Dynamics of Hybrid Overrelaxation in the Gaussian Model

Ulli Wolff
CERN, Theory Division,
CH-1211 Genève 23, Switzerland

Abstract

For gaussian free fields we present an analysis of the dynamics of hybrid overrelaxation algorithms, where \( N \) exactly microcanonical update sweeps alternate with single local heatbath passes through the lattice. Dynamical exponents \( z = 1 \) can be reached with appropriate tuning of \( N \). Special interest in this simulation technique derives from the fact that it represents the most efficient presently known algorithm for pure gauge theory.
Critical slowing down is the main obstacle in removing the ultraviolet cut-off in field theory simulations, at least for systems which are simple enough to use large lattices. Many new Monte Carlo algorithms have been designed to alleviate or eliminate this problem. In this paper we deal with overrelaxation (OR) techniques. Although they are not believed to be capable of reducing dynamical critical exponents $z$ of the slowest modes below one — as has become possible for some spin systems —, they boost the efficiency very significantly. At the same time the class of systems for which OR is available includes nonabelian gauge theory for which it is the optimal known algorithm at present. OR in the original form proposed by Adler [1] (AOR) is applicable to systems with noncompact fields and actions that are quadratic in all individual field variables (multiquadratic actions). As in the analogous methods for solving difference equations, there is an overrelaxation parameter $\omega$ in AOR that has to be tuned to reach $\omega \rightarrow 2$ as the critical point is approached. For free fields the dynamics of AOR has been analyzed first by Adler [2] and later by Neuberger [3] for the both simpler and more realistic case of checkerboard updating on the torus. It is expected that the dynamics of simulations of gaussian systems is not untypical, and we can thus learn important features from their rigorous analysis.

More recently the restriction to multiquadratic actions has been overcome by a variation of AOR, which we propose to call hybrid overrelaxation (HOR). There one makes local update steps which are energy conserving (microcanonical) and correspond to $\omega = 2$. As they by themselves are not ergodic — as is demonstrated by the never changing energy — they are blended with some standard local ergodic algorithm, say heatbath. The tunable parameter in this case is $N$, the number of microcanonical sweeps per heatbath sweep. Such algorithms recently allowed simulations close to criticality for the XY model [4], the $O(3)$ model [5], the $CP^3$ model [6, 7] and, last but not least, gauge theory [8]. To the best of the author’s knowledge, there is however no demonstration of how HOR works in the exactly soluble case of free field theory. This gap is filled by the present note.

The gaussian model is defined by the free field action

$$S(\varphi) = \frac{1}{2} \sum_x \left\{ \sum_\mu (\partial_\mu \varphi(x))^2 + m^2 \varphi(x)^2 \right\}$$

for a real scalar field $\varphi(x)$. We sum over all $L^D$ sites $x$ of a $D$-dimensional torus, and $\partial_\mu$ is the discrete forward difference. Lattice units ($a = 1$) are
used, and criticality is reached for $m^2 \to 0$. In a general local AOR update at a site $x_0$ one replaces $\varphi(x_0)$ by

$$\varphi'(x_0) = (1 - \omega)\varphi(x_0) + \omega \sigma^2 \sum_{|x-x_0|=1} \varphi(x) + \sqrt{\omega(2 - \omega)} \sigma \eta,$$  

(2)

where $\sigma^{-2} = 2D + m^2$, and $\eta$ is a normalized gaussian random number. For a checkerboard ordering such updates are first (independently of each other) carried out on all even sites and then for all odd ones. This is conveniently analyzed in Fourier space. Fields on the even (odd) sublattice are spanned by

$$\hat{\varphi}_\pm(p) = \frac{1}{2} \sum_x (1 \pm e^{i\Pi x}) e^{-ipx} \varphi(x) = \hat{\varphi}_\pm^*(-p)$$  

(3)

with $\Pi = (\pi, \pi, \ldots, \pi)$. To avoid double counting one may restrict momenta $p_\mu = 2\pi n_\mu/L$ to $-\pi/2 \leq p_0 < \pi/2$, $-\pi \leq p_k < \pi$ for $k = 1, \ldots, D - 1$. The updates (2) are blockdiagonal in momentum space such that they only mix the two components $\hat{\varphi}_+(p)$ and $\hat{\varphi}_-(p)$. It is now easy to work out the effect of a full even update followed by an odd one. In a 2-component notation,

$$\chi(p) = \begin{pmatrix} \hat{\varphi}_+(p) \\ \hat{\varphi}_-(p) \end{pmatrix},$$  

(4)

we have

$$\chi'(p) = M(\lambda, \omega) \chi(p) + Q(\lambda, \omega) \xi$$  

(5)

with

$$M = \begin{pmatrix} 1 - \omega & \omega \lambda \\ (1 - \omega)\omega \lambda & 1 - \omega + (\omega \lambda)^2 \end{pmatrix},$$  

(6)

$$Q = \sigma \sqrt{\omega(2 - \omega)} \begin{pmatrix} 1 & 0 \\ \omega \lambda & 1 \end{pmatrix}$$  

(7)

and

$$\lambda(p) = \frac{2 \sum_\mu \cos(p_\mu)}{2D + m^2}.$$  

(8)

The Fourier transformed random numbers in (5) also have two components $\xi_\pm$ and covariance

$$\langle \xi_\alpha \xi_{\alpha'} \rangle = \frac{L^D}{2} \delta_{\alpha\alpha'}.$$  

(9)
As (5) is iterated \(n\) times starting from \(\chi^{(0)}\) we get

\[
\chi^{(n)} = M^n \chi^{(0)} + \sum_{i=1}^{n} M^{n-i} Q \xi^{(i)},
\]

(10)

where \(\xi^{(i)}\) are the independent random numbers that enter. This leads to a 2 \(\times\) 2 autocorrelation matrix for each pair of momenta

\[
G(t) = \lim_{n \to \infty} \langle \chi^{(n+t)} \chi^{t(n)} \rangle = \frac{L^D}{2} M^t \sum_{i=0}^{\infty} M^i Q Q^t M^{ti}.
\]

(11)

Any reference to \(\chi^{(0)}\) has gone as eigenvalues of \(M\) have to be and will be smaller than one in magnitude. This is simplified if we use a relation [9] following form (5) being a legal update that gives an equilibrated \(\chi'\) from an equilibrated \(\chi\). As all quantities are gaussian distributed this is just a statement on covariances

\[
h^{-1}(p) = M h^{-1}(p) M^\dagger + \frac{L^D}{2} Q Q^\dagger,
\]

(12)

where

\[
h(p) = 2L^{-D} \sigma^{-2} \begin{pmatrix} 1 & -\lambda \\ -\lambda & 1 \end{pmatrix}
\]

(13)

follows from writing the action in momentum space

\[
S = \frac{1}{2} \sum_p \chi^\dagger h \chi.
\]

(14)

With the help of (12) the autocorrelations now assume the simple form

\[
G(t) = M^t h^{-1}.
\]

(15)

The decay of autocorrelations is controlled by the modulus of the largest eigenvalue of \(M\) which we write as \(\exp(-1/\tau(p))\). For simple heatbath, \(\omega = 1\) in (6), the matrix is singular with one nonzero eigenvalue giving

\[
\tau_{HB}(p) = -\frac{1}{2 \log(|\lambda|)} \approx \frac{D}{m^2 + p^2} \quad \text{as } p^2 \to 0.
\]

(16)

The slowest mode is at zero momentum, and the growth of \(\tau_{HB}\) with \(m^{-2}\) corresponds to \(z = 2\). For AOR one chooses \(\omega\) such that all \(M\) acquire
complex conjugate pairs of eigenvalues [3]. The smallest maximum of all $\tau(p) (= \tau_{\text{exp}}$, the exponential autocorrelation time) is achieved by tuning

$$4 \frac{\omega - 1}{\omega^2} = \lambda(0)^2, \quad \omega = 2 - \frac{2m}{\sqrt{D}} + O(m^2).$$

(17)

With this, all modes evolve at the same speed

$$\tau_{\text{AOR}} = \frac{1}{\log(\omega - 1)} \simeq \frac{\sqrt{D}}{2m}$$

(18)

which diverges with $z = 1$. This tuning of $\omega$ will be understood from now on when we talk about AOR.

To implement HOR we have to replace

$$M(\lambda, \omega) \rightarrow K(\lambda, N) = M(\lambda, 1)M(\lambda, 2)^N.$$  

(19)

Diagonalization of $M(\lambda, 2)$ gives eigenvalues $\exp(\pm i\theta)$ with

$$|\lambda(p)| = \cos(\theta(p)/2), \quad 0 < \theta \leq \pi, \quad \theta \simeq 2\sqrt{\frac{m^2 + p^2}{D}} \quad \text{as } p^2 \rightarrow 0$$

(20)

and leads to the dyadic form for $K$,

$$K(\lambda, N) = lr^+, \quad \lambda = \frac{1}{\sqrt{\lambda}}, \quad r = \sin(\theta/2)^{-1} \begin{pmatrix} 1 & -\sin(N\theta) \\ \lambda & \sin((N + 1/2)\theta) \end{pmatrix}.$$

(21)

(22)

Clearly, the eigenvalues are zero and

$$\exp(-(N + 1)/\tau_{\text{HOR}}(\theta)) = r^\dagger \cdot l = |\lambda \cos((N + 1/2)\theta)|.$$  

(23)

The extra factor $N + 1$ on the left hand side takes into account that $N + 1$ sweeps are required to realize one iteration with $K$. All our $\tau$ values thus refer to units of comparable computational complexity. We can now keep $\tau_{\text{HOR}}(0)/N$ at $O(1)$ if we keep $N\theta$ fixed as $\theta \propto m \rightarrow 0$, which corresponds to $z = 1$ for this pair of modes. This is precisely the optimization strategy found numerically in [5]. We note however, that on large lattices the possible values for $\theta$ are quasi continuous, and the $|\cos|$ in (23) will be very close
to one for some nonzero momenta. Their autocorrelation time then peaks to $2N\tau_{\text{HB}}$. All the momentum dependent correlation times derived so far are shown in Fig.1 for a representative case. This degradation of HOR for isolated frequencies is linked to the fact that the microcanonical updates move these modes by an integer number of cycles. A similar phenomenon has been observed for hybrid Monte Carlo [10]. The way out suggested there also works for HOR. The effect is washed out if one picks $N$ at random, since the momenta of the “slow” modes depends sensitively on $N$. If we make between 1 and $2N - 1$ sweeps with equal probability, such that we have the same work as before on average, and call this time $\bar{\tau}_{\text{HOR}}$, we find

$$\exp(-(N + 1)/\bar{\tau}_{\text{HOR}}(p)) = |\lambda \cos((N + 1/2)\theta) \frac{\sin((N - 1/2)\theta)}{(2N - 1)\sin(\theta/2)}|. \quad (24)$$
This is shown in Fig.2, and now $p = 0$ is the slowest mode and we effectively have $z = 1$ for $\tau_{\text{exp}}$. It is not clear if this randomization of the number of microcanonical sweeps is necessary or beneficial in non gaussian systems, where modes are coupled. An experiment with the asymptotically free $C P^3$ model [7] at $\xi \approx 9$ on a $64^2$ lattice showed absolutely no difference in autocorrelation times.

We conclude now with results on integrated autocorrelation times. For quantities that couple only to the zero momentum mode, like the average field (magnetization) or the susceptibility, they slow down like the corresponding $\tau(0)$ worked out before. A less trivial observable in this context is the total energy, which couples to all modes. Although it is not universal and therefore not interesting for the continuum limit, it is usually measured and analyzed.
to diagnose algorithms. Working out the relevant correlations gives

\[
\Gamma(t) = \lim_{n \to \infty} \left\langle \sum_{p} \chi_t^{(n+1)} h \chi_t^{(n+1)} \sum_{p'} \chi_t^{(n)} h \chi_t^{(n)} \right\rangle = 2 \sum_{p} \text{tr}(M^t h^{-1} M^t h). \quad (25)
\]

We here define the integrated autocorrelation time as

\[
\tau_{\text{int}, E} = \sum_{t=1}^{\infty} \frac{\Gamma(t)}{\Gamma(0)}.
\quad (26)
\]

Its relevance is that \(1 + 2 \tau_{\text{int}, E}\) is precisely the factor by which the variance has to be multiplied to correct error estimates for autocorrelations. For heatbath it is easy to derive

\[
\tau_{\text{int}, E}^{\text{HB}} = \frac{1}{2} \int_{-\pi}^{\pi} \frac{d^D p}{(2\pi)^D} \frac{\lambda^2}{1 - \lambda^4}
\quad (27)
\]
on the infinite lattice. In the limit of vanishing mass we have

\[
\tau_{\text{int}, E}^{\text{HB}} = \begin{cases} 
- (4\pi)^{-1} \log (m) & \text{for } D = 2 \\
0.1595 & \text{for } D = 3 \\
0.0840 & \text{for } D = 4 \\
(4D)^{-1} (1 + O(D^{-2})) & \text{for } D \to \infty
\end{cases}
\quad (28)
\]

and the corresponding exponent \(z\) hence vanishes. This is a drastic example that exponents referring to integrated autocorrelations can depend on the observable and can differ from the fundamental one for \(\tau_{\text{exp}}\) as has been emphasized [11] on principal grounds. The efficient decorrelation of the energy under heatbath updating is, of course, related to the same fact that makes it less physically interesting, namely that it is dominated by short wavelength modes, which evolve fast. For AOR all modes have the same \(\tau(p)\), and so it is not surprising that one also finds (twice) this time in the energy correlations,

\[
\tau_{\text{int}, E}^{\text{AOR}} \approx \frac{\sqrt{D}}{4m}. \quad (29)
\]

For HOR we derive after some algebra

\[
\tau_{\text{int}, E}^{\text{HOR}}/(N+1) = L^{-D} \sum_{p} \frac{\lambda^2}{1 - \lambda^2 \cos^2((N+1/2)\theta)}
\quad (30)
\]
where $N + 1$ on the left hand side again takes into account the complexity of HOR steps. The positive right hand side can be bounded by replacing the \( \cos \) by unity, and then the corresponding integral converges for $D > 2$. We thus have $z = 1$ for $D \geq 2$ for the energy with HOR updating.

To summarize, we have demonstrated that hybrid overrelaxation works equally well as the original more “coherent” method by Adler in the free field case. In order to move all Fourier modes efficiently it is necessary to randomize the length of the microcanonical sections of the evolution, whose average share relative to the ergodic updates has to grow proportional to the correlation length to achieve $z = 1$. The need of randomization may well be an artefact of free fields, but it is worth trying also in other cases.
References

R. Gupta, G. W Kilcup, A. Patel, S. R. Sharpe and Ph. de Forcrand,
S. P. Booth et al. (UKQCD Collaboration), Towards the continuum limit of SU(2) lattice gauge theory,Liverpool preprint LTH 271 (1991)