# Fourier acceleration in lattice gauge theories. III. Updating field configurations

C. T. H. Davies,* G. G. Batrouni,† G. R. Katz,‡ A. S. Kronfeld,§ G. P. Lepage, P. Rossi,** B. Svetitsky,†† and K. G. Wilson‡‡

*Newman Laboratory of Nuclear Studies, Cornell University, Ithaca, New York 14853*

(Received 26 September 1989)

We describe how the technique of Fourier acceleration can be used to alleviate the problem of critical slowing down when updating gauge field configurations in numerical simulations of a lattice gauge theory. Problems arise in the non-Abelian theory because of gauge freedom. To solve these requires both gauge fixing and the addition of a correction term to the updating algorithm. Some numerical results are presented for SU(3) pure gauge theory on small lattices which indicate that large-scale calculations in full QCD could be considerably faster using Fourier acceleration.

## I. INTRODUCTION

A major problem confronting attempts to do realistic numerical simulations of QCD using lattice techniques is that of critical slowing down. Quantities must be calculated as averages over sets of field configurations and the value obtained will only be reliable if sufficient uncorrelated measurements are made. This becomes more and more difficult as we approach the continuum limit since this is a critical point at which correlation lengths on the lattice diverge and, along with them, the autocorrelation time (on the computer) of measurements of physical quantities.

To take advantage of importance sampling of the Feynman path integral, field configurations on the lattice must be generated with the probability distribution $e^{-S}$, where $S$ is the action. Conventional algorithms produce these configurations one after another in computer time by making local changes to the field. The changes tend to reduce the action but have some noise associated with them to give quantum fluctuations. When the action only links fields at nearby points, the effect of changes to the configuration at one point diffuses across the lattice with the number of iterations of the updating algorithm. Measurements made on successive configurations will tend to be correlated if the quantity measured is sensitive to long-distance fluctuations in the field. Unfortunately, it is precisely these fluctuations which contain the physics of interest in the continuum limit. If we decrease the lattice spacing $a$ (for a fixed physical volume), physical lengths will grow in lattice units (see Fig. 1). This means that more iterations of the updating algorithm will be required to obtain uncorrelated measurements of physical quantities. Since the computational cost of a single iteration also grows at best as the number of lattice sites $V$, taking the continuum limit will prove prohibitively expensive.

Many other iterative algorithms suffer similar problems. They occur when the change of the system under one iteration is governed by a matrix which becomes ill conditioned. The eigenvectors of this matrix evolve separately under the algorithm at a rate which is controlled by their eigenvalues. The eigenvectors with a small eigenvalue lag behind the others and control the effective rate of change. In physical systems these eigenvectors are often, at least approximately, the Fourier modes, and it is possible to identify which momenta are evolving slowly and holding up the algorithm. If the step size of these modes can be increased in a way which does not affect the final result, then critical slowing down can be overcome. This numerical technique, known as Fourier acceleration,[1,2] has been very successful in speeding up algorithms for the simulation of spin models[1,3,4] as well as in a number of other problems requiring numerical optimization.[5–8]

In this paper we describe its application to lattice gauge theories and in particular to the case of the gauge group SU(3), relevant to QCD. We begin in Sec. II by describing how the acceleration method works when the updating algorithm is based on a Langevin equation.[1] The method is readily adapted for use with updating algorithms based upon molecular dynamics or hybrids of molecular dynamics and the Langevin equation. We discuss in Sec. III results for the U(1) group, where the method is particularly simple. In Sec. IV we show how the extra problems associated with the non-Abelian nature of SU(3), particularly the requirement of gauge fixing, can be overcome. Finally, in Sec. V we present some preliminary results which give encouraging indica-

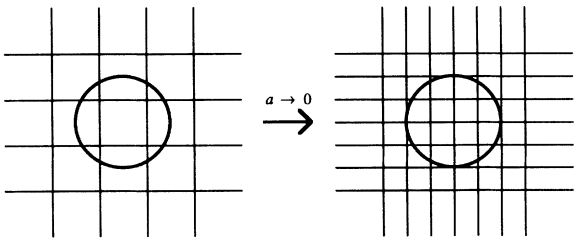


FIG. 1. Physical lengths grow in lattice units as the lattice spacing decreases.

 

tions of the effectiveness of Fourier acceleration for future simulations of QCD with small lattice spacing. Our results are based on a simulation of the pure gauge theory but we discuss their relevance to a full calculation including dynamical quarks in the conclusions.

## II. THE LANGEVIN EQUATION AND FOURIER ACCELERATION

We use Langevin updating[9,1] to generate a set of gauge fields on the links of the lattice with probability distribution $e^{-S}$. The entire lattice is updated simultaneously so that a Fourier transform of the updating step is possible and Fourier acceleration can be applied.

To review critical slowing down and Fourier acceleration it is easiest to focus on the scalar theory. Each configuration is generated from the previous one using a discrete Langevin equation, which in this case is

$$\phi^{(n+1)}(x)=\phi^{(n)}(x)-f_x[\phi^{(n)},\eta^{(n)}] , \tag{2.1}$$

where

$$f_x=\epsilon\frac{\delta S}{\delta\phi^{(n)}(x)}-\sqrt{\epsilon}\,\eta^{(n)}(x) . \tag{2.2}$$

$\epsilon$ is the Langevin time step and $\eta$ is a noise term drawn independently from a Gaussian distribution for each $x$ and $n$:

$$\langle\eta^{(n)}(x)\eta^{(m)}(y)\rangle_\eta=2\delta_{nm}\delta(x-y) . \tag{2.3}$$

For a free scalar theory with the usual local quadratic action Eq. (2.1) can be written in momentum space as

$$\tilde{\phi}^{(n+1)}(p)=\tilde{\phi}^{(n)}(p)-\epsilon(p^2+m^2)\tilde{\phi}^{(n)}(p)+\sqrt{\epsilon}\,\tilde{\eta}^{(n)}(p) , \tag{2.4}$$

where $p^2$ are the eigenvalues of $-\partial^2$. It is quite clear now that, as a result of updating the fields locally, the Fourier modes with small values of $p^2$ have a small effective time step and will hold up the algorithm.

We can put this argument on a more quantitative footing by solving Eq. (2.4):

$$\tilde{\phi}^{(n)}(p)=\sqrt{\epsilon}\sum_{j=1}^{n}[1-\epsilon(p^2+m^2)]^{n-j}\tilde{\eta}^{(j-1)}(p) \quad \text{with } \tilde{\phi}^{(0)}(p)=0 . \tag{2.5}$$

The autocorrelation function of the fields in Langevin time steps is given by

$$\rho_\phi(\tau)=\frac{\langle\tilde{\phi}^{(n)}(p)\tilde{\phi}^{(n+\tau)}(p)\rangle_n-\langle\tilde{\phi}^{(n)}(p)\rangle_n^2}{\langle[\tilde{\phi}^{(n)}(p)]^2\rangle_n-\langle\tilde{\phi}^{(n)}(p)\rangle_n^2} . \tag{2.6}$$

Assuming ergodicity, the averages over $n$ can be replaced by averages over $\eta$ at each time. Then, using Eq. (2.5),

$$\rho_\phi(\tau)=[1-\epsilon(p^2+m^2)]^\tau . \tag{2.7}$$

As $\epsilon\to 0$ and $\tau\to\infty$ with $\epsilon\tau$ fixed we obtain

$$\rho_\phi(\tau)\sim\exp[-\epsilon\tau(p^2+m^2)] . \tag{2.8}$$

The autocorrelation function of products of the fields can be calculated in a similar fashion. For example, the autocorrelation function of the propagator decays as $\exp[-2\epsilon\tau(p^2+m^2)]$.

Equation (2.7) shows that, for stability of the algorithm,

$$\epsilon<\frac{1}{p_{\max}^2+m^2} , \tag{2.9}$$

i.e., that the maximum time step is dictated by the highest-momentum modes. The lowest-momentum modes then have an autocorrelation time, from Eq. (2.8),

$$\tau_c(p_{\min})=\frac{p_{\max}^2+m^2}{p_{\min}^2+m^2}\tau_c(p_{\max}) . \tag{2.10}$$

When the field $\phi$ is defined on a finite lattice of space-time points the lattice spacing $a$ appears in combination with dimensionful quantities in Eqs. (2.4)–(2.10). The operator $-\partial^2a^2$ has a bounded discrete spectrum. We use skew-periodic boundary conditions for which

$$p^2a^2=4\sum_{\mu=1}^{d}\sin^2\left[\frac{l\pi}{L^\mu}\right], \quad l=0,\ldots,L^d-1 \tag{2.11}$$

in $d$ dimensions, but our conclusions are independent of this choice. Clearly $p_{\max}^2a^2$ is roughly constant on large lattices so if we make the lattice spacing smaller keeping the physical volume of the lattice $(La)^d$ fixed then

$$\tau_c(p_{\min})\propto\frac{1}{a^2} . \tag{2.12}$$

This is critical slowing down.

An alternative, but entirely equivalent point of view is that obtained from the Fokker-Planck equation.[9,10] This equation describes the approach of the probability distribution $P[U]$ to the equilibrium $e^{-S}$ after many Langevin time steps. To leading order in $\epsilon$ the change of the probability distribution after one time step is given by

$$\frac{\Delta P}{\epsilon}=\sum_x\frac{\delta}{\delta\phi(x)}\left[\frac{\delta S}{\delta\phi(x)}P+\frac{\delta P}{\delta\phi(x)}\right] . \tag{2.13}$$

By writing $P=\exp(-S/2)\Psi$ this can be transformed to

$$\frac{\Delta\Psi}{\epsilon}=-H\Psi , \tag{2.14}$$

where the eigenvalues of the Fokker-Planck Hamiltonian $H$ control the convergence of the probability distribution. $\Psi$ relaxes eventually to the ground state $[\exp(-S/2)]$ but this will take a large number of steps if the first excited state has a small eigenvalue. For the free scalar theory on a lattice described above the eigenvalues of $H$ are in an harmonic-oscillator-like sequence with separation $p^2a^2+m^2a^2$. The time constant for thermalization of $P$ is controlled by the size of $(p_{\min}^2a^2+m^2a^2)^{-1}$ and will again diverge as $a\to 0$. So, as expected, this thermalization time is related to the decorrelation time for the fields since they both reflect the rate at which configurations evolve.

As stressed in the Introduction, critical slowing down for this theory will occur in any updating algorithm

which changes the fields locally. The arguments above are not specific to the Langevin formalism—it simply reveals the mechanism at work very clearly and suggests a method to rectify the problem.

Fourier acceleration works by allowing the slow modes at low momentum to take large steps. We introduce a momentum-dependent time step where the step size is chosen to be, for example, inversely proportional to $p^2a^2+m^2a^2$. In the case of the free scalar field above, all Fourier modes would now evolve at the same rate and critical slowing down would be cured. The updating algorithm becomes

$$\phi^{(n+1)}(x)=\phi^{(n)}(x)-\hat{F}^{-1}\left[\epsilon(p)\hat{F}\frac{\delta S}{\delta\phi^{(n)}(x)}-\sqrt{\epsilon(p)}\,\hat{F}\eta^{(n)}(x)\right], \tag{2.15}$$

where $\hat{F}$ represents a Fourier transform and

$$\epsilon(p)=\frac{\epsilon(p_{\max}^2a^2+m^2a^2)}{p^2a^2+m^2a^2}\ . \tag{2.16}$$

For more complicated theories with interactions the optimal form of $\epsilon(p)$ may require modification from the one based on the free field propagator.

The accelerated algorithm is equivalent to introducing a time step which is nonlocal in position space, $\epsilon_{xy}$:

$$\epsilon_{xy}=\sum_p e^{ip(x-y)}\epsilon(p)\ . \tag{2.17}$$

A study of the Fokker-Planck equation readily shows that the algorithm yields an equilibrium probability distribution for the fields which, as before, is $e^{-S}$ to leading order in $\epsilon$. The cost of one iteration of the algorithm will have a piece proportional to $V\ln V$ from a fast Fourier transform (FFT) and an additional part proportional to $V$ from the rest of the calculation [provided $\epsilon(p)$ is diagonal in momentum space]. Which dominates at a given value of $V$ will depend to some extent on the computer used. The number of iterations required for a given accuracy is now independent of $V$ if critical slowing down is cured so either dependence is to be preferred to the $VL^2$ growth in the unaccelerated algorithm as the continuum limit is taken as fixed physical volume.

This algorithm was applied successfully to the two-dimensional $XY$ model in Ref. 1 and subsequently to various field theories in Refs. 3 and 4.

## III. SIMULATION OF ABELIAN GAUGE THEORIES

The application of Fourier acceleration to gauge theories is complicated by the additional gauge symmetry which obscures the relationship between momentum and wavelength. We might wonder how to identify and accelerate the long-distance physical modes which cause a critical slowing down of gauge-invariant quantities. For the Abelian theory no such separation of the field components turns out to be necessary[11] because the unphysical modes are decoupled from the physical ones. We simply have to work out what momentum-dependent time step would suit the physical modes and apply this to the whole field.

The free Abelian theory has the Langevin equation

$$A_\mu^{(n+1)}(x)=A_\mu^{(n)}(x)+\epsilon[\partial^2A_\mu^{(n)}(x)-\partial_\mu\partial_\nu A_\nu^{(n)}(x)]+\sqrt{\epsilon}\,\eta_\mu^{(n)}(x) \tag{3.1}$$

with

$$\langle\eta_\mu^{(n)}(x)\eta_\nu^{(m)}(x')\rangle=2\delta_{\mu\nu}\delta_{mn}\delta^4(x-x')\ ,$$

which, in momentum space, becomes

$$\tilde{A}_\mu^{(n+1)}(p)=\tilde{A}_\mu^{(m)}(p)-\epsilon(p^2\delta_{\mu\nu}-p_\mu p_\nu)\tilde{A}_\mu^{(n)}(p)+\sqrt{\epsilon}\,\tilde{\eta}_\mu^{(n)}(p)\ . \tag{3.2}$$

The appearance of the singular matrix $p^2\delta_{\mu\nu}-p_\mu p_\nu$ is evidence of the gauge symmetry. However, the force term in the Langevin equation is invariant under a gauge transformation

$$A_\mu\rightarrow A_\mu+\partial_\mu\phi\ . \tag{3.3}$$

This means that the physical and unphysical modes do not interfere with each other on updating. This is true even if $\epsilon$ becomes nonlocal as $\epsilon(p)$.

To see what form $\epsilon(p)$ should take it is simplest to split $A_\mu$ into $A_\mu^T$ and $A_\mu^L$ where[9]

$$\tilde{A}_\mu^T(p)=\left[\delta_{\mu\nu}-\frac{p_\mu p_\nu}{p^2}\right]\tilde{A}_\nu(p)\ , \tag{3.4}$$

$$\tilde{A}_\mu^L(p)=\frac{p_\mu p_\nu}{p^2}\tilde{A}_\nu(p)\ . \tag{3.5}$$

Then

$$\tilde{A}_\mu^{T(n+1)}(p)=\tilde{A}_\mu^{T(n)}(p)-\epsilon p^2\tilde{A}_\mu^{T(n)}(p)+\sqrt{\epsilon}\,\tilde{\eta}_\mu^{T(n)}(p) \tag{3.6}$$

with

$$\langle\tilde{\eta}_\mu^{T(n)}(p)\tilde{\eta}_\nu^{T(m)}(p')\rangle=2\left[\delta_{\mu\nu}-\frac{p_\mu p_\nu}{p^2}\right]\delta^4(p+p')\delta_{nm} \tag{3.7}$$

and

$$\tilde{A}_\mu^{L(n+1)}(p)=\tilde{A}_\mu^{L(n)}(p)+\sqrt{\epsilon}\,\tilde{\eta}_\mu^{L(n)}(p) \tag{3.8}$$

with

$$\langle\tilde{\eta}_\mu^{L(n)}(p)\tilde{\eta}_\nu^{L(m)}(p')\rangle=2\frac{p_\mu p_\nu}{p^2}\delta^4(p+p')\delta_{nm}\ . \tag{3.9}$$

Equation (3.8) shows that the unphysical longitudinal modes undergo a random walk in gauge space. The physical transverse modes on the other hand do yield convergent expectation values. By solving equations similar to Eqs. (2.4)–(2.8) it is clear that the decorrelation time of modes of momentum $p$ is given by

$$\tau_c(p)\propto\frac{1}{p^2}\ . \tag{3.10}$$

The gauge-invariant quantities that we measure will be related to $A_\mu^T(x)$, so to avoid critical slowing down for these observables we should take

$$\epsilon(p)=\frac{p_{\max}^2}{p^2}\epsilon \quad \text{for } p\neq 0 . \tag{3.11}$$

A more useful example to study numerically is a U(1) gauge theory on a lattice. We take the action to be

$$S=-\beta\sum_{\theta_{\text{plaq}}}\cos(\theta_{\text{plaq}}) , \tag{3.12}$$

where $\theta_{\text{plaq}}$ is the sum around a plaquette of angular variables which sit on the links of the lattice:

$$\theta_{\text{plaq}}=\theta_\mu(x)+\theta_\nu(x+\hat{\mu})-\theta_\mu(x+\hat{\nu})-\theta_\nu(x) \tag{3.13}$$

and the sum over $\theta_{\text{plaq}}$ includes all distinct plaquettes of one orientation. The gradient of the action in the Langevin equation for the $\theta$ field is then

$$\frac{\delta S}{\delta\theta_\mu(x)}=-\beta\sum_{\theta_{\text{plaq}}\supset\theta_\mu(x)}\sin(\theta_{\text{plaq}}) , \tag{3.14}$$

where the sum is now over all plaquettes of both orientations which start on the link $\theta_\mu(x)$. This is gauge invariant under

$$\theta_\mu(x)\to\theta_\mu^g(x)\equiv\theta_\mu(x)+\phi(x)-\phi(x+\hat{\mu}) , \tag{3.15}$$

the discretized version of Eq. (3.3). As discussed above, no gauge fixing is required to implement Fourier acceleration in this case, because mixing force terms from different sites in a nonlocal step will not affect the behavior of the force term under gauge transformations.

For large $\beta$ we expect this theory to behave in a similar way to the noncompact electrodynamics discussed above, with $A=\theta$. We can therefore test numerically the ideas of Fourier acceleration in a setting where correlation lengths can be calculated analytically.

As an observable whose correlation length we will measure we consider the quantity

$$\chi(p)=\sum_\mu\tilde{D}_\mu(p)\tilde{D}_\mu(-p) , \tag{3.16}$$

where $\tilde{D}_\mu(p)$ is the Fourier transform of the gradient of the action [Eq. (3.14)] divided by the coupling constant

$$D_\mu(x)=\sum_{\theta_{\text{plaq}}\supset\theta_\mu(x)}\sin(\theta_{\text{plaq}}) . \tag{3.17}$$

Note that $\chi(p)$ is gauge invariant and $\chi(0)=0$. In the large-$\beta$ limit $\tilde{D}_\mu(p)$ will change from one configuration to the next under the Langevin algorithm in a very similar way to that in which $\tilde{\theta}(p)$ changes. For the unaccelerated algorithm

$$D_\mu^{(n+1)}(x)-D_\mu^{(n)}(x)=-\beta\epsilon[\Delta_\mu\Delta_\nu D_\nu^{(n)}(x)-\Delta^2 D_\mu^{(n)}(x)]+\sqrt{\epsilon}[\Delta_\mu\Delta_\nu\eta_\nu^{(n)}(x)-\Delta^2\eta_\mu^{(n)}(x)] , \tag{3.18}$$

dropping factors of $\cos(\theta_{\text{plaq}})$ which tend to 1 in this limit. $\Delta$ represents the finite-difference operator on the lattice. The mean value of $\chi(p)$ measured over many iterations can then be calculated. To leading order in $\epsilon$ we have

$$\langle\chi(p)\rangle=\frac{3p^2a^2}{\beta} . \tag{3.19}$$

The autocorrelation function in Langevin time steps is given by

$$\rho_\chi(\tau)=\exp(-2\epsilon\tau\beta p^2a^2) \tag{3.20}$$

so that the decorrelation time in sweeps is

$$\tau_c(\chi(p))=\frac{1}{2\epsilon\beta p^2a^2} \tag{3.21}$$

dependent for fixed $p$ only on the product $\beta\epsilon$.

For the accelerated algorithm with $\epsilon(p)=\epsilon p_{\max}^2a^2/p^2a^2$ (the $p^2=0$ mode is left unaccelerated) we should find

$$\tau_c(\chi(p))=\frac{1}{2\epsilon\beta p_{\max}^2a^2} \tag{3.22}$$

independent of $p$.

Results on $2^4$ lattices at $\beta=1.5$, $\epsilon=0.02$ confirm this picture (Fig. 2). The range of nonzero momenta on such a small lattice is not large, varying from $p^2a^2=4.0$ to $p^2a^2=12.2$. Nevertheless it is sufficient to demonstrate the principles of critical slowing down and the efficacy of Fourier acceleration.

The decorrelation time of $\chi(p)$ was measured in various ways. With such a simple theory we were able to do 200 000 sweeps through the lattice, measuring $\chi(p)$ at every sweep, and to get an accurate picture of the autocorrelation function $\rho_\chi(\tau)$, at least out to distances $\tau$ of about 20 sweeps. In this region $\rho_\chi(\tau)$ as well approximated in all cases by an exponential $\exp(-\tau/\tau_c)$. Thus $\tau_c$

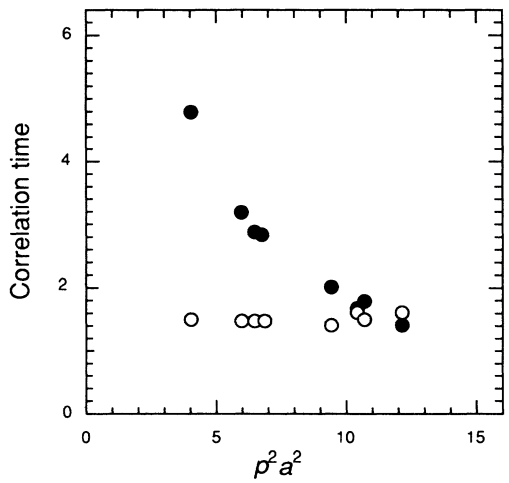


FIG. 2. Correlation times for $\chi(p)$ [Eq. (3.16)] plotted against $p^2a^2$ for a U(1) gauge theory on a $2^4$ lattice at $\beta=1.5$, $\epsilon=0.02$. The correlation time is defined from the slope of the logarithm of the autocorrelation function; ●—without Fourier acceleration; ○—with Fourier acceleration, $\epsilon(p)=\epsilon p_{\max}^2a^2/p^2a^2$. 196 608 sweeps were done without acceleration and 98 304 sweeps were done with acceleration, measuring $\chi(p)$ every sweep.

could be taken either from the slope of $\ln\rho$ vs $\tau$ or from the point where $\rho = e^{-1}$, interpolating in $\tau$.

We also attempted to extract $\tau_c$ from the integrated autocorrelation function, defining

$$\tau_c = \lim_{\tau' \to \infty} \int_0^{\tau'} \rho_\chi(\tau) d\tau \ . \tag{3.23}$$

This gave poor results with no very clear plateau as a function of $\tau'$, so that a value for $\tau_c$ compatible with those above was obtained with large errors. The problem is that $\rho_\chi(\tau)$ is extremely difficult to measure, especially where it is small. Even with 200 000 measurements, $\rho_\chi(\tau)$ was not smooth beyond a few times $\tau_c$ ($\approx 20$ sweeps) but developed oscillations of order the size of the signal. These oscillations lasted for up to 1000 sweeps so they had a large effect in the integrated autocorrelation function even though $\rho_\chi(\tau)$ was less than 0.01 at these values of $\tau$. To measure $\rho_\chi(\tau)$ accurately at time separation $\tau$ requires measurements over many times $\tau$ sweeps. If $\tau$ is so large that $\rho_\chi(\tau)$ is small this is neither feasible nor necessarily desirable. We concluded that the decorrelation time was more sensibly measured from the autocorrelation function itself, without integrating it. When $\rho(\tau)$ is not a good exponential, the more usual case, the simplest and perhaps more meaningful definition of $\tau_c$ is that point at which $\rho(\tau)$ has dropped to some acceptably small level.

For $\chi(p)$ we obtained decorrelation times with the unaccelerated algorithm which varied from 5 sweeps at low momentum to 1.5 sweeps at high momentum, in rough agreement with Eq. (3.21). With Fourier acceleration, all modes of $\chi(p)$ had correlation times of $\approx 1.5$ sweeps as expected. This behavior is shown in Fig. 2.

The improvement provided by Fourier acceleration for a given observable will depend on which momentum modes of the field it couples to most strongly. Measurements of the plaquette on the same lattice as above gave an acceleration factor of $\approx 1.5$ from an unaccelerated correlation length of 3.5 sweeps to an accelerated one of 2.3.

We see from the example above the Fourier acceleration would enable results to be obtained in up to three times fewer sweeps on the $2^4$ lattice (of course, a much larger factor would be expected on larger lattices). It remains to check that the mean values of observables thus obtained are the same as for the uncorrelated case. A complication here is that expectation values obtained with a Langevin algorithm depend on the step size $\epsilon$. This is because the equilibrium action differs from that which would be obtained from a Metropolis algorithm by terms which start at $O(\epsilon)$.

For the unaccelerated algorithm it is possible to show that the $O(\epsilon)$ corrections simply amount to a renormalization of the field and coupling constants.[1] We therefore expect them to be independent of the physical size of the lattice. This is important since if the step size had to decrease as the volume increased it would be disastrous for the use of the algorithm in the long term as we attempt to approach the continuum and (physical) infinite-volume limits. We performed a numerical test by measuring the plaquette very accurately at $\beta = 1.5$ on $2^4$, $4^4$, and $8^4$ lattices as a function of $\epsilon$. The slopes of the curves obtained were consistent with the hypothesis that the $O(\epsilon)$ corrections are the same for all the different lattice sizes (although the $\epsilon = 0$ value for the plaquette were clearly different). This makes it clear that the step size chosen for the simulation does not have to depend on the physical volume. This must also be checked for the accelerated algorithm since any volume dependence could wipe out the benefit to be gained from acceleration.

In fact the unaccelerated algorithm shows a very unbalanced distribution of $O(\epsilon)$ corrections. They are concentrated in the high-momentum modes since these are the ones which change rapidly from iteration to iteration and are sensitive to the step size. The effective step size in the unaccelerated algorithm is very small for the low-momentum modes and so the errors in these modes are unnecessarily small, given the much larger errors in the high-momentum modes. In the accelerated algorithm the corrections are more evenly distributed over the different modes so that those at low momentum now match those at high momentum. We might thus expect the $O(\epsilon)$ corrections to observables to be rather larger with Fourier acceleration than without but only by a factor of about 2 independently of the lattice size. This is borne out by numerical tests in which the plaquette was measured on various size lattices as in the previous paragraph but now with Fourier acceleration. When the value of the plaquette was plotted versus $\epsilon$, the slope of the curves increased with the number of lattice sites initially and then leveled off. It varied from 1.44 on the $2^4$ lattice to 2.12 on the $4^4$ lattice and 2.24 on the $8^4$ lattice. This compares with a slope of 1.08 in the unaccelerated case.

For the Fourier accelerated algorithm it is no longer possible to calculate the $\epsilon = 0$ value for observables such as Wilson loops by a simple rescaling. If this value is required it must be obtained by numerical extrapolation. For the plaquette calculation described above, the extrapolated values for the plaquette with and without Fourier acceleration agreed, as they must, on all the different volumes. The $\epsilon = 0$ value varied from 0.813 on $8^4$ to 0.825 on $2^4$.

The U(1) theory investigated above is not in a physically very interesting region. As discussed earlier, it is behaving essentially as a free theory because the length scale on which there are interactions between the modes is much larger than the size of the lattice. It is an interesting question how well Fourier acceleration will work on larger lattices at smaller values of $\beta$, in particular close to the phase transition at $\beta = 1.0$. Work in Ref. 4 indicates that Fourier acceleration will work close to a phase transition where the behavior is far from that of a free theory. We have tackled instead the problem of critical slowing down in non-Abelian gauge theories, relevant to QCD.

## IV. SIMULATION OF NON-ABELIAN GAUGE THEORIES

The ordinary Langevin equation for updating non-Abelian gauge fields in a numerical simulation takes the form

$$U^{(n+1)}=e^{-if\cdot T}U^{(n)}\ , \tag{4.1}$$

where the gauge fields $U$ are elements of the appropriate Lie group with generators $T_i$ (where $[T_i,T_j]=ic_{ijk}T_k$ and $\mathrm{tr}(T_iT_j)=\delta_{ij}/2$). The force term $f$ by analogy with Eq. (2.2) is given by

$$f_{j\mu x}=\epsilon\partial^{(n)}_{j\mu x}S-\sqrt{\epsilon}\,\eta^{(n)}_{j\mu x} \tag{4.2}$$

in the unaccelerated case. The derivative represents differentiation with respect to $U_\mu(x)$ within the group manifold [defined such that $f(e^{i\delta\cdot T}U)=f(U)+\delta^i\partial_i f+O(\delta^2)$]. The noise term $\eta$ obeys

$$\langle\eta^{(n)}_{i\mu x}\eta^{(m)}_{jvy}\rangle_\eta=2\delta_{ij}\delta_{\mu v}\delta_{xy}\delta_{nm}\ . \tag{4.3}$$

In the case of the standard action for $SU(N_C)$ gauge theories, that is,

$$S[U]=-\frac{\beta}{2N_c}\sum_{\text{plaq}}\mathrm{Tr}(U_{\text{plaq}}+U^\dagger_{\text{plaq}})\ , \tag{4.4}$$

where $U_{\text{plaq}}$ is the product of link variables $U_\mu$ forming a plaquette, the Langevin equation can be written in the form

$$\begin{aligned}&U^{(n+1)}_\mu(x)=e^{-F_\mu(x)}U^{(n)}_\mu(x)\ ,\\ &F_\mu(x)=\epsilon\frac{\beta}{4N_c}D_\mu(x)-\sqrt{\epsilon}\,H_\mu(x)\ .\end{aligned} \tag{4.5}$$

$D_\mu(x)$ is the field derivative of the action

$$D_\mu(x)=\sum_{U_{\text{plaq}}\supset U_\mu}\left[U_{\text{plaq}}-U^\dagger_{\text{plaq}}-\frac{1}{N_c}\mathrm{Tr}(U_{\text{plaq}}-U^\dagger_{\text{plaq}})\right] \tag{4.6}$$

and $H_\mu(x)$ is an anti-Hermitian noise matrix satisfying

$$\begin{aligned}&\langle H^{ab}_\mu(x)\rangle=0\ ,\\ &\langle H^{ab}_\mu(x)H^{cd}_v(y)\rangle=-\left[\delta_{ad}\delta_{bc}-\frac{1}{N_c}\delta_{ab}\delta_{cd}\right]\delta_{xy}\delta_{\mu v}\ .\end{aligned} \tag{4.7}$$

Using a local updating algorithm such as Eq. (4.5) we again expect to see critical slowing down on taking the continuum limit, and again we seek a cure using Fourier acceleration. As usual, an accelerated algorithm is obtained by replacing the step size $\epsilon$ by a nonlocal matrix $\epsilon_{xy}$ defined as the Fourier transform of an appropriate function $\epsilon(p)$, of momentum $p$. Here, however, the algorithm is complicated by the need to preserve local gauge invariance. The steps involved in proceeding from one gauge-field configuration to the next are as follows.

(a) Fix the gauge to a smooth gauge (e.g., Landau gauge). Gauge fixing is essential since Fourier acceleration uses momentum to resolve the different modes of the gauge field, and momentum, being gauge dependent for a charged particle, is meaningless in the absence of gauge fixing. The effect of gauge fixing is to replace the gauge field by another physically equivalent gauge field:

$$U_\mu(x)\to U^g_\mu(x)\equiv G_U(x)U_\mu(x)G^\dagger_U(x+\hat\mu)\ , \tag{4.8}$$

where the gauge transformation matrix is a function of position and a functional of the original gauge field configuration.

(b) Generate an anti-Hermitian noise matrix $H_\mu(x)$ as specified by Eq. (4.7).

(c) Compute $F^g_\mu(x)$ but with $\epsilon$ replaced by $\epsilon_{xy}$:

$$F^g_\mu(x)=\sum_y\left[\epsilon_{xy}\frac{\beta}{4N_C}D^g_\mu(y)-\sqrt{\epsilon_{xy}}H_\mu(y)\right]\ , \tag{4.9}$$

where $D^g_\mu(y)$ is calculated with the gauge-fixed configuration $U^g$. As usual, fast Fourier transforms must be used if this step is to be cost effective. To compute the first term in $F^g_\mu$ [Eq. (4.9)] we Fourier transform $D^g_\mu(x)$ into $p$ space, multiply each matrix $D^g_\mu(p)$ by $\epsilon(p)$, and finally transform the product $\epsilon(p)D^g_\mu(p)$ back to coordinate space. In forming the second term of $F^g_\mu$ we multiply $H_\mu(p)$ by $\sqrt{\epsilon(p)}$ rather than $\epsilon(p)$.

(d) A correction term

$$F^\delta_\mu(x)=\frac{1}{2\delta}\left[\sum_y\epsilon_{xy}H_\mu(y)-G^\dagger_{U^\delta}(x)\sum_y[\epsilon_{xy}G_{U^\delta}(y)H_\mu(y)G^\dagger_{U^\delta}(y)]G_{U^\delta}(x)\right] \tag{4.10}$$

must be added to $F^g_\mu$ to compensate for the field dependence implicit (through the gauge fixing) in the step-size matrix. The configuration $U^\delta$ is computed from $U^g$,

$$U^\delta_\mu(x)=e^{H_\mu(x)\delta}U^g_\mu(x)\ , \tag{4.11}$$

and $G_{U^\delta}$ is the gauge transformation that fixes $U^\delta$ to the same gauge as $U$ [see step (a)]. Parameter $\delta$ must be chosen sufficiently small that the finite step-size errors indicated by $F^\delta_\mu(x)$ are of order $\epsilon$, but not so small that $F^\delta_\mu(x)$ is significantly larger than $F^g_\mu(x)$. The choice $\delta=\sqrt{\epsilon}$ is optimal in this regard. Note that the number of Fourier transforms can be reduced by combining the evaluation of $F^g_\mu$ and $F^\delta_\mu$.

(e) Finally $\exp[-F^g_\mu(x)-F^\delta_\mu(x)]$ is calculated (for example, by expanding to an appropriate power), and applied to $U^g$ to obtain a new configuration.

Now let us describe why the algorithm takes this form. The additional complication in the non-Abelian gauge theory is that the force term is not gauge invariant. Instead, from Eq. (4.6), it transforms under a gauge transformation [Eq. (4.8)] as

$$F_\mu(x)\to G(x)F_\mu(x)G^\dagger(x)\ , \tag{4.12}$$

where $G(x)$ belongs to the gauge group and sits at the site $x$. The force term thus mixes the physical and unphysical modes of the gauge field on updating. (The Gaussian noise in $F$ is gauge transformed to Gaussian noise with the same distribution.)

When the updating is the usual local one in $x$ space the gauge transformations of the force term are just such as

to compensate the gauge transformations of the field. Gauge-equivalent configurations are then updated to gauge-equivalent configurations since

$$\begin{aligned} U_\mu^{g(n+1)}(x) &= e^{-F_\mu^g(x)} U_\mu^{g(n)}(x) \\ &= e^{-G(x)F_\mu(x)G^\dagger(x)} G(x) U_\mu^{(n)}(x) G^\dagger(x+\hat{\mu}) \\ &= G(x) e^{-F_\mu(x)} U_\mu^{(n)}(x) G^\dagger(x+\hat{\mu}) \\ &= G(x) U_\mu^{(n+1)}(x) G^\dagger(x+\hat{\mu}) \ . \end{aligned} \tag{4.13}$$

Gauge-invariant quantities will be the same step by step whether we start from a particular configuration or a gauge transform of it. Such a gauge-covariant algorithm ensures that no terms which break local gauge invariance, such as gluon mass terms, can appear in the equilibrium action.

This is not so easy to ensure when we introduce a momentum-dependent time step. If we simply replace $\epsilon$ by $\epsilon_{xy}$ then the Langevin force term becomes nonlocal:

$$\begin{aligned} F_\mu(x) &\to \sum_y F_\mu(x,y) \ , \\ F_\mu(x,y) &= \epsilon_{xy} \frac{\beta}{4N_c} D_\mu(y) - \sqrt{\epsilon_{xy}} H_\mu(y) \ . \end{aligned} \tag{4.14}$$

It is no longer gauge covariant since different pieces of the force term transform differently under a gauge transformation:

$$\sum_y F_\mu(x,y) \to \sum_y G(y) F_\mu(x,y) G^\dagger(y) \ . \tag{4.15}$$

Evidently all effects due to the nonlocal components of this force will be wiped out over many updates as the updates randomize the gauge at remote sites $y$. One way to avoid this randomization is to completely specify the gauge (up to global transformations) at all sites before updating the gauge field. Then, obviously, gauge-invariant quantities formed from the updated field will be independent from which set of gauge-equivalent configurations we started.

In the case of lattice axial gauges, where $V-1$ links of the lattice are gauge transformed to the unit matrix, one can implement complete gauge fixing by not updating those links which are set to unity. This is not a good procedure, however, because the longitudinal infrared singularities characteristic of axial gauges greatly exacerbate the problem of critical slowing down.[12] Our procedure of alternating gauge-field steps with unconstrained updating steps does not suffer from this problem. It is clear from Eq. (4.15) that, for the local updating scheme, updates and gauge transformations commute, and thus the gauge-fixing steps can have no effect upon the decorrelation times for gauge-invariant quantities.

In the nonlocal case, the combined gauge fixing and update step takes the form

$$U_\mu^{(n+1)}(x) = \exp\left[ -\sum_y G^{(n)}(y) F_\mu^{(n)}(x,y) G^{\dagger(n)}(y) \right] G^{(n)}(x) U_\mu^{(n)}(x) G^{\dagger(n)}(x+\hat{\mu}) \ , \tag{4.16}$$

where the gauge transformation $G^{(n)}(x)$ completely specifies the gauge of configuration $U^{(n)}$ (and is therefore a functional of $U^{(n)}$). In fact this nonlocal update, although gauge invariant, is not correct. To see why, we replace the update equation (4.16) by

$$\begin{aligned} U_\mu^{(n+1)}(x) &= G^{\dagger(n)}(x) \exp\left[ -\sum_y G^{(n)}(y) F_\mu^{(n)}(x,y) G^{\dagger(n)}(y) \right] G^{(n)}(x) U_\mu^{(n)}(x) \\ &= \exp\left[ -\sum_y G^{\dagger(n)}(x) G^{(n)}(y) F_\mu^{(n)}(x,y) G^{\dagger(n)}(y) G^{(n)}(x) \right] U_\mu^{(n)}(x) \ , \end{aligned} \tag{4.17}$$

which differs from the original update only by a gauge transformation, and is thus equivalent for gauge-invariant quantities. Rewriting this equation in the form

$$U_\mu^{(n+1)}(x) = e^{-i f_\mu^{(n)}(x)\cdot T} U_\mu^{(n)}(x) \tag{4.18}$$

we see that the step-size matrix now depends in effect on the field, through the gauge-fixing matrices $G(x)$:

$$f_{j\mu}(x) = \sum_{k,y} [\epsilon_{xy}^{jk} \partial_{k\mu y} S - \sqrt{\epsilon_{xy}^{ik}}\, \eta_{k\mu}(y)] \ , \tag{4.19}$$

where

$$\epsilon_{xy}^{jk}[U] = 2\,\mathrm{Tr}[G_U(x) T^j G_U^\dagger(x) G_U(y) T^k G_U^\dagger(y)] \epsilon_{xy} \tag{4.20}$$

is a functional of the gauge field and where

$$\sum_{k,y} \sqrt{\epsilon_{xy}^{jk}} \sqrt{\epsilon_{yz}^{kl}} = \epsilon_{xz}^{jl} \ . \tag{4.21}$$

The field dependence in $\epsilon_{xy}^{jk}[U]$ changes the equilibrium probability distribution of the configurations generated by Langevin updates. This is evident from the Fokker-Planck equation for the probability distribution. This has the form, to leading order in $\epsilon$,

$$\sum_{\alpha,\beta} [\partial_\alpha \epsilon_{\alpha\beta} (\partial_\beta S) P + \partial_\alpha \partial_\beta \epsilon_{\alpha\beta} P] = 0 \ , \tag{4.22}$$

where indices $\alpha$ and $\beta$ each represent a color, site, and direction index. The Fokker-Planck equation implies that $P \propto e^{-S}$, the desired result, provided that $\epsilon$ is field independent. For $\epsilon[U]$ in Eq. (4.20), however, there is an additional term $(\partial_\beta \epsilon_{\alpha\beta}) P$ and the distribution is changed.[13] This extra term is a leading-order effect and we have observed it numerically. We measured Re(plaquette) and Re(Wilson line) at $\beta = 0.0$ using the force term above and found that they did not vanish as $\epsilon \to 0$.

One way to correct for this complication is to change the force term in the Langevin equation. Schematically, the correct Fokker-Planck equation results if $\epsilon_{\alpha\beta}\partial_\beta S$ is replaced by $\epsilon_{\alpha\beta}\partial_\beta S - \partial_\beta \epsilon_{\alpha\beta}$ in the Langevin force term. A direct calculation of $\partial_\beta \epsilon_{\alpha\beta}$ is quite costly for most choices of gauge fixing. However, a simple stochastic estimator can be used for this term. For example, $f_{j\mu}(x)$ in Eq. (4.19) can be augmented by

$$f^{\delta}_{j\mu}(x)=\frac{1}{2\delta}\sum_{k,y}[\epsilon^{jk}_{xy}(e^{-i\eta\cdot T\delta}U)\eta_{k\mu}(y)-\epsilon^{jk}_{xy}(U)\eta_{k\mu}(y)] \tag{4.23}$$

whose stochastic average is just what is needed (provided $\delta\sim\sqrt{\epsilon}$):

$$\langle f^{\delta}_{j\mu}\rangle=-\sum_{k,y}\partial_{k\mu y}\epsilon^{jk}_{xy}+O(\delta^2\epsilon)\ . \tag{4.24}$$

This is the correction term specified in the general updating procedure given above.[14] We have checked that the correct $\epsilon\to 0$ results for Re(plaquette) and Re(Wilson line) at $\beta=0.0$ are now obtained numerically when this correction term is included.

The correction term adds significantly to the computational overhead for an update. The storage requirements are doubled (relative to the local algorithm) and the time required for an update is longer (four times longer in our not-too-highly-optimized code). However, these costs are roughly constant as the lattice volume grows and $\beta$ increases, while the benefit should grow quadratically with the correlation length. There exists gauges for which the correction term vanishes—axial gauges, for example—but we have yet to find one that performs well for current lattice sizes and parameters. In any case the overhead incurred for Fourier acceleration of gauge-field updates is negligible once dynamical quarks are included.

Recently a gauge-invariant acceleration scheme has been developed and tested on the hybrid updating algorithm.[15] This scheme avoids the use of gauge fixing and the consequent correction term by using an acceleration factor which is the inverse of a covariant derivative. This has the the disadvantage of requiring a numerical inversion, a calculation which will suffer from critical slowing down as we approach the continuum limit, although perhaps not as badly as the updating algorithm it is trying to cure. Our algorithm, on the other hand, can completely remove the critical slowing down associated with the continuum limit once the lattice spacing is small enough (i.e., once $\beta$ is large enough). Regrettably, neither algorithm can deal with the critical slowing down that results when the quark mass is taken to zero. The gauge fixing required in our algorithm is easily implemented. Indeed, apart from its use in the updating of the gauge field, gauge fixing may be seen as necessary for the acceleration of the inversion of the fermion matrix which is required for a simulation of full QCD.

## V. SOME RESULTS WITH SU(3) GAUGE THEORY

To decide what gauge to use for Fourier acceleration of the updating algorithm we studied the effect of different gauges on Fourier acceleration of the inversion of the fermion matrix. (This work is described in Ref. 6.) The two problems are very similar, requiring an optimization in a space of many dimensions, and it is much faster to measure the decay of the residue on a few configurations than to measure decorrelation times over many thousands of sweeps.

Axial gauges did not produce good results in a Fourier accelerated inversion, at least with configurations thermalized at values of $\beta$ around 6.0. For the $A^0=0$ gauge this is because there can be large fluctuations in the field for large values of $k^2$ if $(k^0)^2$ is small. We were unable to find a simple form of $\epsilon(p)$ which would correct for this. We also tried more sophisticated axial gauges in which the maximal tree on which the gauge condition is imposed was more uniformly spread throughout the lattice.[5] The performance of these gauges was a considerable improvement over the lattice temporal gauge on an $8^4$ lattice but was still not good enough to be worthwhile. Perhaps on larger lattices at higher values of $\beta$ these gauges will prove useful for Fourier acceleration.

A gauge which worked very well was the lattice Landau gauge in which, by an iterative process, the average value over the lattice of the trace of the link field is maximized.[5] At this point the lattice version of $\partial^\mu A_\mu=0$ is obeyed. The method for implementing this gauge is also an optimization procedure, which again suffers from critical slowing down. We found that a Fourier accelerated steepest-descent algorithm moved toward Landau gauge very rapidly.[5] Usually only about ten iterations of the algorithm were required to raise the average trace of the link field to a sufficient value for maximal acceleration of the fermion matrix inversion.

For the updating algorithm we require, as previously explained, a complete gauge fixing to be done at each step. To attain $\partial^\mu A_\mu=0$ within machine precision is too time consuming and the gauge fixing provided by a few iterations of the Landau gauge-fixing algorithm, although sufficient to give a smooth enough field for Fourier acceleration, is certainly not complete. We therefore define a new gauge, the *AL* gauge which consists of an axial gauge fixing following by a *given number* of iterations of the Landau gauge-fixing algorithm. This provides a complete gauge fixing and also gives a smooth field configuration. The axial gauge used was either the $A^0=0$ gauge, in which case $\sim 10$ iterations of Landau gauge fixing were required, or a blocked axial gauge for which rather fewer subsequent iterations were needed.[5]

The first test of the accelerated updating algorithm was undertaken on a small $2^4$ lattice at a very weak coupling, $\beta=10.0$. Under these conditions, critical slowing down is very clear and we expect Fourier acceleration to work well, as it did in the Abelian theory, with the very simple form for $\epsilon(p)$ given by Eq. (3.11), $\epsilon(p)=\epsilon p^2_{\max}a^2/p^2a^2$ (for $p\neq 0$). The results bear this out (Fig. 3). As an observable we measure the correlation function of the gradient of the action. By analogy with Eqs. (3.16) and (3.17) this is given by

$$\chi(p)=\sum_{\mu j}\tilde{D}_{\mu j}(p)\tilde{D}_{\mu j}(-p) \tag{5.1}$$

with

$$D_{\mu j}(x)=\mathrm{Tr}[T^j D_\mu(x)] \ . \tag{5.2}$$

$D(x)$ is not gauge invariant and was measured in the same $AL$ gauge as was used when Fourier accelerating the update step. This $AL$ gauge consisted of the lattice temporal gauge followed by 50 hits of the $Z_3$-invariant version of the Landau gauge-fixing algorithm described in Ref. 5. At these values of $\beta$ (above the deconfinement phase transition) we use the $Z_3$-invariant gauge fixing to make sure that all terms in the equilibrium action are $Z_3$ invariant and no $Z_3$ phase is favored over the others.

Figure 3 shows clearly the presence of critical slowing down for the unaccelerated algorithm and its successful treatment by Fourier acceleration. The correlation times plotted correspond to the number of sweeps separation for the autocorrelation function to drop to a value of 0.1. This criterion was chosen as a reasonable one when the autocorrelation function was not a good exponential. The values obtained agree approximately with the expected behavior of an Abelian theory with coupling $\beta/6$ [see Eq. (3.21)], except at $p=0$. Since $\chi(0)$ is not zero in the non-Abelian theory it does have a correlation time which seems to be similar to the correlation time for the lowest nonzero momentum. To obtain good measurements of the autocorrelation function required 50 000 sweeps of the unaccelerated algorithm and 18 000 sweeps with acceleration, measuring $\chi(p)$ at every sweep. A similar autocorrelation time is measured for the plaquette and this is accelerated from 17 sweeps to 12 sweeps. The correlation time for the Wilson line is enormous (several thousand sweeps) without acceleration and with acceleration it is still too large to measure accurately with this amount of data.

Moving to stronger couplings, similar to those in use to today's practical calculations, we find, not unexpectedly, that the pattern of critical slowing down is much less clear as the different momentum modes interact. Fourier acceleration is correspondingly less effective.

Figure 4 shows results at $\beta=5.8$ (Ref. 16) for the correlation times of $\chi(p)$ with and without Fourier acceleration. Without Fourier acceleration there is a clear difference in correlation times for different momenta although it is not as large as at $\beta=10.0$ and the relationship between correlation times and momenta is not as smooth. This is presumably because $\chi(p)$ has a more complicated effective updating equation than Eq. (3.18) now. With Fourier acceleration we found that taking $\epsilon(p)=\epsilon p^2_{\max}a^2/p^2a^2$ as before did not work in this case. Forcing this degree of acceleration on the low-momentum modes caused instability. Autocorrelation times became very long and the mean value for observables were very different from the unaccelerated values. Instead we used

$$\epsilon(p)=\frac{p^2_{\max}a^2+\lambda^2}{p^2a^2+\lambda^2}\epsilon \ , \tag{5.3}$$

which provides a mass term $\lambda^2$ to limit the acceleration at low momentum. This mass may represent some physical scale on which the interactions between modes is large. The algorithm now worked perfectly well with $\lambda^2=4.0$. Acceleration factors were not as large as at $\beta=10.0$ but the correlation times which were largest were substantially reduced. The correlation time of the plaquette ( $\sim 40$

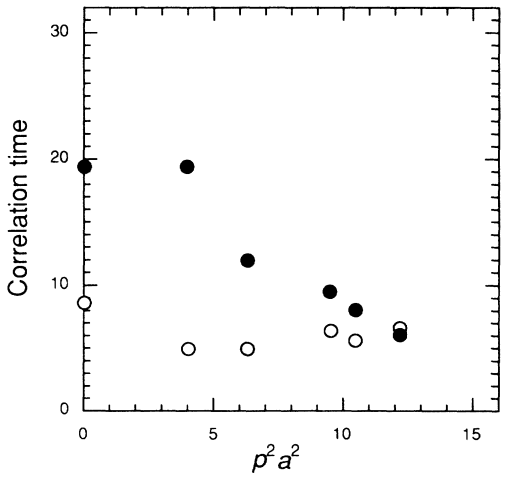


FIG. 3. Correlation times for $\chi(p)$ [Eq. (5.2)] plotted against $p^2a^2$ for an SU(3) gauge theory on a $2^4$ lattice at $\beta=10.0$, $\epsilon=0.01$. The correlation time is defined as the number of sweeps for the autocorrelation function to drop to a value of 0.1; ●—without Fourier acceleration; ○—with Fourier acceleration, $\epsilon(p)=\epsilon p^2_{\max}a^2/p^2a^2$. 50 000 sweeps were done without acceleration and 18 000 sweeps with acceleration, measuring $\chi(p)$ every sweep.

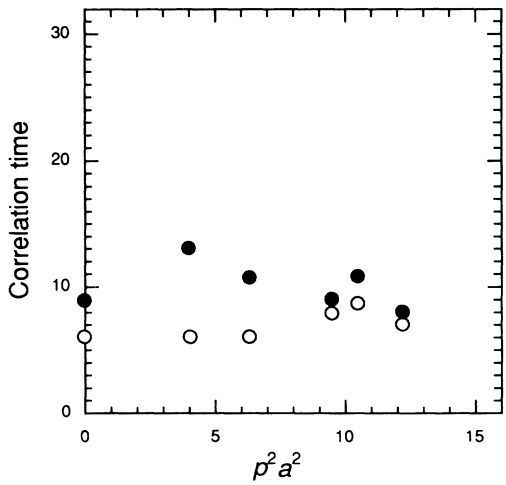


FIG. 4. Correlation times for $\chi(p)$ [Eq. (5.2)] plotted against $p^2a^2$ for an SU(3) gauge theory on a $2^4$ lattice at $\beta=5.9083$, $\epsilon=0.017\,24$. The correlation time is defined as the number of sweeps for the autocorrelation function to drop to a value of 0.1; ●—without Fourier acceleration; ○—with Fourier acceleration, $\epsilon(p)=\epsilon p^2_{\max}a^2+4.0/p^2a^2+4.0$. 50 000 sweeps were done without acceleration and 38 000 sweeps with acceleration, measuring $\chi(p)$ every sweep.

sweeps) is not altered significantly by acceleration. It is possible that better performance could be obtained with a more complicated function for $\epsilon(p)$.

$\chi(p)$ is not a particularly useful quantity to measure although it has the advantage that, in the limit of weak coupling, its correlation times can be compared to analytic calculations. Instead, gauge-invariant quantities such as Wilson loops are measured in practical calculations. It is not clear *a priori* what correlation times to expect for these variables or how the times will depend on the size of the loops since they have both ultraviolet (perimeter) and infrared (area) components.

Measurements at $\beta=5.8$ (Ref. 16) on a $4^4$ lattice showed that correlation times for various square Wilson loops were long but little acceleration could be achieved since the mass cutoff $\lambda^2$ required in the acceleration factor $\epsilon(p)$ was 6.0. The acceleration did not work with a mass cutoff of 4.0 as in the $2^4$ case. This presumably reflects the fact that there are smaller nonzero momenta on a $4^4$ lattice than a $2^4$ one.

At $\beta=6.2$ (Ref. 17) the different loops have more widely separated correlation times and acceleration can be achieved with a mass cutoff of 4.0 in $\epsilon(p)$. Figure 5 shows the results. 70 000 sweeps were performed with the unaccelerated algorithm, measuring square Wilson loops of side 1, 2, and 3 lattice units every sweep. 30 000 sweeps were carried out with the Fourier accelerated algorithm. The gauge fixing used there consisted of the lattice temporal gauge followed by 10 iterations of the $Z_3$ invariant Landau gauge fixing. All the loops show shorter correlation times with acceleration but the largest gain is for the $3\times3$ loop which is accelerated by a factor of 3. The correlation time of the Wilson line (both the real part and the magnitude) is substantially reduced by acceleration from over a thousand sweeps to several hundred sweeps.

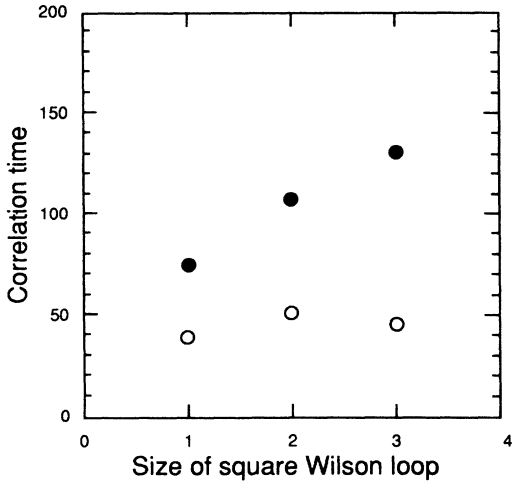


FIG. 5. Correlation times for square Wilson loops of size $1\times1$, $2\times2$, and $3\times3$ for an SU(3) gauge theory on a $4^4$ lattice at $\beta=6.2$, $\epsilon=0.0161$. The correlation time is defined as the number of sweeps for the autocorrelation function to drop to a value of 0.1; ●—without Fourier acceleration; ○—with Fourier acceleration, $\epsilon(p)=\epsilon p^2_{\max}a^2+4.0/p^2a^2+4.0$. 70 000 sweeps were done without acceleration and 30 000 sweeps with acceleration, measuring Wilson loops at every sweep.

The mean values obtained for the Wilson loops differ between the unaccelerated and accelerated algorithms, as expected at finite $\epsilon$. The difference is at most 0.02, consistent with the accelerated algorithm having slightly larger $O(\epsilon)$ corrections to the effective action.

The results described above all indicate that the Fourier-accelerated algorithm does work and can give a considerable gain in speed. The algorithm needs to be tested at weak coupling on larger lattices to demonstrate its full potential. To measure correlation functions sensibly requires, however, a huge amount of data, far more than is normally collected in a practical calculation and so far we have been unable to do this. One shortcut is the possibility of measuring thermalization or cooling times rather than correlation times. We have tested this by looking at the gluon propagator in the $AL$ gauge on $8^4$ lattices. We cool the configuration from a thermalized one at some value of $\beta$ and measure the rate of decay of the propagator as a function of momentum. One might hope that a "cooling time" could be defined that would be related in some way to the decorrelation time. Unfortunately this does not seem to be true. The fact that some components of the field are decaying faster than others seems to lead to odd feedback effects. The low-momentum components of the gluon propagator actually increase with time as the cooling goes on. More work needs to be done before this method can be used to assess correlation times.

## VI. CONCLUSIONS

We have extended the use of Fourier acceleration, described for the $XY$ model in Ref. 1, to lattice gauge theories. For a U(1) gauge theory this is straightforward and we find that the acceleration works exactly as expected at weak coupling. The behavior at stronger couplings near the phase transition needs to be investigated but we would anticipate that a considerable gain in speed could be made using Fourier acceleration in useful calculations. The $O(\epsilon)$ corrections for the accelerated algorithm are shown to be comparable with those of the unaccelerated algorithm.

For non-Abelian gauge theories, there are complications. It is necessary to transform the configurations to a smooth gauge to prevent gauge artifacts from affecting the correlation time of gauge-invariant quantities. This introduces an extra term in the effective action at leading order which must be explicitly removed by the numerical simulation of a correction term. We have shown that, once this is done, the algorithm works as expected. We have tested it on relatively small lattices at weak couplings and found a gain on speed. It seems possible to calculate Wilson loops at $\beta=6.2$ with a factor of 3 fewer sweeps. The acceleration factor in momentum space must be modified from the naive form inversely proportional to the square of the momentum. A mass term is

necessary to limit the acceleration of low momenta and this restricts the gains that can be made. If this mass represents a physical interaction length in lattice units then it should get smaller as $\beta$ is increased.

The price to be paid for this acceleration is the overhead for gauge fixing, Fourier transforming and adding the correction term to the Langevin force. Of these the last one is the most significant. We have tested the algorithm by simulations in pure gauge QCD since we believe that this can answer the question of whether Fourier acceleration will work. The overheads must be thought of, however, in the context of a realistic calculation on large lattices at weak coupling and including dynamical quarks. The cost of including dynamical quarks is so large that any extra calculation required in the pure gauge part of the simulation will be completely negligible. If the gauge field can be made to decorrelate more quickly, however, it will cut down the number of sweeps of the whole algorithm required. Thus, although it is probably true that a pure gauge calculation would find negligible benefit from Fourier acceleration (unless it was on such a large lattice that the acceleration could gain a factor much larger than 4 to overcome the overheads), there is no doubt that a simulation of full QCD could be speeded up. The most significant numerical cost in including dynamical quarks is the calculation of the inverse of the fermion matrix. This can also be accelerated in momentum space provided a smooth gauge fixing has been applied.[6] The cost of the inversion can be cut by a factor of about 3 on $8^4$ lattices at values of $\beta$ around 6.0. This factor will then apply to the cost of one sweep of the full algorithm. Combining this factor of 3 with a factor of 3 fewer sweeps from Fourier acceleration of the updating algorithm could lead to a factor of 10 gain in speed for a simulation of full QCD.

We conclude that Fourier acceleration can provide gains in speed when updating gauge fields with a Langevin equation. We believe it can also be successfully applied to the molecular-dynamics—Langevin-hybrid algorithm,[18] where significant gains have already been made with spin models.[3,4] The more recent hybrid Monte Carlo[19] algorithm also shows acceleration for spin models[20] which we believe can be extended to the simulation of gauge theories.

## ACKNOWLEDGMENTS

This research was supported by the National Science Foundation. Numerical work was performed on the NSF's supercomputer facility at Cornell University. G.P.L. was partially supported by the Alfred P. Sloan Foundation. B.S. thanks the IBM corporation for financial support. Fermilab is operated by Universities Research Association Inc., under contract with the U.S. Department of Energy.

*Present address: Department of Physics and Astronomy, Univeristy of Glasgow, Glasgow, G12 8QQ, UK.

†Present address: Lawrence Livermore National Lab., P.O. Box 808, Livermore, CA 94550.

‡Present address: Naval Research Laboratory, Washington, D.C. 20375.

§Present address: Theoretical Physics Group, Fermi National Accelerator Laboratory, Batavia, IL 60510.

**Present address: Department of Physics, University of California at San Diego, La Jolla, CA 92093.

††Present address: Department of Nuclear Physics, Tel Aviv University, Ramat Aviv 69978, Israel.

‡‡Present address: Department of Physics, Ohio State University, Columbus, OH 43210.

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

[2]G. Parisi, in *Progress in Gauge Field Theory,* NATO Advanced Study Institute, Cargèse, France, 1983, edited by G. 't Hooft *et al.* (NATO ASI Series B: Physics, Vol. 115) (Plenum, New York, 1984).

[3]E. Dagotto and J. B. Kogut, Nucl. Phys. **B290** [FS20], 451 (1987); E. Dagotto and A. Moreo, Illinois Report No. P/87/8/149, 1987 (unpublished); G. G. Batrouni and E. Dagotto, Phys. Rev. B **37**, 9875 (1988).

[4]G. G. Batrouni and B. Svetitsky, Phys. Rev. B **36**, 5647 (1987); E. Dagotto and J. B. Kogut, Phys. Rev. Lett. **58**, 299 (1987).

[5]C. T. H. Davies, G. G. Batrouni, G. R. Katz, A. S. Kronfeld, G. P. Lepage, P. Rossi, B. Svetitsky, and K. G. Wilson, Phys. Rev. D **37**, 1581 (1988).

[6]G. R. Katz, G. G. Batrouni, C. T. H. Davies, A. S. Kronfeld, G. P. Lepage, P. Rossi, B. Svetitsky, and K. G. Wilson, Phys. Rev. D **37**, 1589 (1988).

[7]C. Davies, G. G. Batrouni, G. Katz, A. S. Kronfeld, G. P. Lepage, B. Svetitsky, and K. Wilson, J. Stat. Phys. **43**, 1073 (1986); G. G. Batrouni, Nucl. Phys. **A461**, 351c (1987); A. S. Kronfeld, in *Lattice Gauge Theory: A Challenge in Large-Scale Computing,* proceedings of the NATO Advanced Study Institute, Wuppertal, West Germany, 1985, edited by B. Bunk, K. H. Mütter, and K. Schilling (NATO ASI Series B: Physics, Vol. 140) (Plenum, New York, 1986), p. 67; C. Davies, in *Lattice Gauge Theory '86,* proceedings of the NATO Advanced Research Workshop, Upton, New York, 1986, edited by H. Satz, I. Harrity, and J. Potvin (NATO ASI Series B: Physics, Vol. 159) (Plenum, New York, 1987), p. 63.

[8]G. G. Batrouni, A. Hansen, and M. Nelkin, Phys. Rev. Lett. **57**, 1336 (1986).

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

[10]P. H. Damgaard and H. Hüffel, Phys. Rep. **152**, 277 (1987); I. Drummond, S. Duane, and R. R. Horgan, Nucl. Phys. **B220**, 119 (1983); C. Bender, F. Cooper, and B. Freedman, *ibid.* **B219**, 61 (1983).

[11]S. Duane, R. Kenway, B. J. Pendleton, and D. Roweth, Phys. Lett. B **176**, 143 (1986).

[12]B. Lautrup, M. J. Lavelle, M. P. Tuite, and A. Vladikas, Nucl. Phys. **B290** [FS20], 188 (1987).

[13]The problem here is similar to that which arises for the Langevin equation in curved space-time, see G. G. Batrouni, H. Kawai, and Pietro Rossi, J. Math. Phys. **27**, 1646 (1986).

[14]Of course the correction term specified by Eq. (4.23) is that appropriate to the updating algorithm Eq. (4.19). The algorithm we use, Eq. (4.16), is a gauge transform of this and the correction term must be modified accordingly. This is done in Eq. (4.10).

[15]S. Duane and B. J. Pendleton, Phys. Lett. B **206**, 101 (1988); B.

J. Pendleton, in *Field Theory on the Lattice,* proceedings of the International Symposium, Seillac, France, 1987, edited by A. Billoire *et al.* [Nucl. Phys. B (Proc. Suppl.) **4**, 590 (1988)].

[16]We actually choose $\beta=5.9083$, $\epsilon=0.1724$. This value of $\beta$ is equivalent for the unaccelerated algorithm to a value of $\beta$ of 5.8 used in Metropolis algorithms provided measured observables are also renormalized (see Ref. 1).

[17]We actually choose $\beta=6.3083$, $\epsilon=0.1613$. This value of $\beta$ is equivalent for the unaccelerated algorithm to a value of $\beta$ of 6.2 used in Metropolis algorithms provided measured observables are also renormalized (see Ref. 1), which is not done for the Wilson loop values quoted in the text.

[18]S. Duane, Nucl. Phys. **B257 [FS14]**, 652 (1985).

[19]S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, Phys. Lett. B **195**, 216 (1987).

[20]G. G. Batrouni, work in progress.