PHYSICAL REVIEW B VOLUME 36, NUMBER 16 1 DECEMBER 1987

Hybrid molecular-dynamics algorithm for the numerical simulation of many-electron systems

R. T. Scalettar, D. J. Scalapino, and R. L. Sugar
Department of Physics, University of California, Santa Barbara, California 93106

D. Toussaint
Department of Physics, University of California, La Jolla, California 92093
(Received 29 May 1987)

We present an algorithm for the numerical simulation of many electron systems, which is based upon the hybrid stochastic approach presently being used to study lattice gauge theory. A method for preconditioning the fermion matrices is described which significantly reduces the computing time needed in the simulation. Results are presented for the three-dimensional Hubbard model on lattices ranging in size from 4^3 to 10^3.

I. INTRODUCTION

Numerical simulation has proven to be a very powerful tool for studying systems with many strongly coupled degrees of freedom. This has been particularly true for systems with bosonic or spin degrees of freedom and for fermionic systems in one dimension. However, the simulation of two- and three-dimensional systems with many fermion degrees of freedom has proven to be extremely challenging. In this paper we present an algorithm which we believe has considerable promise for the study of a variety of many electron systems in two and three dimensions. We describe the algorithm in detail in Sec. II. It is based upon the hybrid stochastic method[1-3] which is presently being used with considerable success to study lattice-gauge theory with dynamical quarks.[4,5]

In this paper we illustrate our approach by applying it to the three-dimensional Hubbard model. We have chosen this model because of its intrinsic interest, and because we believe that it provides a particularly stringent test of fermion algorithms. The fermion matrices for the Hubbard model become ill conditioned at low temperatures or at strong coupling, considerably slowing the simulation. In Sec. III we show how this difficulty can be overcome by preconditioning the fermion matrices.

In Sec. IV we discuss tuning of the algorithm and compare our results on a 4^3 lattice with those obtained by Hirsch using an exact updating method. We also present initial results obtained on 4^3, 6^3, 8^3, and 10^3 lattices. Finally in Sec. V we briefly discuss our results and the application of our approach to other models.

The numerical calculations presented in this paper were performed on the San Diego Supercomputer Center Cray-XMP and on our ST-100 array processor.

II. THE ALGORITHM

Our approach can be employed to study a wide variety of many-electron models. However, for definiteness we shall describe its application to the three-dimensional Hubbard model. The Hamiltonian we shall study is

$$H=-t\sum_{\langle i,j\rangle,\sigma}(c_{i\sigma}^{\dagger}c_{j\sigma}+c_{j\sigma}^{\dagger}c_{i\sigma})+U\sum_{i}(n_{i+}-\tfrac{1}{2})(n_{i-}-\tfrac{1}{2})\ . \tag{1}$$

Here $c_{i\sigma}^{\dagger}$ and $c_{i\sigma}$ are the creation and annihilation operators for electrons located on the ith lattice site with z component of spin $\sigma=\pm$, and $n_{i\sigma}=c_{i\sigma}^{\dagger}c_{i\sigma}$. U is the repulsive Coulomb interaction between electrons on the same lattice site, t is the hopping parameter, and $\langle i,j\rangle$ represents a pair of nearest-neighbor sites on the lattice. We shall consider the case of a half-filled band, so we have explicitly written the Hamiltonian in particle-hole symmetric form.

We are interested in performing calculations at finite temperature. The average value of a physical observable, A is given by

$$\langle A\rangle=\frac{tr\,Ae^{-\beta H}}{tre^{-\beta H}}\ . \tag{2}$$

We begin by turning the traces in Eq. (2) into path integrals. To this end we break up the imaginary-time interval, $0\le\tau\le\beta$, into N_τ subintervals of width $\Delta\tau$. Then the partition function can be written in the form

$$\begin{aligned}Z&=tre^{-\beta H}\\&=tr(e^{-\Delta\tau H})^{N_\tau}\\&\simeq tr(e^{-\Delta\tau H_K}e^{-\Delta\tau H_V})^{N_\tau}\ ,\end{aligned} \tag{3}$$

where

$$H_K=-t\sum_{\langle i,j\rangle,\sigma}(c_{i\sigma}^{\dagger}c_{j\sigma}+c_{j\sigma}^{\dagger}c_{i\sigma}) \tag{4}$$

and

$$H_V=U\sum_{i}(n_{i+}-\tfrac{1}{2})(n_{i-}-\tfrac{1}{2})\ . \tag{5}$$

For many models of interest the Hamiltonian is either quadratic in the electron reaction and annihilation operators or can be made so by the introduction of one or more auxiliary fields. For the Hubbard model we introduce a Hubbard-Stratonovich variable $x_{i,l}$ at each lattice site i and time slice l by means of the identity

$$\exp[-\Delta\tau U(n_{i+}-\tfrac{1}{2})(n_{i-}-\tfrac{1}{2})]=(\Delta\tau/\pi)^{1/2}e^{-\Delta\tau U/4}\int_{-\infty}^{+\infty}dx_{i,l}\exp\{-\Delta\tau[x_{i,l}^2+(2U)^{1/2}x_{i,l}(n_{i+}-n_{i-})]\}\ . \tag{6}$$

We can now perform the trace over the fermion coordinates to obtain[6]

$$Z=\int[\delta x]e^{-S_B(x)}\det M_+(x)\det M_-(x)\ , \tag{7}$$

where [ ] indicates a functional integral. The partition function can be written in the general form given in Eq. (7) for a wide class of models. For the particular case of the Hubbard model we have[6]

$$S_B(x)=\Delta\tau\sum_{i,l}x_{i,l}^2 \tag{8}$$

and

$$M_\sigma=\begin{pmatrix} I & 0 & 0 & 0 & \cdots & B_{N_\tau,\sigma} \\ -B_{1,\sigma} & I & 0 & 0 & \cdots & 0 \\ 0 & -B_{2,\sigma} & I & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & -B_{N_\tau-2,\sigma} & I & 0 \\ 0 & 0 & \cdots & 0 & -B_{N_\tau-1,\sigma} & I \end{pmatrix}\ . \tag{9}$$

The entries in the matrix M_σ are themselves $V_s\times V_s$ matrices, where V_s is the number of spatial lattice points. I is the unit matrix, and

$$B_{l,\sigma}=e^{-\Delta\tau K}e^{-\sigma\Delta\tau V_l}\ , \tag{10}$$

where

$$K_{i,j}=\begin{cases}-t & \text{if } i \text{ and } j \text{ are nearest neighbors}\ ,\\ 0 & \text{otherwise}\ .\end{cases}$$
$$(V_l)_{i,j}=\delta_{i,j}(2U)^{1/2}x_{i,l}\ . \tag{11}$$

Making use of a particle-hole transformation on the spin-down electrons, it is easy to see from Eqs. (3) and (6) that[7]

$$\det M_-(x)=\exp\left[-\Delta\tau(2U)^{1/2}\sum_{i,l}x_{i,l}\right]\det M_+(x)\ , \tag{12}$$

so the integrand of Eq. (7) is positive definite, and it is permissible to replace $\det M_+(x)$ and $\det M_-(x)$ by their absolute values. We can therefore eliminate the determinants by the introduction of two auxiliary scalar fields, Φ_+ and Φ_-, and write

$$Z=\int[\delta x\,\delta\Phi_+\delta\Phi_-]\exp\{-[S_B(x)+\Phi_+O_+^{-1}\Phi_+ +\Phi_-O_-^{-1}\Phi_-]\}\ , \tag{13}$$

where

$$O_\sigma=M_\sigma^TM_\sigma\ . \tag{14}$$

We do not choose to employ the standard Metropolis algorithm in our simulation since it would require us to recompute $\Phi_\sigma O_\sigma^{-1}\Phi_\sigma$ each time we updated one component of x. This would be too costly in computer time. We can greatly reduce the number of times we need to recompute $O_\sigma^{-1}\Phi_\sigma$ by making use of the ideas of molecular dynamics. To this end we introduce a set of coordinates $p_{i,l}$ which play the role of momenta conjugate to $x_{i,l}$ and we extend the definition of the partition function to

$$\begin{aligned}Z&=\int[\delta x\,\delta p\,\delta\Phi_+\delta\Phi_-]\\ &\quad\times\exp\left[-\left[\sum_{i,l}p_{i,l}^2+S_B(x)+\Phi_+O_+^{-1}\Phi_++\Phi_-O_-^{-1}\Phi_-\right]\right]\\ &=\int[\delta x\,\delta p\,\delta\Phi_+\delta\Phi_-]e^{-H(x,p,\Phi_\sigma)}\ .\end{aligned} \tag{15}$$

Clearly the introduction of the p field has no effect on correlation functions involving x and the Φ_σ.

Equation (15) is the starting point for our simulation. We wish to obtain a sequence of field configurations $\{x,p,\Phi_\sigma\}$ with a probability distribution proportional to $\exp(-H)$. Physical observables that can be expressed in terms of the $x_{i,l}$ can be easily measured in this set of configurations. In general, measurements of fermion correlation functions will require us to obtain averages of various components of M_σ^{-1} or their products. We will show below how M_σ^{-1} can be simply and rapidly evaluated using an unbiased estimator.

We employ three types of updating steps to obtain the desired distribution of field configurations. The com-

ponents of p and Φ_σ are updated using heat baths. We simply set each $p_{i,l}$ equal to a Gaussian random number with probability distribution proportional to $\exp(-p_{i,l}^2)$. For the Φ_σ fields we generate two vectors of Gaussian random numbers, R_σ, each component of which has a probability distribution proportional to $\exp(-R_{i\sigma}^2)$. We then write

$$\Phi_\sigma = M_\sigma^T R_\sigma \ , \tag{16}$$

from which it follows trivially that the Φ_σ have a probability distribution proportional to $\exp(-\Phi_\sigma O_\sigma^{-1}\Phi_\sigma)$.

The only nontrivial part of the algorithm is the updating of the x field. We vary x and p for fixed Φ_σ according to Hamilton's equations. That is,

$$\begin{aligned} \dot{p}_{i,l} &= -\frac{\partial H}{\partial x_{i,l}} = -\frac{\partial V}{\partial x_{i,l}} \ , \\ \dot{x}_{i,l} &= \frac{\partial H}{\partial p_{i,l}} = 2p_{i,l} \ , \end{aligned} \tag{17}$$

where we have written

$$V(x) = S_B(x) + \sum_\sigma \Phi_\sigma O_\sigma^{-1}\Phi_\sigma \ . \tag{18}$$

The dots over $x_{i,l}$ and $p_{i,l}$ in Eq. (18) indicate differentiation with respect to molecular-dynamics or simulation time t. From Liouville's theorem we know that in integrating Hamilton's equations we move along a trajectory in which both H and the differential volume element in phase space are constant.

If the system is in equilibrium, that is, if the probability distribution of the fields is given by

$$P(x,p,\Phi_\sigma) = Z^{-1} e^{-H(x,p,\Phi_\sigma)} \ , \tag{19}$$

then both the heat-bath updatings and the molecular-dynamics steps will keep it in equilibrium. If the system is not equilibrium, then just as in any Monte Carlo algorithm the heat-bath steps will drive it towards equilibrium. However, the molecular-dynamics steps will be neutral in this regard since they simply amount to a permutation among configurations of equal probability weight.

In order to introduce the auxiliary fields Φ_σ in Eq. (13), it was sufficient that the product of the determinants of the spin-up and spin-down fermion matrices be positive. This was assured by Eq. (12). However, if there were a surface in the space of the $x_{i,l}$ on which the individual determinants vanish, then on this surface $V(x)$ would be infinite. Hamilton's equations would not propagate the system across such an infinite barrier, so it would be necessary to introduce an additional type of updating step to do so. In order to investigate whether this problem arises for the Hubbard model we have generated a number of random and ordered configurations for the x field, and then minimized $\det M_\sigma$ with respect to the $x_{i,l}$.[8] For $\Delta\tau \le 0.05$, $\beta \le 10.0$, we found the minimum of $\det M_\sigma$ to be 1.0 in all cases. In the minimization process the determinants decreased in size by 20–30 orders of magnitude. We were able to find specific configurations of the x field for which the determinant did become negative, but they carried negligible weight in the junctional integral for the range of parameters we studied. Thus the concern about zeros of the determinant is not empty, and must be checked for each new problem.

We integrate Hamilton's equations numerically, which requires the introduction of a finite step size Δt. We make use of the well-known "leap-frog" method, which rests on the trivial observations that

$$\begin{aligned} x_{i,l}(t+\Delta t) - x_{i,l}(t) &= \int_t^{t+\Delta t} dt' 2p_{i,l}(t') \\ &= 2p_{i,l}(t+\tfrac{1}{2}\Delta t)\Delta t + O(\Delta t^3) \end{aligned} \tag{20}$$

and

$$\begin{aligned} p_{i,l}(t+\tfrac{3}{2}\Delta t) - p_{i,l}(t+\tfrac{1}{2}\Delta t) &= -\int_{t+(1/2)\Delta t}^{t+(3/2)\Delta t} dt' \frac{\partial V(t')}{\partial x_{i,l}} \\ &= -\frac{\partial V(t+\Delta t)}{\partial x_{i,l}}\Delta t + O(\Delta t^3) \ . \end{aligned} \tag{21}$$

Equations (20) and (21) suggest that we define the x and Φ_σ fields at the times $t, t+\Delta t, t+2\Delta t, \ldots$ and p at the times $t+\frac{1}{2}\Delta t, t+\frac{3}{2}\Delta t, \ldots$. Then given $x(t)$ and $p(t+\frac{1}{2}\Delta t)$, we compute $x(t+\Delta t)$, $p(t+\frac{3}{2}\Delta t)$, etc. It will be noted that from Eqs. (17) and (18) that to move one time step Δt, it is necessary to compute $O_\sigma^{-1}\Phi_\sigma$ once. However, with this calculation we are able to update all components of x.

We must update Φ_σ at the full rather than the half-time intervals since we need M_σ for this purpose. We also perform the heat-bath updating of p at the full-time intervals. In order to start the leap-frog process we must then propagate p forward by $\frac{1}{2}\Delta t$. This must be done by the forward difference equation

$$p_{i,l}(t+\tfrac{1}{2}\Delta t) = p_{i,l}(t) - \frac{\partial V(t)}{\partial x_{i,l}}\tfrac{1}{2}\Delta t + O(\Delta t^2) \ . \tag{22}$$

However, in order to evolve x and p at approximately the same rate, we perform of order $1/\Delta t$ molecular-dynamics steps for each heat-bath updating of p. Thus, the error for one physical time step, consisting of one heat-bath updating p and $1/\Delta t$ molecular-dynamics steps, is of order Δt^2. This error is reflected in the measured quantities.

The most time-consuming step in our algorithm is the computation of

$$X_\sigma(t) = O_\sigma^{-1}\Phi_\sigma(t) \ . \tag{23}$$

As is discussed in detail in Sec. III, this calculation is performed by the conjugate-gradient method. One of the major advantages of using deterministic equations for x is that X_σ will then evolve smoothly in time. Except for the two steps immediately following the heat-

bath updating of Φ_σ, we use the simple extrapolation formula

$$X_\sigma(t+\Delta t)=2X_\sigma(t)-X_\sigma(t-\Delta t)+O(\Delta t^2) \tag{24}$$

to obtain a good starting point for the conjugate-gradient calculation. At the expense of storing one extra copy of X_σ, we obtain very significant savings of computer time. Quantitative results will be given in Sec. IV.

In order to measure many quantities of physical interest we need an unbiased estimator of the matrix elements of $M_\sigma^{-1}(x)$. It would be extremely costly in computer time to invert M_σ each time we make a measurement. Instead, we note that immediately after an updating of Φ_σ,

$$X_\sigma=O_\sigma^{-1}\Phi_\sigma=M_\sigma^{-1}R_\sigma \ . \tag{25}$$

Thus, $2X_{i\sigma}R_{j\sigma}$ is an unbiased estimator of $M_{ij\sigma}^{-1}$. In order to measure this quantity or $M_{ij+}^{-1}M_{kl-}^{-1}$ we do not need to perform any matrix inversions beyond those already required in the simulation. To calculate $M_{ij\sigma}^{-1}M_{kl\sigma}^{-1}$ requires only one extra conjugate-gradient calculation, etc.

Let us summarize the algorithm. We begin each updating cycle by making a heat-bath change in p. p is then evolved through a time $\frac{1}{2}\Delta t$ using Eq. (22). Next, of order $1/\Delta t$ molecular-dynamics steps are made using Eqs. (20) and (21). Simultaneously with the heat-bath updating of p, and possibly at several other times during the cycle, a heat-bath updating of Φ_σ is made using Eq. (16). Measurements are made after each new value of Φ_σ is generated.

III. CONDITIONING THE FERMION MATRIX

In the algorithm described in Sec. II the overwhelming fraction of the computer time is spent in solving the V_sN_τ coupled linear equations

$$OX=\Phi \ . \tag{26}$$

(In this section we will suppress the spin index, σ.) As a result, it is worth taking some care to expedite this process.

Since O is a positive-definite symmetric matrix, Eq. (26) can be solved iteratively by means of the conjugate-gradient method. If we denote the value of X after the nth iteration by X_n, and the residual after the nth iteration by

$$r_n=\Phi-OX_n \ , \tag{27}$$

then one iteration consists of the steps

$$\begin{aligned} &X_{n+1}=X_n+a_np_n \ , \\ &r_{n+1}=r_n-a_nOp_n \ , \\ &b_n=(r_{n+1},r_{n+1})/(r_n,r_n) \ , \\ &p_{n+1}=r_{n+1}+b_np_n \ , \\ &a_{n+1}=(r_{n+1},r_{n+1})/(p_{n+1},Op_{n+1}) \ , \end{aligned} \tag{28}$$

where we have denoted the scalar product of r with p by (r,p). The process begins by selecting an initial value of X,X_0, calculating $r_0=\Phi-OX_0$, and setting $p_0=r_0$. As we mentioned in Sec. II, in the midst of the molecular-dynamics steps, we use Eq. (24) to give X_0. Immediately after an updating of Φ we have found nothing better to do than take $X_0=0$. In the following step we use a one-term extrapolation formula for X_0.

The constant a_n defined in Eq. (28) is chosen to ensure that the quadratic form $\frac{1}{2}(X,OX)-(\Phi,X)$ is a minimum at each step of the calculation. b_n is fixed by the requirement that $(p_{n+1},Op_n)=0$. In fact, it is possible to show that, for $m\neq n$,

$$(p_m,Op_n)=0 \ . \tag{29}$$

This guarantees that the p_n are linearly independent and that the iterative procedure will converge in a number of steps equal to the dimension of the matrix, V_sN_τ. Of course, the algorithm will be useful only if convergence to the desired accuracy occurs in significantly fewer steps.

The standard conjugate-gradient algorithm gives an acceptable convergence rate only for small values of U and β. The difficulty is that as either of these variables is increased, the matrix O becomes increasingly ill conditioned. That is, it develops a very wide spread in its eigenvalues. As is well known, the convergence of iterative methods for matrix inversion can often be accelerated by preconditioning the matrix. Let $\tilde{O}$ be a positive-definite symmetric matrix which approximates O, but which can be inverted expeditiously. Such a matrix can always be decomposed into the form

$$\tilde{O}=L^TL \ . \tag{30}$$

Equation (26) can therefore be rewritten as

$$O'X'=\Phi' \ , \tag{31}$$

with

$$\begin{aligned} &O'=L^{T-1}OL^{-1} \ , \\ &X'=LX \ , \\ &\Phi'=L^{T-1}\Phi \ , \\ &r'=\Phi'-O'X'=L^{T-1}r \ . \end{aligned} \tag{32}$$

Since O' is also a positive-definite symmetric matrix, the conjugate-gradient method can be applied to Eq. (31). Equation (28) is replaced by

$$\begin{aligned} &X_{n+1}=X_n+a_n'p_n \ , \\ &r_{n+1}=r_n-a_n'Op_n \ , \\ &b_n'=(r_{n+1},\tilde{O}^{-1}r_{n+1})/(r_n,\tilde{O}^{-1}r_n) \ , \\ &p_{n+1}=\tilde{O}^{-1}r_{n+1}+b_n'p_n \ , \\ &a_{n+1}'=(r_{n+1},\tilde{O}^{-1}r_{n+1})/(p_{n+1},Op_{n+1}) \ , \end{aligned} \tag{33}$$

with starting conditions

$$r_0=\Phi-OX_0\ ,$$
$$p_0=\tilde{O}^{-1}r_0\ . \tag{34}$$

We have tried a variety of forms for the conditioning matrix $\tilde{O}$. The most promising is based on a strong-coupling approximation for O. We start by writing

$$\tilde{O}_0=\begin{bmatrix} \alpha I+B_{0,1}^2 & -B_{0,1} & 0 & 0 & \cdots & B_{0,N_\tau} \\ -B_{0,1} & \alpha I+B_{0,2}^2 & -B_{0,2} & 0 & \cdots & 0 \\ 0 & -B_{0,2} & \alpha I+B_{0,3}^2 & -B_{0,3} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \alpha I+B_{0,N_\tau-1}^2 & -B_{0,N_\tau-1} \\ B_{0,N_\tau} & 0 & 0 & \cdots & -B_{0,N_\tau-1} & \alpha I+B_{0,N_\tau}^2 \end{bmatrix}\ . \tag{35}$$

Here, $B_{0,l}=\exp(-\sigma\Delta\tau V_l)$, where σ is the spin of the electron and V_l is given by Eq. (11). For $\alpha=1$, $\tilde{O}_0$ is simply O evaluated at zero value of the hopping parameter. We have found it essential to take α slightly greater than one so that $\tilde{O}_0$ itself does not become ill conditioned. Fortunately, as long as $\alpha\gtrsim 1.025$, the convergence rate is relatively insensitive to α for a wide range of values of U and β. All results reported in this paper are for $\alpha=1.05$.

The matrix $\tilde{O}_0$ defined in Eq. (35) cannot be inverted analytically for $\alpha\neq 1$. We therefore take the conditioning matrix to be the incomplete Cholesky decomposition of $\tilde{O}_0$,

$$\tilde{O}=L^T D L\ , \tag{36}$$

with

$$D=\begin{bmatrix} D_1 & 0 & 0 & \cdots & 0 \\ 0 & D_2 & 0 & \cdots & 0 \\ 0 & 0 & D_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & D_{N_\tau} \end{bmatrix} \tag{37}$$

and

$$L=\begin{bmatrix} I & -L_1 & 0 & 0 & \cdots & L_{N_\tau} \\ 0 & I & -L_2 & 0 & \cdots & 0 \\ 0 & 0 & I & -L_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & I & -L_{N_\tau-1} \\ 0 & 0 & 0 & \cdots & 0 & I \end{bmatrix}\ . \tag{38}$$

D_l and L_l are $V_s\times V_s$ diagonal matrices chosen so that $\tilde{O}_{i,j}=\tilde{O}_{0i,j}$ for those elements of $\tilde{O}_0$ which do not vanish. There is no restriction on the remaining elements of $\tilde{O}$. These conditions are satisfied by setting

$$L_l=D_l^{-1}B_{0,l}\ ,$$
$$D_l=\alpha I+B_{0,l}^2-B_{0,l-1}D_{l-1}^{-1}B_{0,l-1}\ , \tag{39}$$

for $l=1,\ldots,N_{\tau-1}$, and

$$L_{N_\tau}=D_1^{-1}B_{0,N_\tau}\ ,$$
$$D_{N_\tau}=\alpha I+B_{0,N_\tau}^2-B_{0,N_\tau-1}D_{N_\tau-1}^{-1}B_{0,N_\tau-1}$$
$$-B_{0,N_\tau}D_1^{-1}B_{0,N_\tau}\ . \tag{40}$$

It is, of course, trivial to invert $\tilde{O}$, since it is a product of triangular and diagonal matrices. The conditioning increases the computer time needed for one conjugate-gradient iteration by only a few percent.

We have studied the convergence of the conjugate-gradient method with and without conditioning using random values of the x variables with $\langle x_{i,l}\rangle=\frac{1}{2}\Delta\tau$, the value determined by $S_B(x)$ alone. We have found that results obtained with such configurations give a good estimate of the dependence of the convergence rate on the conditioning, and on the coupling, temperature, and spatial volume. However, configurations obtained in the actual simulation tend to have slower convergence rates than the random ones. Our stopping criterion is

$$\epsilon=[(r,r)/V]^{1/2}\leq 0.002\ , \tag{41}$$

where r is the residual vector defined in Eq. (27). This choice of ϵ is typical of the accuracy needed for the simulation.

In Fig. 1 we show the number of conjugate-gradient iterations needed for convergence, $N_{\rm CG}$, on a 4^3 lattice as a function of U for $\beta=2.0$. In all of our plots the triangles are for the unconditioned matrices and the squares for the conditioning algorithm just described. $N_{\rm CG}$ is sensitive to $\Delta\tau$. In general, we have found that if the convergence becomes unacceptably slow at large U or β, it can be improved by decreasing $\Delta\tau$ slightly. One, of course, pays the price of a larger lattice. In Fig. 1 the solid triangles and squares are for $\Delta\tau=0.05$ and the open squares are for $\Delta\tau=0.04$. In Fig. 2 we plot $N_{\rm CG}$ as a function of β for $U=8$ on a 4^3 lattice. The solid triangles and squares are for $\Delta\tau=0.04$, and the open square is for $\Delta\tau=0.03$.

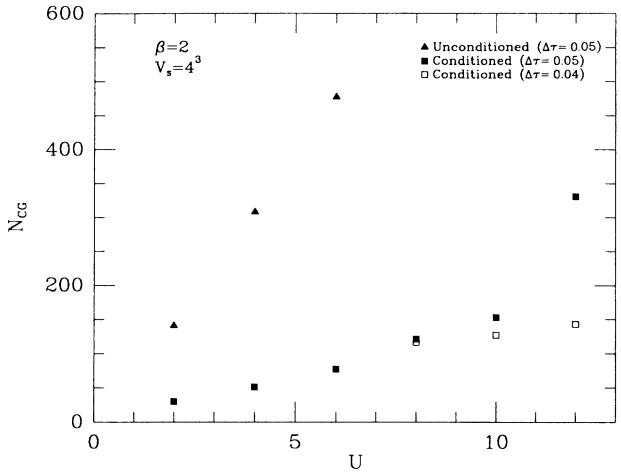

FIG. 1. The number of conjugate-gradient iterations, N_{CG}, needed for convergence as a function of U. Data are for a random configuration of the x field on a 4^3 lattice with $\beta=2.0$. The stopping criterion is given in Eq. (41). Triangles are for the unconditioned fermion matrix and squares for preconditioning with the matrix $\tilde{O}$ described in the text. Solid squares correspond to $\Delta\tau=0.05$ and open squares to $\Delta\tau=0.04$.

N_{CG} is quite insensitive to the spatial volume. In Fig. 3 we plot it as a function of V_s for $U=6$ and $\beta=2.0$ on 4^3, 6^3, 8^3, and 10^3 lattices. All data were taken with $\Delta\tau=0.05$. Finally, in Fig. 4 we plot values of N_{CG} taken in an actual simulation as a function of spatial volume for $U=6.0$ and $\beta=2.0$. The triangles are the average value of N_{CG} immediately after a Φ updating, and the squares the average value of N_{CG} after molecular dynamics steps.

We have also tried the incomplete Cholesky decomposition of O itself as the conditioning matrix. For some values of U and β it gives more rapid convergence than $\tilde{O}$; however, it doubles the computer time needed for

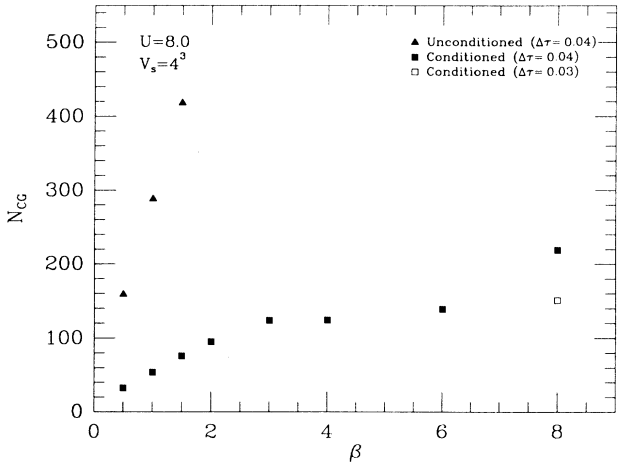

FIG. 2. N_{CG} as a function of β for $U=8.0$. Data are for a random configuration of the x field on a 4^3 lattice. The triangles are for the unconditioned fermion matrix and the squares for preconditioning with the matrix $\tilde{O}$. Solid squares correspond to $\Delta\tau=0.04$ and open squares to $\Delta\tau=0.03$.

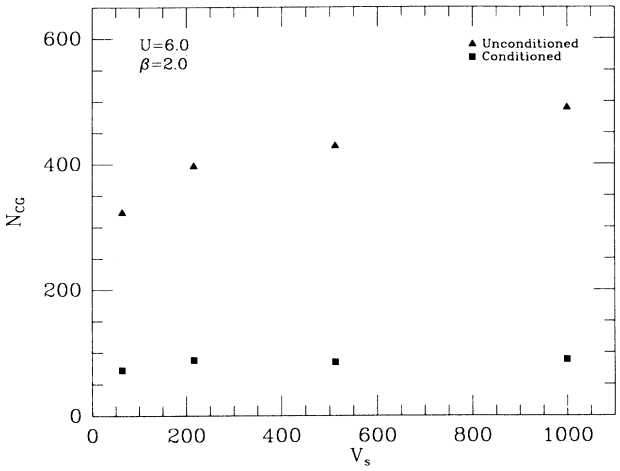

FIG. 3. N_{CG} as a function of spatial volume V_s for $U=6.0$ and $\beta=2.0$. Data are for a random configuration of the x field. The triangles are for the unconditioned fermion matrix and the squares for preconditioning with matrix $\tilde{O}$.

each iteration of the conjugate-gradient algorithm, and it is significantly more sensitive to the size of $\Delta\tau$. This is illustrated in Fig. 5, where we plot N_{CG} as a function of $\Delta\tau$ in a 4^3 lattice for $U=6.0$ and $\beta=2.0$. The solid squares are for the conditioning matrix $\tilde{O}$ and the open squares for the full incomplete Cholesky decomposition.

Finally, we have examined conditioning with a weak-coupling approximation to O; that is, we take the conditioning matrix to be O with U set equal to zero. This is very roughly analogous to the Fourier acceleration by Batrouni *et al.*[9] In Fig. 6 we plot N_{CG} as a function of U on a 4^3 lattice with $\beta=2$ for this weak-coupling conditioning (solid circles), for the unconditioned matrix (solid triangles), and for our standard conditioning matrix $\tilde{O}$ (solid squares). Unlike other conditioning methods we have studied, the weak-coupling one improves as $\Delta\tau$ is increased. The results presented in Fig. 6 are for $\Delta\tau=0.2$, which is somewhat large than the value one

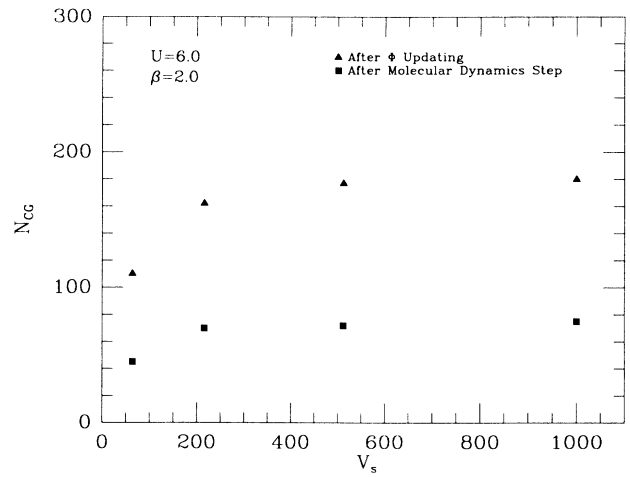

FIG. 4. N_{CG} from an actual simulation as a function of V_s for $U=6.0$ and $\beta=2.0$. The triangles are the average value of N_{CG} immediately after a Φ updating and the squares are the average value of N_{CG} after a microcanonical step.

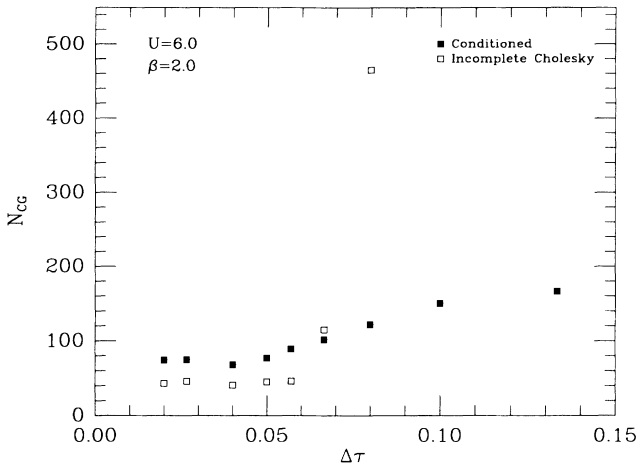

FIG. 5. N_{CG} as a function of $\Delta\tau$ for $U=6.0$ and $\beta=2.0$ on a 4^3 lattice. Data are for a random configuration of the x field. The solid squares are for preconditioning with the matrix $\tilde{O}$ and the open squares are for the preconditioning with the incomplete Cholesky decomposition of the Fermion matrix.

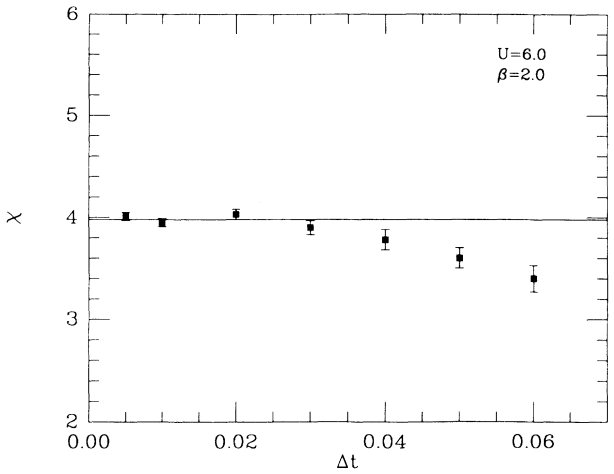

FIG. 7. The staggered susceptibility χ as of function of Δt on a 4^3 lattice for $U=6.0$ and $\beta=2.0$. The error bars indicate statistical errors only. The solid line is our best estimate of the exact answer, and is in agreement with the result of Hirsch. The integration of the molecular-dynamics equations becomes unstable for $\Delta t > 0.06$.

would use in an actual simulation because of the errors introduced in the breakup of the Hamiltonian, Eq. (3). For the unconditioned matrix and conditioning with $\tilde{O}$ we use $\Delta\tau=0.05$. Although the weak-coupling conditioning allows the use of somewhat smaller lattices than other approaches, it takes significantly more computer time per conjugate-gradient iteration because of the Fourier transforms it requires. Even ignoring this fact, we see from Fig. 6 that the weak-coupling conditioning is significantly worse than the conditioning we have adopted at all but the smallest coupling constants. In fact, it becomes worse than no conditioning at all for intermediate values of U. This is not a surprising result.

The Hubbard-Stratonovich field is not dynamical, and there are no gradient terms in the action to damp its high-momentum, large-frequency modes. It is these modes that appear to give rise to the small eigenvalues of O, and they are not well described by a weak-coupling approximation.

IV. NUMERICAL RESULTS

In this section we discuss initial numerical results obtained from applying the algorithm to the three-dimensional Hubbard model. We will present a detailed study of the phase structure of this model at a later time.

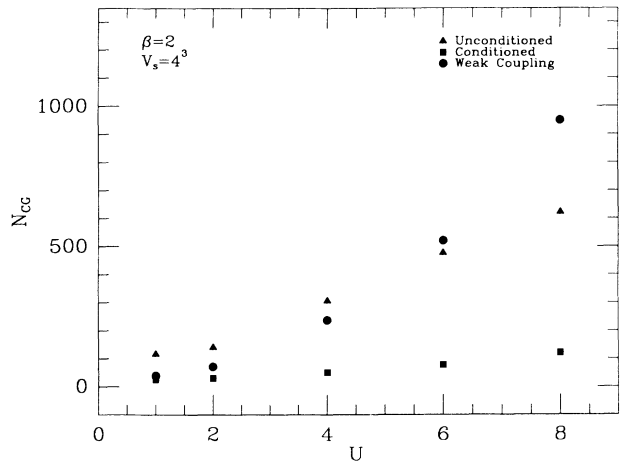

FIG. 6. N_{CG} as a function of U for $\beta=2.0$ on a 4^3 lattice. Data are for a random configuration of the x field. The triangles are for the unconditioned matrix, the squares for conditioning with the matrix $\tilde{O}$, and the circles for weak-coupling conditioning. For the latter we use $\Delta\tau=0.2$ and for the first two $\Delta\tau=0.05$.

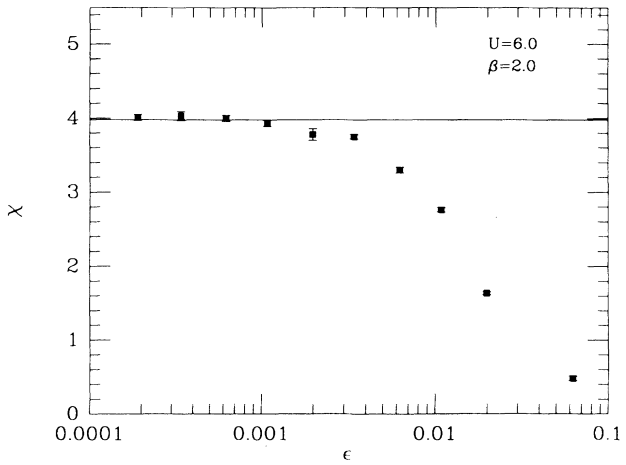

FIG. 8. The staggered susceptibility χ as a function of the conjugate-gradient stopping parameter ϵ on a 4^3 lattice with $U=6.0$ and $\beta=2.0$. The solid line is our best estimate of the exact answer.

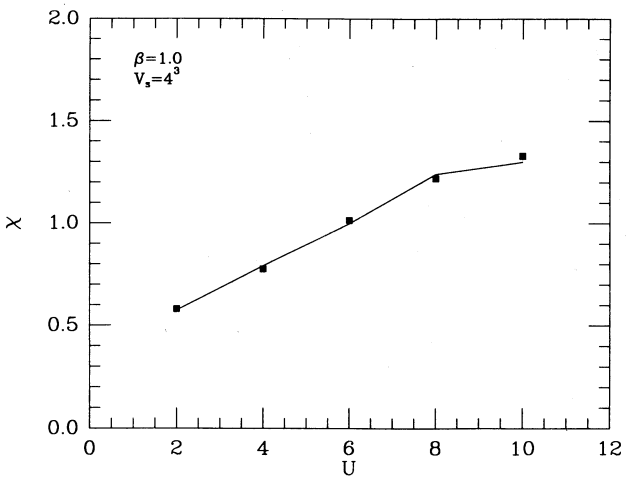

FIG. 9. The staggered susceptibility χ as a function of U on a 4^3 lattice for $\beta=1.0$. The solid squares are our results and the solid line is drawn through those of Hirsch. The error bars on both our data and Hirsch's are smaller than the plotting symbols.

A typical quantity of physical interest for the Hubbard model is the staggered susceptibility,

$$\chi=\frac{1}{V_s}\int_0^\beta d\tau\sum_{i,j}(-1)^{i+j}\langle[c_i^\dagger(\tau)\sigma_z c_i(\tau)]\times[c_j^\dagger(0)\sigma_z c_j(0)]\rangle$$
$$=\frac{1}{V_s}\int_0^\beta d\tau\sum_{i,j}(-1)^{i+j}\langle[n_{i+}(\tau)-n_{i-}(\tau)]\times[n_{j+}(0)-n_{j-}(0)]\rangle\ , \quad (42)$$

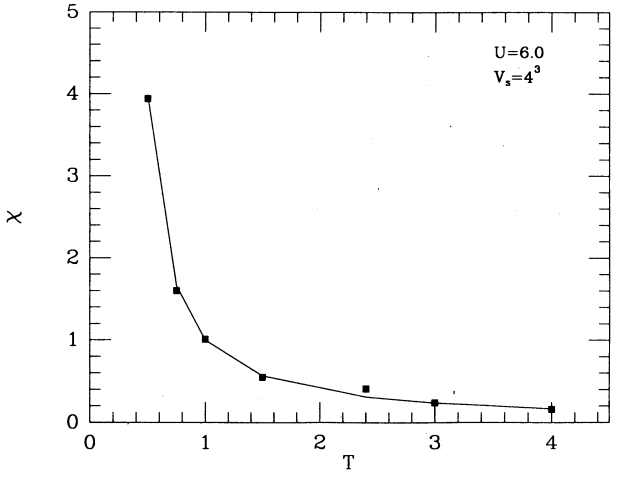

FIG. 10. The staggered susceptibility χ as a function of temperature on a 4^3 lattice for $U=6.0$. The solid squares are our results and the solid line is drawn through those of Hirsch. The error bars on both our data and Hirsch's are smaller than the plotting symbols.

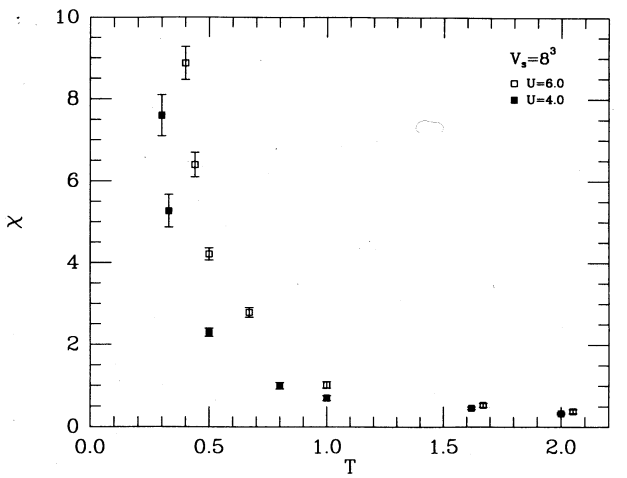

FIG. 11. The staggered susceptibility χ as a function of temperature on an 8^3 lattice. The open squares are for $U=6.0$ and the solid squares are for $U=4.0$.

In writing Eq. (42) we have used the usual z component of the magnetization. Because of rotational invariance one could, of course, have used the x or y components. It turns out that the transverse (x or y) components provide a significantly less noisy estimator for χ,[10] and the results shown below were obtained using

$$\chi=\frac{1}{V_s}\int_0^\beta\sum_{i,j}(-1)^{i+j}\langle[c_i^\dagger(\tau)\sigma_x c_i(\tau)][c_j^\dagger(0)\sigma_x c_j(0)]\rangle\ . \quad (43)$$

Before beginning the simulation, it is essential to determine the values of Δt, $\Delta\tau$, and ϵ necessary to obtain the desired accuracy in the simulation. In Fig. 7 we plot the

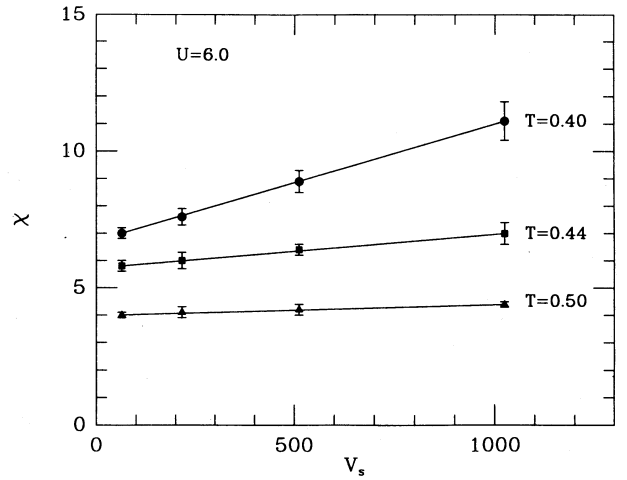

FIG. 12. The staggered susceptibility χ as a function of spatial volume for $U=6.0$. Data are given for three different values of the temperature and for spatial volumes 4^3, 6^3, 8^3, and 10^3. The linear increase in χ with spatial volume for $T=0.40$ and 0.44 indicates that these temperatures are below T_c, while the lack of variation of χ with volume at $T=0.50$ indicates that this temperature is above T_c.

staggered susceptibility as a function of Δt on a 4^3 lattice for $U=6.0$ and $\beta=2.0$. The error bars indicate statistical errors only. As Δt is increased the systematic error remains small almost to the point where the numerical integration becomes unstable.

We have found that it is sufficient to work with $\Delta\tau$ in the range 0.10–0.16 to reduce the errors arising from finite $\Delta\tau$ below our statistical ones. However, we have actually used significantly smaller values, 0.04–0.05, in order to speed the convergence of the conjugate-gradient calculation.

In Fig. 8 we plot the average value of χ as a function of the conjugate-gradient stopping parameter ϵ for a 4^3 lattice with $U=6.0$ and $\beta=2.0$. It should be noted that as the temperature is lowered or the coupling increased, the fermion determinant becomes a more rapidly varying function of time, and it is necessary to decrease both Δt and ϵ.

The algorithm can be tuned by varying the frequency of the two heat-bath–updating steps. This will not change the value of any measured quantity, but it will change the computer time necessary to generate statistically independent quantities. In our simulations we made one heat-bath updating of the p field for every $1/\Delta t$ microcanonical steps. The Φ fields were generally updated every fifth microcanonical step. The more frequent updating of the Φ fields did not significantly affect the autocorrelation time, and it improved the statistics in the measurement of quantities involving the electron Green's functions.

In Fig. 9 we plot the susceptibility χ as a function of U for $\beta=1.0$, and in Fig. 10 χ is shown as a function of β for $U=6.0$. The spatial lattice size is 4^3. The solid squares are data from our present algorithm, and the solid line is drawn through data obtained by Hirsch using an exact updating algorithm and a discrete Hubbard-Stratonovich transformation.[10] The error bars in both our data and Hirsch's are smaller than the plotting symbols. We find equally good agreement for other physical quantities.[10] The excellent agreement with Hirsch's results is our final evidence that we do have good control over the systematic errors in our algorithm.

The advantage of the present algorithm is that it enables one to work on large lattices and at low temperatures. This is illustrated in Fig. 11, where we plot the susceptibility as a function of temperature on a 8^3 lattice for $U=4.0$ and 6.0. The sharp rise in χ as the critical temperature is approached is clearly evident. In Fig. 12 we plot χ as a function of the spatial volume for $U=6.0$ and $\beta=2.0$, 2.25, and 2.50. Data are given for 4^3, 6^3, 8^3, and 10^3 lattices. The lack of variation of χ with volume for $\beta=2.0$ indicates that this temperature is above the critical temperature T_c, whereas the linear increase of χ with volume for $\beta=2.25$ and 2.50 indicates that these temperatures are below T_c. Our results are consistent with Hirsch's conclusion that a phase transition occurs in the neighborhood of $\beta=2.0$. However, we postpone a detailed study of the phase structure of the model to a later time.

V. DISCUSSION

The algorithm we have described can be used to make detailed studies of a wide variety of many electron systems. As long as the Hamiltonian is quadratic in the electron creation and annihilation operators, or can be made so by the introduction of auxiliary fields, the partition function can always be written in the form of Eq. (7). For models of electron-phonon interactions the Hubbard-Stratonovich field is replaced by the phonon field, and S_B contains terms which describe the kinetic energy and self-interaction of the phonons.

The hybrid molecular-dynamics approach offers a number of important advantages. The combination of heat-bath and molecular-dynamics updating steps leads to a relatively rapid evolution of the system through configuration space. The errors introduced by the finite-step size in the integration of the molecular-dynamics equations can be well controlled. In particular, it is possible to integrate these equations correctly to second order in the step size at the cost of only one conjugate-gradient calculation per step. In addition, a mechanism exists for choosing a good starting point for the conjugate-gradient calculation during the molecular-dynamics steps, which leads to a considerable savings in computer time.

The Langevin equation has been used extensively to study analogous problems in lattice-gauge theory. Where direct comparisons have been made,[2,11] it does not lead to as rapid evolution of the system through configuration space as the hybrid method. This is presumably a result of the fact that in contrast to the deterministic molecular-dynamics equations, the Langevin equation leads to a random walk through configuration space. In addition, integration of the Langevin equations to second order in the step size requires at least one extra conjugate-gradient calculation per step. Finally, the noise in the Langevin equation prevents a smooth extrapolation of the quantity $O^{-1}\Phi$ needed as the input into the conjugate-gradient calculation.

We are presently using the algorithm to make a detailed study of the phase structure of the three-dimensional Hubbard model. Work on a model of electron-phonon interaction and on the negative-U Hubbard model in two dimensions is also in progress.

ACKNOWLEDGMENTS

We would like to thank J. B. Bronzan and J. E. Hirsch for helpful conversations. This work was supported in part by the U.S. Department of Energy under Grant Nos. DE-FG03-85ER45197 and DE-AT03-81ER40029, and the National Science Foundation under Grant Nos. DMR83-20423 and PHY82-17853. We gratefully acknowledge Star Technologies, Inc., E. I. DuPont de Nemours and Company, Xerox Corporation, and the San Diego Supercomputer Center for their support.

[1]H. C. Anderson, J. Chem. Phys. **72**, 2384 (1980).
[2]S. Duane, Nucl. Phys. **B257**, 652 (1985); S. Duane and J. B. Kogut, Phys. Rev. Lett. **55**, 2774 (1985).
[3]S. Gottlieb, W. Liu, D. Toussaint, R. L. Renken, and R. L. Sugar, Phys. Rev. D **35**, 2531 (1987).
[4]E. V. E. Kovacs, D. K. Sinclair, and J. B. Kogut, Phys. Rev. Lett. **58**, 751 (1987); J. B. Kogut and D. K. Sinclair, Argonne National Laboratory–University of Illinois Report No. ANL-HEP-PR-86-81/ILL-TH-86-46, 1986 (unpublished); F. Karsch, J. B. Kogut, D. K. Sinclair, and H. W. Wyld, Phys. Lett. **188B**, 353 (1987).
[5]S. Gottlieb, W. Liu, D. Toussaint, R. L. Renken, and R. L. Sugar, Phys. Rev. D **35**, 2531 (1987).
[6]R. Blankenbecler, D. J. Scalapino, and R. L. Sugar, Phys. Rev. D **24**, 2278 (1981).
[7]J. E. Hirsch, Phys. Rev. B **31**, 4403 (1985).
[8]We thank J. B. Bronzan for suggesting this approach to us.
[9]G. G. Batrouni, G. R. Katz, A. S. Kronfeld, G. P. Lepage, B. Svetitsky, and K. G. Wilson, Phys. Rev. D **32**, 2736 (1985).
[10]J. E. Hirsch, Phys. Rev. B **35**, 1851 (1987). We would like to thank Professor Hirsch for making his data available to us for this comparison.
[11]E. Dagotto and J. B. Kogut, Phys. Rev. Lett. **58**, 299 (1987).