Overrelaxed Heat-Bath and Metropolis Algorithms for Accelerating Pure Gauge Monte Carlo Calculations

Frank R. Brown and Thomas J. Woch
Department of Physics, Columbia University, New York, New York 10027
(Received 16 March 1987)

The lattice-gauge-theory generalization of Adler's overrelaxed heat-bath algorithm and an overrelaxed Metropolis update are shown to accelerate the decorrelation of physical observables. The heat-bath's microcanonical limit is especially attractive. Numerical tests for pure gauge SU(3) (4^4 lattice, $\beta=5.6$) show, for example, that overrelaxation reduces the Polyakov-loop magnitude autocorrelation time for the Cabibbo-Marinari algorithm from 28 sweeps to 9, increasing computational efficiency threefold. These approaches are applicable to a large class of physical systems.

PACS numbers: 11.15.Ha, 02.70.+d

The Monte Carlo study of lattice gauge theory is often hindered by slow evolution of the physical variables. When large lattices are used to approximate the continuum limit, large sweep-to-sweep correlations significantly reduce the efficiency of such calculations. Indeed, at a second-order phase transition critical slowing down occurs; the autocorrelation time of the Monte Carlo evolution diverges as the critical point is approached. A simple-minded point of view is that the local-update procedures underlying most Monte Carlo algorithms are hampered by large physical correlation lengths. This mismatch has been addressed directly; Fourier acceleration[1] and multigrid methods[2] employ intrinsically nonlocal updates. We will argue, however, that even minor modifications of conventional local-update algorithms can significantly accelerate decorrelation of physical observables.

In a surprisingly underappreciated paper[3] Adler argued that overrelaxation can accelerate the heat-bath evolution of quadratically coupled fields. (Adler's work has had limited impact no doubt because of its restriction to multiquadratic actions. Nevertheless, as emphasized by Whitmer,[4] it is directly relevant to pseudofermion calculations. In spite of the vast computing resources they have consumed, to the best of our knowledge only one exploratory calculation[5] has taken advantage of Adler's more efficient algorithm.) Whitmer has verified in toy models that autocorrelation times are significantly reduced by overrelaxation.[4] We now present two strategies for overrelaxing more general actions. The first generalizes Adler's algorithm by introducing a change of coordinates which brings the action to quadratic form on an update-by-update basis. The second, which we call the overrelaxed Metropolis algorithm, uses a special Metropolis hit to jump across the minimum of the action. We present an SU(2) overrelaxed heat-bath update and SU(N) overrelaxed Metropolis algorithm, report on SU(3) test calculations, and speculate that overrelaxation might reduce critical slowing down.

Given a multiquadratic action,

$$S(x_i,\{x_{j\neq i}\})=a(x_i-x_i^0)^2+c,$$

where a, x_i^0, and c are functions of $\{x_{j\neq i}\}$, the following overrelaxed heat-bath update[3] (see Fig. 1) satisfies detailed balance with respect to S:

$$x_i'=\omega x_i^0+(1-\omega)x_i+\omega(2-\omega)a^{-1/2}\eta, \tag{1}$$

where η is the unit Gaussian noise. ω is the relaxation parameter; the conventional heat-bath update is recovered when $\omega=1$, while overrelaxation and underrelaxation correspond to ω greater and less than 1. When $\omega=0$, the system does not evolve, while for $\omega=2$, we obtain an interesting microcanonical algorithm with a deterministic, constant-action evolution. (The algorithm diverges if $\omega\notin[0,2]$.) Notice that dropping the noise term, η, yields the standard successive overrelaxation algorithm for linear equations.

To generalize Eq. (1) to SU(2) we map SU(2) onto R^3 in such a way that the SU(2) heat-bath distribution transforms into a three-dimensional unit Gaussian, per-

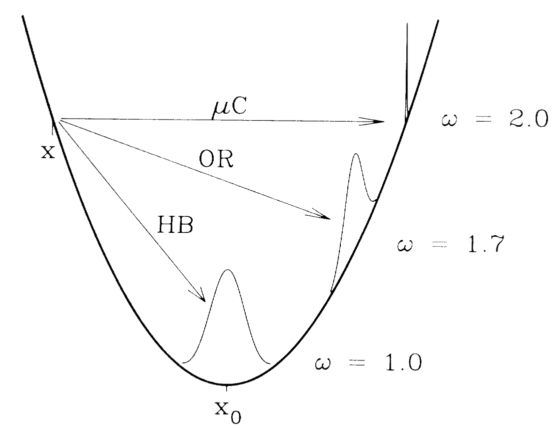

FIG. 1. The overrelaxed heat-bath update. The conventional heat bath, the overrelaxed heat bath, and the microcanonical limit are labeled HB, OR, and μC, respectively.

form the Gaussian overrelaxed heat-bath update, and map back to SU(2). Since the update satisfies detailed balance when expressed in our specially constructed Gaussian coordinates, it also satisfies detailed balance when expressed in conventional coordinates on SU(2). This strategy is clearly quite general, and could be easily carried out for, say, a ϕ^4 theory. Specifically, if the action of a single link $g \in SU(2)$ is $S(g) = -k\,\mathrm{Re\,tr}(g \times g_0^{-1})$, with $k \in R$ and $g_0 \in SU(2)$ functions of the environment, let $g = e^{\theta A} g_0$, with

$$A = (\alpha,\beta,\gamma) \equiv \begin{pmatrix} i\alpha & \beta + i\gamma \\ -\beta + i\gamma & -i\alpha \end{pmatrix},$$

$$A^{\dagger} = -A, \quad \mathrm{tr}(A^{\dagger}A) = 2.$$

Then $\theta \in [0,\pi]$ is the conventional polar angle on $S^3 = SU(2)$, while A yields coordinates on the S^2's of latitude. We now map $S^3 \to R^3$ via $f:\theta \to r$, defined such that

$$Z^{-1}\int_0^{r=f(\theta)} x^2 e^{-x^2/2}\,dx = Z'^{-1}\int_0^{\theta} \sin^2\omega\, e^{k\cos\omega}\,d\omega \qquad (2)$$

(where Z and Z' are the obvious normalizations, given by the integrals from 0 to ∞ and 0 to π, respectively), together with $\mathbf{r} = r\hat{\mathbf{r}}$, where $\hat{\mathbf{r}} = (\alpha,\beta,\gamma)$. This mapping indeed transforms the distribution $\exp[-S(g)]dg$ into a three-dimensional unit Gaussian. We therefore perform the update

$$\mathbf{r}' = (1-\omega)\mathbf{r} + \omega(2-\omega)\boldsymbol{\eta}, \quad \langle \eta_i \rangle = 0, \quad \langle \eta_i \eta_j \rangle = \delta_{ij},$$

and obtain the updated link, g', by mapping back to SU(2),

$$g' = e^{\theta' A'} g_0, \quad \theta' = f^{-1}(r'),$$

$$A' = \begin{pmatrix} i\hat{\mathbf{r}}_1' & \hat{\mathbf{r}}_2' + i\hat{\mathbf{r}}_3' \\ -\hat{\mathbf{r}}_2' + i\hat{\mathbf{r}}_3' & -i\hat{\mathbf{r}}_1' \end{pmatrix}.$$

By the updating of SU(2) subgroups in the manner of Cabibbo and Marinari[6] this algorithm may be applied to arbitrary SU(N) gauge theories. Note that when $\omega = 2$, this update becomes especially easy to implement; given the product of an SU(3) link with its environment, only 450 floating-point operations and one reciprocal suffice to update all three SU(2) subgroups.

The overrelaxed Metropolis algorithm is conceptually even simpler: For arbitrary gauge group calculate g_0, the value of the link which minimizes the action, and propose $g \to g' = g_0 g^{-1} g_0$ as a Metropolis hit to be accepted with probability $\min[1, \exp(S(g) - S(g'))]$, emulating, when accepted, an $\omega = 2$, microcanonical update. Some number of conventional Metropolis hits are then performed. If the overrelaxation hit acceptance rate is large this algorithm behaves much like the overrelaxed heat bath. Omission of the overrelaxation hit corresponds to an underrelaxed heat bath with ω ranging from 0 to 1 as the number of conventional hits is increased from 0 to ∞, and use of the overrelaxation hit corresponds to an overrelaxed heat bath with ω decreasing from 2 to 1 as the number of conventional hits is taken to ∞. In practice g_0 need not be calculated exactly; any approximation to g_0 independent of the initial g yields an algorithm satisfying detailed balance. (The overrelaxation hit acceptance rate will suffer, however, if the approximation is poor.) This strategy is extremely general; it can be applied, for example, as easily to Potts models as it can to SU(N) gauge theories.

SU(3) test calculations were performed on a 4^4 lattice at $\beta = 5.6$ (chosen to maximize sweep-to-sweep correlations) with the conventional action, $S = -\frac{1}{6}\beta \sum \mathrm{tr}(U_{\mathrm{plaq}} + \mathrm{H.c.})$. 25000-sweep overrelaxed [three SU(2) subgroup] heat-bath runs were performed for $\omega = 0.75$, 1.0, 1.25, 1.5, 1.75, and 2.0. For all values of ω except 2.0 (the simpler microcanonical case) Eq. (2) was approximated by an asymptotic expansion through order k^{-2}. The overrelaxed Metropolis data consist of 8600 zero-hit, 10000 one-hit, and 30000 ten-hit overrelaxed metropolis sweeps (i.e., overrelaxation plus zero, one, or ten conventional hits), together with 45000 conventional ten-hit Metropolis sweeps for comparison. The Metropolis table was sampled from SU(3) with the measure $\exp[100\,\mathrm{tr}(U)]dU$, yielding a 63% acceptance rate.

The average action and Polyakov loops in all four directions were measured for each sweep, and analyzed to obtain autocorrelation times following Refs. 7 and 2. Specifically we calculate the (truncated) autocorrelation time

$$\tilde{\rho}_m = \sum_{\Delta=-m}^{m} \rho(\Delta),$$

where $\rho(\Delta)$ is the autocorrelation between sweeps i and $i+\Delta$. This is the correct measure of computational efficiency; $\tilde{\rho}_\infty$ is the number of correlated measurements that yield in effect a single statistically independent measurement. As a statistical compromise[7] we work with $\tilde{\rho}_{150}$ rather than $\tilde{\rho}_\infty$. This analysis was carried out for 5000 sweep blocks, yielding a small number of $\tilde{\rho}_{150}$ measurements for each run, from which errors were estimated. Because of limited statistics our (1σ) error bars are rather crude (and absent for the zero- and one-hit overrelaxed Metropolis results).

Our principal results are given in Fig. 2 and Table I. Figure 2 shows that overrelaxation produces a statistically significant acceleration of the Cabibbo-Marinari heat-bath algorithm. The data suggest that the Polyakov loop (PL) decorrelates most rapidly at $\omega = 2$; the action, which does not evolve in this microcanonical limit, has a smaller optimum ω. Indeed, for the Polyakov-loop magnitude, overrelaxation decreases the autocorrelation time—and hence the cost of the calculation—by a factor of 3. We note that our $\omega = 2$ results seem fully consistent with ergodicity of the deterministic update, which is not *a priori* known to hold.

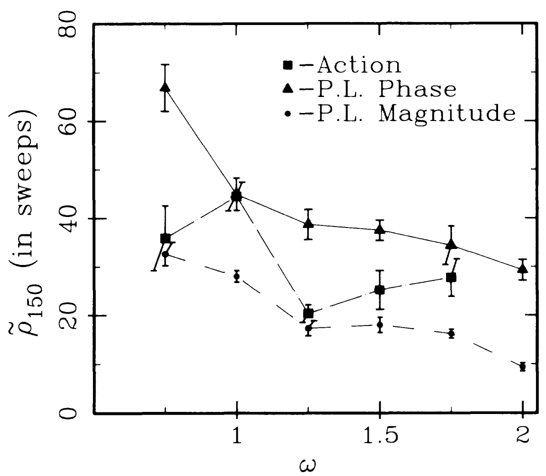

FIG. 2. Overrelaxed heat-bath autocorrelation times ($\tilde{\rho}_{150}$) as a function of ω demonstrate a factor-of-3 gain in computational efficiency.

Table I illustrates similar success for the overrelaxed Metropolis algorithm. Adding the overrelaxation hit to a ten-hit Metropolis update accelerates decorrelation by factors of 2 to 4, and even the one-hit update performs quite well. Also shown are results for the microcanonical-like, zero-hit update, which behaves rather differently from the closely analogous SU(2)-subgroup microcanonical algorithm. First of all, the action does evolve, albeit slowly, because the value of β is felt during the overrelaxation hit (rejection rate is 1.2%). Much more importantly, however, the zero-hit update yields incorrect expectation values [e.g., 0.0801(8) for the Polyakov-loop magnitude versus the correct value of 0.104(6), an error of $> 30\sigma$]. Although certain degrees of freedom evolve very rapidly (hence the small measured autocorrelation times), other degrees of freedom evolve too slowly to be seen in our 8600-sweep sample, or perhaps (since ergodicity is not guaranteed) not at all. Why should the conceptually very similar SU(2)-subgroup microcanonical update and the SU(3) overrelaxation hit behave so differently? We speculate that the noncommutativity of successive SU(2) subgroup hits on the same link naturally leads to greater mixing.

We have demonstrated empirically that overrelaxation accelerates decorrelation of physical measurements—a threefold gain in computational efficiency is seen for interesting observables. Let us speculate briefly on the important question of critical slowing down, not addressed by our 4^4 lattice. Near the continuum one expects autocorrelation times to scale like some power of a typical correlation length, $\tilde{\rho}_\infty \sim \alpha\lambda^\xi$. We advocate the optimistic view that overrelaxation will reduce the critical exponent, ξ (rather than merely the prefactor), for the following reasons. First, a reduced critical exponent can be seen analytically in the purely quadratic case.[8] Secondly, Whitmer[4] sees results suggestive of improved critical slowing down for his anharmonic, multiquadratic model. Finally, microcanonical evolution might reduce critical slowing down, because (one argues) n deterministic sweeps evolve the system a distance n, whereas n stochastic sweeps will only evolve the system a distance $n^{1/2}$. In any event, our overrelaxed updates are still local, and hence unlikely to reduce ξ to zero.

TABLE I. Autocorrelation times ($\tilde{\rho}_{150}$) for a variety of Metropolis updates (first four rows), with heat-bath and microcanonical results for comparison. Overrelaxation dramatically increases efficiency by reducing sweep-to-sweep correlations. (Note: The zero-hit update is presumably not ergodic.)

Method	Action	PL mag.	PL phase
Ten-hit conventional	94 ± 12	89 ± 4	123 ± 8
Ten-hit overrelaxed	37 ± 6	15 ± 1	26 ± 2
One-hit overrelaxed	143	28	19
Zero-hit overrelaxed	288[a]	12	9
Conventional heat bath	44 ± 3	28 ± 1	45 ± 3
Microcanonical	∞	9 ± 1	29 ± 2

[a] $\tilde{\rho}_\infty > 1000$.

Overrelaxation is a simple method for increasing the efficiency of lattice-gauge-theory calculations. On the basis of its performance and simplicity of implementation we recommend the SU(2)-subgroup microcanonical update, hybridized with an occasional conventional heat-bath sweep to ensure ergodicity, as the algorithm of choice for pure gauge SU(3) Monte Carlo calculations. Our approach is very general. Because the overrelaxed heat bath, and especially the overrelaxed Metropolis algorithm, involve such simple additions to the two most common Monte Carlo updates, one can add overrelaxation to most existing Monte Carlo programs with very little effort.

We would like to thank Stephen Adler, Alan Sokal, and Dan Zwanziger for valuable conversations, and Mike Creutz for encouraging us to perform the microcanonical calculations. The work reported here was funded in part by the Department of Energy.

Note added.—Creutz has also investigated SU(2) and SU(3) overrelaxed Metropolis algorithms.[9] As early as 1982 Whitmer had found an SU(2) overrelaxed Metropolis update, and had begun to investigate the possibility of an overrelaxed heat-bath update.[10]

[1] G. G. Batrouni *et al.*, Phys. Rev. D **32**, 2736 (1985).
[2] J. Goodman and A. D. Sokal, Phys. Rev. Lett. **56**, 1015 (1986).
[3] S. L. Adler, Phys. Rev. D **23**, 2901 (1981).
[4] C. Whitmer, Phys. Rev. D **29**, 306 (1984).
[5] P. Rossi and D. Zwanziger, Nucl. Phys. **B261**, 491 (1985).
[6] N. Cabibbo and E. Marinari, Phys. Lett. **119B**, 387 (1982).
[7] G. J. Daniell, A. J. G. Hey, and J. E. Mandula, Phys. Rev. D **30**, 2230 (1984).
[8] We thank Alan Sokal for a careful discussion of this point.
[9] M. Creutz, to be published.
[10] C. Whitmer, private communication.