RAPID COMMUNICATIONS

PHYSICAL REVIEW B VOLUME 40, NUMBER 4 1 AUGUST 1989

Green-function Monte Carlo study of quantum antiferromagnets

Nandini Trivedi
Department of Physics, University of Illinois at Urbana-Champaign, 1110 West Green Street, Urbana, Illinois 61801

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 19 April 1989)

We have studied via Green-function Monte Carlo (GFMC) technique the $S=\frac{1}{2}$ Heisenberg quantum antiferromagnet in two dimensions on a square lattice. GFMC is a $T=0$ stochastic method that projects out the component of the ground state in a given variational wave function. From studies on lattices up to 12×12, we find the ground-state energy per site $E_0/J=-0.6692(2)$. We include the zero-point motion of the elementary excitations in the ground state and show that it produces long-range correlations in the wave function. We obtain a staggered magnetization $m^{\dagger}=0.31(2)$ in units in which the classical Néel state is 0.5. The structure factor at long wavelengths is $S(q)\sim q$ and from the slope we deduce the spin-wave velocity.

The discovery of high-temperature superconductivity has brought about an increased interest in the problem of quantum antiferromagnets.[1] Recently, there have been a number of calculations of the $S=\frac{1}{2}$ Heisenberg model in two dimensions on a square lattice by different methods: exact diagonalization on small clusters[2] up to $N=4\times4$, a variational calculation[3] using trial states with long-range order (LRO) built in, and a finite-temperature path-integral Monte Carlo study.[4,5] Analytically, this problem has been approached by a renormalization-group analysis of the nonlinear σ model as well as by a $1/N$ expansion of the generalized SU(N) Heisenberg model.[6] Based on these studies, there seems to be a growing consensus that the ground state has LRO with a staggered magnetization 60% of the classical value. On the other hand, Liang, Doucot, and Anderson[7] have done a variational calculation using resonating-valence-bond states that do not have any LRO. They show that by including singlet bonds between distant sites they can lower the energy to within 1% of the energy of a state with long-range order.

In this paper, we describe a different method to study the Heisenberg quantum antiferromagnet from those given above, called Green-function Monte Carlo (GFMC).[8] Previously, it has been used successfully to obtain the ground-state and low-lying excited-state properties of *continuum* many-body systems. It has the advantage over finite-temperature Monte Carlo as it does not require an extrapolation to zero temperature or to small time step. It can be applied to fairly large lattices, limited only by computer time considerations, unlike exact diagonalization methods that are limited by memory. GFMC goes beyond variational methods by projecting out of a variational wave function Ψ_T the component of the *true* ground state. We also address the question concerning the nature of correlations present in the ground-state wave function. We argue that to obtain the correct behavior of the staggered magnetization and the structure factor, it is important to include the *zero-point motion* of the spin waves and show that they produce *long-range* correlations between the spins in the ground-state wave function.

The Hamiltonian for the $S=\frac{1}{2}$ Heisenberg quantum antiferromagnet is given by $H=J\sum_{\langle ij\rangle}\mathbf{S}_i\cdot\mathbf{S}_j$ where the coupling J is positive and connects spins on nearest-neighbor sites of a square lattice. Given the commutation relations of the spins, we introduce a new set of boson operators[9] by the following transformation: $S_i^{\dagger}=\epsilon_i b_i^{\dagger}$ and $S_i^z=n_i-\frac{1}{2}$, where $\epsilon_i=\pm1$ on the even (odd) sublattices. Note that since $(b_i^{\dagger})^2=0$, the bosons have a hard core, i.e., a site can be occupied by, at most, one boson. The Heisenberg Hamiltonian is transformed to

$$H=-\frac{J}{2}\sum_{\langle ij\rangle}(b_i^{\dagger}b_j+b_ib_j^{\dagger})+J\sum_{\langle ij\rangle}n_in_j+E_N\,, \tag{1}$$

where $E_N=-JZN/8$ is the energy of the classical Néel state, N is the total number of sites, and $Z=4$ is the coordination number. The first term in Eq. (1) is the kinetic energy for bosons and the second term is a repulsive interaction between bosons on adjacent sites, in addition to the hard-core constraint in the commutation relations. To obtain ground-state properties, we work in the $(S_z)_{\rm tot}=0$ subspace, which implies that the number of bosons is half the number of sites; the computational methods described below are, however, applicable to any value of S_z.

Given a variational wave function $\Psi_T(R)$, where $\{R\}$ is a particular configuration of bosons, the expectation value of an operator is given by $\langle\hat{O}\rangle_V=\langle\Psi_T|\hat{O}|\Psi_T\rangle/\langle\Psi_T|\Psi_T\rangle$. The expectation values require an evaluation of $2N$-dimensional summations, best handled by Monte Carlo methods.

To go beyond the variational method, we define an iterative procedure to obtain the wave function at the nth time step, given its value at the previous time step by

$$\Phi^{(n)}(R)=\sum_{R'}\langle R|G|R'\rangle\Phi^{(n-1)}(R')\,, \tag{2}$$

where $G=1-\tau(H-w)$. The parameters w and τ are chosen such that G projects out the ground state. This is achieved by choosing $w\simeq E_0$, the ground-state energy,

and the time step to be $\tau \leq 2/(E_{\max} - w)$, where $E_{\max}$ is the maximum eigenvalue of H. For this problem, we have $E_{\max} \approx NJ$, corresponding to maximum potential energy, so that $\tau \approx 1/NJ$. This guarantees that all the eigenvalues of G corresponding to the excited states of H have absolute values less than unity. As $n \rightarrow \infty$, the wave function $\Phi^{(n)}$ approaches the ground-state wave function Φ_0. The form of the projection operator in Eq. (2) is well suited for a lattice problem for which there is a cutoff in energy. It does *not* involve a time-step error and is simple to evaluate.

The ground-state wave function for a Bose system is non-negative so that $\Phi^{(n)}(R)$ can be treated as a *probability density*. We begin with a set of about 1000 configurations or walkers distributed according to $\Phi^{(0)}(R') \equiv \Psi_T(R')$. The propagation of a random walker from a point R' in configuration space to R in Eq. (2) involves two basic processes: (i) Diffusion:—these moves are governed by the kinetic energy operator that hops a boson from one site to another. (ii) Branching:—In the course of the random walks, if a configuration evolves into a region with a very high potential energy, such walks are terminated. On the other hand, if a starting configuration evolves into one with a low potential energy, such favorable configurations are multiplied. The process of making multiple copies is called branching. In addition, it is possible to significantly reduce the variance of the energy by importance sampling. The basic idea is to sample $\Phi^{(n)}\Psi_G$, which biases the diffusion of random walks in configuration space according to a simple *guiding function* Ψ_G. The population at the nth time step or generation, given by the total number of walkers in that generation, and its growth and decline is controlled by the constant w. We continue iterating Eq. (2) with importance sampling for approximately 1000 generations until the required variance is achieved. Details of this method will be published elsewhere.[10]

To construct a wave function for the many-body interacting boson system, we use a Jastrow trial state

$$\Psi_{SR}(R) = \prod_{i<j} f(\mathbf{r}_i - \mathbf{r}_j) , \tag{3}$$

where f is typically a short-ranged function of the relative distance between bosons i and j on a lattice. The pair correlated wave function in Eq. (3) is symmetric under the exchange of any two particle coordinates. The hard-core condition requires that $f(0) = 0$, and for large separation between the bosons we take f to be unity. For intermediate values, f is determined by a variational calculation by minimizing the energy or by minimizing the difference between the variational and GFMC estimates of the correlation functions. If $f(\boldsymbol{\delta}) = 0$, where $\boldsymbol{\delta}$ is a vector to a nearest-neighbor site, then one of the sublattices is occupied by bosons and the other one is empty, i.e., we get a state with *diagonal* LRO. In terms of spins, this is a Néel state with sublattice magnetization along the z direction. On the other hand, if $f(\boldsymbol{\delta}) = 1$, the boson wave function has *off-diagonal* LRO, and in terms of spins, corresponds to a Néel state with sublattice magnetization in the XY plane. For guiding the random walks, we use a one-parameter wave function with the optimal value of $f(\boldsymbol{\delta}) \sim 0.58$ in Eq. (3).

By GFMC we obtain the mixed estimate of an operator $\hat{O}$ between the true ground state Φ_0 and a trial state or *importance function* Ψ_T that can be different from the guiding function, given by $\langle \hat{O} \rangle_M = \langle \Phi_0 | \hat{O} | \Psi_T \rangle / \langle \Phi_0 | \Psi_T \rangle$. For the special case of the energy, since Φ_0 is an eigenstate of the Hamiltonian, the mixed estimate gives the true ground-state energy $\langle \hat{H} \rangle_M = E_0$. For a general operator, the mixed estimator must be corrected to give the true expectation value in the ground state. One way to do this is to assume that the difference between the trial wave function and the true ground-state wave function is small,[8] so that $\langle \Phi_0 | \hat{O} | \Phi_0 \rangle = 2\langle \hat{O} \rangle_M - \langle \hat{O} \rangle_V$.

The variational Monte Carlo (VMC) estimates of the energy per site at the optimal value of the parameter $f(\boldsymbol{\delta})$ is shown in Fig. 1. It is approximately 30% lower than the Néel-state energy $E_N = -0.5$ J. Also shown in Fig. 1 are the GFMC results for the energy for lattice sizes up to 12×12 with periodic boundary conditions. We emphasize that even with an importance function as simple as the classical Néel state, we obtain good agreement with the GFMC results in Fig. 1, though as expected, the error bars are larger for the same number of iterations. For a 4×4 lattice our result for the energy ($E_0/J = -0.7018 \pm 0.0002$) agrees very well with the exact results[2] ($E_0/J = -0.701\,780$). The energy is found to scale as L^{-3}, where L is one edge of the lattice and the extrapolated value is

$$E_0/J = -0.6692 \pm 0.0002 . \tag{4}$$

The excess energy $E_0 - E_N$ obtained by GFMC is significantly lower (by $\sim 13\%$) than the one-parameter variational estimate of the energy. Gross, Sanchez-Velasco, and Siggia[5] used a method similar to GFMC,

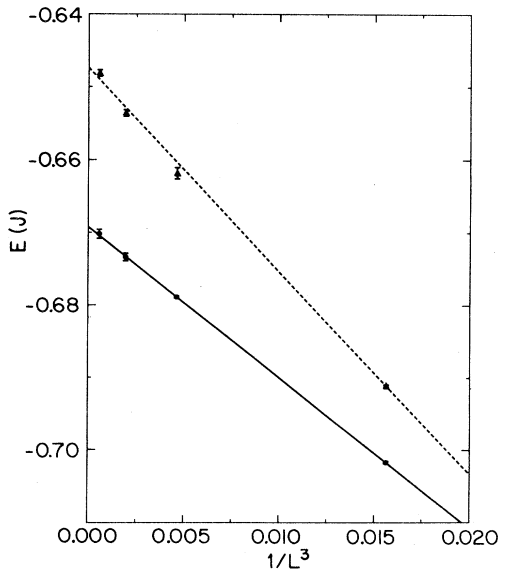

FIG. 1. The GFMC results for the energy per site E as a function of L^{-3}. The GFMC results (solid curve) is compared with the VMC energy values (dashed curve). Both used a short-range one-parameter wave function. The extrapolated excess energy (over the classical Néel state contribution), $E - 0.5$ J obtained by GFMC is lower than the VMC by $\sim 13\%$.

but without importance sampling. We find that our results for lattice sizes up to 12×12 are consistently lower compared to those of Ref. 5 by ~2% (which is about seven standard errors). Perhaps their results have not converged, especially since in the absence of importance sampling there may be large fluctuations in the data. We find that the correlation functions for a 4×4 lattice obtained by the GFMC method are in excellent agreement[10] with the exact results. This shows that even though the trial state has unequal correlations along x, y, and z, the ground-state wave function arrived at with GFMC is very close to a singlet for small lattices.

We give the structure factor for a 12×12 lattice along the [10] direction, using a short-range importance function [Eq. (3)] in the mixed estimate. At long wavelengths, as depicted in Fig. 2, the structure factor is of the form $S_{SR}(q) \sim a + bq^2$. From the Feynman-Bijl theorem, this is not consistent with the elementary excitations being phonons or spin waves. This points to the inadequacy of using importance functions with only short-range correlations and suggests the importance of including the *zero-point motion* of the elementary excitations in the ground-state wave function.[10,11] The Hamiltonian for the collective coordinates representing the sound mode or spin wave is given by

$$H_{LR} = \tfrac{1}{2} \sum_{q<q_c} m_{\mathbf{q}} [\dot{\rho}(\mathbf{q})\dot{\rho}(-\mathbf{q}) + \omega_{\mathbf{q}}^2 \rho(\mathbf{q})\rho(-\mathbf{q})] , \quad (5)$$

where the density fluctuation is related to the displacements by $\rho \sim \nabla \cdot \mathbf{u}$ so that $m_{\mathbf{q}} \sim 1/q^2$. Also $\omega_{\mathbf{q}} = cq$ at long wavelengths (for q less than a cutoff q_c), and c is the velocity of spin waves. The harmonic-oscillator wave function arising from these spin waves is of the form

$$\Psi_{LR} = \prod_{q<q_c} \exp\left[-\frac{m_{\mathbf{q}}\omega_{\mathbf{q}}}{2} \rho(\mathbf{q})\rho(-\mathbf{q}) \right] = \exp\left[-\sum_{i<j} \frac{\tilde{a}}{r_{ij}} \right] , \quad (6)$$

up to a normalization constant, where $\rho(\mathbf{q}) = \sum_i \exp(i\mathbf{q}\cdot\mathbf{r}_i)$. Thus, we find that the inclusion of the zero-point motion of spin waves produces a *long-range* contribution to the wave function.[12] From Eq. (6), the Jastrow factor in Eq. (3) at large separation between the bosons is of the form $f(r) \sim 1 - a/r$, where $a = e^{-\tilde{a}}$. An improved *ansatz* for the importance function is to multiply Eqs. (3) and (6), $\Psi_T = \Psi_{SR}\Psi_{LR}$. A similar wave function was studied by Huse and Elser,[3] described by $f(r) = 1 - a/r^\beta$ at large r. Treating β as a variational parameter, they obtained a value of ~0.7. Here, we give a physical justification for an exponent of $\beta = 1.0$ based on a spin-wave analysis. The structure factor for N_B bosons described by the wave function in Eq. (6) is $S(\mathbf{q}) = (1/N_B)\langle \rho(\mathbf{q})\rho(-\mathbf{q}) \rangle \sim q$ in the long-wavelength limit as seen in Fig. 2 for both lattice sizes 8×8 and 12×12. Note that at large q the structure factor calculated with a short-range importance function merges with that obtained with a long-range importance function, as expected.

The GFMC estimates of the energy in Fig. 1 are not affected by the inclusion of Eq. (6), since the energy is sensitive mainly to short-range correlations in the wave function. However, the spin-spin correlation functions are

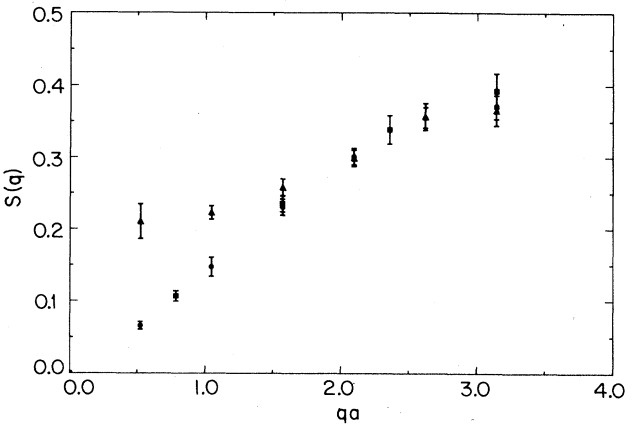

FIG. 2. Structure factor obtained by GFMC. (To obtain the ground-state expectation, the mixed estimate is extrapolated or corrected as described in the text.) $S(\mathbf{q})$ along the [10] direction with only short-range importance function is shown by triangles. Note that it incorrectly goes to a nonzero value at small q. $S(\mathbf{q})$ along [10] for a $N = 8\times8$ lattice (squares) and for a $N = 12\times12$ lattice (circles) with long-range importance function. The spin-wave velocity is extracted from the linear dependence at small q.

dramatically altered. The trial states in Eqs. (3) and (6), at the optimal value of the parameters, possess mainly off-diagonal order measured by $h(l) = (1/N)\sum_i \langle b_i^\dagger b_{i+l} \rangle$. The nonzero value of $h(l)$ at large distances is related to the condensate fraction of the boson superfluid. In terms of spins, $h(l)$ describes the correlation between the x (or y) components of the spins and the constant value at $(L/2, L/2)$ defined as $(m_x^\dagger)^2 + (m_y^\dagger)^2$ is associated with the presence of sublattice magnetization in the x and y directions. We also define the density-density correlation function, or equivalently, the spin-spin correlation function of the staggered z component by

$$g(l) = (1/N)\sum_i \epsilon_i \epsilon_{i+l} \langle n_i n_{i+l} \rangle .$$

$g(l)$ decays to $(m_z^\dagger)^2 \approx 0$ at large distances, showing the lack of diagonal LRO; however, for a finite lattice, we extract a small value for $(m_z^\dagger)^2$. Thus, from the long-distance behavior of $h(l)$ and $g(l)$ we obtain the value of $(m^\dagger)^2$. In Fig. 3, we give the scaling of the staggered magnetization obtained by the VMC and GFMC methods. In the classical Néel state, the value of $m^\dagger = 0.5$; however, for an $S = \frac{1}{2}$ antiferromagnet, quantum fluctuations reduce $m^\dagger$ from it classical value. We find that the extrapolated value obtained by using GFMC with the *short-range* importance function is $m^\dagger = 0.37$, which is lower than the variational results by about 8%. If we now include a *long-range* importance function, we find a GFMC estimate of

$$m^\dagger = 0.31 \pm 0.02 . \quad (7)$$

This is reduced by ~19% compared to the GFMC estimate in short-range importance functions. Our result in Eq. (7) is consistent with spin-wave[13] and finite-temperature Monte Carlo values[4] of $m^\dagger \sim 0.30$.

We obtain a rough estimate of the spin-wave velocity

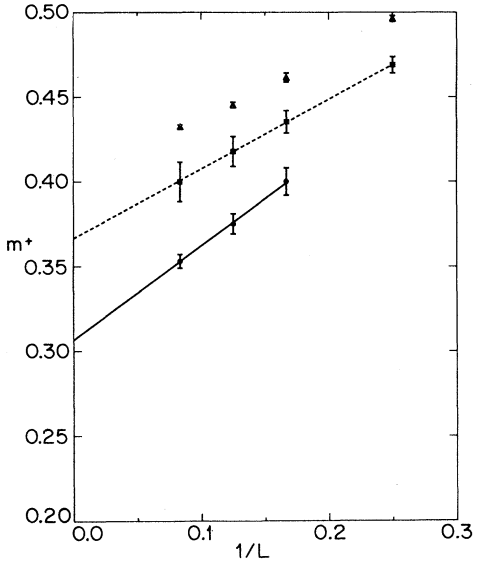

FIG. 3. The staggered magnetization $m^{\dagger}$ vs $1/L$ evaluated at $(L/2,L/2)$ for two different importance functions. Short-range importance function: The GFMC estimate (extrapolated) are the squares and the VMC estimate is shown by the triangles. The dashed line is a least-squares fit to the squares and extrapolates to 0.37 for the infinite lattice. The classical Néel state has $m^{\dagger}=0.5$. Long-range importance function: The GFMC estimate is shown by the 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.

from the structure factor in the ground state by using the f-sum rule. At long wavelengths, the collective excitations exhaust the sum rule and we obtained the well-known Feynman-Bijl formula, $\omega_{\mathbf{q}}=2\langle n(\delta)\rangle\epsilon_{\mathbf{q}}/S_{\mathbf{q}}$. It is somewhat modified on a lattice by the factor $\langle n(\delta)\rangle \approx 0.25$, which is the average amplitude to hop to the nearest-neighbor site.[14] Here $\epsilon_{\mathbf{k}}=J\sum_{\delta}[1-\cos(\mathbf{k}\cdot\delta)]$ for the free-particle energy on a lattice with the sum over the nearest-neighbor sites along $+x$ and $+y$. In Fig. 2, it is seen that $S(\mathbf{q})$ is linear in the long-wavelength region when evaluated using a long-range importance function. The spin-wave velocity is given by

$$Z_c \equiv c/c_0 \sim qa\langle n(\delta)\rangle/[\sqrt{2}S(q)]\,,$$

where $\hbar c_0=2(1-\gamma_q^{\,2})^{1/2}/(qa)\sim\sqrt{2}Ja$ is the classical spin-wave velocity and $\gamma_q=\frac{1}{2}[\cos(q_x a)+\cos(q_y a)]$. We find that Z_c, which is a measure of the renormalization of the spin-wave velocity caused by quantum fluctuations, is $Z_c=1.14\pm0.05$. This should be compared with the spin-wave analysis around the classical Néel state[13] which gives $Z_c=1.158$. A more accurate estimate of the spin-wave velocity can be obtained by a direct study of the excited states via the GFMC method.

In conclusion, we have used the Green-function Monte Carlo method to study the exact ground-state properties of the Heisenberg model by projecting out the excited states in a trial state. We have found that starting with even rather simple trial states, e.g., the classical Néel state, it is possible to obtain very accurate estimates of the ground-state energy for fairly large lattices. Here we have presented results for lattice sizes up to 12×12 but larger lattices are easily possible.[15] To obtain the correct staggered magnetization and excitation spectrum, we show that the zero-point motion of the elementary excitations must be included in the ground state. This produces long-range pair correlations in the wave function and using spin-wave analysis we have obtained the functional form of these correlations.

The success of GFMC for the Heisenberg model opens up the possibility of its application to other problems. Although approximations need to be made to deal with fermions or frustration, some of them have been tested on the electron gas and on ^{3}He with good results.[8] It is therefore encouraging to use the GFMC method to study frustrated antiferromagnets, Hubbard, and extended Hubbard models with doping.

We would like to thank R. M. Martin, E. Fradkin, S. Liang, E. Manousakis, V. Pandharipande, and M. Randeria for several stimulating discussions. This work has been supported by the Division of Materials Research, National Science Foundation Grant No. 86-12860, and by the National Center for Supercomputing Applications.

[1] R. J. Birgeneau and G. Shirane, in *Physical Properties of High Temperature Superconductors,* edited by D. M. Ginsberg (World Scientific, Singapore, 1989); P. W. Anderson, Science **235**, 1196 (1987).

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

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

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

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

[6] S. Chakravarty, B. Halperin, and D. R. Nelson, Phys. Rev. Lett. **60**, 1057 (1988); D. P. Arovas and A. Auerbach, Phys. Rev. B **38**, 316 (1988).

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

[8] 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).

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

[10] N. Trivedi and D. M. Ceperley (unpublished).

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

[12] We use Ewald summation techniques to treat the long-range importance function as described in Ref. 10.

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

[14] In the continuum, $\langle n(\delta)\rangle\rightarrow\langle n(0)\rangle=N_B/N=\frac{1}{2}$ and the usual Feynman-Bijl expression is recovered, where N_B is the number of bosons and N is the number of sites.

[15] The $N=12\times12$ lattice takes about 3 h of CPU time on a Cray XMP/48.