Ground-state correlations of quantum antiferromagnets: A Green-function Monte Carlo study

Nandini Trivedi* and D. M. Ceperley
Department of Physics and National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign, 1110 West Green Street, Urbana, Illinois 61801
(Received 7 August 1989)

We have studied via a Green-function Monte Carlo (GFMC) method the $S=\frac{1}{2}$ Heisenberg quantum antiferromagnet in two dimensions. We use a well-known transformation to map the spin problem onto a system of hard-core bosons that allows us to exploit interesting analogies between magnetism and superfluidity. The GFMC method is a zero-temperature stochastic method that projects out the component of the true ground state in a given variational wave function. This method is complementary to previously used finite-temperature Monte Carlo methods and is well suited to studying the ground state and low-lying excited states. Starting with even a simple wave function, e.g., the classical Néel state, the GFMC method can obtain the short-range correlations very accurately, and we find the ground-state energy per site $E_0/J=-0.6692(2)$. We show that it is important to include the zero-point motion of the elementary excitations in the ground state and by a spin-wave analysis find that it produces long-range correlations in the wave function. Upon inclusion of such long-range correlations, we obtain a staggered magnetization $m^{\dagger}=0.31(2)$ and the structure factor $\mathcal{S}(q)\sim q$ at long wavelengths. Using the Feynman-Bijl relation, from the slope we deduce the renormalization of the spin-wave velocity by quantum fluctuations to be $Z_c\equiv c/c_s=1.14(5)$.

I. INTRODUCTION

The subject of quantum antiferromagnetism has been a challenging problem for several decades. With the recent discovery of high-temperature superconductors there has been a revival of interest in this area.[1] It was suggested by Anderson[2] that the $S=\frac{1}{2}$ Heisenberg quantum antiferromagnet (QAFM) on a square lattice could be used to model the interactions between the electrons in the singly occupied $d_{x^2-y^2}$ orbitals on the copper atoms. In one dimension, it is known that the ground state lacks long-range order (LRO), and the correlation function decays algebraically for the $S=\frac{1}{2}$ chain.[3] In two dimensions, using certain sum rules and bounds on correlation functions, it has been proved[4] that the $S=\frac{1}{2}$ XY model has LRO and the XXZ model has LRO for an anisotropy parameter $\Delta>1.78$. So far, a rigorous result pertaining to LRO in the $S=\frac{1}{2}$ Heisenberg model ($\Delta=1$) in $d=2$ has not been shown.[5] It is known, however, that the ground state of the $S=\frac{1}{2}$ Heisenberg QAFM on a bipartite lattice in any dimension is a *singlet*.[6,7]

Recently, from a study of the Heisenberg QAFM on a square lattice through a variety of numerical techniques, there seems to be a growing consensus that the ground state indeed has a staggered magnetization that is approximately 60% of the classical value. It therefore appears that in two dimensions, quantum fluctuations reduce the moment but do not completely destroy the long-range order. Exact diagonalization has been carried out on small clusters[8] up to $N=4\times4$. These results, while they leave the question about the ground state of the infinite system unresolved, provide a useful check for numerical work done by other approximate methods. Variational calculations have used trial states with LRO built in,[9] Gutzwiller wave functions,[10] and resonating-valence-bond states with both short- and long-range singlet bonds.[11] Finite-temperature path-integral Monte Carlo[12,13] methods have allowed lattices up to $N=24\times24$ to be studied. A perturbation series approach has also been developed around the Ising limit[14,15] to study this model.

In this paper, we use a different method from those reported previously called Green-function Monte Carlo (GFMC) method.[16] Variants of this method have been applied to this system.[13,17,18] It has several features that make it attractive to study the problem at hand: It can be used to study large lattices (cf. exact diagonalization methods), it does not require an extrapolation to zero temperature (cf. finite-temperature quantum Monte Carlo), it is also possible to devise an algorithm without zero-time-step error, and, most importantly, this method takes us beyond variational methods to obtain the *exact* ground-state properties. The basic idea behind the GFMC method is to filter out of a variational wave function Ψ_T the component of the true ground state that is contained in it. This can be done by operating on Ψ_T with an operator like $\hat{G}=\exp(-\tau H)$. To see its effect, expand the variational state in a complete set of basis states of the Hamiltonian $H|\alpha\rangle=E_\alpha|\alpha\rangle$. We then find after applying the operator n times on the variational state

$$e^{-n\tau H}|\Psi_T\rangle\rightarrow a_0 e^{-n\tau E_0}|0\rangle\ . \tag{1.1}$$

In the long time limit the ground state will dominate provided the overlap of the initial variational state with the

ground state is nonzero, i.e., $a_0=\langle 0|\Psi_T\rangle\neq 0$. For a large lattice (say, more than 30 sites) it is not possible to apply the $\hat{G}$ directly to the wave function many times, so one must resort to a Monte Carlo technique. The GFMC technique is a statistical method used to generate a set of configurations of spins that are distributed proportional to the true ground-state wave function.

To implement this GFMC procedure and also to gain physical insight we exploit an exact mapping of the $S=\frac{1}{2}$ spin Hamiltonian onto hard-core bosons on a lattice.[19] An up spin is identified with a boson and a down spin with an empty site. The staggered magnetization is analogous to off-diagonal long-range order in the boson problem. We note two important ingredients in the construction of a good trial wave function: The presence of *short-range* spin-spin correlations are incorporated via a Jastrow pair-correlated function for the bosons. We also show that the zero-point motion of the elementary excitations—spin waves in the antiferromagnet and phonons in the bose system—must be included in the ground state. This is found to lead to *long-range* correlations in the wave function, and we obtain the form of these correlations by a spin-wave analysis. The phase of the wave function on a square lattice is known exactly[6] to be positive on the A sublattice and negative on the B sublattice. The amplitude can be approximated by a product of pair correlations $\Psi_T=\prod_{i<j}f(\mathbf{r}_i-\mathbf{r}_j)$ with $f(0)=0$, where i and j are the positions of the up spins. The spin waves are shown to produce a long-range tail of the form $f(r)\sim 1-\alpha/r$ at large distances, and α is related to the spin-wave velocity.

Our main conclusions for the $S=\frac{1}{2}$ Heisenberg Hamiltonian based on a Green-function Monte Carlo method on lattice sizes up to $N=12\times 12$ and finite-size scaling are as follows.

(i) The ground-state energy per sites is $E_0/J=-0.6692\pm 0.0002$. This result is obtained even when the classical Néel state with an energy $E_N=-0.5J$ is chosen as the initial trial wave function.

(ii) The ground state is found to have long-range order with a staggered magnetization given by $m^{\dagger}=0.31\pm 0.02$ in units in which the classical Néel state has $m^{\dagger}=0.5$. This value is obtained only upon a proper inclusion of the long-range correlations in the wave function.

(iii) We find that the spin-wave velocity is $c/c_s=1.14\pm 0.05$, where $c_s=\sqrt{2}Ja$ is the classical spin-wave velocity about the Néel state. The structure factor at long wave lengths is $\mathcal{S}(\mathbf{q})\sim q$ and from the slope we deduce c by a generalization of the f-sum rule or Feynman-Bijl relation to the lattice

Some of these results were first reported in a short paper.[20]

This paper is organized as follows: In Sec. II, we review the mapping of the $S=\frac{1}{2}$ Heisenberg spin Hamiltonian on a square lattice onto a hard-core boson model. In Sec. III, we discuss trial wave functions for bosons by drawing close analogies with ^{4}He. We start by describing wave functions with short-range pair correlations arising from the hard-core nature of the bosons. We also discuss the effect of the zero-point motion of the elementary excitations in the ground-state wave function. The computational method—variational Monte Carlo (VMC) and Green-function Monte Carlo are discussed in Sec. IV. In particular, we describe a projection operator appropriate for a lattice and the sampling of the Green function. Our algorithm differs from previous GFMC calculations that were adapted to continuum systems. Importance sampling is introduced in the random walks to reduce the variance. The estimators for the energy and correlation functions are also discussed in Sec. IV. Section V contains our results for the ground-state energy, staggered magnetization, and the diagonal and off-diagonal spin-spin correlation function. In Sec. VI we discuss our results for the spin-wave velocity calculated from the finite-size scaling of the ground-state energy and from the long-wavelength behavior of the structure factor and momentum distribution function. We also comment on the modification of the f-sum rule on a lattice. In Sec. VIII, we give a brief discussion of the estimation of errors in the Monte Carlo procedures and conclude in Sec. VIII with a few remarks. In the Appendix, we describe the Ewald summation technique to handle long-range correlations in the wave function.

II. MODEL: TRANSFORMATION TO A BOSON MODEL

The Hamiltonian for the $S=\frac{1}{2}$ Heisenberg quantum antiferromagnetic is given by

$$H=J\sum_{\langle ij\rangle}\mathbf{S}_i\cdot\mathbf{S}_j=\frac{J}{2}\sum_{\langle ij\rangle}(S_i^+S_j^-+S_i^-S_j^+)+J\sum_{\langle ij\rangle}S_i^zS_j^z\ , \tag{2.1}$$

where the coupling J is positive and connects only spins on nearest-neighbor sites of a square lattice. We use periodic boundary conditions along both x and y directions.[21] The classical Néel state (with up spins on the A sublattice and down spins on the B sublattice) minimizes the $S_i^zS_j^z$ term in (2.1), but is not an eigenstate of the Hamiltonian because the xy term mixes spins on the A and B sublattices. We transform the Hamiltonian in (2.1) to a boson model. The spin-raising and -lowering operators on different sites commute and are thus identical to boson operators with the identification $b_i^{\dagger}=S_i^+$ and $b_i=S_i^-$, where $b_i^{\dagger}$ (b_i) is a boson creation (destruction) operator on site i. In addition, note that $S_i^z=S_i^+S_i^--\frac{1}{2}$, so that the z component of the spin gets identified with the boson number operator, i.e., $S_i^z=n_i-\frac{1}{2}$, where $n_i=b_i^{\dagger}b_i$. However, $S_i^{\pm}$ anticommute on the same site, which implies that $(b_i^{\dagger})^2=0$. This shows that the bosons have a hard core whereby a site cannot be occupied by more than one boson. (In general, for a spin S operator, a site can be occupied by at most $2S$ bosons.) This analogy between the spin variables and the hard-core boson variables was first presented in the context of ^{4}He by Matsubara and Matsuda,[19] though in that study a boson system was transformed to a *ferromagnetic* spin problem. The boson transformation discussed above is identical to the Schwinger representation[3] for $S=\frac{1}{2}$. The latter is a more general representation in which each spin is replaced by

two species of bosons a and b and is applicable to arbitrary S with the constraint $S=(n_a+n_b)/2$.

With the aforementioned transformation, the first term in the Hamiltonian in (2.1) that exchanges two spins of opposite signs on adjacent sites is identified as a hopping term for bosons.[22] Usually, we associate a negative sign with the kinetic energy operator because by hopping a boson can get delocalized and thereby reduce its energy. This is achieved by a unitary transformation (that leaves the commutation relations unchanged) in which the spins in the original Hamiltonian (2.1) on the B sublattice are rotated as follows: $S_j^x \rightarrow -S_j^x$, $S_j^y \rightarrow -S_j^y$, and $S_j^z \rightarrow S_j^z$. In terms of the boson operators on the B sublattice, we have, $b_j^\dagger \rightarrow -b_j^\dagger$ and $b_j \rightarrow -b_j$.

Thus with the transformation $S_i^+ = \varepsilon_i b_i^\dagger$ and $S_i^z = n_i - \frac{1}{2}$ to (2.1), where $\varepsilon_i = \pm 1$ on A/B sublattices, we have

$$H = -\frac{J}{2} \sum_{\langle ij \rangle} (b_i^\dagger b_j + b_i b_j^\dagger) + J \sum_{\langle ij \rangle} n_i n_j + E_N \ . \tag{2.2}$$

In this equation, $E_N = -JZN/8$ is the energy of the classical Néel state, where N is the total number of sites and Z is the coordination number for the given lattice ($Z=4$ for a square lattice). The second term in (2.2) is interpreted as a repulsive interaction between bosons on adjacent sites. In the Néel state the total z component of the spin is zero. The xy part of the Hamiltonian introduces quantum fluctuations in the Néel state that must be included to obtain the correct ground state. Since the xy term only exchanges spins on different sublattices, $(S^z)_{\text{tot}}$ is conserved. We will therefore work in the $(S^z)_{\text{tot}}=0$ subspace, which implies that the number of bosons N_b is half the number of sites. The computational methods described in the following sections can be used to study an arbitrary value of S_z. In the next section, we discuss possible variational wave functions for (2.2). The "Marshall sign rule"[6] is a statement about the phase of the ground-state wave function of the Heisenberg Hamiltonian. It implies that the many-body boson wave function can be chosen to be non-negative everywhere in the ground state.[23] We will see later that this property of the wave function will be crucial in applying the GFMC method to this problem.

III. CHOICE OF TRIAL WAVE FUNCTIONS

A. Pair-correlated wave functions

A basic problem confronting a variational calculation is a good functional form for a trial wave function. To construct a wave function for the many-body interaction boson system, we start by including short-range pair correlations. As discussed in Sec. II, the interactions between bosons are of a hard-core nature; therefore the problem cannot be handled by perturbative techniques. We resort to a variational method to study this problem and use trial wave functions first suggested by Bijl and Jastrow[24] for liquid ^{4}He and nuclear physics

$$\Psi_T(R) = \prod_{i<j} f(\mathbf{r}_i - \mathbf{r}_j) \ , \tag{3.1}$$

where f is a function of the relative distance between bosons i and j located on the sites of a square lattice. The pair-correlated wave function in (3.1) is symmetric under the exchange of any two-particle coordinates and also possesses particle-hole symmetry. It has the advantage that it depends on a single function f and is simple to compute with. The hard-core condition requires that $f(0)=0$, and, for large separation between the bosons, f tends to unity by convention. For intermediate values of the separation, the function f can be determined by a variational calculation by minimizing the energy, or, as will be discussed in Sec. VI, by maximizing the overlap between the ground state and trial wave functions.

As a start, the function f is approximated by

$$f_{\text{SR}}(r) = \begin{cases} 0, & \text{if } \mathbf{r}=0 \\ f_1, & \text{if } \mathbf{r}=\vec{\delta} \\ 1, & \text{otherwise} \ , \end{cases} \tag{3.2}$$

with a single variational parameter $f(\delta)=f_1$ at the nearest-neighbor site δ. This choice of f builds in *short-range* correlations in the wave function. It is easy to see that if the amplitude for finding a boson at the nearest-neighbor site is zero, i.e., $f_1=0$ we get a state with perfect *diagonal* LRO in which one of the sublattices is occupied by bosons and the other one is empty. In terms of spins, this a Néel state with sublattice magnetization along the z direction. On the other hand, if $f_1=1$, the state has *off-diagonal* LRO and describes a Néel state in the xy plane. For values of $f(\vec{\delta})$ intermediate between zero and unity, the Jastrow wave function usually describes a superfluid possessing off-diagonal LRO but no diagonal LRO.

The presence of off-diagonal order can be deduced from the expectation value of the amplitude to create a boson at site 0 and destroy a boson at site l. For the wave function in (3.1) it is found that $\langle b_0^\dagger b_l \rangle \rightarrow \text{const} \neq 0$. The nonzero value of the correlation function at large distances is related to the condensate fraction of the superfluid. In terms of spins, the off-diagonal correlations describe the correlation between the x and y components of the spins and the constant is associated with the presence of sublattice magnetization in the xy plane. The depletion of the condensate due to repulsive interactions between the bosons is analogous to a reduction in the sublattice magnetization by quantum fluctuations. The absence of diagonal LRO can be deduced from the behavior of the density-density correlation function that vanishes at large distances between the bosons, i.e., $\langle n_0 n_l \rangle \rightarrow 0$. In terms of spins, it implies that the staggered magnetization in the z direction is zero.

B. Inadequacy of short-range functions

As discussed earlier, the variational wave function in (3.1) can describe a solid or a superfluid, i.e., depending on the value of the parameters, the state may or may not possess diagonal LRO. In either case, the longitudinal excitations of the system are phonons with a linear dispersion. It was suggested by Feynman[25] and Bijl[24] that the excitation spectrum can be obtained from the

structure factor in the *true ground state* by the relation

$$\omega_{\mathbf{q}}=\frac{\varepsilon_{\mathbf{q}}}{\mathcal{S}(\mathbf{q})} , \tag{3.3}$$

where $\varepsilon_{\mathbf{q}}\sim q^2$. Thus at long wavelengths, a linear dispersion $\omega_{\mathbf{q}}\sim q$, implies a linear dependence of $\mathcal{S}(\mathbf{q})$ on the wave vector.

In the state described by (3.1), if the pair function $f(\mathbf{r}_i-\mathbf{r}_j)$ deviates from unity only locally around a given boson, as in (3.2), it can be shown that the structure factor $\mathcal{S}_{\rm SR}$ for such a wave function goes to a nonzero constant at long wavelengths. From the Feynman-Bijl relation this implies that the excitation spectrum is quadratic in q. Such a spectrum, even though gapless, is not consistent with a ground state that supports phononlike excitations or spin waves in the analogous antiferromagnetic spin system.

To see that $\mathcal{S}_{\rm SR}$ indeed shows such an incorrect behavior, note that (3.2) is analogous to the partition function of a classical system of particles interacting via a pair potential $\phi(r_{ij})$ and at a temperature T such that

$$(k_B T)^{-1}\phi(r)=2\ln f(r) .$$

(The factor of 2 arises because the classical partition function is related to the square of the wave function.) Since for most reasonable values of the parameters in (3.1) the system does not show diagonal LRO, we analyze the structure factor of a classical liquid interacting with pair potentials. The structure function can be related to the Ornstein-Zernike direct correlation function $c(\mathbf{q})$, which varies on the scale of the pair potential by[26] $\mathcal{S}(\mathbf{q})=[1-\rho_0 c(\mathbf{q})]^{-1}$. Using the rotational and inversion symmetries of a square, it can be shown that a long wavelengths $c(\mathbf{q})$ must be of the form

$$[\rho_0 c(\mathbf{q})]^{-1}\rightarrow a_0+a_2 q^2+\cdots$$

as $q\rightarrow 0$, where $\rho_0=N_b/A$. This implies that the structure factor is

$$\mathcal{S}_{\rm SR}(q\rightarrow 0)=a+bq^2\cdots , \tag{3.4}$$

where $a=a_0/(1-a_0)$ and $b=a_2/(1-a_0)$. Hence $\mathcal{S}_{\rm SR}(\mathbf{q}=0)\neq 0$ for a quantum bose system described by (3.2).

C. Zero-point motion of elementary excitations

The analysis given shows that the Jastrow wave function with only short-range pair correlations produces a structure factor that is inconsistent with phonons or spin waves as the elementary excitations. This points to the necessity of including the *zero-point motion* of elementary excitations in the ground-state wave function. In the context of ^{4}He, a scheme to include phonons in the wave function was suggested by Reatto and Chester.[27] We give a brief description of their method and apply it to the antiferromagnet in *two* dimensions.

The Hamiltonian for spin waves described as harmonic oscillators is usually written in terms of the normal displacements $\mathbf{u}$ and momenta. Since the density fluctuation $\rho\sim\nabla\cdot\mathbf{u}$, we can write the Hamiltonian for the collective coordinates representing the sound mode by

$$H_{\rm LR}=\tfrac{1}{2}\sum_{q<q_c} m_{\mathbf{q}}[\dot{\rho}(\mathbf{q})\dot{\rho}(-\mathbf{q})+\omega_{\mathbf{q}}^2\rho(\mathbf{q})\rho(-\mathbf{q})] , \tag{3.5}$$

where $\omega_{\mathbf{q}}=cq$ at long wavelengths and c is the velocity of spin waves and $\rho(\mathbf{q})=\sum_{i=1}^{N_b}\exp(i\mathbf{q}\cdot\mathbf{r})$. Also, since the density is related by a divergence to the displacement field, $m_{\mathbf{q}}\sim 1/q^2$. A cutoff in momentum space q_c has been imposed as (3.5) describes the excitations of the boson system only in the long-wavelength limit. The harmonic-oscillator wave function arising from these spin waves is of the form

$$\Psi_{\rm LR}=\prod_{q<q_c}\exp\left[-\frac{m_{\mathbf{q}}\omega_{\mathbf{q}}}{2}\rho(\mathbf{q})\rho(-\mathbf{q})\right] . \tag{3.6}$$

Performing the Fourier transforms in (3.6) in two dimensions, we get

$$\Psi_{\rm LR}=\exp\left[-\sum_{i<j}\frac{\alpha}{|\mathbf{r}_i-\mathbf{r}_j|}\right] . \tag{3.7}$$

In Sec. VI we relate α to the spin-wave velocity. We thus find that the inclusion of the zero-point motion of spin waves produces *long-range* pair correlations in the many-body wave function.

An improved variational ansatz for the ground-state wave function is to multiply (3.2) and (3.7). The function f in (3.1) is now approximated by

$$f_{\rm LR}(\mathbf{r})=\begin{cases}0, & \text{if } \mathbf{r}=0\\ f_1, & \text{if } \mathbf{r}=\vec{\delta}\\ \left[1-\dfrac{\tilde{\alpha}}{r}\right], & \text{otherwise ,}\end{cases} \tag{3.8}$$

where $\tilde{\alpha}=\exp(-\alpha)$. Both wave functions (3.2) and (3.8) include only pair correlations; however, in (3.8) in addition to the short range correlations on the order of a few lattice spacings, we have included a $1/r$ tail arising from the zero-point oscillations of the phonons. For the long-range part, we have approximated

$$f(\mathbf{r}_i-\mathbf{r}_j)\approx f(|\mathbf{r}_i-\mathbf{r}_j|)$$

on the lattice. We emphasize that the $1/r$ tail in the wave function in (3.7) directly implies a linear q dependence in the structure factor. A simple way to see this is to note that the potential energy E_p in (3.5) can be related to the static structure factor

$$\mathcal{S}(\mathbf{q})=(1/N_b)\langle\rho(\mathbf{q})\rho(-\mathbf{q})\rangle$$

by

$$E_p=\tfrac{1}{2}\omega_{\mathbf{q}}^2 m_{\mathbf{q}} N_b \mathcal{S}(\mathbf{q}) . \tag{3.9}$$

However, by the virial theorem, E_p is the ground-state energy $=\hbar\omega_{\mathbf{q}}/2$. This show that $\mathcal{S}(\mathbf{q})\sim q$ in the long-wavelength limit.

A wave function similar to (3.7) was studied by Huse

and Elser[9] with the long-ranged function taken to be of the form $V(r)=\exp(-\alpha/r^{\beta})$. Treating β as a variational parameter they obtained a value of $\simeq 0.7$, which is close to the value of $\beta=1.0$ that we find based on a spin-wave analysis. The Fourier transform of $V(q)$ in two dimensions $V(q)\sim q^{-(2-\beta)}$. Let $\mathcal{S}_{\rm SR}(\mathbf{q})$ be the structure factor in (3.4) of the reference system described by (3.2). The effect of including $V(q)$ on the structure factor is evaluated within linear response[26] to be

$$\frac{1}{\mathcal{S}(\mathbf{q})}=\frac{1}{\mathcal{S}_{\rm SR}(\mathbf{q})}+\frac{V(\mathbf{q})}{2} . \tag{3.10}$$

We see that since $\mathcal{S}_{\rm SR}(\mathbf{q})\rightarrow$ const$\neq 0$ at long wavelengths,

$$\mathcal{S}(\mathbf{q})=2/V(\mathbf{q})\sim q^{2-\beta}$$

as $q\rightarrow 0$. From the compressibility sum rule,[28] we have $\kappa\sim\mathcal{S}(\mathbf{q})/\omega(\mathbf{q})\sim q^{1-\beta}$. Thus for $\beta=1$ the system has a finite spin compressibility. For all other values the system is either incompressible $\kappa=0$ ($\beta<1$) or it diverges ($\beta>1$).

In the following sections, we will present results for the energy and other correlation functions in states both with and without spin waves to study the effect of long-range correlations in the wave function.

IV. COMPUTATIONAL METHODS

A. Variational method

Given a trial wave function for the ground state $\Psi_T(R)$, e.g., the wave function in (3.1) with f defined in (3.2) or (3.8), the expectation value of an operator in this state is given

$$\langle\hat{\mathcal{O}}\rangle_V=\frac{\langle\Psi_T|\hat{\mathcal{O}}|\Psi_T\rangle}{\langle\Psi_T|\Psi_T\rangle}=\sum_R P(R)\sum_{R'}\langle R|\hat{\mathcal{O}}|R'\rangle\frac{\Psi_T(R')}{\Psi_T(R)} , \tag{4.1}$$

where

$$P(R)=\frac{\Psi_T(R)^2}{\sum_R\Psi_T(R)^2} , \tag{4.2}$$

and $\{R=r_1,r_2,\ldots,r_{N_b}\}$ is a particular configuration of bosons. $P(R)$ is interpreted as a probability. The operator $\hat{\mathcal{O}}$ in (4.1) can be a diagonal operator, e.g., the potential energy or the density-density correlation function, or it can be an off-diagonal operator like the kinetic energy. The variational principle guarantees an upper bound for the energy. Notice that the expectation values require an evaluation of N-dimensional summations. If N is large (≥ 30), Monte Carlo methods[29] are required for estimating these summations.

The first step is to generate a collection of boson configurations that are distributed according to the probability in (4.2). We use the Metropolis algorithm to generate the configurations. Starting with an initial configuration of N_b bosons on the sites of an $N=L\times L$ lattice labeled by $\{R\}$, a new configuration R' is generated as follows: A boson is picked sequentially and moved to a neighboring site with equal probability. The move is accepted with probability q given by

$$q=\min\left[1,\frac{\Psi_T^2(R')}{\Psi_T^2(R)}\right] .$$

This process is repeated many times to generate a collection of about 10^4 configurations. The first few hundred configurations are discarded as these may be influenced by the choice of the initial configuration. The equilibrated configurations are used to calculate various averages and correlation functions described in Sec. V.

B. Green-function Monte Carlo

The GFMC method is a very accurate method to obtain the ground-state and low-lying excited-state properties of many-body interacting systems. In the past it has been applied successfully to the ground-state properties of helium,[30] both ^{3}He and ^{4}He. It has also been used to study the interacting electron gas and small molecules.[31]

Green-function Monte Carlo (GFMC) is a general scheme for finding the largest eigenvalue of an operator. A trial eigenvector is subjected to a time evolution whose effect is to enhance those components of the solution with the larger eigenvalues of the operator. To do this, we define an iterative procedure to obtain the wave function at the nth time step, given its value at the preceding time step by

$$\Phi^{(n)}=[1-\tau(H-w)]\Phi^{(n-1)} , \tag{4.3}$$

where w and τ are parameters chosen such that the projection operator G in square brackets satisfies certain properties to be discussed later. In terms of the exact eigenstates of the Hamiltonian $H|\alpha\rangle=E_\alpha|\alpha\rangle$ (4.3) can be written as

$$\Phi^{(n)}=\sum_\alpha[1-\tau(E_\alpha-w)]^n|\alpha\rangle\langle\alpha|\Phi^{(0)}\rangle . \tag{4.4}$$

As the projection operator is applied successively, (4.4) shows that the ground state starts dominating, and the excited states decay exponentially if the constants w and τ have been chosen appropriately. As $n\rightarrow\infty$, $\Phi^{(n)}$ approaches Φ_0, the exact ground-state wave function, provided the overlap of the initial trial state $\Phi^{(0)}$ with the true ground state is nonzero. The projection operator must satisfy the following conditions in order to filter out the excited states after many iterations. (i) The largest eigenvalue of G corresponding to the ground-state energy of H must be equal to unity. This implies that $w\simeq E_0$. (ii) All the other eigenvalues of G corresponding to the excited states of H must have absolute values strictly less than unity. This gives an upper bound on the value of the time step to be $\tau\leq 2/(E_{\rm max}-w)$, where $E_{\rm max}$ is the maximum eigenvalue. For the Heisenberg model $E_{\rm max}=NJ$ corresponding to the maximum potential energy we therefore choose the "time step" $\tau\simeq 1/NJ$. If τ does not satisfy these conditions, the excited-state components of the initial solution, even if present in very small amounts, get amplified and eventually dominate the distribution function.

It is important to note that the choice of G in (4.3) does *not* involve a small time step error. The unique solution to (4.3) is the exact ground state as long as τ is less than its critical value. For a continuum problem, however, it would be necessary to modify G to either $\exp[-\tau(H-w)]$ or $[1+\tau(H+w)]^{-1}$ since there is no maximum eigenvalue. Working with the projection operator in (4.3) for a lattice problem has the obvious advantage that it is simple to evaluate. It might also be thought that the method is slow as $N\to\infty$ since then $\tau\to 0$, but as we will show in the following sections, the computer time per move of a single boson is independent of N.

1. Sampling the Green function

Green-function Monte Carlo is a *stochastic* method to sample the projection operator or Green function in (4.3). It recasts the eigenvalue problem as a random walk process, whereby the most probable regions of space are sampled more often. To see how such a search through configuration space is achieved effectively using Monte Carlo methods, rewrite the integral equation (4.3) in the position basis of the bosons. We then have

$$\Phi^{(n)}(R)=\sum_{R'} G(R,R')\Phi^{(n-1)}(R') , \tag{4.5}$$

where the Green function G that propagates a configuration R' to R is given by

$$G(R,R')=\langle R|[1-\tau(H-w)]|R'\rangle . \tag{4.6}$$

The ground-state wave function for a Bose system and the propagator G are both non-negative by inspection. Thus $\Phi^{(n)}(R)$ can be treated as a *probability density*. We begin with a set of configurations distributed according to $\Phi^{(0)}(R')\equiv\Psi_T(R')$, i.e., the probability of finding a particular configuration R' at time $t=0$ is given by some known trial wave function. Each configuration is treated as a random walker in configuration space and the total number of walkers at each time $t=n\tau$ is described as the population in the nth generation. Our aim now is, given the initial set of random walkers distributed according to $\Phi^{(0)}$, to generate the next generation of walkers distributed according to $\Phi^{(1)}$, where the new generation of walkers is obtained from the preceding generation by evolving each configuration with the Green function in (4.6). The propagation of a random walker from a point R' in configuration space to R involves two basic processes. (a) Diffusion: These moves are governed by the kinetic energy operator that hops a boson from one site to another. (b) Branching: In the course of the random walks if a configuration evolves into a region with a high potential energy, such walks are terminated with some probability. On the other hand, if a starting configuration evolves into a configuration with a low potential energy, such favorable configurations are multiplied.

2. Importance sampling

As described earlier, the dynamics of the Green function allows each particle to move with equal probability to an unoccupied neighboring site. However, if a reasonably good guess for the wave function is available, it is much more efficient to guide the probability density according to this wave function, called a guiding wave function Ψ_G. Such a procedure leads to a significant reduction in the variance as first introduced by Kalos[16] and is known as importance sampling. The diffusion of random walks is thus biased, and regions of configuration space that are deemed "important" according to the guiding function are sampled more often. To introduce importance sampling in (4.5), change the probability density at the nth time step to $F^{(n)}=\Phi^{(n)}\Psi_G$. Equations (4.5) and (4.6) are modified to

$$F^{(n)}(R)=\sum_{R'} G(R,R')F^{(n-1)}(R') , \tag{4.7}$$

where the Green function is now given by

$$G(R,R')=\Psi_G(R)\langle R|[1-\tau(H-w)]|R'\rangle\frac{1}{\Psi_G(R')} . \tag{4.8}$$

As $n\to\infty$, the probability density approaches $F^*=\Phi_0\Psi_G$, where Φ_0 is the true ground-state wave function.

To implement the diffusion and branching processes in the propagation of the random walks, we write G as a product of a transition probability $P(R,R')$ and a multiplicity function $m(R,R')$. The actual multiplicity is $M=\llcorner m(R,R')+\chi\lrcorner$ where $\llcorner\cdots\lrcorner$ denotes that the integer part is to be taken (truncation) and χ is a random variable uniformly distributed in [0,1]. We make M number of copies of configuration R that become members of $F^{(n+1)}$. If M equals zero, that random walk is discarded. The average contribution is

$$\begin{aligned} G(R,R')&=\int_0^1 d\chi P(R,R')\llcorner m(R,R')+\chi\lrcorner \\ &=P(R,R')m(R,R') . \end{aligned} \tag{4.9}$$

Some of the previous work[13] has not included importance sampling in the diffusion of the random walkers leading to higher variance of the ground-state energy, as will be discussed in Sec. V.

3. Optimal choice of multiplicity

As seen from (4.9), only the product of the transition probability and multiplicity is predetermined; we therefore invoke an additional requirement to define these functions. We assume the optimal breakup of G into P and m is that which leads to the smallest mean-squared branching $\langle(m-\langle m\rangle)^2\rangle$. Thus from (4.8) and (4.9) we have for the transition from an old configuration R' to R,

$$m(R,R')=\{1-\tau[E_L(R')-w]\}=\frac{1}{C(R')} \tag{4.10}$$

and

$$P(R,R')=\begin{cases} C(R)\{1-\tau[U(R')-w]\}, & \text{if } R=R' \\ C(R)(\tau J/2)\Psi_G(R)/\Psi_G(R'), & \text{if } R=\mathcal{N}(R'). \end{cases} \tag{4.11}$$

The constant $C(R)$ is determined by normalizing the transition probability, i.e., from the condition $\sum_R P(R,R')=1$. In (4.10) $E_L=H\Psi_G/\Psi_G$ is the local energy with respect to the guiding function. At the optimal value of P, the multiplicity $m(R,R')$ is found to only depend on the old configuration R'. As the guiding function approaches the ground-state wave function the multiplicity approaches unity. In (4.11),

$$U(R')=J\left\langle R' \left| \sum_{\langle ij \rangle} n_i n_j \right| R' \right\rangle$$

is the potential energy in configuration R' and the second line in (4.11) is obtained from the effect of the kinetic energy operator in (2.2). The set $\mathcal{N}(R')$ is the "neighborhood" of R' and consists of all those configurations that can be obtained from R' by moving any of the bosons by one site in each of the four directions. Notice that for $R \neq R'$, the transition probability is proportional to $\Psi_G(R)/\Psi_G(R')$, which shows that the hopping of a boson is in the most likely direction as determined by the guiding function.

4. *Some technical points*

Our sampling of the Green function involves the following specific features.

(i) We start with an initial set of 1000 configurations or walkers $\{R'\}$ sampled from a distribution $F^{(0)}(R')=\Psi_G^2(R')$, where Ψ_G is the guiding function.

(ii) For each configuration in the initial generation, we construct a transition table of the $4N_b+1$ possible new configurations according to (4.11), where N_b is the number of bosons. We then sample P from this table, which gives us the boson to be moved and the direction of the move. After each move, the transition table is updated. To keep the update process efficient, the guiding function must be simple; we use a single-parameter Jastrow function in (3.2) with f_1 chosen at the optimal value.

(iii) The multiplicity of a configuration is calculated according to (4.10). However, we defer branching until later. Multiplicities are accumulated for K moves (i.e., we obtain a product of multiplicities in the K moves) and then perform branching on the accumulated weight. The value of K is chosen such that the rms value of the multiplicity is around 2. A larger rms value of the multiplicity is undesirable, since that implies that one is following walkers with low weight. On the other hand, too small a value would result in slower algorithm because of the overhead involved in branching. We used $K \simeq 10N_b$ so that on average each boson was moved ten times before a new configuration was generated.

(iv) In (4.11) we sample the number of times the random walk will remain at the same position from a Poisson distribution as

$$n=1+\lfloor \ln\chi/\ln P(R,R') \rfloor .$$

An expectation value is then weighted with the number of times the walk remained at R'. This is a considerable improvement in the algorithm, since for large N the walkers remain longer and longer at the same position. Then an off-diagonal move is sampled from (4.11).

(v) The population is defined as the total number of walkers in a generation, and its growth and decline is controlled by the constant w. We adjust the value of w over a sufficiently long period of about 200–300 generations in order to stabilize the population to within 1000 ± 200. Quite often a small initial sample size or a poor trial wave function, can lead to a sharp rise or decline in the population, but with some care a constant population can be quickly achieved.

The iteration of (4.7) is continued for approximately 1000 generations until the required variance is achieved. The first 200 generations, when the population has not yet stabilized, are discarded. The list of configurations obtained after the initial transient period are distributed according to $\Phi_0\Psi_G$, the product of the desired ground-state wave function and a known guiding function, and are used to evaluate the required expectation values discussed in the next section.

C. Ground-state properties

1. *Energy*

There are two methods of calculating the energy: It can be obtained from the behavior of the population $P_n=\sum F^{(n)}(R_i)$ as a function of the time step n, where the summation is over all configurations in the nth generation and $F^{(n)}=\Phi^{(n)}\Psi_G$. The rise or fall of P_n provides a *growth estimate* for the energy $E_g^{(n)}$ in the nth generation. From (4.7) and (4.8), it can be shown that

$$E_g^{(n)}=\frac{1}{\tau}\left[1-\frac{P_{n-1}}{P_n}\right]+w \ . \tag{4.12}$$

This estimate has a larger variance than the mixed estimate described later, however, it provides a useful check on the algorithm. (If the value of w is adjusted while stabilizing the population, that must be taken into account.)

The second method to calculate the energy is from the local or *mixed estimate.* This involves taking the expectation value of an operator between the ground state Φ_0 and a trial state Ψ_T. The mixed estimate is defined by

$$\langle E \rangle_0=\frac{\langle \Phi_0|H|\Psi_T \rangle}{\langle \Phi_0|\Psi_T \rangle} \ . \tag{4.13}$$

By applying the Hamiltonian to the left on Φ_0, we obtain the true ground-state energy. In the nth generation, the mixed estimate is given by

$$E_n[\Psi_T]=\frac{\sum \Phi^{(n)}H\Psi_T}{\sum \Phi^{(n)}\Psi_T}=\frac{\sum F^{(n)}WE_L[\Psi_T]}{\sum F^{(n)}W} \ , \tag{4.14}$$

where the summations are over all configurations, $E_L[\Psi_T]=H\Psi_T/\Psi_T$ is the local energy of Ψ_T. The expression in (4.14) is given for random walks with a guiding function, therefore we also get a weight factor $W(R)=\Psi_T(R)/\Psi_G(R)$. While a simple function Ψ_G is used for guiding the walks. a more accurate and complicated function Ψ_T is used to calculate the averages.

2. Correlation functions

For a general operator $\hat{\mathcal{O}}$, only the mixed estimate can be obtained by the GFMC method given by

$$\langle \mathcal{O} \rangle_M = \frac{\langle \Phi_0 | \hat{\mathcal{O}} | \Psi_T \rangle}{\langle \Phi_0 | \Psi_T \rangle} . \tag{4.15}$$

Note that since the operator is sandwiched between two different states, and since in general Φ_0 is not an eigenstate of $\mathcal{O}$, we have a mixed estimator. The mixed estimate can be corrected to give a true estimate, called the extrapolated estimate. By assuming that the difference between the trial function and the ground-state wave function is small it can be shown[16] that

$$\langle \mathcal{O} \rangle_0 \simeq 2\langle \mathcal{O} \rangle_M - \langle \mathcal{O} \rangle_V + O(\Psi_0 - \Psi_T)^2 ,$$

where the average with respect to V is the variational Monte Carlo result defined in (4.1). As seen from (4.14), if a weight factor is included, it is possible to choose the trial function to be different from the guiding function. We use a simple guiding function with short-range correlations (3.2) so that the configurations can be generated efficiently. More complicated wave functions, e.g., those described in Sec. III C in Eq. (3.8) are chosen, while evaluating the correlation functions so as to minimize the error from the mixed estimator.

The staggered spin operator at a site i is defined by $\tilde{\mathbf{S}}_i \equiv \epsilon_i \mathbf{S}_i$. The total staggered magnetization operator is $M = (1/N)\sum_i \tilde{\mathbf{S}}_i$. In the classical Néel state the value of $M = 0.5$. However, for an $S = \frac{1}{2}$ antiferromagnet the quantum fluctuations are large and the moment $m^\dagger$, which is the magnitude of the expectation value of the staggered magnetization, is reduced from its classical value. A finite value of $m^\dagger$ would indicate the existence of antiferromagnetic LRO in the ground state. We would like to emphasize that this does not contradict the singlet nature of the ground state. Consider, for example, the Néel state with the staggered moment $m^\dagger$ pointing along the z direction. If we measure the spin-spin correlation along the z direction it would decay to 0.25, equal to the square of the staggered moment. As prepared, the correlations along the x and y directions will decay to zero. Now by applying the rotation operator, the Néel state can be made rotationally invariant, and therefore a singlet. If we now measure the correlations, it will be found that the correlations along x, y, and z decay to a constant $= 0.25/3$. Thus it is possible to construct states that are rotationally invariant and at the same time have a broken symmetry.

As described earlier, the Jastrow wave function in (3.1) possess off-diagonal LRO but does not have any diagonal LRO. This introduces an anisotropy in the correlation functions. This is of no real consequence since we are interested in the total magnitude of the moment. An alternative way to understand this is to assume that the rotational symmetry of the magnet is broken by the application of a small symmetry breaking field, and the amplitude of this field is then taken to zero *after* taking the infinite "volume" limit. In the experiments on the lanthanum and yttrium high-temperature superconducting compounds in their antiferromagnetic phases, the moments on the copper sites are observed to lie in the xy plane. Thus in an experimental system, the rotational symmetry is automatically broken.

V. RESULTS

A. Energy

The classical Néel state with an energy per site $E_N = -0.5J$ contributes approximately $\frac{3}{4}$ of the ground-state energy of the Heisenberg Hamiltonian. We next calculate the additional 25% contribution denoted by ΔE by variational Monte Carlo (VMC) and by Green function Monte Carlo (GFMC). The total ground-state energy per site is then $E_0 = E_N + \Delta E$.

1. Variational results

We start with the trial wave function given in (3.1) and (3.2) with a single variational parameter f_1. The energy per site as a function of f_1 is obtained by VMC and is shown in Fig. 1 for a 12×12 system with periodic bound-

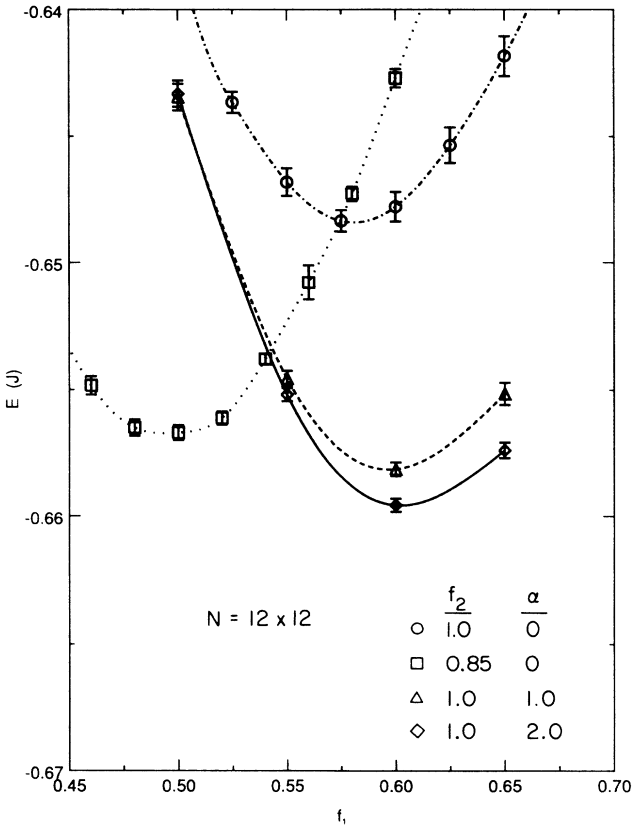

FIG. 1. The energy per site E vs f_1 for $N = 12 \times 12$ obtained using VMC. The circles (dashed-dotted curve) are calculated using (3.1) and (3.2). The energy of the classical Néel state $E_N = 0.5J$ is lowered by approximately 26% at the optimal value of $f_1 \simeq 0.58$. The squares (dotted curve) gives the energy with a second parameter f_2, the amplitude at the next neighbor site, included in (3.2). The effect of long-range correlations in states (3.1) and (3.8) is seen from the triangles ($\alpha = 1$) and diamonds ($\alpha = 2$), where α is related to the coefficient of the $1/r$ term.

ary conditions. As mentioned previously, the classical Néel state in the xy plane is obtained for $f_1=1.0$. The effect of reducing the variational parameter from unity is to introduce pair correlations in the wave function. The energy is the thereby lowered by almost 25% from E_N; at the optimal value of $f_1\simeq 0.58$, the energy is

$$\Delta E/J=-0.1481\pm 0.0004$$

or

$$E_V^{(1)}/J=-0.6481\pm 0.0004 \quad (N=12\times 12)\ .$$

We next introduce a second parameter in (3.2), by defining $f(\mathbf{r}_i-\mathbf{r}_j)=f_2$ for j on the next neighbor sites along the diagonals with respect to i. The optimal values of the parameters in this two-parameter space are $f_1\simeq 0.5$ and $f_2\simeq 0.85$. The energy is now lower than the energy obtained with a single parameter by a few percent. It may seem that by introducing more variational parameters we may eventually get a very good estimate for the energy. However, there are two problems: (i) The search through a multidimensional parameter space very quickly becomes rather involved and it becomes necessary to use steepest descent methods to minimize such a function. (ii) More importantly, as we will see from our GFMC analysis, it is quite unnecessary to put in more parameters, because the GFMC method gives a very accurate energy even when the trial function is a Néel state.

It was argued[13,12] that the energy per site must scale as L^{-3}, where L is one edge of the lattice. The variational estimates of the energy for both one- and two-parameter wave functions are shown in Fig. 2 for lattice sizes up to $N=12\times 12$. When extrapolated to the infinite system, the energy per site for short-range Jastrow functions with one variational parameter is

$$E_V^{(1)}/J=-0.6474\pm 0.0003$$

and with two parameters

$$E_V^{(2)}/J=-0.6559\pm 0.0002\ .$$

There is one curious fact about the size dependence of the energy: As one goes to a larger system the energy per site *increases*. One might expect the opposite trend because it could be argued that the *exact* wave function for a 4×4 system can be treated as a *variational* wave function for an 8×8 system and therefore the exact energy of the 8×8 system must be lower than the 4×4 system. This reasoning is not correct because of periodic boundary conditions. By introducing extra bonds at the boundaries, the potential energy is increased and that leads to a net increase in the energy per site as the system size is increased. We solved, analytically, a simpler problem of 4, 6, and 8 sites in one dimension with periodic and fixed boundary conditions and find that the exact result $E_0=-0.443J$ is approached from above with fixed boundary conditions and from below with periodic boundary conditions.

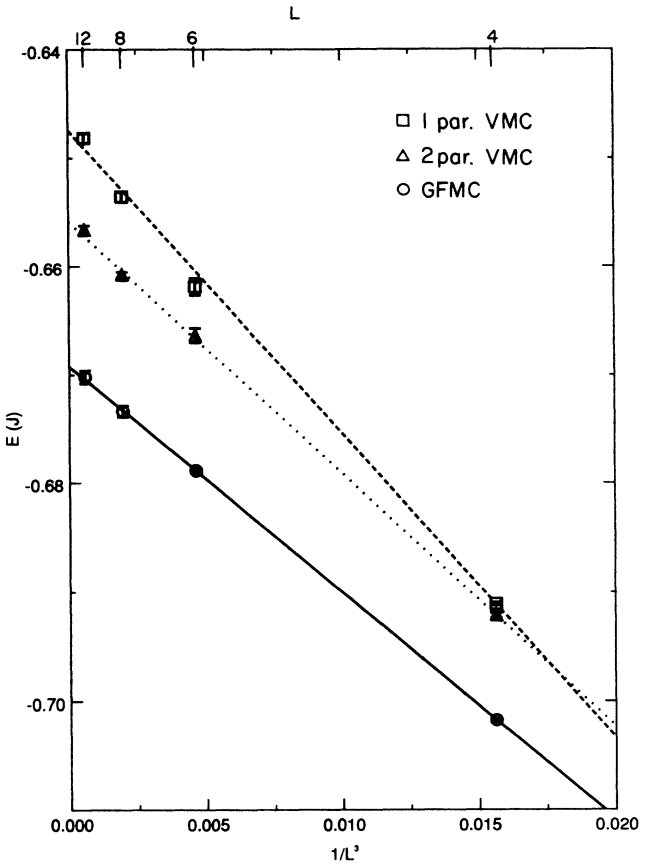

FIG. 2. The ground-state energy per site E as a function of L^{-3}. The squares are the VMC results at the optimal value of f_1 and with $f_2=1.0$, i.e., with only a single variational parameter. The triangles are evaluated at the optimal values of f_1 and f_2 by VMC. The circles are obtained by the GFMC method using (3.2) as the initial and the guiding wave function. The GFMC results are, however, insensitive to the choice of both these functions as discussed in the text. The extrapolated excess energy $E-0.5J$ obtained by the GFMC method is lower than the one-parameter result by $\simeq 13\%$.

2. *Effect of long-range correlations*

Next let us include long-range correlations in the wave function and see its effect on the energy. There are now basically two parameters as seen from (3.8): f_1 characterizes the depletion of probability from regions where two bosons are on adjacent sites and α is related to the coefficient of the $1/r$ correlation in the wave function. The energy as a function of these variational parameters is shown in Fig. 1. Notice a considerable reduction in the one-parameter variational estimate of the energy by $\simeq 7\%$ obtained upon including the long-range part in the wave function. As discussed in the Appendix, we use the Ewald summation technique to treat the long-range term. This involves starting with the integral representation of the $1/r$ term and dividing it into parts: The short-distance part is treated in real space, whereas the long-distance part is written in terms of small momentum contributions and treated in k space. These two regions are separated by a cutoff k_c chosen such that the contribu-

tion in real space is small at the edge of the cell. The energy is not a strong function of k_c, provided the aforementioned condition is satisfied.

3. GFMC results

In Fig. 2, the GFMC results are presented for the energy of lattice sizes up to 12×12 with periodic boundary conditions. The scaling behavior of the energy versus L^{-3} fits very well and the extrapolated value of the energy is found to be

$$E_0/J=-0.6692\pm0.0002\ . \tag{5.1}$$

If $J=0.2$ eV, an accuracy of 10^{-4} corresponds to 0.1 K. The excess energy obtained by the GFMC method is significantly lower by $\simeq13\%$ than the one-parameter variational estimate of the energy. It is important to note that if we were to compare the total energy with the Néel state contribution E_N included in it, we would come to the conclusion that the GFMC result differed from the variational energy by only $\simeq3\%$. However, since there is a large background term in the energy from the Néel state it is much more meaningful to compare the excess energy. The GFMC estimate is also lower than the two-parameter variational energy by almost 7%.

In Table I, we show the values for the ground-state energy obtained by a simple variational calculation using a one-parameter Jastrow wave function and contrast it with the dramatic improvement found by the GFMC method for lattice sizes up to 12×12. In the last column we give the results of Gross *et al.*,[13] obtained by a projection method similar to the GFMC method, and find that our results are consistently lower for all the lattices studied. There importance sampling has not been used, which can lead to big fluctuations in the energies of the different random walkers and it is possible that these fluctuations have not been fully accounted for. As discussed in Sec. VII, correlation between the data points, in general, leads to an underestimate of the error bars. Barnes *et al.*[17] have also used a random walk method to calculate ground-state energies. In their method the weights for each configuration are accumulated for successive time steps *without* branching. This approach carries along unfavorable configurations with weights close to zero and can give poor statistics.

We find that while the variational estimates of the energy are greatly affected by the inclusion of long-range correlations in the wave function, the energy obtained by the GFMC method is independent of the wave function. In other words, the mixed estimator in (4.14) for the energy reported in the fourth column of Table I is independent of whether (3.2) or (3.8) is used for f in the wave function (3.1). This is because the energy is sensitive only to short-range correlations in the wave function that are accurately obtained when the GFMC method is used to project out the excited states in the trial wave function.

In Table II, we compare our results with other methods that give estimates for the ground-state energy. To make meaningful comparison we look at the excess over the Néel state energy. Spin-wave theory gives

$$E_{SW}/J=-2S^2(1+0.158/S)=0.658$$

for $S=\frac{1}{2}$, which is about 7% higher than the GFMC value. The variational energies with short-range trial functions are about 10–15% higher than our GFMC estimate in (5.1) and by allowing for some additional freedom via long-range trial functions of the form (3.7), the energy can be brought to within 3% of the GFMC result. Resonating-valence-bond (RVB) wave functions give surprisingly good energies; though the dimer states are about 40% higher, including just short-range singlet bonds brings the energy to within 3% of the GFMC result. Even though the energy using Jastrow states with long-range correlations and RVB states with short-range dimer bonds are very close, the sublattice magnetization in these states is markedly different—from being around 0.33 in the Jastrow states to 0 in these RVB states (see

TABLE I. Ground-state energy per site E_0/J obtained by variational Monte Carlo using trial wave functions with short-range (SR) correlations (one and two parameters) [Eqs. (3.1) and (3.2)] and with long-range (LR) correlations [Eq. (3.7)]. Also calculated are the energies by the Green function Monte Carlo method using the wave function in Eq. (3.2) as the initial trial and guiding state. For comparison we give the results of Ref. 13 obtained by a projector method without importance sampling.

Size	VMC (SR) one parameter	VMC (SR) two parameter	VMC (LR)	GFMC	Reference 13
4×4	−0.6911	−0.6921		−0.7017	−0.70178
	±0.0002	±0.0002		±0.0002	±0.00009
6×6	−0.6619	−0.6663	−0.6690	−0.6789	−0.6780
	±0.0008	±0.0007	±0.0006	±0.0001	±0.00005
8×8	−0.6536	−0.6608	−0.6640	−0.6734	−0.6715
	±0.0005	±0.0003	±0.0003	±0.0005	±0.0004
12×12	−0.6481	−0.6566	−0.6598	−0.6702	−0.6670
	±0.0004	±0.0004	0.0004	±0.0006	±0.0007
$\infty\times\infty$	−0.6474	−0.6559	−0.6590	−0.6692	−0.6672
	±0.0003	±0.0002	0.0004	±0.0002	±0.0008

TABLE II. Comparison of the ground-state energy per site E_0 of the Heisenberg quantum antiferromagnet calculated by different methods. The last column gives the difference of the excess energy $E_0 - E_N$ with respect to the GFMC result in Eq. (5.1). [$E_N = -0.5J$ is the energy in the classical Néel state.] WF denotes wave function.

Method	E_0/J	$\delta E_0(\%)$
Spin-wave theory, see Anderson[a]	-0.658	$+6.6$
Perturbation around Ising limit, see Singh[b]	-0.6696 ± 0.0003	
Variational short range Jastrow WF		
Huse-Elser[c] [$f_1 \sim 0.58$]	-0.644	$+14.9$
Present work [$f_1 \sim 0.58$]	-0.6474 ± 0.0003	$+12.9$
Huse-Elser[c] [f_1, f_2, f_3]	-0.6550	$+8.4$
Present work [$f_1 \sim 0.5, f_2 \sim 0.85$]	-0.6559 ± 0.0002	$+7.9$
Variational long-range WF [$V(r) = \alpha/r^\beta$]		
Huse-Elser[c] [$f_1 \sim 0.27$, $\alpha \sim 0.95$, $\beta \sim 0.7$]	-0.6638	$+3.2$
Present work [$f_1 \sim 0.6$, $\alpha \sim 1$, $\beta = 1.0$]	-0.6590 ± 0.0004	$+6.0$
Variational-Gutzwiller WF		
Yokoyama-Shiba[d]	-0.642 ± 0.002	$+16.1$
Variational RVB WF		
Liang, Doucot, and Anderson[e]		
dimmer	-0.604	$+38.5$
short-range RVB	-0.6650	$+2.5$
long-range RVB	-0.6688 ± 0.0004	0.3
Finite-temperature Monte Carlo		
Reger, Riera, and Young[f]	-0.670 ± 0.0005	-0.5
Projector method		
Barnes, Kotchan, and Swanson[g]	-0.669 ± 0.001	-2.1
Gros, Sanchez, and Siggia[h]	-0.6662 ± 0.0005	$+1.8$
Carlson[i]	-0.6692 ± 0.0001	
Present Work (see also Trivedi and Ceperley)[j]	-0.6692 ± 0.0002	

[a]Reference 32.
[b]Reference 15.
[c]Reference 9.
[d]Reference 10.
[e]Reference 11.
[f]Reference 42.
[g]Reference 17. This is an improvement on a previous estimate of -0.672 ± 0.001 given in Ref. 17.
[h]Reference 13. We obtain this value by a least-squares fit to the energy of 6×6, 8×8, and 12×12 lattices reported in Ref. 13, which differs from their extrapolated value of -0.6672 ± 0.0008.
[i]Reference 18.
[j]Reference 20.

Table IV). Thus states with and without LRO can have energies that are very close. Upon including long-range bonds the energy is only 0.3% higher than GFMC however, $m^\dagger \simeq 0.23$, which is lower than the accepted value of $\simeq 0.3$. The finite-temperature Monte Carlo appears to give energies that are about 0.5% lower than the GFMC results. Our results agree well with the series estimate obtained by keeping terms up to (J_{xy}/J_z) (Ref. 10) in the perturbation around the Ising limit. It is also heartening to see that the results for the energy of Carlson[18] obtained by a GFMC calculation using trial wave functions in Ref. 9 are identical to ours. However, since he has a lower value for the exponent of the long-range correlations, he finds a higher sublattice magnetization.

4. *GFMC on classical Néel trial state*

Thus it is clear that by a judicious choice of the variational wave function, a very good estimate of the ground-state energy can be obtained. We next want to explore the efficiency of the GFMC method by using the simplest possible state, the classical Néel state, as the guiding function. We iterate (4.7) approximately 1000 times with Ψ_G equal to the Néel state and obtain a ground-state energy for an 8×8 system to be -0.669 ± 0.004. This agrees with the result obtained with an optimized wave function of -0.6734 ± 0.0005 that includes short-range pair correlations, i.e., (3.2) with the optimal value of f_1. Notice that for the same number

of iterations, the error bars in the former case are larger, as expected. This study in fact points to a great strength of the GFMC method, i.e., to obtain the ground-state energy even a very simple wave function is sufficient, provided it has a nonzero overlap with the ground state. This is true since the Néel state, written in the boson language is nodeless.

B. Correlation functions

We can define the off-diagonal correlation function between the bosons, which is a measure of the amplitude of a boson to hop from a site i to a site $i+l$ by

$$h(l)=\frac{1}{N}\sum_i \langle b_i^\dagger b_{i+l}\rangle \ . \tag{5.2}$$

Equation (5.2) is equivalent to

$$(1/N)\sum_i \langle \tilde{S}_i^x \tilde{S}_{i+l}^x+\tilde{S}_i^y \tilde{S}_{i+l}^y\rangle \ ,$$

the spin-spin correlations in the xy plane. We also study the diagonal correlation function defined by

$$g(l)=\frac{1}{N}\sum_i \langle \tilde{n}_i \tilde{n}_{i+l}\rangle-\frac{\varepsilon_l}{4} \ , \tag{5.3}$$

the density response of the bose system. Equivalently

$$g(l)=(1/N)\sum_i \langle \tilde{S}_i^z \tilde{S}_{i+l}^z\rangle$$

is the correlation function of the staggered z component of the spin.

1. Comparison with exact diagonalization results

The correlation functions $h(l)$ and $g(l)$ obtained by GFMC for a 4×4 system are shown in Fig. 3 and Table III and compared with the exact diagonalization values. There is excellent agreement with exact results. Even though the walks are guided by (3.2), which is *not* a singlet (the correlation functions along x, y, and z are not equal), after projecting out the excited states the GFMC method is able to achieve equal correlation along all directions. This is seen from Fig. 3, where $h(l)\simeq 2g(l)$. For larger systems the computer time needed to restore this symmetry becomes prohibitive.

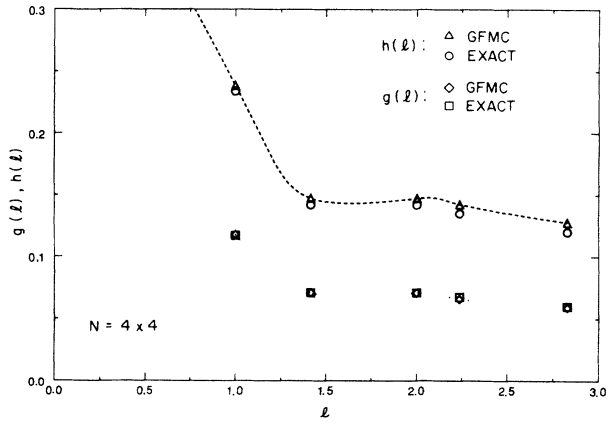

FIG. 3. Comparison of correlation functions for a 4×4 system obtained by the GFMC method and by exact diagonalization (Ref. 8). The off-diagonal correlation $h(l)$ is shown by triangles (GFMC) and circles (exact) and the diagonal correlation $g(l)$ is shown by diamonds (GFMC) and squares (exact). The error bars are of the size of the symbols.

TABLE III. Spin-spin correlation function $\langle S_i^z \cdot S_{i+\lambda}^z\rangle$ for a 4×4 lattice. Comparison of results obtained by GFMC method using the short-range trial and guiding functions in Eq. (3.2) with exact diagonalization results (Ref. 8).

λ	GFMC	Exact
1	−0.1177±0.0018	−0.117
$\sqrt{2}$	0.0710±0.0021	0.071
2	0.0710±0.0023	0.071
$\sqrt{5}$	−0.0660±0.0019	−0.0675
$2\sqrt{2}$	0.0591±0.0022	0.06

2. Off-diagonal correlations

As mentioned previously, the GFMC results are obtained for a random walk guided by the wave function in (3.2), which has all the staggered moment tipped in the xy plane. This introduces a symmetry breaking field that leads to anisotropy between $h(l)$ and $g(l)$ for large lattices. The results for $h(l)$ are given in Fig. 4 for a lattice size 8×8. The wave functions used are (3.1) and (3.2) and the VMC result is calculated using (4.1). For the GFMC result, the mixed estimate is found from (4.15) also with a short-range function in (3.2). The solid line in Fig. 4 is the corrected or extrapolated GFMC result obtained as explained in Sec. IV C. Over a very short distance of the order of the lattice spacing, $h(l)$ decays to a nonzero con-

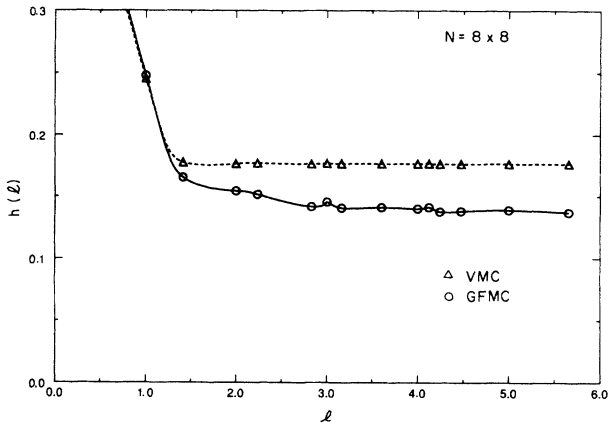

FIG. 4. The off-diagonal correlations $h(l)$ for an 8×8 system. The dashed curve is the result of a VMC calculation in states (3.2). The solid curve is the extrapolated GFMC result with the mixed average in (4.15) obtained using the trial function (3.2). The GFMC method tends to lower the magnetization in the xy plane compared to the VMC result and also introduces angular dependence in $h(l)$. The error bars are of the size of the symbols.

stant. For noninteraction bosons $h(l)=(N_b/N)^2=0.25$ at large l; the reduction seen in the figure is a measure of the depletion of the condensate because of repulsive interaction between the bosons. The GFMC value for $h(l)$ at large distances is lower than the VMC by about 13%. This shows that in the ground state the bosons experience enhanced repulsive interactions compared to the Jastrow wave function (3.2). Alternatively, the ground state can be viewed as including additional quantum fluctuations that leads to a reduction of the staggered magnetization in the xy plane that is reflected in the difference between the VMC and GFMC results.

It is interesting to note that while the VMC result for $h(l)$ is rather featureless, the GFMC result shows wiggles as a function of the distance. This is due to the fact that the spin-spin correlation function between two sites i and j depends not only on the distance between these sites, but also on the direction.

Using the same set of configurations generated by the guiding function (3.2), we evaluate the mixed estimator in (4.15) with a long-ranged trial function (3.8). The effect of such long-range correlations in the wave functions is twofold: First, the difference between the VMC and GFMC method in $h(l)$ is smaller than shown in Fig. 4 and second, we find that spin waves lower the staggered moment further compared to the GFMC result in wave functions with only short-range correlations, as would have been expected. These indicate that (3.8) is perhaps closer to the ground state than (3.2)

3. *Diagonal correlations*

The behavior of $g(l)$ for two different lattice sizes is shown in Fig. 5 and is seen to decay almost to zero at large distances, showing a lack of diagonal LRO. The GFMC tends to reduce the staggered moment along xy, but increase the moment along z (compare Figs. 4 and 5).

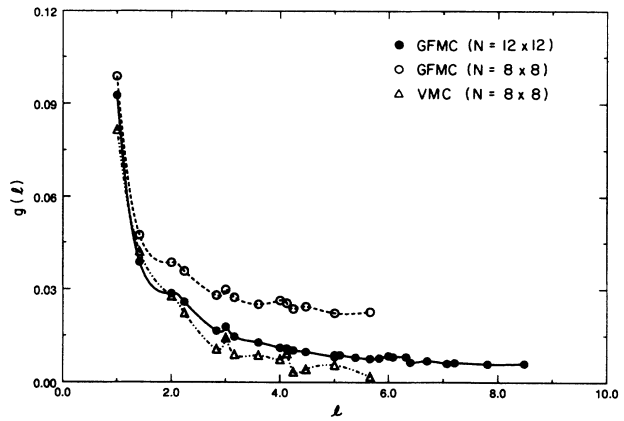

FIG. 5. The extrapolated GFMC results for $g(l)$, using states (3.2) in the mixed estimator, is shown for two lattice sizes: $N=12\times12$ (solid circles) and $N=8\times8$ (open circles). The triangles show the VMC results for an 8×8 system. In contrast to Fig. 4, the GFMC method tends to increase the moment along z compared to the VMC result. The wave function in (3.1) has all the magnetization in the xy plane, so $g(l)$ rapidly decreases to zero as the lattice size is increased.

The GFMC is attempting to obtain a singlet ground state with $h(l)=2g(l)$, however, it is unable to do that since the rotational symmetry has been broken by the guiding wave function. Once again, similar to the behavior of $h(l)$, the inclusion of the zero-point motion of spin waves shows $g(l)$ decaying to a smaller constant, implying that the moment in the z direction is reduced.

4. *Staggered magnetization*

At large distances $h(l)$ decays to the staggered magnetization in the xy plane defined as $(m_x^\dagger)^2+(m_y^\dagger)^2$. The function $g(l)$, on the other hand decays to $(m_z^\dagger)^2$. Thus from the long-distance behavior of (5.2) and (5.3) we obtain the value of the staggered moment $(m^\dagger)^2$. The scaling of the staggered moment, extracted from the correlation functions, is depicted in Fig. 6. We compare in this figure, the VMC and corrected GFMC results for a trial function with only short-range correlations as in (3.2).

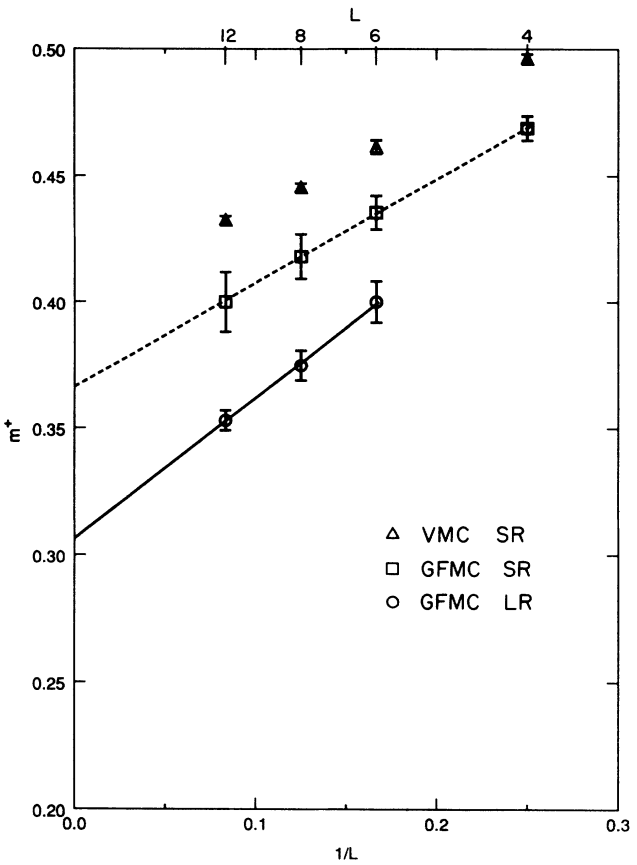

FIG. 6. The staggered magnetization $m^\dagger$ vs $1/L$ evaluated at $(L/2,L/2)$ for two different trial functions. (In the classical Néel state $m^\dagger=0.5$). (i) Short-range (SR) trial function Eq. (3.2): The extrapolated GFMC estimate is shown by the squares and the VMC estimate is shown by the triangles. The dashed line is a least-squares fit to the squares and extrapolates to 0.37 for the infinite lattice. (ii) Long-range (LR) trial function Eq. (3.8): The GFMC estimate is shown by circles. The variational estimate (not shown) is within a few percent of the GFMC estimate. The solid line is a least-squares fit to the circles and extrapolates to 0.31 for the infinite lattice.

The value of $m^\dagger$ extrapolated to the infinite lattice obtained by GFMC is $m^\dagger=0.37$, which is about 8% lower from the variational results. We also give the corrected GFMC results for $m^\dagger$ for the long-range trial function in (3.8). We find

$$m^\dagger=0.31\pm0.02\ , \tag{5.4}$$

which is lower than the GFMC estimate with short-range trial functions by $\simeq 19\%$. This result is consistent with spin wave[32] and finite-temperature Monte Carlo[12] of $m^\dagger\simeq 0.30$. It also agrees with neutron scattering experiments[1] on pure La_2CuO_4 that show evidence for a two-dimensional antiferromagnetic ordering of the spins on the copper atoms at low temperatures, stabilized by the inter-planar coupling. The moment[1] is found to be pointing along the diagonals of the square lattice and have a magnitude of about 0.25 ± 0.08.

VI. SPIN-WAVE VELOCITY

The spin-wave velocity can be obtained from the scaling of the ground-state energy as a function of the system size[33] from the relation

$$E_0(L)/N=E_0/N-1.4372\hbar c/L^3\ .$$

From Fig. 2, we obtain a slope of 2.0806 ± 0.02, which implies that $\hbar c=(1.45\pm0.01)Ja$. If we define $Z_c\equiv c/c_s$ to be the renormalization of the spin-wave velocity by quantum fluctuations, where $\hbar c_s=\sqrt{2}Ja$ is the classical spin-wave velocity, we get $Z_c=1.024\pm0.007$.

It is also possible to get an estimate of the spin-wave velocity from the structure factor in the *ground state* by using the f-sum rule, as we describe later. The structure factor is the Fourier transform of $g(l)$ and is given by

$$\mathcal{S}(\mathbf{q})=\frac{1}{N_b}\langle S^z(\mathbf{q})S^z(-\mathbf{q})\rangle\ . \tag{6.1}$$

In the case of antiferromagnetic LRO along z there is a δ function peak at $\mathbf{q}_0=(\pi,\pi)$. From (6.1), it can be seen that $\mathcal{S}_\pi\equiv\mathcal{S}(\mathbf{q}_0)=\langle M_z^2\rangle$. We digress to derive the f-sum rule on the lattice as it involves some modifications of its counterpart in the continuum. We start by considering the commutator of the Hamiltonian in (2.2) with $\rho(\mathbf{q})$. It can be shown[34] that

$$[H,\rho(\mathbf{q})]=\sum_{\mathbf{k}} b^\dagger_{\mathbf{k}+\mathbf{q}}b_{\mathbf{k}}(\varepsilon_{\mathbf{k}+\mathbf{q}}-\varepsilon_{\mathbf{k}}) \tag{6.2}$$

and

$$[[H,\rho(\mathbf{q})],\rho(-\mathbf{q})]=\sum_{\mathbf{k}} n_{\mathbf{k}}(2\varepsilon_{\mathbf{k}}-\varepsilon_{\mathbf{k}+\mathbf{q}}-\varepsilon_{\mathbf{k}-\mathbf{q}})\ , \tag{6.3}$$

where $\rho(\mathbf{q})=\sum_{\mathbf{k}}b^\dagger_{\mathbf{k}+\mathbf{q}}b_{\mathbf{k}}$ and $\varepsilon_{\mathbf{k}}=J\sum_{\vec{\delta}}[1-\cos(\mathbf{k}\cdot\vec{\delta})]$ for the free particle energy on a lattice and the sum is over the nearest-neighbor sites along $+x$ and $+y$. If we take the expected value of (6.3) in the ground state, we find

$$\langle[[H,\rho(\mathbf{q})],\rho(-\mathbf{q})]\rangle=-N\varepsilon_q|T_k|\ . \tag{6.4}$$

We have used

$$\sum_{\mathbf{k}}\cos(\mathbf{k}\cdot\vec{\delta})\langle n_{\mathbf{k}}\rangle=-NT_k/2\ ,$$

where $T_k\approx-0.5$ is the kinetic energy per site in units of J and is related to the off-diagonal correlation function at the nearest-neighbor site $h(1)=-T_k/2$. It can also be shown, by inserting exact eigenstates of the Hamiltonian in the left-hand side (LHS) of (6.4) and comparing with the usual definition of the structure factor in (6.1), the f-sum rule is modified on a lattice to

$$\int_0^\infty d\omega\omega\mathcal{S}(\mathbf{q},\omega)=N\varepsilon_{\mathbf{q}}|T_k|/2\ . \tag{6.5}$$

At long wavelengths, the collective excitations exhaust the sum rule and the single quasiparticle-quasihole excitations and multipair excitations do not have much oscillator strength. We then obtain the well-known Feynman-Bijl formula[25,24] from (6.5) by writing $\mathcal{S}(\mathbf{q},\omega)\sim\delta(\omega-\omega_{\mathbf{q}})$, given by

$$\omega_{\mathbf{q}}=\frac{\varepsilon_{\mathbf{q}}}{\mathcal{S}(\mathbf{q})}|T_k|\ . \tag{6.6}$$

Thus from the behavior of the structure factor at small q it is possible to understand the nature of the excitation spectrum. In the continuum (6.6) reduces to the usual expression without the kinetic energy factor.

In Fig. 7 we show the structure factor for a 12×12 system along [10] obtained by calculating the extrapolated estimator by the GFMC method using trial functions that have both short- and long-range correlations. It can be clearly seen that $\mathcal{S}(\mathbf{q})$ is linear in the long wavelength region only in the presence of long-range correlations. We have used a *maximum overlap method*[35] to obtain the parameters in the trial function. The basic idea is to minimize the difference between the variational results and the mixed estimate from the GFMC method. In other

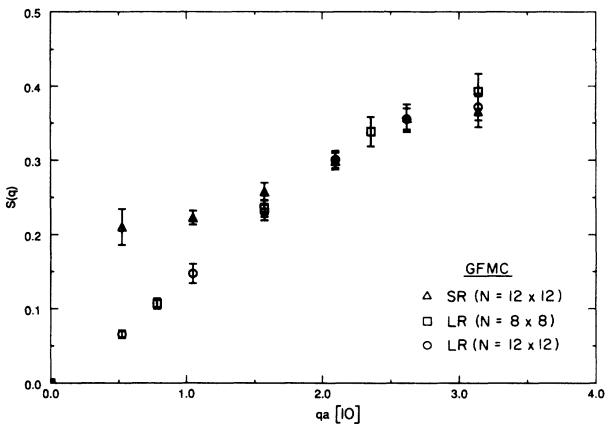

FIG. 7. Structure factor along [10] obtained by the GFMC method. (i) $\mathcal{S}(\mathbf{q})$ for a 12×12 lattice (triangles) calculated with a short-range trial function. Note that it incorrectly goes to a nonzero value at small q. (ii) $\mathcal{S}(\mathbf{q})$ for a 8×8 lattice (squares) and for a 12×12 lattice (circles) with long-range trial function. The spin-wave velocity is extracted from the linear dependence at small q. There are important and noticeable differences at small q between (i) and (ii), whereas at large q all the curves merge, since the short-range correlations are the same.

words, we have not used the values of the parameters, f_1 and α in (3.8) that minimize the energy, instead we have used the criteria that the best parameters for the structure function are those that minimize that difference between the variational and the mixed estimates of $\mathcal{S}(q)$, since the energy is not really sensitive to the long-range structure in the wave function. In this way we maximize the overlap between the variational and the ground-state wave function.

From (6.6), given the dispersion on the lattice at long wavelengths, the spin-wave velocity is given by

$$Z_c \equiv \frac{c}{c_s} \simeq \frac{qa}{2\sqrt{2}\mathcal{S}(q)} |T_k| , \tag{6.7}$$

where

$$\hbar c_s = 2(1-\gamma_q^2)^{1/2}/(qa) \sim \sqrt{2}Ja$$

is the classical spin-wave velocity and

$$\gamma_q = (\tfrac{1}{2})[\cos(q_x a) + \cos(q_y a)] .$$

We obtain a value for $Z_c = 1.14 \pm 0.05$. This is a measure of the renormalization of the spin-wave velocity caused by quantum fluctuations and can be compared with the spin-wave analysis around the classical Néel state,[32] which evaluates the first correction of order $1/S$, and finds $Z_c = 1.158$.

In the preceding analysis, we have obtained the spin-wave velocity from the long-wavelength behavior of the structure factor. It is also interesting to relate α, the coefficient of the $1/r$ term in the wave function in (3.7), to the velocity. Then by using the optimal α that minimizes the energy, a value for c can be obtained by a variational calculation. As discussed in Sec. III, within linear response $1/\mathcal{S}(q) = V(\mathbf{q})/2$ at long wavelengths since $\mathcal{S}_{\rm SR}(\mathbf{q}) \rightarrow$ const$\neq 0$ from Fig. 7 as $q \rightarrow 0$. Now substitute for $\mathcal{S}(\mathbf{q})$ from (6.6) and use $V(q) = 2\pi\alpha/q$ in two dimensions to obtain

$$\hbar c = \frac{\pi\alpha |T_k|}{2} Ja . \tag{6.8}$$

At the optimal value of $\alpha \approx 2$ from Fig. 1, $Z_c \simeq 1.1$ from (6.8), which implies a $\simeq 10\%$ renormalization of the spin-wave velocity. In an independent spin-wave analysis, Manousakis[36] has found that $V(r) = 4\sqrt{2}/(\pi r)$ for larger r, so that he has a parameter-free wave function. His wave function contains the symmetry of the lattice, however, by not allowing for a free parameter, the spin-wave velocity has its classical value.

A direct value of the spin-wave velocity can be obtained by the GFMC method by finding the difference of the energy of the excited state at a finite value of the wave vector and the ground state. If we take a trial function $\Psi_{\rm ex} = \rho(\mathbf{q})\Psi_0$, as a product of the density operator and the ground-state wave function, using the GFMC method it is possible to project onto the subspace with a nonzero q, and thereby obtain the excited-state energy. Such a method should give a very accurate value of the excited-state energy and the spin wave velocity.

It is also instructive to study the long-wavelength behavior of the momentum distribution function, which is

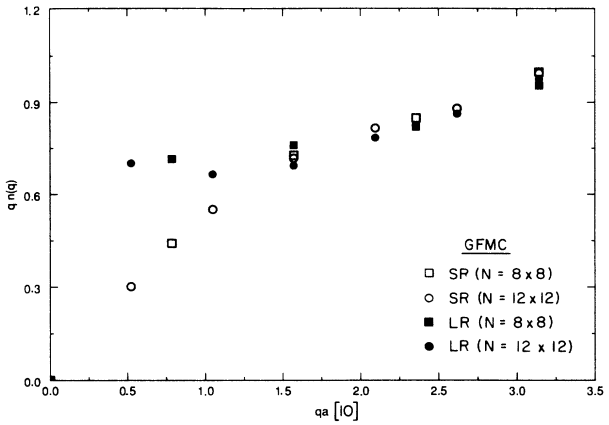

FIG. 8. Momentum distribution function $q[n(\mathbf{q}) - n(0)]$ vs qa along [10] obtained by extrapolated GFMC results. The short-range (SR) trial function in (3.2) give open squares ($N = 8\times 8$) and open circles ($N = 12\times 12$). The long-range (LR) trial function in (3.8) give solid squares ($N = 8\times 8$) and solid circles ($N = 12\times 12$). As $q \rightarrow 0$, it approaches a constant related to the spin-wave velocity.

the Fourier transform of $h(l)$ given by

$$\delta n(\mathbf{q}) = \frac{1}{N_b} \langle b^\dagger(\mathbf{q}) b(-\mathbf{q}) \rangle , \tag{6.9}$$

where $\delta n(\mathbf{q}) = n(\mathbf{q}) - n(0)$. The presence of a condensate fraction or equivalently, a staggered magnetization in the xy plane, is indicated by a peak in the zero momentum state. At long wavelengths it can be shown by using sum rules[37] that $\delta n(\mathbf{q}) \sim 1/q$ and the prefactor are related to the condensate fraction and the spin-wave velocity. Once again, it is seen from Fig. 8, that the correct behavior is found only when long-range trial functions are used to calculate the momentum distribution function. For short-range trial functions, we get an incorrect result of $\delta n(\mathbf{q}) \sim 1/q^2$. Thus from the requirement that certain sum rules be obeyed, we find that the ground-state wave function must have long-range pair correlations of the form given in (3.8).

VII. SOURCES OF ERROR

For the Green-function Monte Carlo calculation performed here, we can identify basically four sources of error.

Statistical error: In calculating any quantity of interest, one necessarily has a finite set of configurations. In general, since the configurations tend to be correlated, the naive variance of the energy E given by $V_1 = \langle (E - E_0)^2 \rangle$, where E_0 is the ground-state energy, will underestimate the true variance.[38] To see this, consider the extreme limit when all the data are identical, a case of perfect correlation, which will indeed produce a zero variance. Correlation between the configurations can arise from branching and also because successive states are close together.

A simple way to estimate the error bars when the data is correlated, is to divide the total random walk into a

number of separate "blocks"; in our case we used approximately ten blocks each containing about 100 generations. The basic criteria is that the length of the blocks be longer than the correlation times, in which case the block average can be assumed to be an independent random variable. We then determine the standard deviation from the dispersion of the block averages. If the length of the blocks is indeed longer than the correlation time, the error bars will not change if the block size is varied.

Error from use of mixed estimator: As discussed in Sec. IV, in the GFMC method one can naturally obtain the mixed estimator, i.e., the expectation value of any operator $\hat{\mathcal{O}}$ sandwiched between the true ground-state wave function and some trial wave function. (For the special case of $\hat{\mathcal{O}}=H$, there is no error from the mixed estimator.) There are two possible ways to obtain the true estimate: one is to maximize the overlap between the ground-state wave function and the trial state by varying the parameters in the latter state, as we have done for the structure factor in Sec. VI. Another possibility is to assume that the true wave function is close to the variational wave function, and by a perturbation expansion obtain an extrapolated estimate as we have done for the spin-spin correlation functions. Though the above-mentioned schemes are possible ways to improve the estimate, the mixed estimator nevertheless remains a source of error. A different algorithm exists[39] where one removes the effect of the trial function on the expectation value by using the asymptotic branching of a given configuration. It involves the filtering of both bra and ket in any expectation value.

Convergence error: If the trial function is in the right phase, the effect of a convergence error is fairly well controlled since one can vary the equilibration time (the number of walkers discarded before computing averages) until the averages are independent of the equilibration. The better the trial function, the shorter the convergence time. However, if the trial function is not chosen in the right phase the averages will show large fluctuations and after many iterations it might incorrectly appear that the system has reached "equilibrium." A long equilibration time should be used as an indicator for more thought and improvement of the trial function.

Finite system error: We have accounted for finite-size effects by scaling. The energy per site $E/N\sim L^{-3}$ and the staggered magnetization $m^{\dagger}\sim L^{-1}$ in agreement with spin-wave theory. However, in case the number of data points is small and an incorrect scaling behavior is assumed, the error from finite-size effects is important. Also, certain physical effects that appear on very long length scales (e.g., spiral phases with a long pitch) may be missed if the lattice size is too small.

VIII. CONCLUSIONS

In this paper we have studied the $S=\frac{1}{2}$ Heisenberg antiferromagnet by the Green-function Monte Carlo method. We had two aims: first, to calculate various ground-state and low-lying excited-state properties and second, from a comparison of the results with other methods reported in the literature, to develop confidence in the method and to understand its strengths and weaknesses. This should then allow us to use the GFMC method in more complicated problems. If we compare our GFMC (GF) results with spin-wave (SW) theory—ground-state energy $E_{\rm GF}/J=-0.6692\pm0.0002$ and $E_{\rm SW}/J=-0.658$; staggered magnetization $m^{\dagger}_{\rm GF}=-0.31\pm0.02$ and $m^{\dagger}_{\rm SW}=-0.303$; and renormalization of the spin-wave velocity $Z_c^{\rm GF}=1.14\pm0.05$ and $Z_c^{\rm SW}=1.158$ we conclude that the two-dimensional $S=\frac{1}{2}$ Heisenberg antiferromagnet is surprisingly well described by lowest-order spin-wave theory. (For a comparison of our GFMC results with other numerical techniques, see Tables II and IV.) We also emphasized the importance of including the zero-point motion of the spin waves in the ground state and showed that it altered the approach of the wave function to its asymptotic behavior. This was found to be crucial in obtaining the correct long-wavelength properties of the correlation functions consistent with sum rules.

TABLE IV. Comparison of the staggered magnetization $m^{\dagger}$ for the Heisenberg AFM in two dimensions calculated by different methods. The values are given in units in which the Néel state staggered moment is 0.5.

Method	$m^{\dagger}$
Spin-wave theory[a]	0.303
Series[b] $x=(J_{xy}/J_z)$	
Terms up to x^6	0.32
Terms up to x^{10}	0.3025±0.008
Variational	
Jastrow WF, short range	
Huse-Elser[c]	0.42
Present work	0.40±0.02
Jastrow WF, long range	
Huse-Elser	0.36
Present work	0.32±0.02
Gutzwiller related WF[d]	0.43
RVB WF[e]	
short-range bonds	0.0
long-range bonds	0.23
Finite-temperature Monte Carlo	
Reger-Young[f]	0.31
Reger, Riera, and Young[g]	0.308±0.015
Gros, Sanchez, and Siggia[h]	0.285±0.025
Tang and Hirsch[i]	0.25
Projector	
Carlson[j]	0.34±0.01
Present work (see also Trivedi and Ceperley)[k]	
short-range importance fn.	0.37±0.02
long-range importance fn.	0.31±0.02

[a]Reference 32.
[b]References 14 and 15.
[c]Reference 9.
[d]Reference 10.
[e]Reference 11.
[f]Reference 12.
[g]Reference 42.
[h]Reference 13.
[i]Reference 43.
[j]Reference 18.
[k]Reference 20.

In our analysis we used the interesting analogy between two-dimensional bosons and the Heisenberg model. From the rotational symmetry of the magnetization (singlet) and the result that the ground state has antiferromagnetic long-range order, we conclude that the boson model has both momentum condensation (off-diagonal long-range order) and infinite range density waves (diagonal long-range order), i.e., it is a "supersolid."[40] On the other hand, our trial function does not have diagonal long-range order and is therefore not a singlet. We are currently exploring a modified form of the trial function with all the known symmetry properties of the ground state.

As summarized in the Introduction, the GFMC method has several attractive features. It can handle larger lattices are certainly feasible), as opposed to exact diagonalization methods that are limited to 4×4 lattices because of memory considerations. This method is well suited to studying the ground-state and low-lying excited-state properties and is thus complementary to the finite-temperature path integral quantum Monte Carlo techniques. The GFMC method goes beyond variational methods by projecting out of a trial state its component in the ground state. For bosons (e.g., the Heisenberg Hamiltonian), the GFMC method gives the exact ground-state energy with only statistical errors. It is also possible to evaluate the overlap of the variational and the true ground-state wave function. This information can then be used as a feedback to improve the trial wave functions. One disadvantage, at least at this point, is that except for the energy, the GFMC method gives mixed estimators so that for correlation functions there is a definite bias from the wave function used. It would be better if the true estimate were obtained directly.

For the future we are attempting a GFMC calculation for strongly correlated fermionic systems. It would also be interesting to study the $S=\frac{1}{2}$ spin models with frustration. For these models, it is well known that the finite-temperature methods have "sign problems," i.e., the Monte Carlo weights can become negative at low temperatures. This limits calculations for the Hubbard model at finite doping to a temperature roughly $T\simeq t/5$, where t is the hopping energy. The GFMC method for fermions is also more difficult than its bose counterpart discussed earlier because of the lack of knowledge of the nodes of the many-body wave function. However, it is possible to calculate properties of interest within a fixed node approximation,[16] which gives an upper bound on the ground-state energy (in continuum problems). This is similar in spirit to the variational method and is therefore well controlled, but has not yet been applied to fermion lattice models.

ACKNOWLEDGMENTS

We would like to thank E. Fradkin, C. Henley, M. Ma, R. M. Martin, M. Randeria, and R. Singh for several very useful and interesting discussions. This work has been supported by the National Science Foundation–Department of Materials Research (NSR-DMR) under Grant No. NSF-DMR88-08126 and by a grant of computer time on the CRAY XMP/48 at the National Center for Supercomputing Applications at the University of Illinois.

APPENDIX: EWALD SUMMATION METHOD TO HANDLE LONG-RANGE CORRELATIONS

We have shown in Sec. III that the inclusion of the zero-point motion of the elementary excitations leads to long-range correlations in the ground-state wave function given by (3.7). Since we are applying periodic boundary conditions on the Hamiltonian, the long-range nature of correlations can lead to serious difficulties. This is solved by using the Ewald summation technique.[41] The basic idea is as follows: let us assume that a cell of size $L\times L$ is repeated periodically to fill all space. The wave function in (3.7) requires a summation for a site i of the form

$$\sum_{L}\sum_{j\in\text{cell}}{}' 1/|\mathbf{r}_i-\mathbf{r}_j+\mathbf{L}| ,$$

where the prime on the summation implies that the $j=i$ term is not included. We drop constants that merely contribute to the normalization of the wave function. By a well-known identity

$$\sum_{\mathbf{L}}\frac{1}{|\mathbf{r}+\mathbf{L}|}=\frac{2}{\sqrt{\pi}}\sum_{\mathbf{L}}\int_0^\infty dk\,\exp(-k^2|\mathbf{r}+\mathbf{L}|^2) , \tag{A1}$$

where $\mathbf{r}$ lies within a cell. The integral in (A1) is divided into two parts by a cutoff k_c and it can be shown that the second part of the integral is related to the error function by

$$\int_{k_c}^\infty=\frac{\sqrt{\pi}}{2|\mathbf{r}+\mathbf{L}|}\,\text{erfc}(k_c|\mathbf{r}+\mathbf{L}|) . \tag{A2}$$

For the first part of the integral, define a function

$$\mathcal{F}(\mathbf{r})=\frac{2}{\sqrt{\pi}}\sum_{\mathbf{L}}e^{-k^2|\mathbf{r}+\mathbf{L}|^2}=\sum_{\mathbf{K}}\mathcal{F}_{\mathbf{K}}e^{i\mathbf{k}\cdot\mathbf{r}} . \tag{A3}$$

Note that $\mathcal{F}(\mathbf{r})$ has the periodicity of $\mathbf{L}$ and it can therefore be expanded in a Fourier series, where $\mathbf{K}=(2\pi/L)(m_x\hat{x}+m_y\hat{y})$. Now $\mathcal{F}_{\mathbf{K}}$ can be evaluated by the inverse transform. The summation over cells simplifies greatly, since by shifting the origin to each cell, all the terms become identical and the result is a simple Gaussian integral that gives

$$\mathcal{F}_{\mathbf{K}}=\frac{2\sqrt{\pi}}{L^2k^2}e^{-K^2/4k^2} . \tag{A4}$$

The integral $\int_0^{k_c}$ can be easily done and also yields an error function. The final result is

$$\sum_{\mathbf{L}}\frac{1}{|\mathbf{r}+\mathbf{L}|}=\sum_{\mathbf{L}}\frac{1}{|\mathbf{r}+\mathbf{L}|}\,\text{erfc}(k_c|\mathbf{r}+\mathbf{L}|)+\frac{1}{\Omega}\sum_{\mathbf{K}}\frac{2\pi}{K}e^{i\mathbf{K}\cdot\mathbf{r}}\,\text{erfc}(K/2k_c) . \tag{A5}$$

In (A5) a separation of the short-range and the long-wavelength parts of the correlations has been achieved. We neglect the short-range contribution, as that is already included in the Jastrow function in (3.2), but carefully include the small moment contributions. The number of K modes in the summation is determined by the choice of the cutoff. We choose k_c large enough so that the first term in (A5) is smaller than $\varepsilon \simeq 10^{-3}$ at the center of the cell. The larger k_c is, the more terms we need to keep in the second summation. We keep all modes for which the second term is greater than ε. Provided these conditions are satisfied, the results are not very sensitive to the specific values of the cutoff.

*Present address: Department of Physics, S.U.N.Y. at Stony Brook, Stony Brook, NY 11794-3800.

[1]R. J. Birgeneau and G. Shirane, in *Physical Properties of High Temperature Superconductors,* edited by D. M. Ginsberg (World Scientific, Singapore, 1989); for information specifically on the insulating state, a short description is given in D. Vaknin, S. K. Sinha, D. E. Moncton, D. C. Johnston, J. M. Newsam, C. R. Safinya, and H. E. King Jr., Phys. Rev. Lett. **58**, 2802 (1987).

[2]P. W. Anderson, Science **235**, 1196 (1987); (unpublished).

[3]D. C. Mattis, *The Theory of Magnetism* (Springer-Verlag, New York, 1981).

[4]T. Kennedy, E. H. Lieb, and B. S. Shastry, Phys. Rev. Lett. **61**, 2582 (1988); K. Kubo and T. Kishi, *ibid.* **61**, 2585 (1988).

[5]It is known from the work of N. D. Mermin and H. Wagner, Phys. Rev. Lett. **17**, 1133 (1966) that the Heisenberg model in two dimensions is disordered at any nonzero temperature.

[6]E. Lieb, T. Schulz, and D. Mattis, Ann. Phys. **16**, 407 (1961); W. Marshall, Proc. R. Soc. London, Ser. A **232**, 48 (1955).

[7]The singlet nature of the ground state does not preclude LRO. Although no direction is picked out by the ground state, the spin-spin correlation function can be finite at large distances, implying a net staggered magnetization.

[8]J. Oitmaa and D. D. Betts, Can. J. Phys. **56**, 897 (1978); E. Dagotto and A. Moreo, Phys. Rev. B **38**, 5087 (1988).

[9]D. A. Huse and V. Elser, Phys. Rev. Lett. **60**, 2531 (1988).

[10]H. Yokoyama and H. Shiba, J. Phys. Soc. Jpn. **56**, 3570 (1987).

[11]S. Liang, B. Doucot, and P. W. Anderson, Phys. Rev. Lett. **61**, 365 (1988).

[12]J. D. Reger and A. P. Young, Phys. Rev. B **37**, 5978 (1988).

[13]M. Gross, E. Sanchez-Velasco, and E. Siggia, Phys. Rev. B **39**, 2484 (1989).

[14]D. A. Huse, Phys. Rev. B **37**, 2380 (1988).

[15]R. R. P. Singh, Phys. Rev. B **39**, 9760 (1989).

[16]D. M. Ceperley and M. H. Kalos, in *Monte Carlo Methods in Statistical Physics,* edited by K. Binder (Springer-Verlag, Berlin, 1979); K. E. Schmidt and M. H. Kalos, in *Applications of the Monte Carlo Method in Statistical Physics,* 2nd ed., edited by K. Binder (Springer-Verlag, Berlin, 1984).

[17]T. Barnes, D. Kotchan, and E. S. Swanson, Phys. Rev. B **39**, 4357 (1989).

[18]J. Carlson, Phys. Rev. B **40**, 846 (1989).

[19]T. Matsubara and H. Matsuda, Prog. Theor. Phys. **16**, 569 (1956).

[20]N. Trivedi and D. Ceperley, Phys. Rev. B **40**, 2737 (1989).

[21]See Sec. V for the effects of other boundary conditions.

[22]It is easy to see that even though the kinetic energy operator has a positive sign, it acquires a negative expectation value from the alternating nature of the phase of the wave function.

[23]R. P. Feynman, *Statistical Mechanics* (Benjamin, New York, 1972).

[24]A. Bijl, Physica (Utrecht) **7**, 860 (1940); R. Jastrow, Phys. Rev. **98**, 1479 (1955).

[25]R. P. Feynman, Phys. Rev. **94**, 262 (1954).

[26]J. P. Hansen and I. R. McDonald, *Theory of Simple Liquids* (Academic, New York, 1976).

[27]G. V. Chester and L. Reatto, Phys. Lett. **22**, 276 (1966).

[28]D. Pines and P. Nozieres, *The Theory of Quantum Liquids* (Addison-Wesley, Reading, Mass., 1989), Vol. I.

[29]S. E. Koonin, *Computational Methods* (Addison-Wesley, Reading, Mass., 1986); M. H. Kalos and P. A. Whitlock, *Monte Carlo Methods: The Basics* (Wiley, New York, 1986).

[30]P. A. Whitlock, M. H. Kalos, G. V. Chester, and D. M. Ceperley, Phys. Rev. B **19**, 5598 (1979).

[31]D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980); J. Chem. Phys. **81**, 5833 (1984).

[32]P. W. Anderson, Phys. Rev. **86**, 694 (1952); T. Oguchi, *ibid.* **117**, 117 (1960).

[33]H. Neuberger and T. Ziman, Phys. Rev. B **39**, 2608 (1989).

[34]Note that in (2.2) the operators b and $b^\dagger$ anticommute at a site that is equivalent to a hard-core condition. In momentum space this gives $[b_{\mathbf{k}}, b^\dagger_{\mathbf{k}'}] = \delta_{\mathbf{k},\mathbf{k}'} - 2n(\mathbf{k}-\mathbf{k}')$, which is modified from the usual commutation relation for bosons. However, it is also possible to treat the bosons as ordinary species and include the hard-core condition in the commutation relations as an infinitely repulsive potential. This on-site potential will also commute with $\rho(\mathbf{q})$ and will not contribute to the f-sum rule.

[35]L. Reatto, Phys. Rev. B **26**, 130 (1982).

[36]E. Manousakis (unpublished).

[37]A. Griffin, Phys. Rev. B **30**, 5057 (1984).

[38]D. M. Ceperley, J. Stat. Phys. **43**, 815 (1986).

[39]K. S. Liu, M. H. Kalos, and G. V. Chester, Phys. Rev. A **10**, 303 (1974).

[40]A. J. Leggett, Phys. Rev. Lett. **25**, 1543 (1970).

[41]J. M. Ziman, *Principles of the Theory of Solids,* 2nd ed. (Cambridge University Press, Cambridge, England, 1972).

[42]J. D. Reger, J. A. Riera, and A. P. Young, J. Phys. Cond. Matter **1**, 1855 (1989).

[43]S. Tang and J. Hirsch (unpublished).