

Hybrid-molecular-dynamics algorithms for the numerical simulation of quantum chromodynamics

Steven Gottlieb

Department of Physics, Indiana University, Bloomington, Indiana 47401

W. Liu and D. Toussaint

Department of Physics, University of California, La Jolla, California 92093

R. L. Renken and R. L. Sugar

Department of Physics, University of California, Santa Barbara, California 93106

(Received 24 November 1986)

We discuss two algorithms for the numerical simulation of SU(3) lattice gauge theory with dynamical quarks. Both are based on the hybrid stochastic method of Duane and Kogut. They provide a relatively rapid evolution of the gauge fields through configuration space and good control of errors. One of the algorithms allows the simulation of arbitrary numbers of quarks. Tests of the algorithms are presented as well as initial data from a study of the thermodynamics of quarks and gluons with Kogut-Susskind fermions.

I. INTRODUCTION

Including the effects of dynamical fermions is presently the most important and challenging problem in the numerical study of lattice gauge theory. A number of approaches are being actively pursued.¹⁻⁴ In this paper we present first results from two algorithms which we think have great promise. Our approach is based upon the hybrid stochastic method suggested by Duane and Kogut.¹ This method provides a relatively rapid evolution of the gauge fields through configuration space and allows good control of the errors arising from finite step size. Our algorithms differ sufficiently from those of Duane and Kogut that we think it worthwhile to set them out in detail.

In Sec. II we discuss an algorithm in which the fermion determinant is eliminated by the introduction of a pseudofermion field, and in Sec. III one in which the fermions are represented by noise as suggested by Batrouni *et al.*³ We shall refer to these as the Φ and R algorithms, respectively. Both algorithms appear to provide a rapid evolution of the fields through configuration space. In both algorithms errors in measured quantities vanish as the square of the step size used in numerically integrating the molecular-dynamics equations, and in both one needs only a single conjugate-gradient calculation for each molecular-dynamics step. The Φ algorithm also contains a mechanism for minimizing the number of conjugate-gradient iterations needed in the molecular-dynamics steps. On the other hand, the Φ algorithm is applicable only to four equal-mass flavors of Kogut-Susskind fermions or two equal-mass flavors of Wilson fermions, while the R algorithm is applicable to any number (even noninteger) of either type of fermion. We compare results obtained from these algorithms with calculations of other authors, and give first results from a study of the thermodynamics of four staggered fermions on an $8^3 \times 4$ lattice.

Finally in Sec. IV we briefly discuss our results and prospects for future work.

II. THE Φ ALGORITHM

Our starting point is the path integral for the partition function

$$\begin{aligned} Z &= \int [\delta U] e^{-S_W(U)} \det M(U) \\ &= \int [\delta U \delta \Phi] e^{-[S_W(U) + \Phi^* (M^\dagger M)^{-1} \Phi]} \\ &= \int [\delta U \delta \Phi] e^{-S_{\text{eff}}}. \end{aligned} \quad (1)$$

Here $S_W(U)$ is the Wilson action for the pure gauge theory. We use staggered fermions, so the matrix $M(U)$ is given by

$$M(U)_{i,j} = 2m \delta_{i,j} + \sum_{\mu} \eta_{i,\mu} (U_{i,\mu} \delta_{i,j-\mu} - U_{i-\mu,\mu}^\dagger \delta_{i,j+\mu}). \quad (2)$$

Here i and j refer to lattice points and μ is a unit vector on the lattice. $U_{i,\mu}$ is the SU(3) matrix associated with the link leaving the i th lattice point in the μ direction. m is the quark mass and $\eta_{i,\mu}$ are the usual staggered fermion phases. The phases $\eta_{i,\mu}$ can be absorbed into the link matrices U at the cost of changing the sign of the gauge action (and remembering that some link matrices have determinant -1 when reunitarizing.) This will be assumed in the following. As is well known,⁵ $M^\dagger M$ has no matrix elements connecting even and odd lattice sites, and the determinants of its submatrices on the even and odd sites are equal. As a result, we can avoid a redoubling of flavors from using the square of the fermion matrix by defining the pseudofermion field Φ only on even lattice sites. As it is, the action of Eq. (1) represents four flavors

of quarks. Our approach can be used unchanged for Wilson fermions with an even number of flavors. In Sec. III we will discuss the modifications necessary to study arbitrary numbers of flavors.

We shall use molecular dynamics to determine the evolution of the gauge fields.⁶ We expect these deterministic equations to lead to a more rapid movement of the fields through configuration space than would be obtained with small-step-size algorithms based on random walks, because many successive steps move in the same direction through phase space. In order for the matrix $U_{j,\mu}$ to remain an element of SU(3), its equation of motion must have the form

$$\dot{U}_{j,\mu} = iH_{j,\mu}U_{j,\mu}, \quad (3)$$

where $H_{j,\mu}$ is a traceless Hermitian matrix. \dot{U} is the derivative of U with respect to the “molecular dynamics” time, the time through which the fields evolve according to the molecular dynamics equations. The matrices $H_{j,\mu}$ play a role analogous to the canonical momenta in an ordinary molecular dynamics simulation. We therefore introduce an auxiliary field into the partition function by writing

$$\begin{aligned} Z &= \int [\delta U \delta \Phi \delta H] \exp[-(\frac{1}{2} \text{tr} H^2 + S_{\text{eff}})] \\ &= \int [\delta U \delta \Phi \delta H] e^{-\mathcal{H}}, \end{aligned} \quad (4)$$

where

$$\text{tr} H^2 = \sum_{j,\mu} \text{tr} H_{j,\mu}^2. \quad (5)$$

Clearly the introduction of the H field will have no effect on correlation functions involving U and Φ .

We wish to generate a set of field configurations with a

probability distribution proportional to $\exp(-\mathcal{H})$. To this end we employ three types of updating steps. The H and Φ fields are updated using heat baths.^{1,3,7} We write the H field in the form

$$H_{j,\mu} = \sum_{\alpha} \lambda_{\alpha} h_{j,\mu}^{\alpha}, \quad (6)$$

where λ_{α} are the generators of SU(3), normalized so that $\text{tr}(\lambda_a \lambda_b) = 2\delta_{ab}$. We then simply equate each $h_{j,\mu}^{\alpha}$ to a complex Gaussian random number with $\langle |h_{j,\mu}^{\alpha}|^2 \rangle = 1$. To update the Φ field we generate a complex vector of Gaussian random numbers R with a probability distribution proportional to $\exp(-R^* \cdot R)$. We then form the vector

$$\Phi = M^{\dagger}(U) \cdot R. \quad (7)$$

The probability distribution for Φ is then proportional to $\exp[-\Phi^*(M^{\dagger}M)^{-1}\Phi]$. Since M^{\dagger} has matrix elements between even and odd sites as well as diagonal elements, it is necessary to define R on all lattices sites. However, the even and odd components of Φ each have the distribution required by Eq. (1), so we can simply ignore the odd components or, more accurately, average over them.

The final type of updating is a molecular dynamics step in which we vary H and U for fixed Φ in such a way as to keep \mathcal{H} and the differential volume element in configuration space fixed. As will be clear from what follows, it suffices to find an equation of motion with these properties. It is not necessary to explicitly sort out the canonical conjugacy relations among the momenta and the unconstrained degrees of freedom in the link matrices.

The equation of motion for the $U_{j,\mu}$ is given in Eq. (3). To obtain an equation of motion for H we require that \mathcal{H} be a constant of the motion. So

$$\begin{aligned} \dot{\mathcal{H}} &= \sum_{j,\mu} \text{tr} \left[\dot{H}_{j,\mu} H_{j,\mu} + \frac{\beta}{6} (\dot{U}_{j,\mu} V_{j,\mu} + V_{j,\mu}^{\dagger} \dot{U}_{j,\mu}^{\dagger}) \right] - \sum_{j,\mu} \Phi^* (M^{\dagger}M)^{-1} \left[M^{\dagger} \frac{\partial M}{\partial U_{j,\mu}} \dot{U}_{j,\mu} + \frac{\partial M^{\dagger}}{\partial U_{j,\mu}} \dot{U}_{j,\mu} M + \text{H.c.} \right] (M^{\dagger}M)^{-1} \Phi \\ &= 0, \end{aligned} \quad (8)$$

where H.c. arises from differentiation with respect to $U_{j,\mu}^{\dagger}$. $V_{j,\mu}$ is the sum of the staples, or products of the other three matrices in the plaquettes containing $U_{j,\mu}$. To handle the pseudofermion part of \mathcal{H} it is convenient to introduce the projection operator

$$P_{ij} = X_i X_j^*, \quad (9)$$

where

$$X = (M^{\dagger}M)^{-1} \Phi. \quad (10)$$

Since X and Φ are defined only on even sites, P_{ij} is taken to be zero unless i and j are both even sites. We use $\dot{U}_{j,\mu} = iH_{j,\mu}U_{j,\mu}$ to eliminate $\dot{U}_{j,\mu}$ from Eq. (8). P_{ij} is introduced into the pseudofermion term in \mathcal{H} to allow us to factor out $H_{j,\mu}$. We then find

$$\begin{aligned} \dot{\mathcal{H}} &= \sum_{j,\mu} \text{tr} \left[\dot{H}_{j,\mu} H_{j,\mu} + \frac{\beta}{6} (iH_{j,\mu} U_{j,\mu} V_{j,\mu} + \text{H.c.}) \right] - \sum_{j,\mu} \text{tr} \left[iH_{j,\mu} U_{j,\mu} P M^{\dagger} \frac{\partial M}{\partial U_{j,\mu}} + iH_{j,\mu} U_{j,\mu} M P \frac{\partial M^{\dagger}}{\partial U_{j,\mu}} + \text{H.c.} \right] \\ &= 0. \end{aligned} \quad (11)$$

This leads to an equation of the form

$$\dot{\mathcal{H}}=0=\sum_{j,\mu}\text{tr}\{H_{j,\mu}[\dot{H}_{j,\mu}+iF_{j,\mu}(U)]\}. \quad (12)$$

Since $H_{j,\mu}$ is traceless, Eq. (12) will be satisfied if the quantity in square brackets is a multiple of the unit matrix. The requirement that $H_{j,\mu}$ stay traceless, $\text{tr}(\dot{H}_{j,\mu})=0$, determines the remaining constant.⁸ Because $P_{i,j}$ only connects even sites, the expression for $\dot{H}_{j,\mu}$ takes a different form for even and odd sites:

$$i\dot{H}_{j,\mu}=\left[\frac{\beta}{3}U_{j,\mu}V_{j,\mu}-2U_{j,\mu}\left[\sum_{\nu}U_{j+\hat{\mu},\nu}P_{j+\hat{\mu}+\hat{\nu},j}-\sum_{\nu\neq\mu}U_{j+\hat{\mu}-\hat{\nu},\nu}P_{j+\hat{\mu}+\hat{\nu},j}\right]\right]_{\text{TA}} \quad (13)$$

for even sites and

$$i\dot{H}_{j,\mu}=\left[\frac{\beta}{3}U_{j,\mu}V_{j,\mu}-2U_{j,\mu}\left[\sum_{\nu}P_{j+\hat{\mu},j-\hat{\nu}}U_{j-\hat{\nu},\nu}-\sum_{\nu\neq\mu}P_{j+\hat{\mu},j+\hat{\nu}}U_{j,\nu}^{\dagger}\right]\right]_{\text{TA}} \quad (14)$$

for odd sites. Here the subscript TA indicates the traceless anti-Hermitian part of the matrix:

$$A_{\text{TA}}=\frac{1}{2}(A-A^{\dagger})-\frac{1}{6}\text{Tr}(A-A^{\dagger}). \quad (15)$$

We remind the reader that the signs associated with the links $\eta_{j,\mu}$ have been absorbed into the U 's.

The differential volume element in configuration space as well as \mathcal{H} remains constant under integration of Eqs. (3), (13), and (14). That is, the flow defined by these equations satisfies Liouville's theorem. This is obvious for δH from the form of Eqs. (13) and (14), and it follows for δU because Eq. (3) simply corresponds to a group rotation of $U_{j,\mu}$, and the Haar measure is invariant under such a rotation. This is necessary since it assures us that if the probability distribution is correct $\mathcal{P}(U,H,\Phi)=Z^{-1}e^{-\mathcal{H}}$, or if, in fact, the probability has any distribution depending only on \mathcal{H} , then the probability distribution will remain constant as we evolve through "molecular dynamics time."

It is easy to see that if the system is in the equilibrium configuration $\mathcal{P}_{\text{eq}}=Z^{-1}\exp(-\mathcal{H})$, it will stay in equilibrium. The heat-bath steps clearly maintain the equilibrium distribution. Since the molecular-dynamics steps move the system along paths of equal \mathcal{H} while maintaining the volume of phase-space elements, they also leave \mathcal{P}_{eq} unchanged. On the other hand, if the system is not in equilibrium, the heat-bath steps will drive it towards equilibrium as in a conventional Monte Carlo calculation. The molecular dynamics steps are neutral because they simply exchange configurations of equal \mathcal{H} . The formal argument of Duane and Kogut² proving convergence of

their form of the hybrid algorithm can easily be extended to ours.

The behavior of $H_{j,\mu}$ under a gauge transformation can be read off from Eq. (3). Under a gauge transformation G_j at site j , $U_{j,\mu}\rightarrow G_jU_{j,\mu}$. So, $H_{j,\mu}\rightarrow G_jH_{j,\mu}G_j^{-1}$. On the other hand, under a gauge transformation, $G_{j+\mu}$ at site $j+\mu$, $U_{j,\mu}\rightarrow U_{j,\mu}G_{j+\mu}^{-1}$, so $H_{j,\mu}$ remains unchanged. Thus the matrices $H_{j,\mu}$ are to be associated with sites rather than links. It is now clear that Eqs. (13) and (14) have the proper behavior under gauge transformations. [Of course, we could equally well have placed the H 's at the other end of the links by beginning with $\dot{U}=iUH$ instead of Eq. (3).]

In order to carry out the integration of Eqs. (3), (13), and (14) numerically we make the equations discrete by introducing a finite time step Δt . We then use the "leap frog" method to calculate the change in the U and H fields in each time interval. This method depends on the trivial observations that

$$H_{j,\mu}(t+\frac{1}{2}\Delta t)-H_{j,\mu}(t-\frac{1}{2}\Delta t)=\int_{t-\Delta t/2}^{t+\Delta t/2}dt'H_{j,\mu}(t')=\dot{H}_{j,\mu}(t)\Delta t+O(\Delta t^3) \quad (16)$$

and

$$U_{j,\mu}(t+\Delta t)=T\exp\left[i\int_t^{t+\Delta t}dt'H(t')\right]U_{j,\mu}(t)=e^{iH(t+\Delta t/2)\Delta t}U_{j,\mu}(t)+O(\Delta t^3). \quad (17)$$

These equations suggest that we define the U and Φ fields at the times $t, t+\Delta t, t+2\Delta t, \dots$, and the H field at the times $t+\frac{1}{2}\Delta t, t+\frac{3}{2}\Delta t, \dots$. Given U and Φ at time t and H at time $t+\frac{1}{2}\Delta t$, we use Eq. (17) to obtain U at time $t+\Delta t$, then Eq. (16) to obtain H at time $t+\frac{3}{2}\Delta t$, and so forth (Φ remains fixed during the molecular dynamics steps). Note that in each of these calculations the error is of order Δt^3 .

In calculating the change in H we need $(M^{\dagger}M)^{-1}\Phi$, which we obtain using the conjugate-gradient method. Even though the integration is second order in Δt , we need only perform one conjugate-gradient calculation per time interval. The number of conjugate-gradient sweeps to produce a given accuracy can be significantly reduced by taking advantage of the fact that the quantity $X(t)=(M^{\dagger}M)^{-1}\Phi$ is a relatively smooth function of t during the molecular-dynamics updating. We can therefore obtain a rather good approximation to $X(t+\Delta t)$ through the extrapolation formula

$$X(t+\Delta t)=2X(t)-X(t-\Delta t)+O(\Delta t^2). \quad (18)$$

For $\Delta t=0.01$ using Eq. (18) to determine the starting point for the conjugate-gradient calculation at $t+\Delta t$ reduces the number of conjugate-gradient sweeps as much as a factor of 5 from the number needed when starting with $X=0$ as is done after the heat-bath steps. The savings will of course depend on the couplings, the step size, and the desired accuracy of the conjugate-gradient calculation. In general we expect the interpolation to help most

with small-frequency modes of $X(t)$, which are just the ones for which the conjugate gradient has the most difficulty.

In this method the heat-bath updatings of the H and Φ fields are performed simultaneously. After such an updating we know the value of all three fields at the same time t . Before beginning the molecular-dynamics steps we most propagate the H field forward one-half of a time interval. We set

$$H_{j,\mu}(t + \frac{1}{2}\Delta t) = H_{j,\mu}(t) + \frac{1}{2}\Delta t \dot{H}_{j,\mu}(t). \quad (19)$$

The error in this step is of order Δt^2 . However, since we perform of order $1/\Delta t$ molecular-dynamics steps for each heat bath in order to obtain approximately equal motion of the fields through configuration space, the errors from the two types of steps are comparable. Indeed, errors in

the measured quantities are of order Δt^2 , as we shall see in the next section.

Alternatively, we could have declared the H configuration obtained from the heat-bath updating to be H at time $t + \Delta t/2$. Then in calculating $U(t + \Delta t)$ it would have been necessary to integrate for half a time step using the old H and for half a time step using the new H .

There is one technical point that is worth mentioning. In updating the U 's we expand $\exp(iH\Delta t)$ in a power series in Δt . Given the other errors in the calculation, one might think that it would be sufficient to retain only the first three terms in this series. However, if one does, then it is necessary to make inordinately frequent reunitarizations of the U 's. This problem is solved by taking two extra terms in the series expansion of the exponential, which costs only two more matrix multiplications per link update. We use

$$e^{i\Delta t H_{j,\mu}} U_{j,\mu} \simeq U_{j,\mu} + i\Delta t H_{j,\mu} \{ U_{j,\mu} + i\frac{1}{2}\Delta t H_{j,\mu} [U_{j,\mu} + i\frac{1}{3}\Delta t H_{j,\mu} (U_{j,\mu} + i\frac{1}{4}\Delta t H_{j,\mu} U_{j,\mu})] \} \quad (20)$$

to update $U_{j,\mu}$.

We now turn to numerical results obtained with this algorithm. For the gauge fields we use periodic boundary conditions in all four directions, while for the fermions we use periodic boundary conditions in space and antiperiodic boundary conditions in imaginary time. In Fig. 1 we plot the average value of the plaquette versus the step size Δt for the pure gauge theory on a 4^4 lattice with $\beta = 4.80$. The open circles are results from the present algorithm without the fermion term while the filled circle at $\Delta t = 0$ was obtained from a standard Metropolis Monte Carlo calculation. The quadratic variation of the error with step size is evident.

In Figs. 2(a) and 2(b) we plot the average plaquette and $\langle \bar{\psi}\psi \rangle$ versus step size for four flavors of staggered fer-

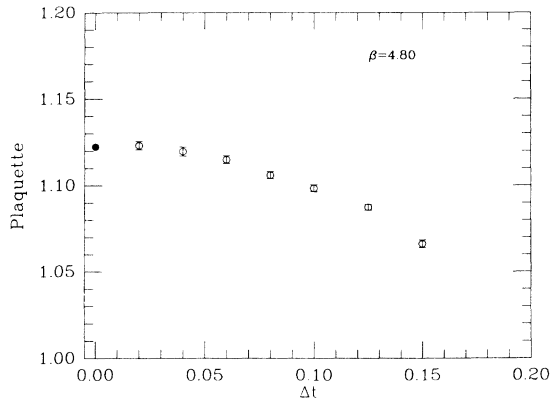


FIG. 1. The average value of the plaquette vs step size Δt for pure SU(3) gauge theory on a 4^4 lattice at $\beta = 4.80$. The open circles are results from the hybrid algorithm, while the filled circle at $\Delta t = 0$ was obtained from a standard Metropolis Monte Carlo calculation.

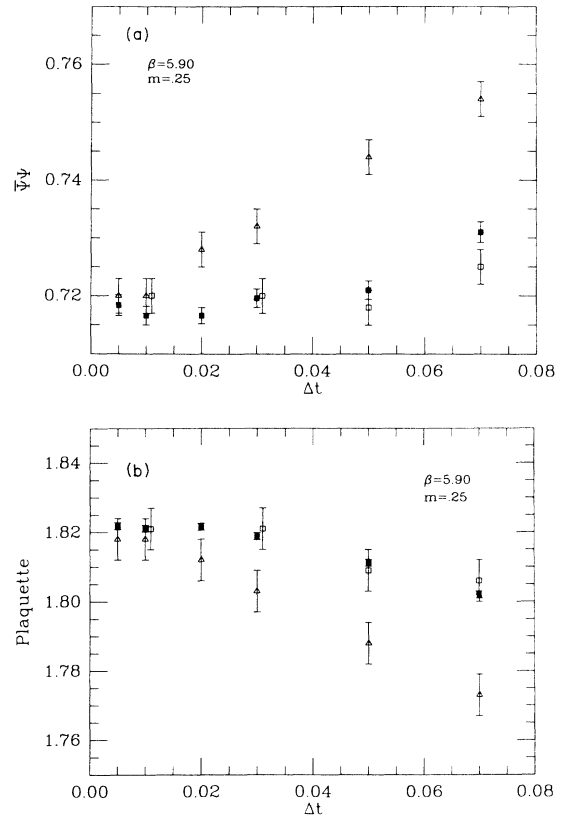


FIG. 2. (a) $\langle \bar{\psi}\psi \rangle$ and (b) the average plaquette for four flavors of staggered fermions on a 4^4 lattice with $\beta = 5.90$ and $m = 0.25$. The filled squares were obtained with the Φ algorithm. The open triangles and squares are results of Duane and Kogut obtained from algorithms in which molecular-dynamics equations are integrated to first and second order in Δt , respectively.

mions on a 4^4 lattice with $\beta=5.90$ and $m=0.25$. The filled squares are our results and the open triangles and squares are those of Duane and Kogut¹ obtained from algorithms in which molecular dynamics equations are integrated to first and second order in Δt , respectively. We find equally good agreement with results reported by Kogut for $\beta=5.90$ and $m=0.10$ on a $8^3 \times 4$ lattice.¹

We have also collected data on a 4^4 lattice at $\beta=4.80$ and $m=0.10$ in order to compare with the results of Gocksch, Gavai, and Heller⁹ (GGH), who employ an exact algorithm. Our results for this run were in agreement with those obtained with our own exact algorithm,¹⁰ but were in disagreement with GGH. For example, we obtained 1.214 ± 0.002 for the average plaquette while GGH report 1.248 ± 0.009 , while for $\langle \bar{\psi}\psi \rangle$ we find 0.623 ± 0.002 vs 0.606 ± 0.007 . We find similar disagreement at $m=0.05$. The discrepancy may be due to the fact that we use significantly more conjugate gradient sweeps in both this algorithm and our exact one than are used by GGH. Finally we note the Gavai, Potvin, and Sanielevici¹¹ have recently reported a calculation at $\beta=4.80$ and $m=0.10$ using the Langevin equation. If one ignores the point at $\epsilon=0.001$, which appears to include only a small number of statistically independent configurations, their result for the plaquette appears to extrapolate to ours.

This algorithm can be tuned by varying the number of molecular dynamics steps N_{md} between heat-bath updatings of the H and Φ fields. In Fig. 3 we plot the correlation time versus N_{md} for $\beta=5.90$, $m=0.10$, and $\Delta t=0.01$ on a 4^4 lattice. We obtain similar results at $m=0.05$ and $m=0.25$.

We measure the autocorrelation time in the following way. If we make measurements at physical time separations δt , then we can obtain the autocorrelation function $C(t)$ at times $0, \delta t, 2\delta t, 3\delta t, \dots$. We then define an autocorrelation time T_j for the j th time interval through

$$C((j+1)\delta t)/C(j\delta t) = \exp(-\delta t/T_j).$$

We find that except for T_0 , the T_j are approximately constant until j is large enough so that $C(j\delta t)$ has fallen off

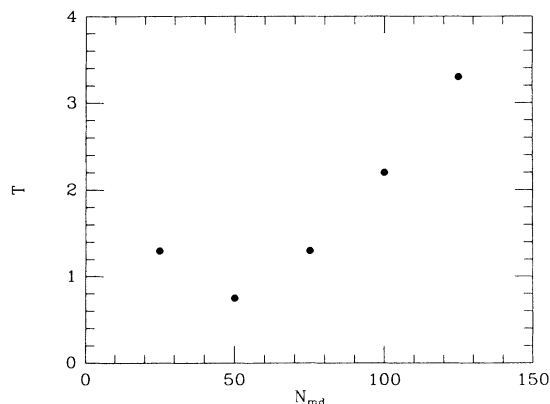


FIG. 3. The autocorrelation time T as a function of the number of molecular-dynamics steps between heat-bath updatings N_{md} . Data are for the Φ algorithm on a 4^4 lattice with $\beta=5.90$ and $m=0.10$.

by approximately 1 order of magnitude. At that point we no longer have enough data to make an accurate measurement of $C(t)$. We take as our value of the autocorrelation time the average of the T_j over the range where they are approximately constant. We would get very similar results by defining it to be the physical time for $C(t)$ to fall off by $1/e$. Of course by this procedure we are only obtaining a lower bound on the autocorrelation time.

As we mentioned before, one of the advantages of working with deterministic equations is that we can smoothly interpolate the vector $X(t) = (M^\dagger M)^{-1} \Phi$, thus obtaining a good starting point for the conjugate-gradient calculations. Clearly the interpolation becomes more accurate as Δt decreases. Since the computation of X dominates the computer time for small masses, a figure of merit is the number of conjugate sweeps N_{cg} needed to move one unit of physical time during the molecular-dynamics updating. We plot N_{cg} vs Δt in Fig. 4. It is evident that the penalty in computer time for going to smaller values of Δt is not as great as one might naively expect.

We shall present details of a study of the thermodynamics of four flavors of Kogut-Susskind fermions on a $8^3 \times 4$ lattice elsewhere. To indicate the type of results we are able to obtain with this algorithm, we show in Fig. 5 a plot of the average value of the real part of the Polyakov loop versus β for $m=0.10$. Figure 6 shows the time history of the real part of the Polyakov loop at $\beta=5.13$. The two different curves result from performing 50 warmup sweeps with β above and below 5.13. The last 150 trajectories from these two runs were used to obtain the two points at $\beta=5.13$ in Fig. 5. These two figures provide clear evidence for a first-order phase transition.¹²

III. THE R ALGORITHM

The major drawback of the algorithm discussed in the last section is that it is applicable only to sets of four equal-mass Kogut-Susskind fermions or two equal-mass Wilson fermions. In this section we present an algorithm

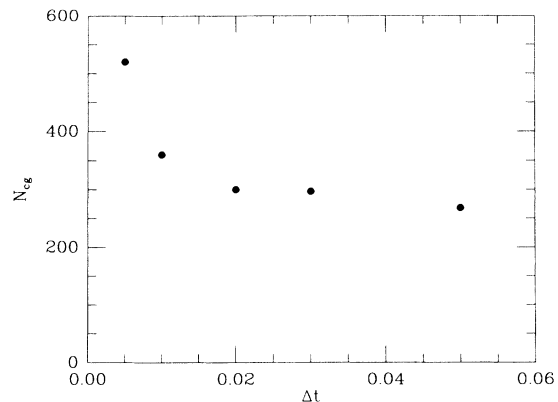


FIG. 4. The number of conjugate gradient sweeps N_{cg} needed to move one unit of physical time during the molecular-dynamics steps as a function of Δt . Data are for the same lattice and parameters as in Fig. 3.

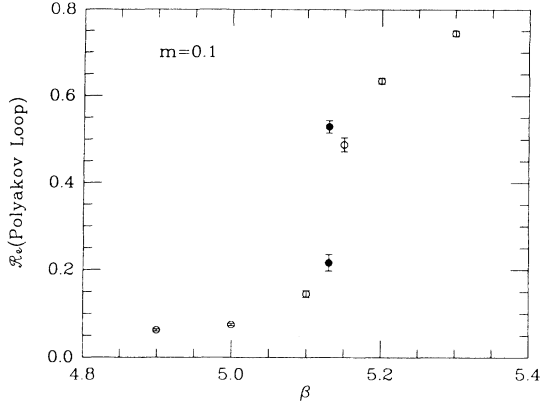


FIG. 5. Average value of the real part of the Polyakov loop as a function of β . Data were taken with the Φ algorithm on an $8^3 \times 4$ lattice with $m=0.10$. The two filled circles at $\beta=5.13$ are averages taken in the two coexisting phases.

which is applicable to any number of fermions of either variety.

In order to simplify the presentation we begin by discussing the interaction of a real scalar field q with N_f flavors of staggered fermions. We denote by q_i the component of q on the i th lattice site and by p_i the momentum conjugate to q_i . The partition function can be written in the form

$$Z = \int [\delta p \delta q] e^{-[p^2/2 + S_0(q)]} \det(M^\dagger M)^{N_f/4} \\ = \int [\delta p \delta q] \exp\left\{-\left[\frac{1}{2}p^2 + S_0(q) - \frac{1}{4}N_f \text{tr} \ln M^\dagger(q)M(q)\right]\right\}. \quad (21)$$

Here $p^2 = \sum_i p_i^2$, and $M^\dagger M$ is defined only on even lattice sites. The factor of $N_f/4$ occurs in the fermion term because $\det(M^\dagger M)$ represents four flavors of fermions. In discussing the molecular dynamics equations it will be convenient to write the effective Hamiltonian in the form

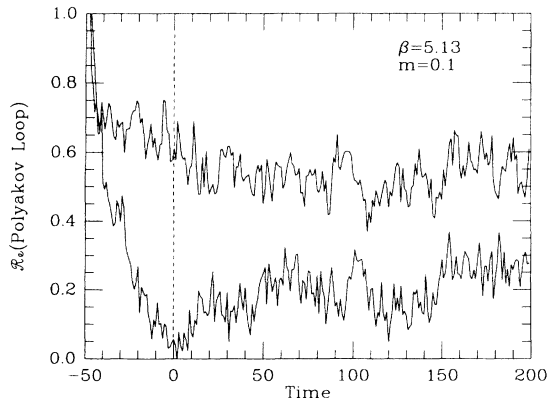


FIG. 6. The time history of the real part of the Polyakov loop with hot and cold starts at $\beta=5.13$.

$$\mathcal{H} = \frac{1}{2}p^2 + S_0(q) - \frac{N_f}{4} \text{tr} \ln(M^\dagger M) \\ = \frac{1}{2}p^2 + V(q). \quad (22)$$

We have in mind alternately performing heat-bath updates of the p field and molecular dynamics updates of the p and q fields which keep \mathcal{H} fixed. The molecular dynamics equations arising from the Hamiltonian of Eq. (22) are

$$\dot{q}_i = p_i, \quad (23)$$

$$\dot{p}_i = -\frac{\partial}{\partial q_i} V(q) \\ = -\frac{\partial}{\partial q_i} S_0(q) + \frac{N_f}{4} \text{tr} \left[\frac{1}{M^\dagger(q)M(q)} \frac{\partial}{\partial q_i} [M^\dagger(q)M(q)] \right].$$

It is clearly impractical to integrate these equations numerically since at each time step we would have to calculate all elements of the inverse of the matrix $M^\dagger M$ rather than just the inverse applied to a single vector. To avoid this difficulty we replace $(M^\dagger M)^{-1}$ in the last line of Eq. (23) by a noisy estimator.^{3,2} We begin by considering the straightforward replacement of $(M^\dagger M)^{-1}$ by the projection operator

$$P_{i,j}^{R_0} = X_i^{R_0} X_j^{R_0*} \quad (24)$$

with

$$X^{R_0} = \frac{1}{M^\dagger(q)M(q)} M^\dagger(q)R = \frac{1}{M(q)} R. \quad (25)$$

As always R is a vector of Gaussian random numbers. With this choice the estimator for \dot{p}_i is

$$\dot{p}_i^{R_0} = -\frac{\partial}{\partial q_i} S_0(q) + \frac{N_f}{4} X_i^{R_0*} \frac{\partial}{\partial q_i} [M^\dagger(q)M(q)] X_i^{R_0} \\ = -\frac{\partial}{\partial q_i} V^{R_0}(q). \quad (26)$$

It should be noted that the expression for $\dot{p}_i^{R_0}$ is very similar to that which we would have obtained with the Φ algorithm. The only substantive difference is that instead of holding $\Phi = M^\dagger(q)R$ fixed during the molecular-dynamics steps, we update it at each step, and we evaluate it using the instantaneous value of $M^\dagger(q)R$. Although it is still possible to integrate the molecular dynamics equations with an error of order Δt^3 at each time step, an error of order Δt^2 is introduced by keeping the noise vector R fixed throughout the time step. We shall see that this error can be eliminated by a trivial modification of X^{R_0} .

Let us begin by considering the integration of the molecular dynamics equations. Suppose we know $q(t)$ and $p(t)$. We wish to calculate these fields at time $t + \Delta t$ for a given value of R with an error of order Δt^3 . Replacing Eq. (23) by a standard difference approximation

$$q_i(t + \Delta t) = q_i(t) + \Delta t p_i(t + \frac{1}{2}\Delta t) + O(\Delta t^3), \quad (27) \\ p_i(t + \Delta t) = p_i(t) - \Delta t \frac{\partial}{\partial q_i} V^{R_0}(t + \frac{1}{2}\Delta t) + O(\Delta t^3).$$

The equation for $p(t + \Delta t)$ requires $q(t + \frac{1}{2}\Delta t)$ with an error of order Δt^2 . A simple forward difference approximation to Eq. (23) gives

$$q_i(t + \frac{1}{2}\Delta t) = q_i(t) + \frac{1}{2}\Delta t p_i(t) + O(\Delta t^2). \quad (28)$$

We can calculate $p(t + \Delta t)$ from Eq. (27) after making the only conjugate gradient calculation of the time step to obtain

$$X^{R_0}(t + \frac{1}{2}\Delta t) = \frac{1}{M^\dagger(t + \frac{1}{2}\Delta t)M(t + \frac{1}{2}\Delta t)} M^\dagger(t + \frac{1}{2}\Delta t)R. \quad (29)$$

In order to obtain $q(t + \Delta t)$ we approximate $p(t + \frac{1}{2}\Delta t)$ by the average of $p(t)$ and $p(t + \Delta t)$, the latter of which we have just calculated.

$$\begin{aligned} q_i(t + \Delta t) &= q_i(t) + \frac{1}{2}\Delta t [p_i(t) + p_i(t + \Delta t)] + O(\Delta t^3) \\ &= q_i(t + \frac{1}{2}\Delta t) + \frac{1}{2}\Delta t p_i(t + \Delta t) + O(\Delta t^3). \end{aligned} \quad (30)$$

Notice that in the next time step we will use $p(t + \Delta t)$ to calculate $q(t + \frac{3}{2}\Delta t)$ from $q(t + \Delta t)$. So, we use $p(t + \Delta t)$ for the entire time interval $t + \frac{1}{2}\Delta t$ to $t + \frac{3}{2}\Delta t$. Thus except for the edge effects in the intervals immediately before and after the heat-bath updating of the p field, we are back to the leap-frog method.

We are now in a position to discuss the errors introduced by holding the noise vector R constant during a molecular-dynamics step. We denote the probability distribution of the q and p fields at time t by $\mathcal{P}(q, p, t)$. Following Batrouni *et al.*³ we write the Fokker-Planck equation for the change in \mathcal{P} during a single time step in the form

$$\begin{aligned} \mathcal{P}(q, p, t + \Delta t) &= \int [\delta q' \delta p'] \mathcal{P}(q', p', t) \\ &\quad \times \prod_i \delta(q_i - q'_i - \Delta q_i) \delta(p_i - p'_i - \Delta p_i). \end{aligned} \quad (31)$$

From Eqs. (27) and (30) we see that to order Δt^2

$$\begin{aligned} \Delta q_i &= \Delta t p'_i + \frac{1}{2}\Delta t \Delta p_i \\ &= \Delta t p'_i - \frac{1}{2}\Delta t^2 \frac{\partial}{\partial q'_i} V^{R_0}(q'), \end{aligned} \quad (32)$$

and

$$\Delta p_i = -\Delta t \frac{\partial}{\partial q'_i} V^{R_0}(q'_i) - \frac{1}{2}\Delta t^2 \sum_j p_j \frac{\partial^2}{\partial q'_i \partial q'_j} V^{R_0}(q'). \quad (33)$$

Expanding Eq. (31) as a power series in Δq_i and Δp_i and keeping terms through second order gives

$$\begin{aligned} \Delta \mathcal{P} &= \mathcal{P}(q, p, t + \Delta t) - \mathcal{P}(q, p, t) \\ &= \sum_i \left[-\frac{\partial}{\partial q_i} (\mathcal{P} \Delta q_i) - \frac{\partial}{\partial p_i} (\mathcal{P} \Delta p_i) \right] \\ &\quad + \sum_{i,j} \left[\frac{1}{2} \frac{\partial^2}{\partial q_i \partial q_j} (\mathcal{P} \Delta q_i \Delta q_j) + \frac{1}{2} \frac{\partial^2}{\partial p_i \partial p_j} (\mathcal{P} \Delta p_i \Delta p_j) \right] + \sum_{i,j} \frac{\partial^2}{\partial q_i \partial p_j} (\mathcal{P} \Delta q_i \Delta p_j) + O(\Delta t^3). \end{aligned} \quad (34)$$

Suppose that at time t the system is in equilibrium. That is, $\mathcal{P}(q, p, t)$ is proportional to $\exp(-\mathcal{H})$. Then making use of Eqs. (32) and (33) we find that

$$\begin{aligned} \Delta \mathcal{P} &= \mathcal{P} \left\{ \sum_i \left[\frac{\partial V}{\partial q_i} - \frac{\partial V^{R_0}}{\partial q_i} \right] \left[\Delta t p_i + \frac{1}{2}\Delta t^2 \frac{\partial V^{R_0}}{\partial q_i} \right] \right. \\ &\quad \left. + \frac{1}{2}\Delta t^2 \sum_{i,j} p_i p_j \left[\frac{\partial^2 V^{R_0}}{\partial q_i \partial q_j} - \frac{\partial^2 V}{\partial q_i \partial q_j} + \left[\frac{\partial V}{\partial q_i} - \frac{\partial V^{R_0}}{\partial q_i} \right] \left[\frac{\partial V}{\partial q_j} - \frac{\partial V^{R_0}}{\partial q_j} \right] \right] \right\}. \end{aligned} \quad (35)$$

We now average over the noise. Denoting the average over R by $\langle \rangle_R$ and using the facts that

$$\left\langle \frac{\partial V^{R_0}}{\partial q_i} \right\rangle_R = \frac{\partial V}{\partial q_i} \quad \text{and} \quad \left\langle \frac{\partial^2 V^{R_0}}{\partial q_i \partial q_j} \right\rangle_R = \frac{\partial^2 V}{\partial q_i \partial q_j},$$

we see that

$$\langle \Delta \mathcal{P} \rangle_R = \frac{1}{2}\Delta t^2 \mathcal{P} \sum_{i,j} (p_i p_j - \delta_{i,j}) \left\langle \frac{\partial V^{R_0}}{\partial q_i} \frac{\partial V^{R_0}}{\partial q_j} - \frac{\partial V}{\partial q_i} \frac{\partial V}{\partial q_j} \right\rangle_R. \quad (36)$$

The right-hand side of Eq. (36) fails to vanish because of the terms which are quartic in the components of R . We see from Eqs. (25) and (26) that

$$\langle \Delta \mathcal{P} \rangle_R = \frac{1}{2} \Delta t^2 \mathcal{P} \left[\frac{N_f}{4} \right]^2 \sum_{i,j} (p_i p_j - \delta_{i,j}) \text{tr} \left[\frac{1}{M^\dagger M} \frac{\partial}{\partial q_i} (M^\dagger M) \frac{1}{M^\dagger M} \frac{\partial}{\partial q_j} (M^\dagger M) \right]. \quad (37)$$

We can make $\langle \Delta \mathcal{P} \rangle_R$ vanish through order Δt^2 by means of a trivial change in our expression for Δp_i . To see this consider the Φ algorithm for $N_f=4$. For this algorithm $\Delta \mathcal{P}$ vanishes to order Δt^2 even if we update Φ at the beginning of every time step. The only change is that $X^{R_0}(t + \frac{1}{2} \Delta t)$, defined in Eq. (28), is replaced by

$$\begin{aligned} X^\Phi(t + \frac{1}{2} \Delta t) &= \frac{1}{M^\dagger(t + \frac{1}{2} \Delta t) M(t + \frac{1}{2} \Delta t)} M^\dagger(t) R \\ &= X^{R_0}(t + \frac{1}{2} \Delta t) - \frac{1}{M^\dagger(t + \frac{1}{2} \Delta t) M(t + \frac{1}{2} \Delta t)} \sum_i \Delta q_i \frac{\partial M^\dagger}{\partial q_i} R + O(\Delta t^2), \end{aligned} \quad (38)$$

where $\Delta q_i = \frac{1}{2} \Delta t p_i(t)$ and we have used the fact that M is a function of q only. Denoting the right-hand side of Eq. (33) by $\Delta p_i^{R_0}$, we see that this quantity is replaced in the Φ algorithm by

$$\begin{aligned} \Delta p_i^\Phi &= \Delta p_i^{R_0} - \frac{1}{2} \Delta t^2 \sum_j p_j \left[R^* M(q') \frac{1}{M^\dagger(q') M(q')} \frac{\partial}{\partial q'_i} [M^\dagger(q') M(q')] \frac{1}{M^\dagger(q') M(q')} \frac{\partial M^\dagger(q')}{\partial q'_j} R + \text{c.c.} \right] \\ &= \Delta p_i^{R_0} - \sum_j p_j f_{i,j}(q'). \end{aligned} \quad (39)$$

The only term in Eq. (34) that is changed through order Δt^2 is

$$- \sum_i \frac{\partial}{\partial p_i} (\mathcal{P} \Delta p_i^\Phi) = - \sum_i \frac{\partial}{\partial p_i} (\mathcal{P} \Delta p_i^{R_0}) - \sum_{i,j} \mathcal{P} (p_i p_j - \delta_{i,j}) f_{i,j}(q). \quad (40)$$

After the average over R is taken, the second term on the right-hand side of Eq. (40) cancels the right-hand side of Eq. (37), and we have $\langle \Delta \mathcal{P} \rangle_R = 0$ through order Δt^2 .

With this hint from the Φ algorithm it is clear how to proceed for arbitrary values of N_f . In the expression for $\dot{p}_i(t + \frac{1}{2} \Delta t)$ we simply replace $(M^\dagger M)^{-1}$ by the projection operator

$$P_{i,j}^R(t + \frac{1}{2} \Delta t) = X_i^R(t + \frac{1}{2} \Delta t) X_j^{R*}(t + \frac{1}{2} \Delta t) \quad (41)$$

with

$$X^R(t + \frac{1}{2} \Delta t) = \frac{1}{M^\dagger(t + \frac{1}{2} \Delta t) M(t + \frac{1}{2} \Delta t)} M^\dagger(t + (\frac{1}{2} - \frac{1}{8} N_f) \Delta t) R, \quad (42)$$

which gives

$$\begin{aligned} \dot{p}_i^R(t + \frac{1}{2} \Delta t) &= - \frac{\partial}{\partial q_i} S_0(t + \frac{1}{2} \Delta t) + \frac{N_f}{4} X^{R*}(t + \frac{1}{2} \Delta t) \frac{\partial}{\partial q_i} [M^\dagger(t + \frac{1}{2} \Delta t) M(t + \frac{1}{2} \Delta t)] X^R(t + \frac{1}{2} \Delta t) \\ &= - \frac{\partial}{\partial q_i} V^R(t + \frac{1}{2} \Delta t). \end{aligned} \quad (43)$$

With this choice for the estimator, $\langle \Delta \mathcal{P} \rangle_R = O(\Delta t^3)$ for all N_f (Ref. 13). The procedure followed here is not quite the same as that used by the Batrouni *et al.*³ or by Duane and Kogut,² who solve for the actual probability distribution. However, it suffices for our purpose which has been to obtain a second-order algorithm.

The same analysis goes through for gauge theories. We give the details in the Appendix. Here we will simply state the algorithm with the final formulas. The effective Hamiltonian is

$$\mathcal{H} = \frac{1}{2} \text{tr} H^2 + S_W - \frac{1}{4} N_f \text{tr} \ln(M^\dagger M). \quad (44)$$

We update the H field using the heat-bath method described in Sec. II. For each heat-bath step there are approximately $1/\Delta t$ molecular-dynamics steps. For the latter we use the leap-frog method except for the half-time steps immediately before and after the heat-bath step. To be specific $H_{j,\mu}$ is updated through the formula

$$H_{j,\mu}(t + \Delta t) = H_{j,\mu}(t) + \dot{H}_{j,\mu}^R(t + \frac{1}{2} \Delta t) \Delta t \quad (45)$$

with

$$i\dot{H}_{j,\mu}^R = \left[\frac{\beta}{3} U_{j,\mu} V_{j,\mu} - \frac{1}{2} N_f U_{j,\mu} \left[\sum_v U_{j+\hat{\mu},v} P_{j+\hat{\mu}+\hat{v},j}^R - \sum_{v \neq \mu} U_{j+\hat{\mu}+\hat{v},v}^{\dagger} P_{j+\hat{\mu}-\hat{v},j}^R \right] \right]_{\text{TA}} \quad (46)$$

for even sites and

$$i\dot{H}_{j,\mu}^R = \left[\frac{\beta}{3} U_{j,\mu} V_{j,\mu} - \frac{1}{2} N_f U_{j,\mu} \left[\sum_v P_{j+\hat{\mu},j-\hat{v}}^R U_{j-\hat{v},v} - \sum_{v \neq \mu} P_{j+\hat{\mu},j+\hat{v}}^R U_{j,v}^{\dagger} \right] \right]_{\text{TA}} \quad (47)$$

for odd sites. P^R is defined in Eqs. (41) and (42).

Now let us begin with $U(t)$ and a newly refreshed $H(t)$, and compute $U(t+N\Delta t)$ and $H(t+N\Delta t)$, using N molecular-dynamics steps.

(1) Generate an intermediate U :

$$U_{j,\mu} \left[t + \Delta t \left[\frac{1}{2} - \frac{N_f}{8} \right] \right] = \exp \left[i \Delta t \left[\frac{1}{2} - \frac{N_f}{8} \right] H_{j,\mu}(t) \right] U_{j,\mu}(t). \quad (48)$$

(2) Generate a Gaussian random vector R and an intermediate vector Φ :

$$\Phi = M^{\dagger} \left[t + \Delta t \left[\frac{1}{2} - \frac{N_f}{8} \right] \right] R. \quad (49)$$

(3) Compute U at the midpoint:

$$U_{j,\mu}(t + \frac{1}{2} \Delta t) = e^{i \Delta t (N_f/8) H_{j,\mu}(t)} \times U_{j,\mu} \left[t + \Delta t \left[\frac{1}{2} - \frac{N_f}{8} \right] \right]. \quad (50)$$

(4) Compute X :

$$X = \frac{1}{M^{\dagger}(t + \frac{1}{2} \Delta t) M(t + \frac{1}{2} \Delta t)} \Phi. \quad (51)$$

(5) Use Eqs. (46) and (47) to compute $\dot{H}_{j,\mu}(t + \frac{1}{2} \Delta t)$.

(6) Compute $H_{j,\mu}(t + \Delta t)$:

$$H_{j,\mu}(t + \Delta t) = H_{j,\mu}(t) + \Delta t \dot{H}_{j,\mu}^R(t + \frac{1}{2} \Delta t). \quad (52)$$

(7) Unless this is the last time step, compute U at the next intermediate point:

$$U_{j,\mu} \left[t + \Delta t \left[\frac{1}{2} - \frac{N_f}{8} \right] + \Delta t \right] = \exp \left[i \Delta t \left[1 - \frac{N_f}{8} \right] H_{j,\mu}(t + \Delta t) \right] U_{j,\mu}(t + \frac{1}{2} \Delta t). \quad (53)$$

Go back to step 2, increasing all times by Δt and repeat steps 2–7 N times, leaving the loop after step 6 in the last iteration. We have now computed $H(t+N\Delta t)$ and $U[t + (N - \frac{1}{2})\Delta t]$. Finish things up neatly by computing $U(t+N\Delta t)$:

$$U_{j,\mu}(t+N\Delta t) = e^{i(\Delta t/2) H_{j,\mu}(t+N\Delta t)} U_{j,\mu}(t + (N - \frac{1}{2})\Delta t). \quad (54)$$

Notice we need to update U twice in each step. However, this is negligible compared to the updating of H and to the conjugate gradient, both of which are done only once per step. Also, no additional storage is needed.

We have tested these algorithms for the case of $N_f=2$. Our prediction from Eq. (36) is that the order Δt^2 error in the integration of the molecular-dynamics steps for the R_0 algorithm should translate into an order Δt error in measured quantities. Our analysis of the Fokker-Planck equation indicates that for $N_f=2$ the Φ algorithm (with updating of the Φ field at each time step) should have a Δt error in measured quantities of the same magnitude as the R_0 algorithm, but opposite in sign. The R algorithm of course is expected to have errors of order Δt^2 in measured quantities. In Fig. 7 we show the average value of

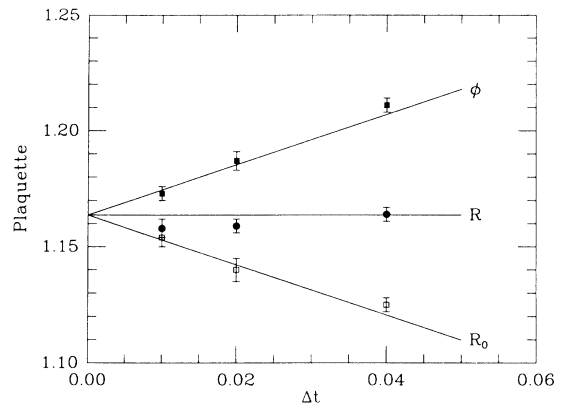


FIG. 7. The average value of the plaquette as a function of Δt for two flavors of staggered fermions on a 4^4 lattice with $\beta=4.80$ and $m=0.25$. The open squares are from the R_0 algorithm, the filled squares from the Φ algorithm, and the filled circles from the R algorithm. The straight lines are a least-squares fit in which it was assumed that the R_0 and Φ algorithms have errors of order Δt which are equal in magnitude and opposite in sign, and that the R algorithm has errors of order Δt^2 .

the plaquette for these three algorithms as a function of step size for a 4^4 lattice with $\beta=4.80$. The straight lines are a least-squares fit incorporating our prediction, which is well borne out. We have also run the R algorithm at $N_f=4$ in order to compare with previous results obtained with the Φ algorithm. One might expect to obtain shorter autocorrelation times by keeping the Φ field fixed throughout the molecular-dynamics steps, since then the U and H fields will evolve in the presence of a fixed external field rather than in a noisy one. This does not prove to be the case. We detect no appreciable difference in correlation times between the two algorithms. The one advantage of the Φ algorithm is that during the molecular-dynamics steps one can make good choices for the starting point of the conjugate-gradient calculation by using Eq. (18). No such extrapolation of X^R is possible in the R algorithm since the random field is changed at each time step. Of course the Φ algorithm is applicable only to four flavors of Kogut-Susskind fermions or two flavors of Wilson fermions, while the R algorithm is applicable to any number of flavors.

IV. CONCLUSION

We have discussed two algorithms for the numerical simulation of QCD based upon the hybrid stochastic method developed by Duane and Kogut. Both algorithms provide a relatively rapid evolution of the gauge fields through configuration space, have errors of second order in the molecular-dynamics step size, and require only one conjugate-gradient calculation per molecular-dynamics step. The Φ algorithm has the advantage of allowing us to minimize the number of conjugate gradient iterations, while the R algorithm has the advantage of being applicable to arbitrary numbers of flavors. We are presently applying the Φ algorithm to the study of the thermodynamics of four flavors of Kogut-Susskind fermions, and the R algorithm to the thermodynamics of two flavors of Kogut-Susskind fermions. Results of these calculations will be presented separately. Our preliminary results are

encouraging enough to make us believe that we can apply these algorithms to more challenging problems such as the study of the mass spectrum on moderate sized lattices.

ACKNOWLEDGMENTS

We would like to thank F. Niedermayer for helpful conversations. This work was supported in part by National Science Foundation Grants No. PHY-83-13324 and DMR 83-20423 and by Department of Energy Grants Nos. DE-AT03-81ER40029 and DE-AC02-84ER40125. We gratefully acknowledge E. I. Dupont de Nemours and Co. and Xerox Corp. for their support.

APPENDIX

In this appendix we present the details of the error analysis of the R algorithm for gauge theories. Our discussion parallels the one given for the scalar field in Sec. III.

It is convenient to introduce $D(U)$ the covariant derivative with respect to the $SU(3)$ matrix U through the relation

$$\lim_{\Delta t \rightarrow 0} \frac{f(e^{i\Delta t H} U) - f(U)}{\Delta t} = \sum_{\alpha, \beta} i H^{\alpha, \beta} D_{\beta, \alpha}(U) f(U) \\ = \text{Tr}[i H D(U)] f(U). \quad (\text{A1})$$

Here H is an arbitrary traceless, Hermitian matrix, and α and β are $SU(3)$ indices. Throughout this appendix Tr will refer to a trace over $SU(3)$ indices only. We will continue to use tr to indicate a trace over lattice sites and $SU(3)$ indices. It follows from the invariance of the Haar measure that one can integrate by parts with this definition of the derivative.

The effective Hamiltonian is

$$\mathcal{H} = \frac{1}{2} \text{tr} H^2 + S_W - \frac{1}{4} N_f \text{tr} \ln(M^\dagger M). \quad (\text{A2})$$

Let $\mathcal{P}(H, U, t)$ be the probability distribution of the H and U fields at the time t . The Fokker-Planck equation for the change in \mathcal{P} during a single time step is

$$\mathcal{P}(U, H, t + \Delta t) = \int [\delta U' \delta H'] \mathcal{P}(U', H', t) \\ \times \prod_{j, \mu} \delta(H_{j, \mu} - H'_{j, \mu} - \Delta H_{j, \mu}^R) \prod_{j, \mu} \delta(U_{j, \mu} - \exp[i \Delta t (H'_{j, \mu} + \frac{1}{2} \Delta H_{j, \mu}^R)] U'_{j, \mu}), \quad (\text{A3})$$

where

$$\Delta H_{j, \mu}^R = \Delta t \dot{H}_{j, \mu}^R(U(t + \frac{1}{2} \Delta t)) \\ = \Delta t \dot{H}_{j, \mu}^R(U') + \frac{1}{2} \Delta t^2 \sum_{k, \nu} \text{Tr}[i H'_{k, \nu} D(U'_{k, \nu})] \dot{H}_{j, \mu}^R(U') + O(\Delta t^3), \quad (\text{A4})$$

and $\dot{H}_{j, \mu}^R$ is given by Eqs. (41), (42), (46), and (47).

Expanding the right-hand side of Eq. (A3) in powers of Δt gives

$$\begin{aligned}
&= \sum_{j,\mu} \left[-\frac{\partial}{\partial H_{j,\mu}} (\mathcal{P} \Delta H_{j,\mu}^R) - \text{Tr} \{ D(U_{j,\mu}) [\mathcal{P} i \Delta t (H_{j,\mu} + \frac{1}{2} \Delta t \dot{H}_{j,\mu}^R)] \} \right] + \frac{1}{2} \sum_{\substack{j,\mu \\ k,\nu}} \frac{\partial^2}{\partial H_{j,\mu} \partial H_{k,\nu}} (\mathcal{P} \Delta H_{j,\mu}^R \Delta H_{k,\nu}^R) \\
&+ \sum_{\substack{j,\mu \\ k,\nu}} \sum_{\substack{\alpha,\beta \\ \gamma,\delta}} D_{\beta,\alpha}(U_{j,\mu}) D_{\delta,\gamma}(U_{k,\nu}) [\mathcal{P} (i \Delta t)^2 H_{j,\mu}^{\alpha,\beta} H_{k,\nu}^{\gamma,\delta}] + \sum_{\substack{j,\mu \\ k,\nu}} \sum_{\gamma,\delta} \frac{\partial}{\partial H_{j,\mu}} D_{\delta,\gamma}(U_{k,\nu}) (\mathcal{P} \Delta H_{j,\mu}^R i \Delta t H_{k,\nu}^{\gamma,\delta}) + O(\Delta t^3) . \quad (\text{A5})
\end{aligned}$$

In obtaining Eq. (A5) from Eq. (A3) we have performed the integrations over U and H by means of the δ functions. Therefore in Eq. (A5) $\Delta H_{j,\mu}^R$ is to be evaluated using the second line of Eq. (A4) with $U'_{j,\mu} = U_{j,\mu}$.

We shall assume that $\mathcal{P}(U, H, t) = Z^{-1} \exp(-\mathcal{H})$. We shall then show that after averaging over R the distribution function is unchanged at time $t + \Delta t$ to within errors of order Δt^3 . That is, we have the correct equilibrium distribution to within the stated errors. With our present notation

$$\begin{aligned}
\dot{H}_{j,\mu} &= \frac{1}{2} i [D(U_{j,\mu}) \mathcal{H}]_{\text{TA}} \\
&= i \left[\frac{\beta}{3} U_{j,\mu} V_{j,\mu} - \frac{1}{2} N_f \text{tr} \left[\frac{1}{M^\dagger M} D(U_{j,\mu}) (M^\dagger M) \right] \right]_{\text{TA}} . \quad (\text{A6})
\end{aligned}$$

In addition, since $\dot{H}_{j,\mu}$ is a function of U , but not of H

$$\ddot{H}_{j,\mu} = \sum_{k,\mu} \text{Tr} [i H_{k,\nu} D(U_{k,\nu})] \dot{H}_{j,\mu} . \quad (\text{A7})$$

The same is true of $\dot{H}_{j,\mu}^R$ to leading order in Δt , so

$$\ddot{H}_{j,\mu}^R = \sum_{k,\mu} \text{Tr} [i H_{k,\nu} D(U_{k,\nu})] \dot{H}_{j,\mu}^R + O(\Delta t) . \quad (\text{A8})$$

Substituting these results into Eq. (A5) yields

$$\begin{aligned}
\Delta \mathcal{P} &= \mathcal{P} \left[\Delta t \sum_{j,\mu} \text{Tr} \left[H_{j,\mu} (\dot{H}_{j,\mu} - \dot{H}_{j,\mu}^R) - \frac{\partial}{\partial H_{j,\mu}} \dot{H}_{j,\mu}^R \right] \right. \\
&\quad \left. + \frac{1}{2} \Delta t^2 \sum_{j,\mu} \text{Tr} [H_{j,\mu} (\ddot{H}_{j,\mu} - \ddot{H}_{j,\mu}^R) + \dot{H}_{j,\mu}^R (\dot{H}_{j,\mu} - \dot{H}_{j,\mu}^R)] + \frac{1}{2} \Delta t^2 \left[\sum_{j,\mu} \text{Tr} [H_{j,\mu} (\dot{H}_{j,\mu} - \dot{H}_{j,\mu}^R)] \right]^2 \right] + O(\Delta t^3) . \quad (\text{A9})
\end{aligned}$$

Referring to Eq. (42) we see that

$$X^R(t + \frac{1}{2} \Delta t) = \frac{1}{M^\dagger(t + \frac{1}{2} \Delta t) M(t + \frac{1}{2} \Delta t)} \left[M^\dagger(t + \frac{1}{2} \Delta t) - \frac{1}{8} N_f \Delta t \sum_{j,\mu} \text{Tr} [i H_{j,\mu} D(U_{j,\mu})] M^\dagger(t + \frac{1}{2} \Delta t) \right] R , \quad (\text{A10})$$

from which we learn that

$$\langle \dot{H}_{j,\mu}^R \rangle_R = \dot{H}_{j,\mu} + \frac{1}{8} N_f \Delta t \sum_{k,\nu} \text{tr} \left[\frac{1}{M^\dagger M} i D(U_{j,\mu}) (M^\dagger M) - \frac{1}{M^\dagger M} \text{Tr} [i H_{k,\nu} D(U_{k,\nu})] (M^\dagger M) \right] . \quad (\text{A11})$$

Finally,

$$\langle \ddot{H}_{j,\mu}^R \rangle_R = \ddot{H}_{j,\mu} + O(\Delta t) . \quad (\text{A12})$$

Substituting back into Eq. (A9) we find

$$\langle \Delta \mathcal{P} \rangle_R = O(\Delta t^3) , \quad (\text{A13})$$

the desired result.

¹S. Duane, Nucl. Phys. **B257**, 652 (1985); S. Duane and J. Kogut, Phys. Rev. Lett. **55**, 2774 (1985).

²S. Duane and J. Kogut, Nucl. Phys. **B275**, 398 (1986).

³G. G. Batrouni, G. R. Katz, A. S. Kronfeld, G. P. Lepage, B. Svetitsky, and K. G. Wilson, Phys. Rev. D **32**, 2736 (1985).

⁴F. Fucito, E. Marinari, G. Parisi, and C. Rebbi, Nucl. Phys.

B180, 369 (1981).

⁵J. Polonyi, H. W. Wyld, J. B. Kogut, J. Shigemitsu, and D. K. Sinclair, Phys. Rev. Lett. **53**, 644 (1984); O. Martin and S. Otto, Phys. Rev. D **31**, 435 (1985).

⁶D. Callaway and A. Rahman, Phys. Rev. Lett. **49**, 613 (1982).

⁷R. T. Scalettar, D. J. Scalapino, and R. L. Sugar, Phys. Rev. B

- 34**, 7911 (1986).
- ⁸F. Niedermayer (private communication).
- ⁹R. V. Gagai and A. Gocksch, BNL Report No. BNL-38204 (unpublished); R. V. Gagai, A. Gocksch, and U. M. Heller, ITP Report No. NSF-ITP-86-89 (unpublished).
- ¹⁰S. A. Gottlieb, W. Liu, D. Toussaint, and R. L. Sugar, this issue, Phys. Rev. D **35**, 2611 (1987).
- ¹¹R. V. Gagai, J. Potvin, and S. Sanielevici, BNL Report No. BNL-38526 (unpublished).
- ¹²M. Fukugita and A. Ukawa, Phys. Rev. Lett. **57**, 503 (1986); R. Gupta, G. Guralnik, G. Kilcup, A. Patel, and S. Sharpe, *ibid.* **57**, 2621 (1986).
- ¹³The same idea can be used in the Langevin equation if the fermions are replaced by noise. It would appear that one can perform a second-order integration of the equations with only two conjugate-gradient calculations per step rather than three as in the algorithm of Batrouni [G. Batrouni, Phys. Rev. D **33**, 1815 (1986)].