Biased Monte Carlo algorithms on unitary groups

Michael Creutz, H. Gausterer, and S. Sanielevici
Physics Department, Brookhaven National Laboratory, Upton, New York 11973
(Received 17 August 1988)

We introduce a general updating scheme for the simulation of physical systems defined on unitary groups, which eliminates the systematic errors due to inexact exponentiation of algebra elements. The essence is to work directly with group elements for the stochastic noise. Particular cases of the scheme include the algorithm of Metropolis *et al.*, overrelaxation algorithms, and globally corrected Langevin and hybrid algorithms. The latter are studied numerically for the case of SU(3) theory.

I. INTRODUCTION

Monte Carlo–type algorithms have become important tools for evaluating the multidimensional integrals which appear in lattice field theory and in statistical physics. The most demanding problems for this approach are those involving fermionic systems and systems near criticality. In order to treat such systems with acceptable accuracy within the constraints of presently available computer resources, one needs to devise simulation algorithms which are both fast and exact. Indeed, the fastest algorithms which are presently available for updating fermionic systems suffer from systematic errors.[1–3]

Several ideas for improved algorithms have been put forward recently. It has been proposed that an efficient way to correct for the systematic error in the fermionic algorithms is to accept or reject trial configurations by means of a global Metropolis-type step.[4–8] On the other hand, overrelaxation algorithms[9,10] have been shown to improve the performance of bosonic simulations, even near criticality.

When the globally corrected algorithms of Refs. 4–8 are applied to systems evolving on SU(N) manifolds, there is an additional complication for $N \geq 3$. This is because these algorithms use elements of the group algebra (tangent vectors) to move around phase space. In order to accomplish the updating, the algebra element must be exponentiated to yield an element of the group. Exact exponentiation is straightforward for U(1) and SU(2), but it is nontrivial for $N \geq 3$. If the exponentiation is only done to a finite order, detailed balance will be violated despite the final accept/reject step[5–7] and there will be a residual systematic error. For SU(3) theory, it has been checked[5,6] that this residual error is numerically small, but the main point of the globally corrected scheme is to remove systematic errors altogether.

We will introduce and discuss a general updating scheme from which one can derive an algorithm which remains exact (satisfies detailed balance) even when the exponentiation is not done exactly. The essence of the approach is to work directly with group elements for the stochastic noise. Variations on the general scheme lead to the classic algorithm of Metropolis *et al.*[11] or to the overrelaxation algorithm of Ref. 10.

The updating scheme and some examples of algorithms which can be derived from it will be introduced in Sec. II. Section III presents some results obtained while testing the modified globally corrected hybrid algorithm for SU(3) theory. Some conclusions will be drawn in Sec. IV.

II. THE UPDATING SCHEME

Consider a set of group elements $U_i \in \mathrm{SU}(N)$, $i=1,\ldots,I$ and a desired probability distribution $P(\{U\}) \sim \exp[-S(\{U\})]$, where the function $S(\{U\})$ is bounded from below. The distribution is defined with respect to the invariant group measure. Physically, the function S could be the action or the classical Hamilton function of some system defined on the unitary group. To achieve this distribution, starting from some initial configuration of U variables, let us propose the trial changes

$$\begin{aligned} U_i' &= U_i F_i(\{U\}) V_i \ , \\ V_i' &= F_i(\{U\}) V_i F_i(\{U'\}) \ . \end{aligned} \tag{1}$$

Here the function $F_i(\{U\})$ takes values in SU(N) and $\{V\}$ is an auxiliary set of SU(N) matrices with the following distribution:

$$P(\{V\}) = \prod_i \tilde{P}(V_i), \quad \tilde{P}(V_i) \sim e^{-\tilde{S}(V_i)} , \tag{2}$$

where $\tilde{S}(V_i)$ is an arbitrary group function, with the property $\tilde{S}(V_i) = \tilde{S}(V_i^{-1})$.

Let us show that this updating scheme is area preserving ($dU_i' dV_i' = dU_i dV_i$). Without loss of generality, we can prove this for only one pair of (U, V). Consider an arbitrary function $h(U, V)$. Then, by the properties of the group-invariant measure:

$$\begin{aligned} \int dU\, dV\, h(U,V) &= \int dU\, dV\, h(U, VF(U)) \\ &= \int dU\, dV\, h(UV, VF(UV)) = \int dU\, dV\, h(UF(U)V, F(U)VF(UF(U)V)) \\ &= \int dU\, dV\, h(U', V') = \int dU' dV' h(U', V') \ . \end{aligned} \tag{3}$$

Note that the updating scheme of Eqs. (1) is also invertible by taking V to its inverse and iterating Eqs. (1). Using the property $\tilde{P}(V_i)=\tilde{P}(V_i^{-1})$, as well as Eq. (3), the following statement is readily verified: if we accept the set $(\{U'\},\{V'^{\dagger}\})$ with the probability

$$P_{\rm acc}=\min[1,\exp(-H'+H)] , \tag{4}$$

where

$$H=S(\{U\})+\sum_i \tilde{S}(V_i) , \tag{5}$$

then the overall process satisfies detailed balance. As a consequence, the process approaches equilibrium. Note that, if a trial update is accepted, V is replaced by $V'^{\dagger}$ and not by V'.

Various updating algorithms follow from this general scheme depending on the choice for $F_i(\{U\})$ and $\tilde{S}(V_i)$. Some interesting local choices for $F_i(\{U\})$ are

$$F_i(\{U\})=1 \quad \forall i \tag{6}$$

(the unit matrix), which yields the usual implementation of the algorithm of Metropolis *et al.* (Ref. 11) and

$$F_i(\{U\})=U_i^{\dagger}U_{i0}U_i^{\dagger}U_{i0} \tag{7}$$

(at each site, the element U_{i0} is chosen to minimize the action locally), which defines the overrelaxation algorithm of Ref. 10.

To obtain versions of the Langevin and hybrid algorithms, consider the choice

$$F_i(\{U\})=\exp[-i\epsilon\nabla_{i,\alpha}S(\{U\})T^{\alpha}] , \tag{8}$$

where T^{α} are the generators of the Lie algebra, $\nabla_{i,\alpha}$ denotes the derivatives with respect to the group parameters, and ϵ determines the size of each updating step. This form is inconvenient from a practical point of view for $N\geq 3$, since it requires exact exponentiation. We note, however, that achieving detailed balance with the updating scheme of Eqs. (1) only demands that $F_i(\{U\})$ be group elements. Therefore, the scheme remains exact if we redefine $F_i(\{U\})$ by only keeping the terms up to $O(\epsilon)$ in the expansion of the exponential and by then using a Gram-Schmidt procedure to enforce unitarity.

To relate this choice of $F_i(\{U\})$ to Langevin[2,12] and hybrid[3-8] updating, we should have

$$\begin{aligned} &V_i=\exp(i\sqrt{\epsilon}\eta_{i,\alpha}T^{\alpha})+O(\epsilon^{3/2}) , \\ &\langle\eta_{i,\alpha}\eta_{j,\beta}\rangle=2\delta_{\alpha\beta}\delta_{ij} , \end{aligned} \tag{9}$$

where the $\eta_{i,\alpha}$ are Gaussian random numbers. Thus, the algebra elements $p_i=\eta_{i,\alpha}T^{\alpha}$ are distributed as

$$P(p_i)\sim\exp[-\tfrac{1}{2}\mathrm{tr}(p_i^{\dagger}p_i)] . \tag{10}$$

To obtain the corresponding distribution of the V_i to lowest order in ϵ, we express p_i from Eq. (9) and expand in powers of ϵ. This gives

$$\tilde{P}(V_i)\sim\exp[k\ \mathrm{Re}(\mathrm{tr}V_i)] , \tag{11}$$

with

$$k=\epsilon^{-1}[1+O(\epsilon)] . \tag{12}$$

We have shown how to relate the general updating scheme of Eqs. (1) to Langevin and hybrid updating. The essential difference is that our scheme works entirely in SU(N) group space, without any need for exact exponentiation of algebra elements. This relationship is established by choosing the distribution of the V_i as in Eq. (11). Intuitively, one might expect that being close to this choice of k means being close to the maximal acceptance [with respect to the global accept/reject step in Eq. (4)] which is possible for a given step size ϵ.

For the case of an uncoupled harmonic oscillator $S=\frac{1}{2}x^2$ (one Gaussian random number: $\langle x^2\rangle=1$), one can actually show analytically that a suitable modification of the distribution of p leads to 100% acceptance. Taking

$$\begin{aligned} &x'=x-\frac{1}{2}\frac{dS}{dx}\epsilon+p\sqrt{\epsilon} , \\ &P(p)\sim\exp\left[-\frac{\kappa}{2}p^2\right] , \end{aligned} \tag{13}$$

will lead to trial updates which are always accepted, if

$$\kappa=(1-\epsilon/4)^{-1} . \tag{14}$$

Since this possibility of choosing κ to achieve 100% acceptance for any given step size is destroyed by anharmonic perturbations, we cannot hope to tune k in the case of SU(N) such as to obtain perfect acceptance for nonzero ϵ. Rather, past experience[4-8] suggests that the possible maximum acceptance should decrease with ϵ. Finding the proper relationship between k and ϵ is in fact a way to tune the performance of the algorithm. This is equivalent to the proposal studied in Refs. 4–8, which sought to maximize acceptance by shifting the couplings in the action used to generate the trial updates. In the next section we show a numerical study of the dependence of the acceptance on k for various values of ϵ, in SU(3) theory.

Another numerical experiment we will report in Sec. III concerns the dependence of the acceptance on the number of iterations of Eqs. (1) between the global accept/reject steps. The results obtained in Ref. 13 imply that, for the uncoupled harmonic oscillator, the acceptance does not decrease with the number of "molecular-dynamics" iterations, but oscillates around a value determined by the step size. For a line of coupled harmonic oscillators, a numerical study[5] which used a globally corrected hybrid algorithm showed oscillations in the acceptance, superposed onto a slow decrease. Using this algorithm for SU(3), one sees an initial drop in the acceptance as one increases the number of steps from 1 (the Langevin case), followed by a significantly flatter region. It is interesting to compare the performance of our present, modified algorithm with these results.

III. RESULTS

We have addressed these questions by simulating pure SU(3) theory with the standard Wilson action. As dis-

cussed above, we took $F_i(\{U\})$ to be the reunitarized first-order expansion of Eq. (8). We used two different methods of generating the distribution $P(V_i)$. In the first variant, we start from an initial pair of configurations $(\{U\},\{V\})$. After iterating Eqs. (1) *nmol* times, we accept or reject the resulting pair of configurations $(\{U'\},\{V'^{\dagger}\})$ according to Eq. (4). If the trial update is accepted, the old pair $(\{U\},\{V\})$ is replaced by the new one. The V_i are arranged in a periodic order, and, after each trial change, we reset the starting element V_1 at random.

In the second variant, which is designed to be analogous to the hybrid algorithms of Refs. 3–8, we first bring the V's into their correct distribution (11) by means of a standard procedure of Metropolis *et al.* This is very efficient because the V_i are independent random matrices. Combining the V's thus obtained with an initial choice for the U's, we generate a trial update by Eqs. (1) and accept or reject it by Eq. (4). But even if the update is accepted, we only replace the configuration $\{U\}$: the configuration $\{V'^{\dagger}\}$ is always discarded. After each global accept/reject step, a new configuration $\{V\}$ is generated from the old equilibrated sample by a multiple-hit procedure of Metropolis *et al.* Multiple hits are used to ensure that the new V_i are independent of the old ones (otherwise, there would be a systematic error in the whole algorithm).

We have tested both variants on a 4^4 lattice at $\beta=4.8$. The two methods lead to results which coincide within our statistical errors. As expected, the first variant is slightly faster. The results presented in Figs. 1 and 2 are obtained with the second variant.

Figure 1 shows the dependence of the acceptance rate on the distribution parameter k, for two values of ϵ. Remarkably, the acceptance peaks for both these values are located at $k \approx \epsilon^{-1}+4$. Thus, the numerical study has allowed us to determine the magnitude of the ϵ-independent shift away from $k=\epsilon^{-1}$, which was predicted by Eq. (12). The height of the peak is seen to drop as ϵ increases, as expected from previous studies of globally corrected hybrid algorithms.[4–8]

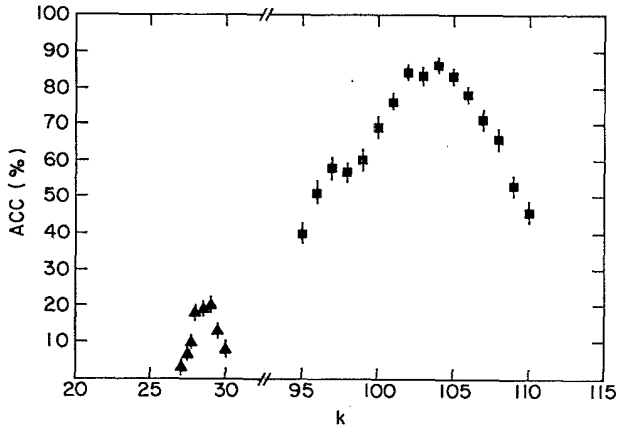

FIG. 1. Dependence of the acceptance rate on the distribution of the V_i for two values of ϵ. The parameter k is defined by the first line in Eq. (12). Squares refer to $\epsilon=0.01$ and triangles refer to $\epsilon=0.04$.

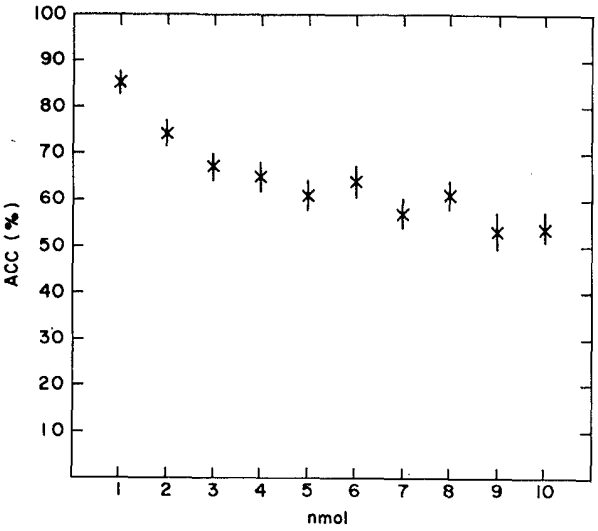

FIG. 2. Dependence of the acceptance rate on *nmol*, the number of iterations of Eqs. (1) between accept/reject steps. The data are obtained for $\epsilon=0.01$ and $k=104$.

Figure 2 exhibits the variation of acceptance with the number of iterations of Eqs. (1) between accept/reject steps. The initial drop in acceptance is seen to be far less pronounced than was typical of the previous globally corrected hybrid algorithms.[5,6] There is a hint of slight oscillations superposed onto a slow decrease [the acceptance becomes (52±3)% for $nmol=15$ and (43±3)% for $nmol=30$].

IV. CONCLUSIONS

We have shown that various interesting and popular algorithms for Monte Carlo–type simulations of physical systems defined on unitary group spaces can be conveniently summarized in a general updating scheme. From a computational point of view, this scheme has the advantage of being formulated in terms of group elements only. In particular, it is possible to define an algorithm which can be related to globally corrected Langevin and hybrid updating,[4–8] but which does not require the exact exponentiation of algebra elements. Besides removing the residual systematic error due to the violation of detailed balance by inexact exponentiation, the new algorithm appears to give a more advantageous dependence of the acceptance on the number of iterations between accept/reject steps. In addition, the algorithm is considerably more computationally efficient than if exact exponentiations had to be performed at each step. The inclusion of fermions into this algorithm is straightforward, since only the form of F_i must be appropriately modified.

ACKNOWLEDGMENTS

This work was supported by the U.S. Department of Energy under Contract No. DE-AC02-76CH00016. The computations were done on the Cray X-MP 22 at the National Magnetic Fusion Energy Computing Center, Lawrence Livermore National Laboratory. H.G. thanks the Fonds zur Förderung der wissenschaftlichen Forschung in Österreich (J0310P) for financial support.

[1]F. Fucito, E. Marinari, G. Parisi, and C. Rebbi, Nucl. Phys. **B180**, 369 (1981).

[2]G. G. Batrouni, G. R. Katz, A. S. Kronfeld, G. P. Lepage, B. Svetitsky, and K. G. Wilson, Phys. Rev. D **32**, 2736 (1985); A. Ukawa and M. Fukugita, Phys. Rev. Lett. **55**, 1854 (1985).

[3]S. Duane, Nucl. Phys. **B257**, 652 (1985); S. Duane and J. Kogut, Phys. Rev. Lett. **55**, 2774 (1985); S. Gottlieb, W. Liu, D. Toussaint, R. L. Renken, and R. L. Sugar, Phys. Rev. D **35**, 2531 (1987).

[4]R. Scalettar, D. Scalapino, and R. Sugar, Phys. Rev. B **34**, 7911 (1986); S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, Phys. Lett. B **195**, 216 (1987); M. Creutz, Nucl. Phys. B (Proc. Suppl.) **4**, 547 (1988).

[5]H. Gausterer and S. Sanielevici, Phys. Rev. D **38**, 1220 (1988).

[6]R. Gupta, Nucl. Phys. B (Proc. Suppl.) **4**, 562 (1988); R. Gupta, G. W. Kilcup, and S. Sharpe, Phys. Rev. D **38**, 1278 (1988).

[7]K. Bitar, A. D. Kennedy, R. Horsley, S. Meyer, and P. Rossi, Report No. FSU-SCRI-88-14, 1988 (unpublished).

[8]M. Creutz, Phys. Rev. D **38**, 1228 (1988).

[9]S. L. Adler, Phys. Rev. D **23**, 2901 (1981).

[10]F. R. Brown and T. J. Woch, Phys. Rev. Lett. **58**, 2394 (1987); M. Creutz, Phys. Rev. D **36**, 515 (1987).

[11]N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).

[12]A. Guha and S. C. Lee, Phys. Rev. D **27**, 2412 (1983).

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