PHYSICAL REVIEW D VOLUME 32, NUMBER 10 15 NOVEMBER 1985

Langevin simulations of lattice field theories

G. G. Batrouni, G. R. Katz, A. S. Kronfeld, G. P. Lepage, B. Svetitsky,* and K. G. Wilson
Newman Laboratory of Nuclear Studies, Cornell University, Ithaca, New York 14853
(Received 17 May 1985)

We present a new analysis of Langevin simulation techniques for lattice field theories, including a general discussion of errors and algorithm speed. We introduce Fourier techniques that greatly accelerate simulations on large lattices. We also introduce a new technique for including quark vacuum-polarization corrections that admits any number of flavors, odd or even, without the need for nested Monte Carlo calculations. Our analysis is supported by a variety of numerical experiments.

I. INTRODUCTION

Considerable progress has been made over the past few years in the numerical simulation of quantum chromodynamics (QCD). With rapid developments in computing technology we can expect major new advances in the near future, leading to definitive tests of the full theory. To realize this goal the next generation of QCD simulations will have to address several critical issues:

the efficient inclusion of effects due to quark vacuum polarization;

the critical slowing down of existing simulation algorithms with larger lattices, resulting in computing needs that can grow quadratically or worse with the lattice volume;

the need for economical methods for high-statistics measurements of such things as the spectrum of the light hadrons, the glueball spectrum, or the heavy quark potential;

the need for a range of checks on the systematic errors in such measurements.

In this paper we demonstrate how the first two problems—those relating to the generation of field configurations—can be tackled using stochastic quantization via the Langevin equation.[1] Langevin updating proves to be ideal here, largely because the entire lattice is updated simultaneously, rather than link-by-link or site-by-site. We will discuss strategies for high statistics measurements on these configurations in future papers.

The first problem, that of including fermion determinants, can be solved by adding a new nonlocal noise term to the Langevin equation for gauge fields. As we demonstrate in Sec. IV, the only price paid for the nonlocality is that linear equations of the form

$$M\psi = \phi \tag{1.1}$$

must be solved for each update, where $M=\gamma_5(D\cdot\gamma+m)$ is the inverse fermion propagator. This is a reasonable price if by "each update" one means an update of the entire field configuration, as in Langevin updating. By contrast, a heat-bath or Metropolis algorithm demands that similar equations be solved once for the update of each separate link and site—an extremely costly venture. The algorithm we describe in Sec. IV has several other attractive features; most notably, it allows the simulation of any number of quark flavors, even or odd, without the need for a Monte Carlo calculation within a Monte Carlo calculation.

Langevin updating also offers a new approach to the problem of critical slowing-down. A crucial weakness in most updating algorithms is that short-wavelength structure tends to evolve far more quickly than structure at long wavelengths—typically faster by a factor of ξ^2, where ξ is the longest correlation length of the theory (measured in units of the lattice spacing). Since one must limit the size of a single update to maintain stability at short distances, the evolution of large-scale features is greatly slowed. Indeed the number of sweeps of the lattice required for appreciable change at these wavelengths grows as ξ^2, which tends to infinity in the continuum limit. In Sec. III we show how this problem can be remedied by updating field configurations in momentum space rather than in coordinate space.[2] Fast Fourier transforms (FFT's) are used to relate the two descriptions, and again the procedure is only feasible when the entire configuration is updated simultaneously (two FFT's are required per update). Using momentum-space updating for several two-dimensional spin models, we found that simulation speed can be increased by factors of 20—50 on 16×16 lattices. Similar gains seem attainable with QCD.

Momentum-space updating of gauge fields is only useful if calculations are done in a smooth gauge such as $A^0=0$ gauge; otherwise, the gauge freedom obscures the relationship between momentum and wavelength. With Langevin updating the cost of such gauge fixing is almost insignificant, once again because it is required only once for each update of the entire field.

Using Fourier transforms one can also accelerate the convergence of the iterative matrix methods needed to solve Eq. (1.1) or to compute fermion propagators. The problems here are very similar to those just discussed for updating. For instance, large-scale structure develops $O(\xi)$ times more slowly than short-wavelength structure when solving Eq. (1.1) with the conjugate gradient algo-

rithm. The remedy is also very similar, as we illustrate in Sec. IV. Fourier acceleration makes possible QCD algorithms for which the amount of computation needed grows only linearly (up to logarithms) with the number of lattice points as the mesh is refined.

An analysis of algorithm speed requires an understanding of the errors inherent in Langevin updating. Errors arise because the differential Langevin equation must be replaced by a finite-difference equation for the purposes of a computer simulation. In scalar field theory, for example, the differential Langevin equation is[1]

$$\frac{\partial\phi(x,\tau)}{\partial\tau} = -\frac{\delta S[\phi]}{\delta\phi(x,\tau)} + \bar{\eta}(x,\tau), \tag{1.2}$$

where τ is a fictitious time labeling successive field configurations and $\bar{\eta}$ is a Gaussian random noise normalized by

$$\langle\bar{\eta}(x,\tau)\bar{\eta}(x',\tau')\rangle_{\bar{\eta}} = 2\,\delta(x'-x)\,\delta(\tau'-\tau).$$

Once τ is sufficiently large, the field configurations $\{\phi(x,\tau)\}$ for different τ's are distributed with probability density $e^{-S[\phi]}$, as desired for a simulation of the field theory. A finite-difference form of this equation, suitable for computer simulations, can be obtained by making τ discrete with step size $d\tau=\epsilon$, e.g.,

$$\phi(x,\tau_{n+1}) = \phi(x,\tau_n) - f_x[\phi,\eta]\,, \tag{1.3a}$$

where

$$f_x = \bar{\epsilon}\frac{\delta S}{\delta\phi(x,\tau_n)} + \sqrt{\bar{\epsilon}}\,\eta(x,\tau_n) \tag{1.3b}$$

and where now η is normalized by

$$\langle\eta(x_i,\tau_i)\,\eta(x_j,\tau_j)\rangle_\eta = 2\,\delta_{x_i,x_j}\delta_{\tau_i,\tau_j}\,.$$

Discrete equations of this type need not be regarded solely as approximations to Eq. (1.2), valid only as $\bar{\epsilon}\to 0$. Rather, they are in themselves viable stochastic equations for any value of $\bar{\epsilon}$. However, for τ sufficiently large, the field configurations generated by the discrete equation (1.3) are distributed with a probability density $e^{-\bar{S}[\phi]}$, where now the equilibrium action $\bar{S}$ differs from S in $O(\bar{\epsilon})$ and beyond:

$$\bar{S}[\phi] = S[\phi] + \bar{\epsilon}S_1[\phi] + \cdots\,. \tag{1.4}$$

The question of errors centers upon the significance of these new additions to the action.

In Sec. II we examine the leading correction, $\bar{\epsilon}S_1$, for local field theories. We find that these corrections simply renormalize existing couplings in S, either directly or through quantum fluctuations. This pattern seems likely to persist in $O(\epsilon^2)$ and beyond, indicating that $\bar{S}$ and S represent the same continuum theory, at least if $\bar{\epsilon}$ is not too large. There is no error associated with using (1.3) in place of (1.2).

More generally, the step size $\bar{\epsilon}$ in Eq. (1.3b) can be replaced by a matrix $\epsilon_{x,y}$.[3] In fact, the key to Fourier acceleration is the proper choice of this matrix. The errors in this case are not merely renormalization effects, as above. Nevertheless, they are easily understood and bounded, as we illustrate in Sec. II. We find that algorithm speed (see Sec. III) depends somewhat upon the accuracy desired, but most importantly, it is independent of the correlation length ξ when $\epsilon_{x,y}$ is well chosen. Furthermore, one can reduce such errors substantially by employing higher-order difference equations, like the Runge-Kutta scheme of Sec. II C, in place of (1.3). None of these conclusions is altered when fermion loops are included as in Sec. IV.

This paper concludes, in Sec. V, with a brief summary of our findings, particularly as they apply to simulations of QCD.

II. DISCRETE LANGEVIN PROCESSES

In this section, we analyze the errors associated with the use of a discrete Langevin equation in place of the differential equation. Although most of our results are quite general, we focus here on theories for which the continuum action is local. Simulations involving the nonlocal effects of fermion loops are discussed in Sec. IV.

A. The effective equilibrium action—scalar fields

When the step size is replaced by a matrix Eq. (1.3) becomes

$$\begin{aligned}\phi(x,\tau_{n+1}) &= \phi(x,\tau_n) - f_x[\phi,\eta]\,,\\ f_x &= \sum_y \left[\epsilon_{x,y}\frac{\delta S}{\delta\phi(y,\tau_n)} + \sqrt{\epsilon_{x,y}}\,\eta(y,\tau_n)\right],\end{aligned} \tag{2.1}$$

where η is Gaussian noise with $\langle\eta^2\rangle = 2$. As $n\to\infty$, this equation generates an ensemble of field configurations from the ensemble of random variables $\eta(x,\tau_n)$. Here we wish to study the equilibrium distribution of these asymptotic field configurations.

The Fokker-Planck equation determines the evolution of the probability density, $P(\{\phi\},\tau_n)$, for obtaining a configuration $\{\phi\}$ after Langevin time τ_n:

$$P(\tau_{n+1}) - P(\tau_n) = \sum_{n=1}^{\infty}\sum_{x_1\cdots x_n}\frac{\delta}{\delta\phi(x_1)}\cdots\frac{\delta}{\delta\phi(x_n)} \times[\,\Delta_{x_1\cdots x_n}P(\tau_n)\,]\,, \tag{2.2a}$$

where

$$\Delta_{x_1\cdots x_n} = \frac{1}{n!}\langle f_{x_1}\cdots f_{x_n}\rangle_\eta\,. \tag{2.2b}$$

This equation is readily derived from the evolution equation

$$P(\{\phi\},\tau_{n+1}) = \left\langle \int d\phi\; P(\{\phi\},\tau_n) \times\prod_x \delta(\phi(x,\tau_{n+1}) - \phi(x,\tau_n) + f_x)\right\rangle_\eta \tag{2.3}$$

by Taylor expanding the δ functions in powers of f and averaging over the noise, η. Equation (2.2a) can be

analyzed order-by-order in ϵ since only $2n$ terms on the right-hand side contribute through $O(\epsilon^n)$. The left-hand side vanishes as $\tau\to\infty$, leading to an equation for the asymptotic distribution $\bar{P}(\{\phi\})\equiv P(\{\phi\},\tau\to\infty)$. To $O(\epsilon)$, at least, Eq. (2.2a) then becomes[4]

$$-\frac{\delta}{\delta\phi(x)}\left[\frac{\delta\bar{P}}{\delta\phi(x)}+\bar{S}_x\bar{P}\right]=0, \tag{2.4}$$

where the effective Langevin force $\bar{S}_x$ can be written as a gradient, $\delta\bar{S}/\delta\phi(x)$. Consequently the asymptotic distribution is

$$\bar{P}(\{\phi\})=e^{-\bar{S}[\phi]}, \tag{2.5a}$$

where the equilibrium action is

$$\bar{S}[\phi]=S[\phi]+\tfrac{1}{4}\sum_{x,y}\epsilon_{x,y}\left[2\frac{\delta^2S[\phi]}{\delta\phi(x)\delta\phi(y)}-\frac{\delta S[\phi]}{\delta\phi(x)}\frac{\delta S[\phi]}{\delta\phi(y)}\right]+\cdots. \tag{2.5b}$$

It is $\bar{S}$ rather than S that determines the physics of the discrete-time simulation.

The impact of the $O(\epsilon)$ terms in $\bar{S}$ is most transparent when $\epsilon_{x,y}$ is completely local:

$$\epsilon_{x,y}=\bar{\epsilon}\,\delta_{x,y}\ . \tag{2.6}$$

Then the correction terms in $\bar{S}$ are local—in a $\lambda\phi^4$ theory, for example, the $O(\epsilon)$ corrections include terms like $\bar{\epsilon}\lambda\phi^2$, $\bar{\epsilon}(\partial^2\phi)^2$, and $\bar{\epsilon}\lambda^2\phi^6$. Since the new interactions generally respect the internal symmetries of the original action, they either renormalize existing couplings in S, or introduce nonrenormalizable interactions. In either case the continuum limit of the theory is completely unaffected. One need only shift the bare couplings in S, by amounts of relative order $\bar{\epsilon}$, to compensate for all effects due to the new terms in $\bar{S}$. Of course, if $\bar{\epsilon}$ is too large (i.e., ~ 1), such shifts might drastically alter the nature of S (e.g., $\lambda\to\bar{\lambda}<0$), thereby destabilizing the simulation. Assuming that this is not the case, $\bar{S}$ is no less fundamental or correct than S as a lattice approximation to the continuum action.

The correction terms in $\bar{S}$ are more complicated when $\epsilon_{x,y}$ is nonlocal. In $\lambda\phi^4$ theory, for example, one obtains new interaction terms like

$$\lambda^2\sum_{x,y}\phi(x)^3\epsilon_{x,y}\phi(y)^3 \tag{2.7a}$$

and

$$\lambda\sum_x\phi(x)^2\epsilon_{x,x}. \tag{2.7b}$$

These are terms of the kind one would obtain in S if there were another scalar "particle" ζ, with propagator $\epsilon_{x,y}$, coupled to the ϕ field via interactions like $\lambda\phi(x)^3\zeta(x)$ and $\lambda\phi(x)^2\zeta(x)^2$. In this picture, the local $\epsilon_{x,y}$ considered above [Eq. (2.6)] is just the infinite-mass limit for the ζ, and so it is not surprising that the ζ corrections in effect simply renormalize the bare couplings of S. As we shall see in Sec. III, a better choice for $\epsilon_{x,y}$ is

$$\epsilon_{x,y}=\bar{\epsilon}G_{x,y}\ , \tag{2.8}$$

where $G_{x,y}$ is the propagator for a free scalar particle. In this case, the fictitious ζ particle behaves very much like the ϕ particle, and its contribution to physical measurables typically will be $\bar{\epsilon}$ times that from the ϕ field. Therefore to simulate S accurately to 10%, for example, one requires $\bar{\epsilon}\sim 0.1$.

Curiously, if one works only to $O(\epsilon)$, all of the new nonrenormalizable and nonlocal interactions in $\bar{S}$ can be removed, leaving behind terms that can be completely absorbed into S simply by shifting the coupling constants by calculable amounts. This is accomplished by changing field variables to

$$\bar{\phi}(x)=\phi(x)-\tfrac{1}{4}\sum_y\epsilon_{x,y}\frac{\delta S}{\delta\phi(y)}. \tag{2.9}$$

Then the action can be reexpressed in terms of $\bar{\phi}$,

$$S[\phi]=S[\bar{\phi}]+\tfrac{1}{4}\sum_{x,y}\frac{\delta S}{\delta\phi(x)}\epsilon_{x,y}\frac{\delta S}{\delta\phi(y)}+\cdots, \tag{2.10}$$

while the measure becomes

$$\prod_x d\phi(x)=\prod_x d\bar{\phi}(x)\left[1+\tfrac{1}{4}\sum_y\epsilon_{x,y}\frac{\delta^2S}{\delta\phi(x)\delta\phi(y)}+\cdots\right]=\left[\prod_x d\bar{\phi}(x)\right]\exp\left[\tfrac{1}{4}\sum_{x,y}\epsilon_{x,y}\frac{\delta^2S}{\delta\phi(x)\delta\phi(y)}+\cdots\right]. \tag{2.11}$$

With these changes the equilibrium action is

$$\bar{S}=S[\bar{\phi}]+\tfrac{1}{4}\sum_{x,y}\epsilon_{x,y}\frac{\delta^2S}{\delta\phi(x)\delta\phi(y)}+\cdots \tag{2.12}$$

which, for most theories, differs from $S[\phi]$ only by calculable shifts of the bare couplings already present in S.

In $\lambda\phi^4$ theory with $\epsilon_{x,y}=\bar{\epsilon}\delta_{x,y}$, for example, the only modification is $m^2\to m^2+\bar{\epsilon}\lambda/4$. Moreover, when ϵ is local like this the change of variables has no effect on the matrix elements of any operator linear in the $\phi(x)$'s, i.e.,

$$\langle f[\bar{\phi}]\rangle=\langle f[\phi]\rangle+O(\epsilon^2)\ , \tag{2.13}$$

where $f[\phi]$ is any multilinear functional. This result fol-

lows directly from the equation of motion

$$\int d\phi\, e^{-S[\phi]} \frac{\delta S}{\delta\phi(x)} \frac{\delta f}{\delta\phi(x)} = 0. \tag{2.14}$$

This technique for removing $O(\epsilon)$ corrections is an effective way of speeding up simulations with a nonlocal $\epsilon_{x,y}$, as in Eq. (2.8), since larger $\bar{\epsilon}$'s can be employed with no loss of accuracy. The added cost of shifting couplings and fields is generally negligible.

We have not yet investigated the $O(\epsilon^2)$ corrections that arise in the discrete formulation of the Langevin equation. The significance of these corrections can be assessed by analyzing their effect on ordinary perturbation theory. A preliminary study of this sort indicates that corrections in higher order are similar in nature to those in $O(\epsilon)$.

B. Non-Abelian fields

The analysis for scalar fields can be generalized readily to include non-Abelian variables such as the link field in QCD or the meson field in a chiral model.[5] All that one needs is to define differentiation with respect to the non-Abelian variable, or, more precisely, differentiation within its group manifold. Assuming that the non-Abelian variable, U, is an element of a Lie group in some representation with generators T_i (where $[T^i, T^j] = ic_{ijk}T^k$ and $\mathrm{tr}(T^iT^j) = \delta_{ij}/2$), a standard definition for this derivative is obtained from

$$f(e^{i\delta\cdot T}U) = f(U) + \delta^i\partial_i f + O(\delta^2). \tag{2.15}$$

This choice allows integration by parts—the critical operation in passing from Eq. (2.3) to Eq.(2.2)—under the invariant Haar measure of the group. Given this definition, the most convenient discrete Langevin equation is

$$U(\tau_{n+1}) = e^{-if\cdot T}U(\tau_n), \tag{2.16}$$

where $f_j = f_j(U(\tau_n), \eta)$ is the driving function. The simplest choice for the driving function is to take $\epsilon_{x,y}$ local,

$$f_j(U(\tau_n), \eta) = \bar{\epsilon}\partial_j S[U] + \sqrt{\bar{\epsilon}}\eta_j, \tag{2.17}$$

leading to results very similar to those obtained from Eqs. (1.3) for scalar fields. In manipulating these equations one must be careful because the derivatives no longer commute. Rather, they obey the algebra of the Lie group:

$$[\partial_i, \partial_j] = -c_{ijk}\partial_k. \tag{2.18}$$

Consequently, the equilibrium action is

$$\bar{S}[U] = \left[1 + \frac{\bar{\epsilon}}{12}C_A\right]S[U] + \frac{\bar{\epsilon}}{4}\sum_j \{2\partial_j^2 S[U] - (\partial_j S[U])^2\} + \cdots \tag{2.19}$$

which differs from that for a scalar theory in the term involving C_A, the Casimir invariant of the Lie group's adjoint representation [$C_A = n$ for SU(n)].

These equations generalize in the obvious way when $\epsilon_{x,y}$ is nonlocal, and the significance of the $O(\epsilon)$ terms here is much the same as for the scalar field theories discussed above. If $\epsilon_{x,y}$ is local, the main effect of these terms is to renormalize the bare couplings in the original theory. Corrections involving a nonlocal $\epsilon_{x,y}$ resemble terms one might obtain from a new particle with propagator $\epsilon_{x,y}$ and with the quantum numbers of the non-Abelian field.

The use of a nonlocal $\epsilon_{x,y}$ in simulations of gauge theories (such as QCD) is complicated by the need to preserve local gauge invariance. Any breaking of this invariance tends to generate a large mass for the gauge bosons, either directly or through quadratic (in four dimensions) ultraviolet divergences in the quantum fluctuations. In QCD, this problem can be avoided by introducing gauge-invariant field variables. This is easily done through a gauge transformation. After each update of a configuration one applies the gauge transformation

$$U_\mu(n) \to G(n)\,U_\mu(n)\,G^\dagger(n+\hat{\mu}), \tag{2.20}$$

where $G(n)$ is a product of links along some path connecting site n to the origin. Such a transformation completely fixes the gauge, leaving only a global symmetry which is preserved for any choice of $\epsilon_{x,y}$. The axial gauge $A^0 = 0$ is an example of such a gauge. (Note that if $\epsilon_{x,y} = \bar{\epsilon}\delta_{x,y}$, the gauge fixing has no effect whatsoever on the evolution, in Langevin time, of gauge invariant quantities.) An alternative to the use of such gauges is to make $\bar{\epsilon}$ very small—$O(1/\xi^2)$ in four dimensions—so that the boson mass and other symmetry-breaking effects are negligible. Another possibility, rather similar to gauge fixing, is to reformulate the theory directly in terms of gauge-invariant fields—e.g., the corner variables of Ref. 6.

To gain experience with Langevin algorithms we applied them to SU(3) gauge theory in four dimensions, using the standard plaquette action

$$S[U] = -\frac{\beta}{2n}\sum_p \mathrm{tr}(U_p + U_p^\dagger), \tag{2.21}$$

where U_p is the plaquette operator. For this theory, the discrete Langevin equations (2.16) and (2.17) take the form

$$U_\mu(\tau_{n+1}) = e^{-F(\tau_n)}U_\mu(\tau_n), \tag{2.22a}$$

where

$$F(\tau_n) = \frac{\bar{\epsilon}\beta}{4n}\sum_{U_p \supset U_\mu}(U_p - U_p^\dagger) - \frac{\bar{\epsilon}\beta}{4n^2}\sum_{U_p \supset U_\mu}\mathrm{tr}(U_p - U_p^\dagger) + i\sqrt{\bar{\epsilon}}H(\tau_n), \tag{2.22b}$$

and $H(\tau_n)$ is a traceless Hermitian noise matrix obeying

$$\langle H_{ik}(x,\tau)H_{lm}(x',\tau')\rangle_H = \left[\delta_{il}\delta_{km} - \frac{1}{n}\delta_{ik}\delta_{lm}\right]\delta_{x,x'}\delta_{\tau,\tau'}. \tag{2.22c}$$

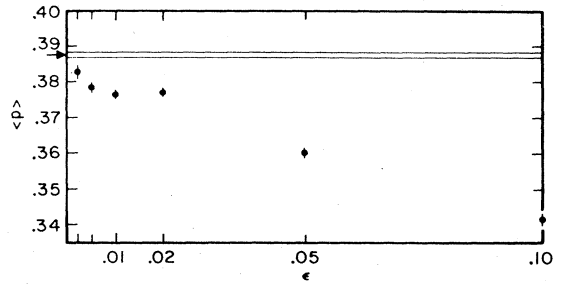

FIG. 1. Plot of the average plaquette vs ϵ at $\beta=4.9$ for the simple Euler algorithm. The horizontal lines near $\langle p\rangle=0.39$ indicate the Monte Carlo values. 1200 sweeps at $\epsilon=0.1$.

We iterated these equations with $\beta=4.9$ on a 2×5^3 lattice. The data for the average plaquette are plotted versus $\bar{\epsilon}$ in Fig. 1. (Note that the plaquette is dominated by ultraviolet contributions, and thus is quite sensitive to the differences between S and $\bar{S}$.) As $\bar{\epsilon}$ increases, the links tend to become more random, giving a smaller value for the average plaquette. At $\bar{\epsilon}=0.1$ our result differs from the Monte Carlo value by about 11%, as might have been expected from our error analysis.

One can remove $O(\epsilon)$ terms in $\bar{S}$ for non-Abelian variables, as for scalar variables, by a change of variables,

$$U \rightarrow \bar{U} = e^{-i(\bar{\epsilon}/4)\partial_j S T^j} U, \tag{2.23}$$

leaving an equilibrium action

$$\bar{S}[U] = \left[1+\frac{\bar{\epsilon}}{12}C_A\right] S[U] + \frac{\bar{\epsilon}}{4}\sum_j \partial_j^2 S[U] + \cdots . \tag{2.24}$$

(This works for a nonlocal $\epsilon_{x,y}$, as well.) For the plaquette action Eq. (2.21), $\bar{S}$ is then identical to S, but with β replaced by

$$\bar{\beta} = \beta[1-\bar{\epsilon}(C_l - C_A/12)], \tag{2.25}$$

where C_l is the Casimir invariant for the representation in which the U matrices are taken [e.g., $C_F=(n^2-1)/2n$ for the fundamental representation in SU(n)]. We redid the simulations described above, but now adjusting β so that $\bar{\beta}=4.9$ for all $\bar{\epsilon}$. The results, plotted in Fig. 2, are significantly closer to the Monte Carlo result, suggesting that $O(\epsilon)$ terms have indeed been removed. In computing the plaquette expectation values, we used the original U's, rather than the shifted $\bar{U}$'s, since in general

$$\langle f(\bar{U})\rangle = (1+\bar{\epsilon}C_l/4)^{\#\text{ links}}\langle f(U)\rangle + O(\bar{\epsilon}^2) \tag{2.26}$$

when f is any linear functional of the U's. Consequently the additional effort required to remove all $O(\bar{\epsilon})$ effects was negligible.

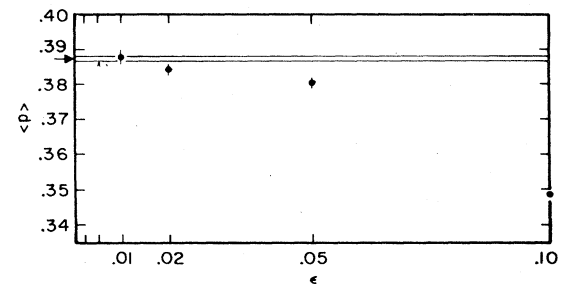

FIG. 2. Plot of the average plaquette vs ϵ at $\beta=4.9$ for the simple Euler algorithm with shifted coupling. 3200 sweeps at $\epsilon=0.1$.

C. Runge-Kutta algorithms

Equation (1.3) represents the simplest of an infinite number of discrete Langevin equations. In particular, there is a group of "Runge-Kutta" algorithms that model the differential Langevin equation (1.2) more closely. For example, one can model the action $S[\phi]$ correctly up to $O(\epsilon^2)$ by using the first-order Runge-Kutta equation [7,8]

$$\phi(\tau_{n+1}) = \phi(\tau_n) - \frac{\bar{\epsilon}}{2}\left[\frac{\delta S[\phi(\tau_n)]}{\delta\phi(\tau_n)} + \frac{\delta S[\tilde{\phi}(\tau_{n+1})]}{\delta\tilde{\phi}(\tau_{n+1})}\right] + \sqrt{\bar{\epsilon}}\,\eta(\tau_n), \tag{2.27}$$

where $\tilde{\phi}$ is a tentative update using the lowest-order algorithm,

$$\tilde{\phi}(\tau_{n+1}) = \phi(\tau_n) - \bar{\epsilon}\frac{\delta S[\phi(\tau_n)]}{\delta\phi} - \sqrt{\bar{\epsilon}}\,\eta(\tau_n).$$

The analogous equation for non-Abelian variables is obtained by replacing f_j in Eq. (2.17) with

$$f_j = \frac{\bar{\epsilon}}{2}\left[1+\frac{C_A\bar{\epsilon}}{6}\right]\{\partial_j S[U(\tau_n)] + \partial_j S[\tilde{U}(\tau_{n+1})]\} + \sqrt{\bar{\epsilon}}\,\eta_j(\tau_n), \tag{2.28}$$

where, as in the scalar case, $\tilde{U}(\tau_{n+1})$ is obtained from $U(\tau_n)$ using the leading order equations (2.16) and (2.17) with the same random noise variable $\eta_j(\tau_n)$. Such higher-order algorithms allow a much larger step size $\bar{\epsilon}$ before the equilibrium action $\bar{S}$ deviates much from the original action S. (They also generalize for nonlocal $\epsilon_{x,y}$.)

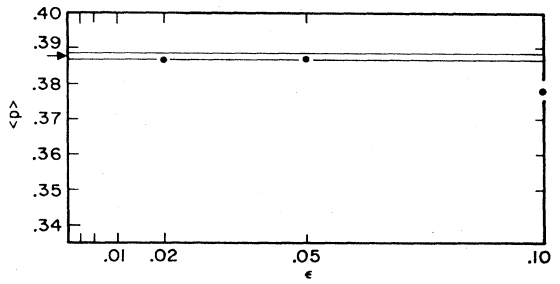

FIG. 3. Plot of the average plaquette vs ϵ at $\beta=4.9$ for the Runge-Kutta method. 1000 sweeps at $\epsilon=0.1$.

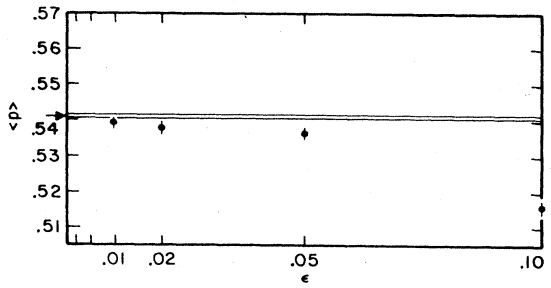

FIG. 4. Plot of the average plaquette vs ϵ at $\beta=5.5$ for the Runge-Kutta method. 1000 sweeps at $\epsilon=0.1$.

On the other hand, there is a penalty of roughly a factor of 2 in both computation and data storage for these particular Runge-Kutta algorithms.

We again simulated the pure gauge theory on a 2×5^3 lattice with $\beta=4.9$ and 5.5, this time using the first order Runge-Kutta scheme just described. Our data for the plaquette are plotted for different $\bar{\epsilon}$'s in Figs. 3 and 4. This algorithm ran 20—50 times faster than the simplest algorithm [Eqs. (2.16) and (2.17)] for similar precision. It also outperforms the algorithm with shifted couplings [Eq. (2.25)].

III. ALGORITHM SPEED AND FOURIER ACCELERATION

A. Free field theory

A critical figure of merit related to the speed of simulation algorithms is the correlation "time" N_c, which is the number of updates of the (whole) field configuration required to obtain a new, statistically uncorrelated configuration. Making the step size $\epsilon_{x,y}$ [Eq. (2.1)] larger reduces N_c, and is desirable. However, one is limited to $|\epsilon_{x,y}| \lesssim 1$ if the simulation is to remain stable (Sec. II A). If $\epsilon_{x,y}$ is local [Eq. (2.6)], this usually means that the correlation time grows quadratically,

$$N_c \sim \xi^2, \tag{3.1}$$

as ξ, the longest correlation length of the theory (in lattice units), increases (see below). With such behavior, the amount of computation required to achieve a given statistical accuracy grows as $V^{3/2}$ when the number of lattice points V is increased. (This assumes that the *physical* volume of space-time covered by the lattice remains constant, as will usually be the case.)

This behavior is evident in free field theory, with the quadratic action $S=\frac{1}{2}\phi M\phi$ where $M=-\partial^2+m^2$. The discrete Langevin equation is

$$\phi(x,\tau_{n+1}) = (1-\bar{\epsilon}M)\phi(x,\tau_n) + \sqrt{\bar{\epsilon}}\,\eta(\tau_n), \tag{3.2}$$

and it leads to an equilibrium action

$$\bar{S} = \tfrac{1}{2}\phi M\left[1-\frac{\bar{\epsilon}}{2}M\right]\phi + \text{constants}. \tag{3.3}$$

This equilibrium theory is obviously unstable unless

$$|\bar{\epsilon}M| < 2, \tag{3.4}$$

which is true if

$$\bar{\epsilon} \lesssim \frac{1}{p_{\max}^2+m^2}, \tag{3.5}$$

where $p_{\max}\sim 1$ is the maximum momentum (in lattice units) on the lattice. Thus, as expected, we require that $\bar{\epsilon}\lesssim 1$ for a stable simulation as $\xi\to\infty$ $(m\to 0)$.

To see how this restriction affects the correlation time N_c, we can solve the evolution equation (3.2) in momentum space, where[9] $M\to p^2+m^2$. With the initial condition $\phi(p,\tau=0)=0$, the solution is

$$\phi(p,\tau_n) = \sqrt{\bar{\epsilon}}\sum_{j=1}^{n}(1-\bar{\epsilon}[p^2+m^2])^{n-j}\eta(p,\tau_{j-1}). \tag{3.6}$$

Since $|1-\bar{\epsilon}[p^2+m^2]|<1$ from Eq. (3.5), $\phi(p,\tau_n)$ depends less and less upon $\eta(p,\tau_j)$ as $n-j$ grows. Roughly then, we can define the correlation time $N_c(p)$ for a given $\bar{\epsilon}$ and p by the relation

$$\log_e\{1-\bar{\epsilon}[p^2+m^2])^{N_c(p)}\}\sim -1$$

which implies

$$N_c(p) \sim \frac{1}{\bar{\epsilon}[p^2+m^2]}. \tag{3.7}$$

Choosing $\bar{\epsilon}\sim\bar{\epsilon}_{\max}\sim[p_{\max}^2+m^2]^{-1}$ implies that the high-momentum components of the field will evolve with a very short correlation time: $N_c(p_{\max})\sim 1$. However, the correlation time for the low momentum structure usually will be much longer:

$$N_c(p\sim 0) \sim \frac{1}{m^2} \sim \xi^2. \tag{3.8}$$

Thus $O(\xi^2)$ times as many updates of the field configuration are required to obtain the same statistics as for high momenta. Unfortunately, it is these long-wavelength components that are usually the most important physically.

This analysis illustrates a general problem with the local Langevin algorithm, and indeed with most other Monte Carlo algorithms used today. The evolution of long-wavelength structure is held back by the need for stability at short wavelengths. Unfortunately, the problem becomes more and more severe as the grid is refined and the correlation length ξ increases.

Such critical slowing down can only be remedied by resolving different length scales and dealing with them separately. Our analysis of free field theory suggests that this can be accomplished by updating the fields in momentum space, using different ϵ's for different momenta in such a way as to speed up the evolution at low momenta without destabilizing the high momenta.[2] This means employing a nonlocal $\epsilon_{x,y}$:

$$\epsilon_{x,y} = \sum_p e^{ip\cdot(x-y)}\epsilon(p).$$

Equations (3.5) and (3.7) indicate that the choice

$$\epsilon(p) \sim \frac{1}{p^2+m^2} \tag{3.9}$$

is optimal, at least for free field theory. Then the stability criterion [Eq. (3.5)] is satisfied, while the correlation time [Eq. (3.7)] is $N_c \sim 1$ for all momenta, not just for $p \sim p_{\max}$. Consequently, the number of field updates required for a simulation is reduced by a factor of $1/\xi^2$, and computing power need only grow linearly with the size of the lattice, rather than as $V^{3/2}$.

B. Interacting field theories

The basic strategy for accelerating simulations of an interacting field theory should be similar to that for free field theory. In particular, working in momentum space provides a way of separating the different length scales. Introducing a momentum-dependent step size is both straightforward and not particularly expensive with Langevin updating: Eqs. (1.3) are replaced by

$$\Delta\phi(x) = -\hat{F}\left[\epsilon(p)\hat{F}^{-1}\frac{\delta S}{\delta\phi(x)} + \sqrt{\epsilon(p)}\eta(p)\right], \tag{3.10}$$

where $\hat{F}$ is an operator representing a Fourier transform. The computation required for fast Fourier transforms is roughly comparable to that needed to evaluate $\delta S/\delta\phi$, since only two transforms are required for each update of the entire field configuration. A similar strategy would be impossible with Metropolis or heat-bath updating, where a new Fourier transform would be necessary for each hit at each site.

The optimal $\epsilon(p)$ for an interacting theory again should be

$$\epsilon(p) = \frac{\bar{\epsilon}}{p^2+m^2}, \tag{3.11}$$

where $\bar{\epsilon}$ determines the accuracy of the simulation (Sec. II) and where m will usually[10] be of order $1/\xi$.

In some theories ultraviolet and infrared behavior are strongly coupled. For example, the bare and renormalized masses in a four-dimensional scalar field theory differ by $O(\xi^2)$ due to ultraviolet fluctuations. One might worry about speeding up evolution at low momenta for such a theory, since the high-momentum structure must be sufficiently well developed to allow a very accurate estimate of the $O(\xi^2)$ renormalizations. However, such effects are largely determined by structure at distances of order a lattice spacing, and thus one has $O(\xi^4)$ independent samples in a single configuration from which they can be determined. Consequently the mass renormalization, for example, is specified to an accuracy of $\xi^2/(\xi^4)^{1/2} \sim 1$ by a single configuration—just adequate for our purposes. Such linkage between high momenta and low momenta generally does not seem to be a problem.

This discussion of scalar field theories applies to non-Abelian theories as well. In particular, Fourier transforms can be used just as they are for scalar theories. The transformations used in Eq. (3.10) can be applied without change to Eq. (2.17), used in updating non-Abelian fields. In gauge theories it is also important that the gauge be fixed as otherwise long-wavelength physics can be disguised by high-frequency fluctuations due to crinkly gauges. The $A^0=0$ gauge mentioned in Sec. II is sufficiently smooth.

C. A numerical example

To investigate the utility of Fourier acceleration we used it in simulating the two-dimensional XY model:[11]

$$S = -\beta\sum_{n,\mu}\cos(\theta_n-\theta_{n+\mu}) - h\sum_n\cos(\theta_n). \tag{3.12}$$

We shifted β, h, and the field θ_n so that all $O(\epsilon)$ corrections due to discrete Langevin updating cancelled [Eqs. (2.9)—(2.12)]. We then ran on a 16×16 lattice with both a p-dependent and a p-independent step size:

$$\epsilon(p) = \begin{cases} 0.02 \\ 0.02\,\dfrac{\left[\beta\sum_\mu[1-\cos(2\pi p_\mu/N)]+h\right]_{\max}}{\beta\sum_\mu[1-\cos(2\pi p_\mu/N)]+h} \end{cases}, \tag{3.13}$$

where the p dependence in the second case is that of the free propagator for θ. Our results for the momentum-space Green's function $\langle|\theta(p)|^2\rangle$, with $\beta=1.5$ and $h=0.23$, are given in Table I. As expected, performance is greatly improved by Fourier acceleration. At low momenta 400 sweeps with Fourier acceleration give results

TABLE I. $\langle|\theta(p)|^2\rangle$ for the XY action of Eq. (3.12). Values of the momentum-space propagator $\langle|\theta(p)|^2\rangle$ for the two-dimensional XY model on a 16^2 lattice. Columns 1 and 2 are results from perturbation theory, 3 and 4 are Langevin results with Fourier acceleration, 5 and 6 are the same but without Fourier acceleration. Runs were done at $\beta=1.5$ and $h=0.23$.

	Perturbation		FFT		No FFT	
$p^2/p^2_{\max}$	0th order	1 loop	400 steps	6000 steps	400 steps	6000 steps
$\frac{1}{32}$	2.093	2.512	2.372(172)	2.568(36)	2.373(710)	2.654(180)
$\frac{1}{16}$	0.798	0.958	0.971 (49)	0.973(11)	1.023(127)	0.998 (35)
$\frac{1}{8}$	0.338	0.405	0.443 (17)	0.421 (6)	0.356 (37)	0.402 (10)
$\frac{1}{4}$	0.190	0.228	0.257 (18)	0.239 (3)	0.263 (19)	0.232 (4)
$\frac{1}{2}$	0.105	0.126	0.131 (7)	0.129 (2)	0.125 (8)	0.131 (2)
1	0.082	0.098	0.095 (18)	0.097 (4)	0.067 (19)	0.098 (4)

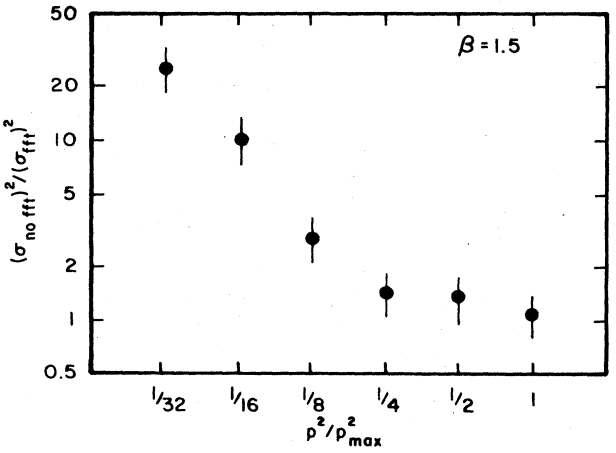

FIG. 5. Improvement due to Fourier acceleration in measurements of the momentum Green's functions for the XY model at $\beta = 1.5$ and $h = 0.23$. The squares of the standard deviation with and without acceleration are compared for different momenta.

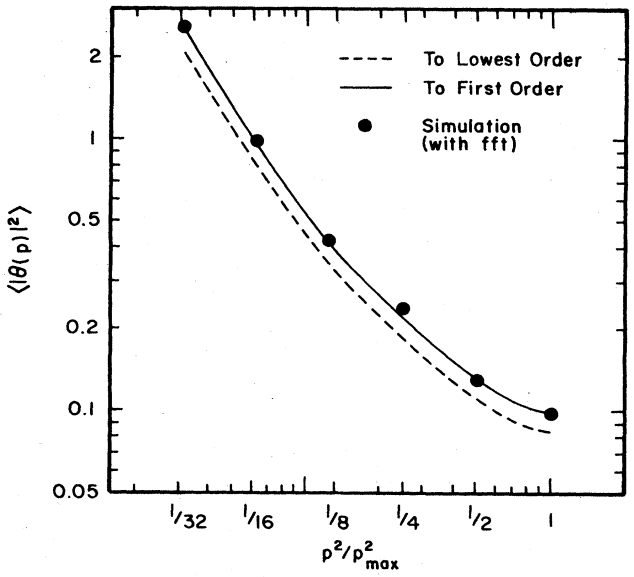

FIG. 7. The Green's function $\langle\,|\,\theta(p)\,|^2\rangle$ for the XY model with $\beta = 1.5$ and $h = 0.23$. Results from lowest and first order perturbation theory are compared with results from our simulation (with Fourier acceleration).

superior to those from 6000 sweeps without acceleration. In Fig. 5 we plot the square of the ratio of the standard deviations with and without FFT's versus momentum. The standard deviations were estimated by measuring fluctuations among 16 samples of 400 measurements each. At high momenta there is little difference, but the simulation at $p \sim 0$ was 20—25 times faster with Fourier acceleration. We also measured the correlation time N_c by binning our data in various sized bins and measuring correlations between successive bins. The results, in Fig. 6, indicate a correlation time of just a few sweeps of the lattice for *any* momentum with Fourier acceleration, while without FFT's the correlation time grows dramatically at low momenta. The values we obtained for the propagator agree well with perturbation theory when one-loop corrections are included—these shift β to 1.25 and h to 0.2. Our data for the propagator are shown in Fig. 7.

We have also simulated the XY model in its nonperturbative phase, running at $\beta = 1.0$ and 0.7. Again Fourier acceleration speeds the simulation by a factor of order ξ^2,

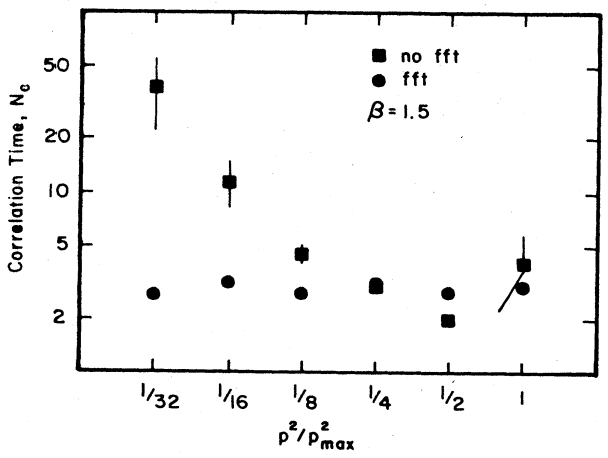

FIG. 6. The number of sweeps required for statistically independent measurements of the Green's function at different momenta in the XY model at $\beta = 1.5$ and $h = 0.23$. Results are from runs with Fourier acceleration (FFT) and without acceleration (no FFT).

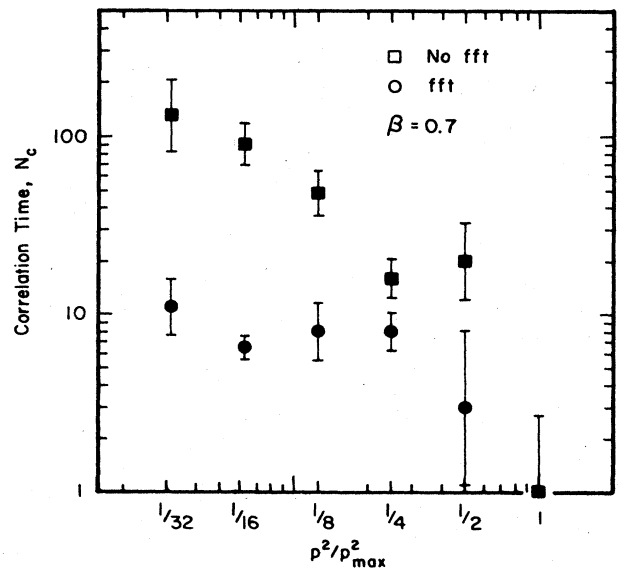

FIG. 8. The number of sweeps required for statistically independent measurements of the Green's function at different momenta in the XY model at $\beta = 0.7$ and $h = 0.11$. Results are from runs with Fourier acceleration (FFT) and without acceleration (no FFT).

even though many vortex-antivortex pairs are present. Lacking perturbation theory, one can determine the optimal p dependence of $\epsilon(p)$ from measurements of the correlation time for different momenta (cf. Fig. 6); $\epsilon(p)$ should be roughly proportional to $N_c(p)$. We show the correlation times, with and without Fourier acceleration, for $\beta = 0.7$ in Fig. 8. The improvement due to Fourier acceleration is substantial, though smaller than at $\beta = 1.5$ since the correlation length is shorter. Clearly, Fourier acceleration is effective whenever different wavelengths evolve at very different speeds, whether or not the theory is in a perturbative domain.

IV. LANGEVIN SIMULATIONS OF FERMIONIC THEORIES

A. Fermion loops

One of the major hurdles in simulating QCD has been the inclusion of effects due to quark vacuum polarization. In principle, quark loops can be included by adding a term $-\mathrm{Tr}\ln M[U]$ for each flavor to the gauge boson action $S_g[U]$. Here, $M = \gamma_5(D\cdot\gamma + m)$ is the inverse propagator of the quark, together with an additional factor of γ_5 which makes M Hermitian without affecting the fermion determinant. Thus for a single flavor, the full action is

$$S = S_g[U] - \mathrm{Tr}\ln M[U]. \tag{4.1}$$

Remarkably, such an action is readily simulated using the Langevin equation. The fermion term can be generated simply by adding a new bilinear noise term:

$$U(\tau_{n+1}) = e^{-if\cdot T}U(\tau_n), \tag{4.2a}$$

$$f_j = \bar{\epsilon}\left[\partial_j S_g - \tfrac{1}{2}\mathrm{Re}\left[\eta_q^\dagger \frac{1}{M}\partial_j M \eta_q\right]\right] + \sqrt{\bar{\epsilon}}\eta_j, \tag{4.2b}$$

where η_q and η_j are Gaussian random numbers (with $\langle|\eta|^2\rangle = 2$). That this equation yields fields distributed according to the action (4.1) is evident from the Fokker-Planck equation. To $O(\epsilon)$ in the Fokker-Planck equation, f_j is always averaged over the bilinear noise η_q:

$$\langle f_j\rangle = \bar{\epsilon}\left[\partial_j S_g - \mathrm{Tr}\left[\frac{1}{M}\partial_j M\right]\right] + \sqrt{\bar{\epsilon}}\eta_j$$

$$= \bar{\epsilon}\partial_j(S_g - \mathrm{Tr}\ln M) + \sqrt{\bar{\epsilon}}\eta_j.$$

As this is just the f_j one would use for the fermionic action (4.1), the fermion contribution is properly included. Note that the derivative of $\mathrm{Tr}\ln M$ has no imaginary part, and so we can drop the Re from (4.2b) when averaging. The only significant cost for including the fermions is that $\psi = M^{-1}\eta_q$ must be determined once for each update by solving the linear equations

$$M[U]\psi = \eta_q; \tag{4.3}$$

in terms of ψ, the driving function f_j is local:

$$f_j = \bar{\epsilon}[\partial_j S_g - \mathrm{Re}(\psi^\dagger \partial_j M \eta_q)] + \sqrt{\bar{\epsilon}}\eta_j. \tag{4.2c}$$

The generalization for nonlocal $\epsilon_{x,y}$ is obvious.

This procedure has several attractive features. Most importantly, it does not involve a Monte Carlo calculation within a Monte Carlo calculation, as do most pseudofermion methods. Also, any number of flavors, odd or even, may be simulated by including a bilinear noise term for each quark, or if the masses are the same, by multiplying the existing term by n_f, the number of flavors.[12] This follows because our pseudofermions enter only as noise, not as dynamical fields. Finally, the analysis developed in the preceding sections provides tight control over errors.

The $O(\epsilon)$ corrections here are similar to those for local theories (Sec. II). In deriving the Fokker-Planck equation one averages over both η_q and η_j. The asymptotic probability density then satisfies

$$\partial_i(\partial_i \bar{P} + \bar{S}_i P) = 0, \tag{4.4a}$$

where

$$\bar{S}_i = \partial_i \bar{S} + \frac{\bar{\epsilon}}{8}(\partial_j - \partial_j S)\,\mathrm{Tr}\left[\frac{1}{M^2}\partial_i M^2 \frac{1}{M^2}\partial_j M^2\right] + \cdots \tag{4.4b}$$

and $\bar{S}$ is given by Eq. (2.19) with S as in Eq. (4.1). For a general $\epsilon_{x,y}$ the new terms in $\bar{S}$ resemble corrections one might expect from a new gauge field with propagator $\epsilon_{x,y}$. For example, the term

$$\sum_{x,y}\mathrm{Tr}\left[\frac{1}{M}\partial_{i,x}M\frac{1}{M}\partial_{i,y}M\right]\epsilon_{x,y}$$

is just a self-energy correction to the quark propagator (in a background U field). Thus, as before, these terms introduce no new problems.

The second term in Eq. (4.4b) is not an exact differential, and therefore cannot be absorbed easily into a redefinition of $\bar{S}$. One might worry that this term introduces unwanted infrared sensitivity due to the rather singular factors of $1/M^2$ (as opposed to $1/M$) that it contains. However, we have examined contributions from this term to various n-point Green's functions in perturbation theory, and we find that almost all such contributions cancel. (This follows from the equations of motion: $\langle(\partial_j - \partial_j S)f[U]\rangle = 0$ for any functional f of the fields.) The few terms that remain seem no more singular than contributions from ordinary perturbation theory, and are therefore suppressed by $\bar{\epsilon}$. Consequently, the performance of the QCD algorithms discussed in Sec. II is unaffected by the inclusion of fermion loops.

We have tested algorithm (4.2) on a 2^4 lattice for QCD with Wilson fermions, for which

$$\gamma_5 M\psi = \psi(n) - \kappa\sum_\mu[(1-\gamma_\mu)U_\mu(n)\psi(n+\hat{\mu}) + (1+\gamma_\mu)U_\mu^\dagger(n-\hat{\mu})\psi(n-\hat{\mu})]. \tag{4.5}$$

We tried several values of β, κ, and n_f. Our results for $\beta=4$ and $\kappa=0.15$ with two flavors are displayed in Table II, together with recent results by Weingarten.[13] While our algorithm requires roughly the same number of sweeps over the lattice as Weingarten's Metropolis algorithm, it is much faster than his because the matrix equation (4.3) is solved only once per sweep, not once per link-update. Our algorithm is also far superior to the pseudofermion method he analyzes. Note that we use skew boundary conditions for three of the four dimensions and consequently our results should differ slightly from Weingarten's. On large lattices our algorithm's performance can be further enhanced by Fourier acceleration.

We have also investigated a number of related algorithms for including fermion loops. One is to simulate a theory with a new auxiliary scalar field ϕ,

$$S[U,\phi] = S_g[U] + \phi^\dagger \frac{1}{M^2}\phi, \tag{4.6}$$

that introduces fermion loops for two flavors. An update of the gauge fields for this theory is identical to Eqs. (4.2), but with the bilinear noise term replaced by $\psi^\dagger(\partial_j M^2)\psi$, where $M^2\psi=\phi$. The scalar field is updated using the Langevin equation[14]

$$\Delta\phi = -\bar{\epsilon}_q\phi - (\bar{\epsilon}_q)^{1/2} M\,\eta_q. \tag{4.7}$$

This procedure is comparable in speed to the one discussed above (in fact, they become identical when $\bar{\epsilon}_q \to 1$), but it is limited to even numbers of flavors since M is not positive definite.

B. Fermion propagators

Another important consideration relating to fermions is the speed with which linear equations like Eq. (4.3) can be solved. Such computations are critical both for including fermion loops and for studying fermionic observables.

Most iterative matrix methods work better when the matrix is positive definite, and so we consider the equation

$$M^2\psi = \phi \tag{4.8}$$

TABLE II. Data from three different algorithms for simulating QCD: Langevin [(Eq. 4.2)], Metropolis, and pseudofermion (Ref. 13). Runs were done on a 2^4 lattice, $\beta=4$, $\kappa=0.15$.

	Number of sweeps	Tr[line]	Tr[plaquette]
Eqs. (4.2)			
$\epsilon = 0.05$	6000	−0.15(2)	0.29(1)
$\epsilon = 0.01$	12000	−0.15(2)	0.31(1)
$\epsilon = 0.005$	24000	−0.15(2)	0.30(1)
Ref. 13			
Metropolis	2400	−0.16(2)	0.36(2)
pseudofermion	85000	−0.15(1)	0.31(3)

in place of (4.3). The speed with which these algorithms converge is intimately related to the spectrum of M^2. This is illustrated by the simplest of iterative procedures, the Jacobi algorithm, where the $(n+1)$st estimate is obtained from

$$\psi^{(n+1)} = (1-\delta M^2)\psi^{(n)} + \delta\phi. \tag{4.9}$$

The parameter δ controls the speed and stability of the algorithm. This equation is easily solved formally:

$$\psi^{(n)} = \delta \sum_{j=0}^{n} (1-\delta M^2)^{j-1}\phi. \tag{4.10}$$

Note the similarity of these equations to the stochastic equations used for updating scalar fields [Eqs. (3.2) and (3.6)]. As in that case, this algorithm converges provided δ is chosen small enough,

$$\delta \lesssim \frac{1}{|M^2|_{\max}} \sim \frac{1}{p^2_{\max}+m^2}, \tag{4.11}$$

while the number of iterations required to obtain a fixed accuracy is [cf. Eq. (3.7)]

$$N_{\text{inv}} \sim \frac{1}{\delta\,|M^2|_{\min}} \sim \xi^2. \tag{4.12}$$

Again one must wait for the long wavelength components of ψ to develop; the short wavelength components converge after only a few iterations.

The solution to this problem is identical to that for the updating: i.e., we solve the equations in momentum space, making the step size δ momentum dependent. For the Jacobi method, this means replacing Eq. (4.9) by

$$\Delta\psi = \hat{F}\,\delta(p)\,\hat{F}^{-1}(-M^2\psi^{(n)}+\phi), \tag{4.13}$$

where again the optimal $\delta(p)$ is something like

$$\delta(p) = \frac{\bar{\delta}}{p^2+m^2}. \tag{4.14}$$

With this sort of Fourier acceleration, requiring only two FFT's per iteration, all wavelengths evolve at roughly the same rate. Then, only $N_{\text{inv}}\sim 1$ iterations are required for a complete solution ψ.

Fourier acceleration can be used with any matrix algorithm that, like the Jacobi algorithm, updates the whole field ψ simultaneously. In particular, it is compatible with the conjugate gradient algorithm.[15] On the other hand, it is unlikely that this acceleration technique can be used with methods like the Gauss-Seidel algorithm that update one site at a time.

When using this method for gauge theories, it is again essential to fix the gauge so that high momentum does in fact correspond to short distances (and not just a crinkly gauge function). Gauges like the $A^0=0$ gauge discussed in Sec. II B are suitable. However, if for some reason one is using Fourier acceleration only in matrix inversion and not for updating the gauge fields, other gauges can be employed. For example, one can apply the gauge transformation

$$U_\mu(n) \to G(n)\,U_\mu(n)\,G(n+\hat{\mu}),$$
$$G(n) = \frac{1}{\alpha}\sum_\nu \left[U_\nu(n-\hat{\nu}) - U_\nu(n) - \frac{1}{n}\mathrm{Tr}[U_\nu(n-\hat{\nu}) - U_\nu(n)]\right] \tag{4.15}$$

before each update. In an Abelian gauge theory, this is equivalent to adding a gauge fixing term $-(\partial\cdot A)^2/\bar{\alpha}$ to the continuum action. The effect is similar for non-Abelian fields, except that these gauges bear no simple relation to common perturbative gauges (witness the absence of ghost fields here).

We tested Fourier acceleration of the conjugate gradient algorithm for QCD on a small lattice (4^4 sites). Using Wilson fermions with $n_f=2$ and $\beta=6$ we solved Eq. (4.3) for ψ as part of the updating algorithm for the gauge fields. [Actually, we multiply Eq. (4.3) by M on each side since the conjugate gradient method requires a positive matrix.] Near the critical κ, $\kappa_c \sim 0.158$, Fourier acceleration reduced the number of sweeps needed by a factor of $\frac{1}{3}$, roughly as expected for such small lattices. The algorithm functioned well for all κ's, both above and below κ_c. We also examined the residual vector, $M^2\psi - \phi$, as a function of momentum. Again as expected, the error was distributed uniformly in momentum with Fourier acceleration, while it was dominated by low momenta without FFT's. We also found that Fourier acceleration had no effect without gauge fixing: there was then no correlation between momentum and error size. We are currently testing Fourier acceleration on larger lattices where the gain should be substantially larger, particularly with small quark masses.

V. SUMMARY

In this paper we have introduced a variety of new algorithms designed to speed the simulation of field theories like QCD. Fourier acceleration, both for updating and for matrix inversion, promises to reduce or even to eliminate critical slowing down. Introducing fermions as a Langevin noise not only avoids the need for a Monte Carlo calculation within a Monte Carlo calculation, but allows simulations with any number of flavors (even, odd, fractional, negative,...). Runge-Kutta versions of the discrete Langevin equation allow much larger steps to be taken for each update. In addition, the Langevin formulation opens up new possibilities such as the use of complex actions, or the direct calculation of connected Green's functions. These may be very important in dealing with both systematic and statistical errors.

The tight control we have over errors is one of the most attractive features of these algorithms. The nature of the errors is readily apparent from the equilibrium action, and their significance is easily estimated, both analytically and numerically. This remains true even with Fourier acceleration and fermion loops.

Our theoretical analysis is amply supported by a variety of numerical experiments, on small lattices or in low dimensions. Needless to say, however, the extent to which our techniques affect QCD simulations will become apparent only after high-statistics tests on large lattices. Results from such tests will be forthcoming.

ACKNOWLEDGMENTS

We are indebted to many of our colleagues for their comments and suggestions throughout the course of this work. In particular, we want to thank Ian Drummond, Ron Horgan, Giorgio Parisi, Pietro Rossi, and Don Weingarten. G.P.L. thanks the Department of Applied Mathematics and Theoretical Physics at Cambridge for their hospitality during the early stages of this work. His work is also partially supported by the Alfred P. Sloan Foundation. B.S. thanks the I.B.M. Corporation for support.

---

*Present address: Center for Theoretical Physics, Laboratory for Nuclear Science, Massachusetts Institute of Technology, Cambridge, MA 02139.

[1]G. Parisi and Wu Yongshi, Sci. Sin. **24**, 483 (1981).

[2]Similar ideas are discussed by G. Parisi, in *Progress in Gauge Field Theory*, edited by G. 't Hooft *et al.* (Plenum, New York, 1984).

[3]This is also true for the Langevin equation in the continuous time limit Eq. (1.2). See, for example, P. H. Damgaard and K. Tsokos, Nucl. Phys. **B235**, 75 (1984).

[4]This result is obvious to lowest order in ϵ. One can show that it still applies in first order by using the lowest-order solution to simplify the right-hand side of Eq. (2.2a). Note that the discrete-time Fokker-Planck equation then takes the form $\Delta P = -Q^\dagger Q P$, where the operator Q has a zero eigenvalue. The analysis of this equation for $\tau \to \infty$ is quite similar to that used in the continuous-time Langevin formalism.

[5]M. B. Halpern, Nucl. Phys. **B228**, 173 (1983); I. T. Drummond, S. Duane, and R. R. Horgan, *ibid.* **B220** [FS8], 119 (1983); A. Guha and S. C. Lee, Phys. Rev. D **27**, 2412 (1983); H. Hamber and U. Heller, Phys. Rev. D **29**, 928 (1984).

[6]I. Bars, Nucl. Phys. **B148**, 445 (1979); **B149**, 39 (1979); G. G. Batrouni and M. B. Halpern, Phys. Rev. D **30**, 1782 (1984).

[7]This is easy to demonstrate using the Fokker-Planck equation, (2.2).

[8]E. Helfand, Bell Syst. Tech. J. **58**, 2289 (1979); H. S. Greenside and E. Helfand, *ibid.* **60**, 1927 (1981).

[9]We are being somewhat schematic here. On a lattice, p^2+m^2 is replaced by the inverse of a lattice propagator.

[10]Insofar as a nonlocal $\epsilon_{x,y}$ mimics a new field in the effective action, the optimal m may be somewhat different from $1/\xi$ due to renormalization.

[11]J. Tobochnik and G. V. Chester, Phys. Rev. B **20**, 3761 (1979).

[12]In fact, n_f could also be fractional, which might be useful in

studying the phase structure of QCD.

[13]D. Weingarten, IBM report, 1985 (unpublished).

[14]This is the standard Langevin equation for action (4.6) but with $\epsilon_{xy}=\bar{\epsilon}M^2$. This choice minimizes errors due to finite step size while at the same time avoiding all critical slowing down in the ϕ fields. This algorithm (but with $\epsilon_{xy}=\epsilon$ rather than $\bar{\epsilon}M^2$) has also been suggested by A. Ukawa and M. Fukugita, KEK report (1985), and by O. Martin, S. Otto, and J. Flower, Caltech report (1985). The Metropolis version of this algorithm is discussed in D. Weingarten and D. Petcher, Phys. Lett. **99B**, 333 (1981).

[15]M. R. Hestenes and E. Stiefel, J. Res. Natl. Bur. Stand. **49**, 409 (1952); L. A. Hageman and D. M. Young, *Applied Iterative Methods* (Academic, New York, 1981).