Phys.Rev. D45 (1992) 4372-4375

arXiv:hep-lat/9204016 v1 30 Apr 1992

Acceptance Rates in Multigrid Monte Carlo

Martin Grabenstein[1] **and Klaus Pinn**[2]

[1] II. Institut für Theoretische Physik, Universität Hamburg,
Luruper Chaussee 149, D-2000 Hamburg 50, Germany

[2] Institut für Theoretische Physik I, Universität Münster,
Wilhelm-Klemm-Str. 9, D-4400 Münster, Germany

Abstract

An approximation formula is derived for acceptance rates of nonlocal Metropolis updates in simulations of lattice field theories. The predictions of the formula agree quite well with Monte Carlo simulations of 2-dimensional Sine Gordon, XY and ϕ^4 models. The results are consistent with the following rule: For a critical model with a fundamental Hamiltonian $\mathcal{H}(\phi)$ sufficiently high acceptance rates for a complete elimination of critical slowing down can only be expected if the expansion of $\langle \mathcal{H}(\phi+\psi) \rangle$ in terms of the shift ψ contains no relevant term (mass term).

Monte Carlo simulations of critical or nearly critical statistical mechanical systems with local algorithms suffer from critical slowing down. Roughly speaking, the autocorrelation time in the Markov chain scales like $\tau \sim \xi^z$, where ξ denotes the correlation length, and z is the dynamical critical exponent. For local algorithms, $z \approx 2$. To overcome this problem, various nonlocal Monte Carlo algorithms have been developed.

Cluster algorithms [1] are successful in overcoming critical slowing down for a large class of models. The alternative is multigrid Monte Carlo [2, 3]. In this letter, every algorithm that updates stochastic variables on a hierarchy of length scales is called multigrid Monte Carlo algorithm. There are models where no successful cluster algorithms have been found whereas multigrid Monte Carlo works [4]. Multigrid Monte Carlo is a candidate for noncritical simulations of nonabelian lattice gauge theory [3, 5].

Presently, the only generally applicable method to study algorithms for interacting models is numerical experiment. It is however important to have some theoretical understanding that helps to predict which algorithms will have a chance to overcome critical slowing down in simulations of a given model. As a contribution to the research in this direction we present a study of the kinematics of multigrid Monte Carlo algorithms. Sufficiently high acceptance rates are considered as necessary for successful multigrid Monte Carlo procedures. In this letter, we derive and discuss an approximation formula for the scale (block size) dependence of acceptance rates for nonlocal Metropolis updates.

We consider models with partition functions

$$Z = \int \prod_{x \in \Lambda_0} d\phi_x \, \exp(-\mathcal{H}(\phi)) \tag{1}$$

on cubic d-dimensional lattices Λ_0. We shall use dimensionless spin variables. Nonlocal Monte Carlo updates are defined as follows: Consider the fundamental lattice Λ_0 as divided in cubic blocks of size l^d. This defines a block lattice Λ_1. By iterating this procedure one obtains a whole hierarchy of block lattices $\Lambda_0, \Lambda_1, \ldots, \Lambda_K$. Let us denote block lattice points in Λ_k by x'. Block spins $\Phi_{x'}$ are defined on block lattices Λ_k. They are averages of the fundamental field ϕ over blocks of side length $L_B = l^k$:

$$\Phi_{x'} = L_B^{(d-2)/2} \, L_B^{-d} \sum_{x \in x'} \phi_x \, . \tag{2}$$

The L_B-dependent factor in front of the average comes from the fact that the corresponding dimensionful block spins are measured in units of the block lattice spacing. A nonlocal change of the configuration ϕ consists of a shift

$$\phi_x \to \phi_x + s \, \psi_x, \tag{3}$$

$$L_B^{-d} \sum_{x \in x'} \psi_x = L_B^{(2-d)/2} \delta_{x', x'_o} \, . \tag{4}$$

s is a real parameter. Note that $\Phi_{x'} \to \Phi_{x'} + s$ for $x' = x'_o$, and remains unchanged on the other blocks. The simplest choice of the "coarse-to-fine interpolation kernel"

ψ that obeys the constraint (4) is a piecewise constant kernel: $\psi_x = L_B^{(2-d)/2}$, if $x \in x'_0$, and 0 else. Other kernels are smooth and thus avoid large energy costs from the block boundaries. A systematic study of different kernels will be reported on elsewhere [6].

The s-dependent Metropolis acceptance rate for such proposals is given by

$$\Omega(s) = \langle \min[1, \exp(-\Delta\mathcal{H})] \rangle \,, \tag{5}$$

where

$$\Delta\mathcal{H} = \mathcal{H}(\phi + s\psi) - \mathcal{H}(\phi) \,. \tag{6}$$

$\Omega(s)$ has an integral representation

$$\Omega(s) = \int du \ \min(1, \mathrm{e}^{-u}) \int \frac{dp}{2\pi} \, \mathrm{e}^{-ipu} \, \langle \mathrm{e}^{ip\Delta\mathcal{H}} \rangle \,. \tag{7}$$

Assuming that the probability distribution of $\Delta\mathcal{H}$ is approximately Gaussian, we are led to

$$\langle \mathrm{e}^{ip\Delta\mathcal{H}} \rangle = \mathrm{e}^{ih_1 p - h_2 p^2} \, [1 + o(p^3)] \,, \tag{8}$$

with $h_1 = \langle \Delta\mathcal{H} \rangle$ and $h_2 = \frac{1}{2}(\langle \Delta\mathcal{H}^2 \rangle - \langle \Delta\mathcal{H} \rangle^2)$. Then the integrations in (7) can be done exactly:

$$\Omega(s) \approx \tfrac{1}{2} \left(\mathrm{erfc}(\frac{h_1}{2\sqrt{h_2}}) + \mathrm{e}^{h_2 - h_1} \mathrm{erfc}(\frac{2h_2 - h_1}{2\sqrt{h_2}}) \right) \,, \tag{9}$$

with $\mathrm{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty dt \, \exp(-t^2)$. Using invariance properties of the measure $\mathcal{D}\phi$ one can show that the difference of h_1 and h_2 is of order s^4. We shall see below that the approximation $h_1 \approx h_2$ is very good. In this case the acceptance rate prediction simplifies further,

$$\Omega(s) \approx \mathrm{erfc}(\tfrac{1}{2}\sqrt{h_1}) \,. \tag{10}$$

(For an analog result in the context of hybrid Monte Carlo see [7].) For free massless field theory with action $\mathcal{H}(\phi) = \frac{1}{2}(\phi, -\Delta\phi)$, we obtain the exact result

$$\Omega(s) = \mathrm{erfc}(\sqrt{\frac{\alpha}{8}}|s|) \,, \tag{11}$$

with $\alpha = (\psi, -\Delta\psi)$. Note that in this case $h_1 = h_2 = \frac{1}{2}\alpha \, s^2$. In d dimensions one finds [6]

$$\begin{array}{lcll} \alpha & = & 2dL_B & \text{for piecewise constant kernels} \,, \\ \alpha & \underset{L_B \gg 1}{\longrightarrow} & \text{const} & \text{for smooth kernels} \,. \end{array}$$

As a consequence, in massless free field theory, to maintain a constant acceptance rate (of, say, 50 percent) the changes s have to be scaled down like $L_B^{-1/2}$ for piecewise constant kernels, whereas for smooth kernels the acceptance rates do not depend on the block size. At least for free field theory, the disadvantage of the piecewise constant kernels can be compensated for by using a W-cycle instead of a V-cycle. Smooth kernels can be used only in V-cycle algorithms.

In the following we shall apply formula (10) in the discussion of multigrid procedures for three different spin models in two dimensions: the Sine Gordon model, the XY model, and the single-component ϕ^4 theory.

The 2-dimensional Sine Gordon model is defined by the Hamiltonian

$$\mathcal{H}(\phi) = \frac{1}{2\beta}(\phi, -\Delta\phi) - \zeta \sum_x \cos\phi_x \,. \tag{12}$$

The model undergoes a (Kosterlitz-Thouless) phase transition at β_c, and $\beta_c \to 8\pi$ for $\zeta \to 0$. For $\beta > \beta_c$, the flow of the effective Hamiltonian (in the sense of the block spin renormalization group) converges to that of a massless free field theory. The long distance behavior of the theory is therefore that of a Gaussian model. One might naively conclude that multigrid should be the right method to fight critical slowing down in the simulation of the Sine Gordon model in the massless phase. But this is not so. For h_1 we find the expression

$$h_1 = \frac{\alpha}{2\beta}s^2 + \zeta C \sum_x [1 - \cos(s\psi_x)] \,, \tag{13}$$

with $C = \langle \cos\phi_x \rangle$. The essential point is that the second term in (13) is proportional to the block volume L_B^2 for piecewise constant and for smooth kernels. One therefore has to face a dramatic decrease of acceptance when the blocks become large, even for small fugacity ζ. A constant acceptance rate is achieved only when the proposed steps are scaled down like L_B^{-1}. It is therefore unlikely that any multigrid algorithm - based on nonlocal updates of the type discussed in this letter - will be successful for this model. We demonstrate the validity of formula (10) (using a Monte Carlo estimate for C) by comparing with Monte Carlo results at $\beta = 39.478$ and $\zeta = 1$. This point is in the massless phase. In figure 1 we show both the numerical and analytical results for $\Omega(s)$ for $L_B = 4, 8, 16, 32$ on lattices of size $16^2, 32^2, 64^2, 128^2$, respectively.

We now discuss the 2-dimensional XY model, defined by the partition function

$$Z = \int \mathcal{D}\Theta \, \exp(\beta \sum_{<x,y>} \cos(\Theta_x - \Theta_y)) \,. \tag{14}$$

As the Sine Gordon model, the XY model has a massless (spin wave) and a massive phase. Nonlocal updates are defined by $\Theta_x \to \Theta_x + s\psi_x$, with ψ obeying again the normalization condition (4). Note that such an update changes the block spin (defined as the block average of the spins in the unit vector representation) by an amount $\approx s$ only when the spins inside the block are sufficiently aligned. This will be the case in the spin wave phase for large enough β. h_1 is given by

$$h_1 = \beta E \sum_{<x,y>} [1 - \cos(s(\psi_x - \psi_y))] \,, \tag{15}$$

with $E = \langle \cos(\Theta_x - \Theta_y) \rangle$, x and y nearest neighbors. For piecewise constant kernels, h_1 is proportional to L_B. For smooth kernels h_1 will become independent of L_B for large enough blocks. For small s,

$$h_1 \approx \tfrac{1}{2}s^2\beta E \sum_{<x,y>} (\psi_x - \psi_y)^2 = \tfrac{1}{2}s^2\beta E\alpha \,. \tag{16}$$

As above, $\alpha = (\psi, -\Delta\psi)$. This quantity becomes nearly independent of L_B already for L_B larger than 16 [6]. From the point of view of acceptance rates the XY model therefore behaves like massless free field theory. In fact, a dynamical critical exponent z consistent with zero was observed in the massless phase [8]. The failure of multigrid Monte Carlo in the massive phase ($z \approx 1.4$ for piecewise constant kernels [8]) is an example for the fact that good acceptance rates are not sufficient to overcome critical slowing down.

We again checked the accuracy of formula (10) by comparing with Monte Carlo results at $\beta = 1.2$ (which is in the spin wave phase). The only numerical input for the analytical formula was the link expectation value E. The results are displayed in figure 2. One can do a similar discussion for the $O(N)$ nonlinear σ-model with $N > 2$, leading to the same prediction for the scale dependence of the acceptance rates. This behavior was already observed in multigrid Monte Carlo simulations of the 2-dimensional $O(3)$ nonlinear σ-model with smooth and piecewise constant kernels [9].

Let us finally turn to single-component d-dimensional ϕ^4 theory, defined by the action

$$\mathcal{H}(\phi) = \tfrac{1}{2}(\phi, -\Delta\phi) + \frac{m_o^2}{2}\sum_x \phi_x^2 + \frac{\lambda_o}{4!}\sum_x \phi_x^4 \,. \tag{17}$$

For h_1 one finds

$$h_1 = s^2 \left\{ \tfrac{1}{2}\alpha + [\frac{m_o^2}{2} + \frac{\lambda_o}{4}P] \sum_x \psi_x^2 \right\} + s^4 \frac{\lambda_o}{4!}\sum_x \psi_x^4 \,, \tag{18}$$

where $P = \langle \phi_x^2 \rangle$. We have used that expectation values of operators which are odd in ϕ vanish on finite lattices. Because of the normalization condition (4), $\sum_x \psi_x^2$ increases with L_B^2, independent of d, whereas $\sum_x \psi_x^4$ scales like L_B^{4-d}, for smooth and for piecewise constant kernels. Therefore also in this model we have to face rapidly decreasing acceptance rates when the blocks become large. As in the case of the Sine Gordon model, s is to be rescaled like L_B^{-1} in order to maintain constant acceptance rates. (Note that this discussion is only of relevance when the correlation length ξ is larger than the block size L_B. The essential point is that the acceptance rates should not become too small before L_B is of the order of ξ.) From our point of view therefore there is little hope that multigrid algorithms of the type considered in this letter can overcome critical slowing down in the 1-component ϕ^4 model. In numerical simulations of 2-dimensional ϕ^4 theory, a dynamical critical behavior is found that is consistent with $z \approx 2$ for piecewise constant and for smooth kernels [10, 11]. In four dimensions, there is no definite estimate for z [12].

Figure 3 shows a comparison of Monte Carlo results for 2-dimensional ϕ^4 theory with the theoretical prediction based on the numerical evaluation of P. The simulations were done in the symmetric phase at $m_o^2 = -0.56$ and $\lambda_o = 2.4$. The correlation length at this point is $\xi \approx 15$ [11].

In this letter, we have presented a simple yet accurate formula that expresses acceptance rates for nonlocal update algorithms in terms of one single parameter (entering h_1), which is easy to compute, e.g. by Monte Carlo simulations on a small lattice.

We encountered two classes of models. For Sine Gordon and ϕ^4 theory s had to be rescaled like L_B^{-1} for piecewise constant and for smooth kernels, whereas for the XY model and the $O(N)$ nonlinear σ-model one can achieve L_B-independent acceptance rates (choosing smooth kernels).

The results are consistent with the following rule: Sufficiently high acceptance rates for a complete elimination of critical slowing down can only be expected if $h_1 = \langle \mathcal{H}(\phi + s\psi) - \mathcal{H}(\phi) \rangle$ contains no relevant operator in ψ. The (superficial) degree r of relevance of a local operator is defined by $r = d + m(2-d)/2 - n$. m is the number of fields, and n counts the derivatives. An operator is called relevant if $r > 0$. A mass term always has $r = 2$, and a ϕ^4-term has $r = 4 - d$. A kinetic term $\alpha = (\psi, -\Delta\psi)$ has $r = 0$. The above statement is formulated for smooth kernels. For piecewise constant some modifications are necessary. The main difference is that $\alpha \propto L$ for piecewise constant kernels.

With the help of the rule stated above it is possible to decide whether a given statistical model is a natural candidate for multigrid Monte Carlo or not. We believe that our formula can be useful to study different choices of interpolation kernels ψ in multigrid Monte Carlo for nonabelian gauge theories [3, 5].

Acknowledgment

We would like to thank G. Mack and S. Meyer for useful discussions. All numerical computations reported on in this paper were performed on the NEC SX-3 of the University of Cologne and the CRAY Y-MP of the HLRZ in Jülich.

References

[1] R.H. Swendsen, J.S. Wang, Phys. Rev. Lett. 58 (1987) 86;
U. Wolff, Phys. Rev. Lett. 62 (1989) 361.

[2] G. Parisi, in: Progress in Gauge Field Theory, 1983 Cargèse Lectures, G. t'Hooft et al. (eds.), Plenum Press, New York, 1984;
J. Goodman, A.D. Sokal, Phys. Rev. Lett. 56 (1986) 1015.

[3] G. Mack, in: Nonperturbative Quantum Field Theory, 1987 Cargèse Lectures, G. t'Hooft et al. (eds.), Plenum Press, New York, 1988.

[4] U. Wolff, CERN-TH 6126/91;
M. Hasenbusch, S. Meyer, Kaiserslautern preprint Th 91-16.

[5] T. Kalkreuter, DESY 90-158, December 1990, to appear in Nucl. Phys. B.

[6] M. Grabenstein, K. Pinn, in preparation.

[7] S. Gupta, A. Irbäck, F. Karsch, B. Petersson, Phys. Lett. B242 (1990) 437.

[8] R.G. Edwards, J. Goodman, A.D. Sokal, Nucl. Phys. B354 (1991) 289;
A. Hulsebos, J. Smit, J.C. Vink, Nucl. Phys. B356 (1991) 775.

[9] M. Hasenbusch, G. Mack, S. Meyer, Nucl. Phys. B (Proc. Suppl.) 20 (1991) 110.

[10] J. Goodman, A.D. Sokal, Phys. Rev. Lett. 56 (1986) 1015;
J. Goodman, A.D. Sokal, Phys. Rev. D 40 (1989) 2035.

[11] J. Linn, diploma thesis, Kaiserslautern 1991.

[12] G. Mack, S. Meyer, Nucl. Phys. B (Proc. Suppl.) 17 (1990) 293.

Figure Captions

FIG.1. $\Omega(s)$ in the 2-dimensional Sine Gordon model, $\beta = 39.478$, $\zeta = 1$. From top to bottom: $L_B = 4, 8, 16, 32$ on a $16^2, 32^2, 64^2, 128^2$ lattice, respectively. Points with error bars: Monte Carlo results, lines: analytical results.

FIG.2. $\Omega(s)$ in the 2-dimensional XY model, $\beta = 1.2$. From top to bottom: $L_B = 4, 8, 16$ on a $16^2, 32^2, 64^2$ lattice, respectively. Points with error bars: Monte Carlo results, lines: analytical results.

FIG.3. $\Omega(s)$ in the 2-dimensional ϕ^4 theory, $m_o^2 = -0.56$, $\lambda_o = 2.4$. From top to bottom: $L_B = 4, 8, 16$ on a $16^2, 32^2, 64^2$ lattice, respectively. Points with error bars: Monte Carlo results, lines: analytical results.