpha(1-\cos\alpha)\exp(-i\alpha\mathbf{\Omega}\cdot\mathbf{S}_{\rm tot})\ , \tag{19}$$

where $\mathbf{S}_{\rm tot}=\sum_i\mathbf{S}_i$. The integrals may be readily worked out and the result is

$$\psi_{\rm NS}(\mathcal{S})\equiv\langle\mathcal{S}|\mathcal{P}|\text{Néel}\rangle=(-1)^{m(\mathcal{S})}\Phi(m(\mathcal{S}))\ . \tag{20}$$

The factor in front is simply the Marshall sign, that is, the phase introduced by the unitary transformation of Eq. (2). $m(\mathcal{S})$ is the number of up spins on one of the sublattices. $\Phi(m)$ is given by

$$\Phi(m)=\frac{n!\,m!}{(n+m+1)!}$$

with $n=(N/2)-m$.

Clearly $\psi_{\rm NS}$ is a very simple wave function since it does not depend at all on how the m up spins are *arranged* on the sublattice. Variationally, neither the energy nor LRO expectation values are changed (to order $1/N$) from the state $|\text{Néel}\rangle$. Nevertheless, it is a perfect singlet and has some useful properties. The lowest-lying excited states of the Hamiltonian [Eq. (1)] are states with nonzero total spin angular momentum (i.e., triplet and higher).[33] Since $\psi_{\rm NS}$ is rigorously orthogonal to those states the forward-walking projection time is shorter. The improvement is seen in Fig. 1. In fact, a factor of L is gained in convergence time since only the spin-wave-like excited states [with $(E_1-E_0)/N\sim c/L^3$ where c is the spin-wave velocity] need to be projected out. The triplet and higher total spin states have $(E_1-E_0)/N\sim1/L^4$. As mentioned above, the error in the forward-walking method increases dramatically with increased projection time since more "families" have died off. Unfortunately, the gain in efficiency from the use of $\psi_{\rm NS}$ was no more than a factor of 2 over that from ψ_G. The poor performance is probably due to the fact that $\psi_{\rm NS}$, while a perfect singlet, is a relatively poor wave function, and so there is much fluc-

tuation of the weights in Eq. (17). It is worthwhile to note that since ψ_{NS} has too much LRO and ψ_G has too little, one has a useful check of the algorithm in that they converge to the same values, as Fig. 1 demonstrates for $L=8$.

It would be useful both computationally and theoretically to develop more accurate singlet functions ψ_T. It is straightforward to apply the singlet projection operator $\mathcal{P}$ to Manousakis' long-range wave function ψ_{LR}, however I have not found a way to evaluate $\langle \mathcal{S}|\mathcal{P}|\psi_{LR}\rangle$ in a number of operations less than the size of the Hilbert space ($\sim 2^N$).[34] It is, however, possible to evaluate $\langle \mathcal{S}|RVB\rangle$ in $O(N^3)$ operations where $|RVB\rangle$ is the nearest-neighbor resonating valence bond state and is a singlet.[10] The evaluation of $\langle \mathcal{S}|RVB\rangle$ is equivalent to a dimer counting problem that may be performed via a Pfafian technique[35] (in other words, it is equivalent to evaluating the determinant of an $N\times N$ matrix that depends on $\mathcal{S}$). The forward-walking measurements may be spaced at intervals $\sim N$ generations (or more) and so the computation of $\langle \mathcal{S}|RVB\rangle$ may not be too costly. This possibility is presently being explored. Finally, it would be interesting to explore whether the *long*-ranged RVB wave functions may be implemented. These functions have been shown to yield the most accurate variational representation of the ground state to date.[10]

V. ANALYSIS OF RESULTS

The staggered magnetization from the forward-walking calculations are presented in Table I. I note that the value for the 4×6 lattice is in agreement with the recently corrected diagonalization result of Dagotto and Moreo.[36] The 4×6 result in Table I is close to the average of the 4×4 and 6×6 values, which is quite reasonable.[37,38] What now remains is to extrapolate the data in Table I to the thermodynamic limit.

By mapping the long-wavelength properties of quantum antiferromagnets onto the nonlinear σ model (presumably correct when long-range order does exist), Neuberger and Ziman[37] have derived a result for the leading finite-size correction of the staggered magnetization,

$$(m_L^\dagger)^2=(m^\dagger)^2\left[1+\kappa\frac{\alpha}{L}\right]+a\frac{(\ln L)^\gamma}{L^2}+\cdots, \qquad (21)$$

where $m^\dagger$ is the $L\to\infty$ limit of the finite system ground-state value $m_L^\dagger$. α is determined from (the large L limit of) various lattice sums and is a function of the aspect ratio. For $L\times L$ lattices the result is[37] $\alpha=0.6208$. κ is related to the finite-size dependence of the ground-state and higher spin excitation energies via,[37,39]

$$\frac{E_L(S)}{N}=\frac{E_0}{N}-\frac{1.438c}{L^3}+\frac{1}{2\chi}\frac{S(S+1)}{L^4}+\cdots. \qquad (22)$$

Neuberger and Ziman's result is $\kappa=1/c\chi$. In Eq. (22) S denotes the total spin of the system, and reflects the well-known fact that the absolute ground state is a perfect singlet. c is the spin-wave velocity and χ is the uniform (or perpendicular) susceptibility.

By performing GFMC calculations of the energy for various L and S I have extracted c and χ and find that $\kappa=9.20\pm 0.26$. The details of the computation will be published elsewhere. The calculation of the energy does not involve forward walking and so in general the statisticla error is smaller than for $(m_L^\dagger)^2$. The use of the constraint from Eq. (21) reduces the statistical error in the extrapolation $L\to\infty$ by at least a factor of 10 because it provides information about very large systems. The intercept of the fit, $m^\dagger$, is thus not as free to fluctuate as it would without the constraint. A least-squares fit of $(m_L^\dagger)^2$ to the form $A+B/L+C/L^2$ along with the constraint yields $m^\dagger=0.3075\pm 0.0025$.[40] The fit is plotted in Fig. 2 along with the modified spin-wave results and exact results for small lattices. The value I have found is consistent with the recent series expansion results of Singh and Huse[5] (0.303 ± 0.007) and of Zheng, Oitmaa, and Hamer[7] (0.307 ± 0.001).

The modified spin-wave theory (MSWT) of Arovas and Auerbach,[41] Takahashi,[42] and Hirsch and Tang[14] is the

TABLE I. Results of the forward-walking Green's function Monte Carlo staggered magnetization $m^\dagger$ for the spin-$\frac{1}{2}$, square-lattice Heisenberg antiferromagnet.

Lattice	$m^\dagger$
4×6	0.4841 ± 0.0008
6×6	0.4581 ± 0.0002
8×8	0.420 ± 0.001
10×10	0.397 ± 0.003
12×12	0.378 ± 0.014

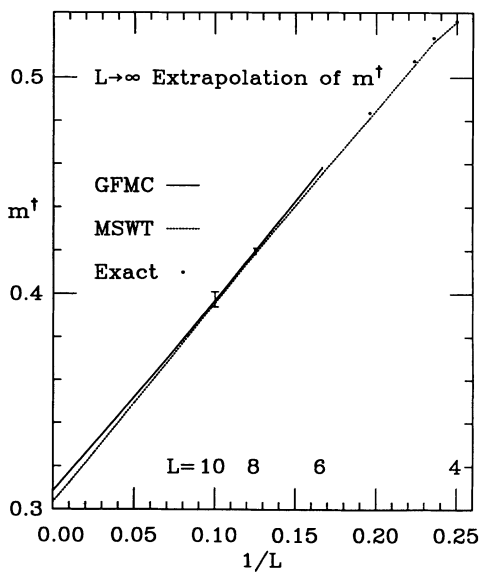

FIG. 2. Extrapolation of the staggered magnetization $m^\dagger$ to the thermodynamic limit. Circles are exact diagonalization results for small lattices. The dotted curve is the result of the modified spin-wave theory. The solid curve is the least-squares quadratic fit to the Green's-function Monte Carlo results of this paper. In these units the classical Néel state has $m^\dagger=\frac{1}{2}$.

usual second-order spin-wave calculation[43] except that a coupling of the $\mathbf{k}=\mathbf{0}$ magnons[44] to a staggered field is used to restore known symmetries of the problem, in particular, $\langle M_z^{\dagger}\rangle=0$. What results is a remarkably accurate theory for *finite* lattices and is, in fact, exact for systems up to $N=8$ sites. As $N\rightarrow\infty$ the MSWT reproduces the standard second-order spin-wave results. For lattices sizes in the range $16<N<30$ Hirsch and Tang observed that the MSWT yielded values of $m^{\dagger}$ about 1% lower than the exact diagonalization results. The value I find for $L=6$ ($N=36$) also deviates from the MSWT value by about the same amount. Hirsch and Tang speculated[14] that the small difference would persist in the thermodynamic limit. The value of $m^{\dagger}$ for $L=\infty$ that I have derived is consistent with their prediction.

Fitting the known $L=4$ value and the data in Table I to the form $A+B/L+C/L^2$ leads to a poor fit ($\chi^2\sim 15$). This result suggests the significance of higher-order terms, at least for $L\leq 4$. Fitting the data with an additional term D/L^3 allows one to compute a rough estimate of the magnitude of the higher-order corrections. Roughly, the result is $D=0.2$. Thus, the D/L^3 term is evidently quite significant for $L=2$ and so brings into question the validity (at least as far as *quantitative* predictions are concerned) of extrapolations[45] that set $D=0$ and include data with $L=2$–4.

I have obtained quite accurate values for the 6×6 lattice: $m^{\dagger}=0.4581(2)$, $E_0=-0.678\,872(8)$, and $E_1=-0.670\,878(10)$, where the number in parentheses denotes the error in the last digit(s). E_0 and E_1 are the ground-state energies per site in the singlet and triplet sectors, respectively. The 6×6 lattice may be exactly diagonalized soon by other workers and so these values should provide useful checks. With an accurate 6×6 result in hand one may examine how it may affect the extrapolation of the exact diagonalization values to $N=\infty$. The extrapolation of the exact diagonalization data (with $4\leq N\leq 30$) to $N=\infty$ tends to give a value of the staggered magnetization that is too low:[45] $m^{\dagger}=0.25\pm 0.03$. I have found that the inclusion of Lin's $N=32$ value[46] and my own $N=36$ ($L=6$) GFMC result does not improve the situation much at all. I believe the reason the exact diagonalization data lead to a small value for $m^{\dagger}$ is that the nonsquare lattices (that is, $N\neq L\times L$) tend to have *larger* $m^{\dagger}$ than do the perfect square lattices. By this I mean the values for the nonsquare lattices (especially for $10\leq N\leq 20$) lie mostly above the "smooth curve" passing through the $L\times L$ values. The tendency is displayed in Fig. 3. Considering $(m^{\dagger})^2$ as a function of $1/\sqrt{N}$ (the standard extrapolation technique[45]), one sees that the large values for the nonsquare lattices add negative curvature to the fit, which thereby lowers the $N=\infty$ intercept. Using *only* the perfect square lattices ($L=2,4,6$) and fitting to $A+B/L+C/L^2$ yields $m^{\dagger}=0.297$, which is in much better agreement with the GFMC and the series expansion techniques. However, given the conclusion of the previous paragraph that the D/L^3 term is significant, the good agreement is probably fortuitous.[47] Nevertheless, the trend is clear. It is possible that the lattice-shape-dependent $1/L$ coefficient [α in Eq. (22)] derived by Neuberger and Ziman[37] would alleviate the problems encountered in using the nonsquare lattices. However, since the $O(1/L^2)$ contribution is fairly large for these lattices, it may also be necessary to have some knowledge of how the higher-order terms depend on lattice shape.

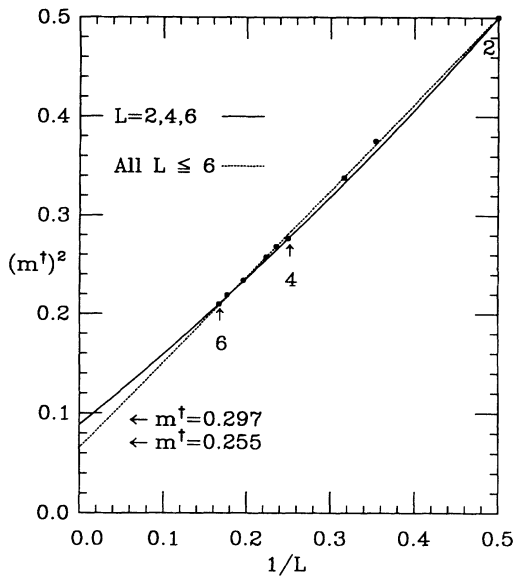

FIG. 3. Extrapolation of $(m^{\dagger})^2$ to the thermodynamic limit using small lattices only. $L=6$ is the present GFMC result, while the smaller system sizes are from exact diagonalization. The dotted curve is a quadratic fit to all of these system sizes, while the solid curve is a fit to the perfect square lattices ($L=2,4,6$) only. The former extrapolates to $m^{\dagger}=0.255$ and the latter to $m^{\dagger}=0.297$. The intercept from the larger lattice GFMC simulations is $m^{\dagger}=0.3085$ (see Fig. 2).

In summary, I have reported unbiased GFMC staggered magnetization data for the square-lattice spin-$\frac{1}{2}$ Heisenberg antiferromagnet on lattices ranging from $L=6$ to 12. The extrapolation of these data to $L=\infty$ indicates that there is long-range antiferromagnetic order in the ground state, thereby confirming the conclusions of previous numerical treatments. GFMC estimations of the spin-wave velocity and uniform susceptibility have been used to constrain the finite-size fit and lead to an accurate value for the staggered magnetization, with $m^{\dagger}=0.3075\pm 0.0025$. This value is slightly above the second-order spin-wave theory value of $m^{\dagger}_{\mathrm{SWT}}=0.3034$. Although it may be quite difficult to compute,[7] it would be interesting to see whether the prediction of third-order spin-wave theory moves closer to the "exact" numerical results or moves away indicating, perhaps, that the spin-wave series is actually asymptotic.

Note added in proof. After this work was completed, I received work prior to publication by H. J. Schulz and T. A. L. Ziman in whch they have diagonalized exactly the 6×6 lattice Their results for E_0, E_1, and $m^{\dagger}$ are in agreement with the values presented in this paper.

ACKNOWLEDGMENTS

I would like to thank M. P. Nightingale, J. Carlson, E. Manousakis, K. E. Schmidt, N. Trivedi, and V. Elser for

useful discussions on two-dimensional antiferromagnets and on the GFMC algorithm. This research was supported by the Cornell National Supercomputer Facility, a resource of the Cornell Theory Center, which receives major funding from the National Science Foundation and IBM Corporation, with additional support from New York State and members of its Corporate Research Institute. The work was also performed by the Lawrence Livermore National Laboratory under U.S. Department of Energy Contract No. W-4705-Eng-48.

APPENDIX

In this appendix I present a heuristic argument on how the statistical error σ_m (for a fixed amount of computing time) of the staggered magnetization $m^\dagger$ should depend on system size in the forward-walking algorithm.

Suppose a simulation of an $L \times L$ lattice is performed for $N_{\rm gen}$ generations with a base population of $N_{\rm pop}$ walkers. As mentioend in Sec. II, the computer time required to do this is independent of L. (If this were not the case, say by the use of a complicated ψ_G or ψ_T, a factor to take account of the extra cost must be included in σ_m.) If each walker at each generation were statistically independent from all others then σ_m^2 would be $\sigma_v^2/(NN_{\rm gen}N_{\rm pop})$. σ_v^2 is the mean-square deviation with a factor $1/N$ extracted. The factor $1/N$ is from the self averaging that usually occurs for intensive thermodynamic quantities such as $m^\dagger$. Due to the gap structure in the spectrum for finite systems, there should be an autocorrelation time $\tau \sim 1/\text{gap} \sim L^4$. τ is also the time required to project the trial state into the ground state via forward walking. Because of the autocorrelation only $\sim N_{\rm gen}/\tau$ of the generations are "independent." At least for the system studied here, and quite possibly generally, it has been shown[32] that the number of distinct families (a family is the set of walkers descended from the same predecessor) is very nearly inversely proportional to the generation time. So during a fluctuation lasting of length τ a walker passes the fluctuation on to its successors. Hence the number of independent walkers at any given time is really $\sim N_{\rm pop}/\tau$ (since the family basically "shares" the fluctuation). Finally, the *forward walk* suffers from diminished statistics due to the death of families (in addition to that described in the previous sentence) and an additional factor of τ must be included in σ_m^2. Thus,

$$\sigma_m^2 \sim \frac{1}{NN_{\rm gen}N_{\rm pop}}\tau^3 \sim L^{10}/N_{\rm gen}N_{\rm pop}$$

or[48] $\sigma_m \sim L^5$. For a quantity that does not suffer from the forward-walk error enhancement (such as the energy) the error scales as L^3, due to the removal of one factor of τ. These forms fit the observed errors for $L = 4-12$ very well. The rapid decrease of the number of independent observations $N_{\rm ind}$ with increasing system size suggests that great care should be exercised in the interpretation of simulation results, as they may quickly become meaningless (i.e., biased) when $N_{\rm ind}$ is not sufficiently large. The forward-walking algorithm yields good results for the HAF, and so it seems likely that it should be even more efficient for problems with a "larger" gap, for example, anisotropic systems. Work along these lines is presently underway.

[1] J. D. Reger and A. P. Young, Phys. Rev. B **37**, 5978 (1988).
[2] M. Gross, E. Sanchez-Velasco, and E. D. Siggia, Phys. Rev. B **40**, 11 328 (1989).
[3] N. Trivedi and D. M. Ceperley, Phys. Rev. B **40**, 2737 (1989); **41**, 4552 (1990).
[4] J. Carlson, Phys. Rev. B **40**, 846 (1989).
[5] R. R. P. Singh, Phys. Rev. B **39**, 9760 (1989); R. R. P. Singh and D. A. Huse, *ibid.* **40**, 7247 (1989).
[6] E. Manousakis, Rev. Mod. Phys. **63**, 1 (1991).
[7] Zheng Weihong, J. Oitmaa, and C. J. Hamer, Phys. Rev. B **43**, 8321 (1991).
[8] S. Miyashita, in *Quantum Simulations of Condensed Matter Phenomena,* edited by J. D. Doll and J. E. Gubernatis (World Scientific, Singapore, 1990), p. 228.
[9] N. D. Mermin and H. Wagner, Phys. Rev. Lett. **17**, 1133 (1966).
[10] S. Liang, B. Doucot, and P. W. Anderson, Phys. Rev. Lett. **61**, 365 (1988).
[11] M. H. Kalos, Computat. Phys. **1**, 257 (1966); K. S. Liu, G. V. Chester, and M. H. Kalos, Phys. Rev. A **10**, 303 (1974); P. A. Whitlock, D. M. Ceperley, G. V. Chester, and M. H. Kalos, Phys. Rev. B **19**, 5598 (1979).
[12] P. W. Anderson, Science **235**, 1196 (1987); Mater. Res. Bull. **8**, 153 (1973).
[13] That is, the LRO is destroyed at arbitrarily small but finite temperature in two dimensions.
[14] J. E. Hirsch and S. Tang, Phys. Rev. B **40**, 4769 (1989); S. Tang. M. E. Lazzouni, and J. E. Hirsch, *ibid.* **40**, 5000 (1989).
[15] In fact, the method can handle *any* function of the z components of spin, e.g., $V(\{S_i^z\})$.
[16] This fact follows from the Perron-Frobenius theorem of positive matrices. See, for example, F. R. Gantmacher, *The Theory of Matrices II* (Chelsea, New York, 1974).
[17] As opposed to the *square* of the ground-state wave function.
[18] D. M. Ceperley and M. H. Kalos, in *Monte Carlo Methods in Statistical Physics,* edited by K. Binder (Springer-Verlag, Berlin, 1979), Chap. 3.
[19] M. P. Nightingale and H. W. Blote, Phys. Rev. B **33**, 659 (1986).
[20] J. H. Hetherington, Phys. Rev. A **30**, 2713 (1984).
[21] For example, if two walkers have weights w_1, $w_2 < \frac{1}{2}$, then walker i $(=1,2)$ is selected with probability $w_i/(w_1+w_2)$, given weight w_1+w_2, and the other walker is deleted.
[22] M. P. Nightingale, in *Finite Size Scaling and Numerical Simulation of Statistical Systems,* edited by V. Privman (World Scientific, Singapore, 1990).
[23] M. C. Gutzwiller, Phys. Rev. Lett. **10**, 159 (1963).
[24] I have found that in nearly all cases minimizing the energy is essentially indistinguishable from minimizing the variance of $E_{\rm loc}$. In some applications the latter may be a better prescription.
[25] However, since the gaps in the excitation spectrum go to zero as the system size increases more generations are required to project into the ground state and breakup autocorrelations

for the larger systems. See the Appendix.
[26]E. Manousakis, Phys. Rev. B **40**, 4904 (1989); Z. Liu and E. Manousakis, *ibid.* **40**, 11 437 (1989).
[27]D. A. Huse and V. Elser, Phys. Rev. Lett. **60**, 2531 (1988).
[28]J. Carlson (private communication).
[29]Quantitatively this is the sum $\sum w_\alpha$ over the "family" that sprung M generations after the initial measurement.
[30]M. H. Kalos and P. A. Whitlock, *Monte Carlo Methods* (Wiley New York, 1986), Vol. I, pp. 102-109.
[31]M. Gross, E. Sanchez-Velasco, and E. D. Siggia, Phys. Rev. B **39**, 2484 (1989).
[32]K. J. Runge and R. J. Runge, in *Quantum Simulations of Condensed Matter Phenomena* (Ref. 8), p. 300. Please note that in this reference Eq. (34) should be corrected by replacing s by s_0 in the numerator and s_0 by 0 in the denominator.
[33]Note that, from the rotational invariance of the Hamiltonian, the higher spin states also reside in the $S^z_{\rm tot}=0$ subspace (i.e., where the GFMC is performed). To see this simply apply the total lowering operator $\sum_i S_i^-$ to the state until $S^z_{\rm tot}=0$ is reached.
[34]Stochastic evaluation of $W(S)$ over a sampling of the Hilbert space appears to introduce negative weights. It is not known how large a problem this introduces.
[35]M. E. Fisher, Phys. Rev. **124**, 1664 (1961).
[36]E. Dagotto and A. Moreo, Phys. Rev. B **38**, 5087 (1988); **44**, 5396(E) (1991).
[37]H. Neuberger and T. Ziman, Phys. Rev. B **39**, 2608 (1989).
[38]I thank R. J. Gooding for pointing out to me the $m^\dagger$ value for the 4×6 lattice in Ref. 36.
[39]D. S. Fisher, Phys. Rev. B **39**, 11 783 (1989).
[40]Replacing the third term in the fit by $C\ln(L)/L^2$ yields a value consistent with the C/L^2 fit. The result is $m^\dagger=0.3060\pm 0.0035$. Because the range of variation of $\ln(L)$ is small over the range of lattice sizes examined, its significance is questionable.
[41]D. P. Arovas and A. Auerbach, Phys. Rev. B **38**, 316 (1988).
[42]M. Takahashi, Phys. Rev. B **40**, 2494 (1989); J. Phys. Soc. Jpn. **58**, 1524, 1989.
[43]T. Oguchi, Phys. Rev. **117**, 117 (1960).
[44]In the standard spin-wave treatment the $\mathbf{k}=\mathbf{0}$ modes are completely ignored. The rotational symmetry of the Hamiltonian means that these modes experience no restoring force and hence have infinite mean-square deviations. In the MSWT the role of the staggered field is to regularize the divergence.
[45]S. Tang and H. Q. Lin, Phys. Rev. B **38**, 6863 (1988); S. Tang and J. E. Hirsch, *ibid.* **39**, 4548 (1989).
[46]H. Q. Lin, Phys. Rev. B **42**, 6561 (1990).
[47]Also, with only three data points ($L=2,4,6$) it is difficult to estimate the size of the systematic error.
[48]For a perfect singlet ψ_T one would evidently have $\sigma_m \sim L^{4.5}$.