12-17-2001

Testing a Fourier Accelerated Hybrid Monte Carlo Algorithm

Simon Catterall
Syracuse University

Sergey Karamov
Syracuse University

February 1, 2008

Testing a Fourier Accelerated Hybrid Monte Carlo algorithm.

Simon Catterall and Sergey Karamov
Physics Department, Syracuse University, Syracuse, NY 13244

Abstract

We describe a Fourier accelerated Hybrid Monte Carlo algorithm suitable for dynamical fermion simulations of non-gauge models. We test the algorithm in supersymmetric quantum mechanics viewed as a one-dimensional Euclidean lattice field theory. We find dramatic reductions in the autocorrelation time of the algorithm in comparison to standard HMC.

[0]Corresponding author: Simon Catterall, email: smc@physics.syr.edu

Introduction

Dynamical fermion algorithms play a crucial role in the simulation of lattice field theories. The favorite algorithm for an even number of fermion species is the Hybrid Monte Carlo HMC algorithm [1]. In this paper we introduce and test an improved version of this algorithm in the case of supersymmetric quantum mechanics. We show that the improved algorithm is far superior to the usual HMC procedure in combating the effects of critical slowing down.

We first discuss the usual HMC algorithm and show how it may be generalized to allow for Fourier acceleration. The idea is closely related to the usual Fourier acceleration used for Langevin simulations - both the fields and their conjugate momenta are evolved in momentum space throughout an individual HMC trajectory - the crucial improvement being to choose a wavelength dependent timestep. The momentum dependence of this timestep is chosen so as to render the magnitude of the resulting update of a given Fourier component insensitive to its wavevector index. We test this algorithm in supersymmetric quantum mechanics treated as a Euclidean lattice theory. This model is formulated on the lattice in such way as to leave intact an exact subgroup of the continuum supersymmetry. We are able to demonstrate that the improved algorithm reduces the dynamical critical exponent to values close to zero on lattices as large as $L = 256$ and correlation lengths $\xi \sim 16$.

Hybrid Monte Carlo algorithm.

We will be concerned with simulations of models involving scalar and fermion fields. The typical action we will be discussing takes the general form:

$$S(x, \psi, \overline{\psi}) = S_B(x) + \overline{\psi} M(x) \psi \tag{1}$$

containing a bosonic field x and a Dirac field ψ defined on a lattice in Euclidean space. The fermion matrix M will contain lattice derivative terms together with couplings to the bosonic field x. The partition function for this system is then just

$$Z = \int Dx D\overline{\psi}\psi e^{-S(x,\overline{\psi},\psi)} \tag{2}$$

In order to simulate this action we first replace the fermion field by a bosonic pseudofermion field ϕ whose action is

$$\sum_{ij} \frac{1}{2} \phi_i \left(M^T M \right)^{-1}_{ij} \phi_j$$

This is an exact representation of the original boson effective action in the case where $\det M > 0$ (which will be the case for SUSY QM). The resultant nonlocal action $S(x,\phi)$ can be simulated using the Hybrid Monte Carlo (HMC) algorithm [1]. In the HMC scheme *momentum* fields (p,π) conjugate to (x,ϕ) are added and a *Hamiltonian* H constructed from the original action and additional terms depending on the momenta:

$$H = S + \Delta S$$

$$\Delta S = \sum_i \frac{1}{2}\left(p_i^2 + \pi_i^2\right)$$

The corresponding partition function $Z' = e^{-H}$ is, up to a constant factor, identical to the original Z. The augmented system (x,p,ϕ,π) is now associated with a classical dynamics depending on an auxiliary time variable t

$$\frac{\partial x}{\partial t} = p \quad , \quad \frac{\partial p}{\partial t} = -\frac{\partial H}{\partial x}$$
$$\frac{\partial \phi}{\partial t} = \pi \quad , \quad \frac{\partial \pi}{\partial t} = -\frac{\partial H}{\partial \phi}$$

Introducing a finite time step Δt allows us to simulate this classical evolution and to produce a sequence of configurations $(x(t),\phi(t))$.

If $\Delta t = 0$ then H would be conserved along such a trajectory. As Δt is finite H is not exactly conserved. However a finite length of such an approximate trajectory can be used as a *global* move on the fields (x,ϕ) which may then be subject to a metropolis step based on ΔH. Provided the discrete, classical dynamics is reversible and care is taken to ensure ergodicity the resulting move satisfies detailed balance and hence this dynamics will provide a simulation of Z' and hence also of the original partition function Z.

Ergodicity is taken care of by drawing new momenta from a Gaussian distribution after each trajectory. The reversibility criterion can be satisfied by using a leapfrog integration scheme. Its general form for a field x with associated momentum p can be written as [2]

$$\begin{aligned} x_{\mathbf{r}}(\delta t) &= x_{\mathbf{r}}(0) + \delta t A_{\mathbf{r}\mathbf{r}'} p_{\mathbf{r}'}(0) + \frac{(\delta t)^2}{2} A_{\mathbf{r}\mathbf{r}''} A^T_{\mathbf{r}''\mathbf{r}'} F_{\mathbf{r}'}(0) \\ p_{\mathbf{r}}(\delta t) &= p_{\mathbf{r}}(0) + \frac{\delta t}{2} A^T_{\mathbf{r}\mathbf{r}'} \left(F_{\mathbf{r}'}(0) + F_{\mathbf{r}'}(\delta t)\right) \end{aligned} \tag{3}$$

where $F = -\partial H/\partial x$ is an associated force and A is an arbitrary matrix. Notice our notation – the fields are indexed by a integer vector giving their

lattice position. This update indeed satisfies the time reversibility criterion: inverting the sign of momentum as $p_n(\delta t) = -p_n(\delta t)$ and updating the resultant configuration leads to $\phi_n(2\delta t) = \phi_n(0)$, $p_n(2\delta t) = -p_n(0)$.

To implement our improved algorithm we choose a matrix $A_{\mathbf{r}\mathbf{r}'}$ whose elements depend only on the difference of the lattice vectors between two sites $\mathbf{r}$ and $\mathbf{r}'$. $A_{\mathbf{r}\mathbf{r}'} = a\left(\mathbf{r} - \mathbf{r}'\right)$

We then take the lattice Fourier transform of equation 3. to arrive at an update equation for the Fourier components $\hat{x}_{\mathbf{r}}$ and $\hat{p}_{\mathbf{r}}$. The Fourier components of the field $x_{\mathbf{r}}$ are given by the discrete transform

$$\hat{x}_{\mathbf{k}} = \frac{1}{\sqrt{V}} \sum_{\mathbf{r}} e^{i\mathbf{k}.\mathbf{r}} x_{\mathbf{r}} \tag{4}$$

with similar expressions for $\hat{p}_{\mathbf{k}}$ and $\hat{a}_{\mathbf{k}}$. Because of the lattice convolution theorem the dependence of this momentum space update on the Fourier coefficients $\hat{a}_{\mathbf{k}}$ may then be completely absorbed into the definition of a wavelength dependent time step $\delta t_{\mathbf{k}} = \delta t \hat{a}_{\mathbf{k}}$. The final update equations take the form

$$\begin{aligned} \hat{x}_{\mathbf{k}}(t + \delta t_{\mathbf{k}}) &= \hat{x}_{\mathbf{k}}(t) + \delta t_{\mathbf{k}} \hat{p}_k(t) + \frac{(\delta t_{\mathbf{k}})^2}{2} \hat{F}_{\mathbf{k}}(t) \\ \hat{p}_{\mathbf{k}}(t + \delta t_{\mathbf{k}}) &= \hat{p}_{\mathbf{k}}(t) + \frac{\delta t_{\mathbf{k}}}{2} \left(\hat{F}_{\mathbf{k}}(t) + \hat{F}_{\mathbf{k}}(t + \delta t_{\mathbf{k}}) \right) \end{aligned} \tag{5}$$

The final step in applying Fourier acceleration to this update is to choose the acceleration kernel $\overline{a}_{\mathbf{k}}$ so as eliminate the wavelength dependence of the update in the free theory [3]. For the update of a bosonic field this implies that we should utilize the square root of the momentum space lattice propagator. Such a propagator contains a mass parameter m_{acc} which can be tuned to optimize the update. In the free theory it should clearly be set to the bare lattice mass, but in a general interacting theory it is left as a parameter. In this paper we provide evidence that the autocorrelation time is optimized by setting m_{acc} to the approximate position of the massgap. The fermionic kernel is then chosen to be the inverse of the bosonic kernel.

Notice the occurrence of the square root of the momentum space lattice propagator here - rather than the propagator itself which would be usual in Fourier accelerated Langevin algorithms - this reflects the fact that here the HMC update corresponds to a discrete second order differential equation in time unlike the first order Langevin equation. This choice of the square root propagator was *not* made in [2] and may explain, in part, why acceleration led to only small reductions in the autocorrelation time.

Model

A simple way to demonstrate the effectiveness of this update is to consider supersymmetric quantum mechanics [4] which contains a real scalar field x and two independent real fermionic fields ψ and $\overline{\psi}$ defined on a one-dimensional lattice of L sites with periodic boundary conditions imposed on both scalar and fermion fields:

$$S = \frac{1}{2}\sum_{ij}\left(D_{ij}x_j + P_i\right)\left(D_{ij}x_j + P_i\right) + \frac{1}{2}\sum_{ij}\overline{\psi}_i\left(D_{ij} + P'_{ij}\right)\psi_j \qquad (6)$$

For our simulations the quantity P_i and its derivative are chosen as:

$$P_i = \sum_j K_{ij}x_j + gx_i^3$$

$$P'_{ij} = K_{ij} + 3gx_i^2\delta_{ij}$$

The symmetric difference operator D_{ij} and the Wilson mass matrix K_{ij} are defined as:

$$D_{ij} = \frac{1}{2}\left[\delta_{j,i+1} - \delta_{j,i-1}\right]$$

$$K_{ij} = m\delta_{ij} - \frac{1}{2}\left(\delta_{i,j+1} + \delta_{i,j-1} - 2\delta_{ij}\right)$$

Here dimensionless lattice units are $m = m_{\rm phys}a$, $g = g_{\rm phys}a^2$ and $x = a^{-\frac{1}{2}}x_{\rm phys}$ and the discrete momenta take on the values $2\pi k/L$ with $k = 0\ldots L-1$. Notice that this model employs a non-standard boson action containing not the usual scalar lattice Laplacian but the square of the symmetric difference operator. This is done in order to treat the fermions and bosons in a symmetric manner - indeed because of this the action is invariant under a single SUSY-like symmetry.

$$\begin{aligned}\delta x_i &= \psi_i\xi \\ \delta\overline{\psi}_i &= (D_{ij}x_j + P_i)\xi \\ \delta\psi_i &= 0\end{aligned}$$

Doubles in both bosonic and fermionic sectors are eliminated by means of the Wilson mass term K. The physics results from this study were published in [5].

For bosonic and pseudo-fermionic field updates respectively we use the following timesteps which are simple inverses of each other.

$$\delta t_k^B = \Delta t(m_{acc}+2)/\sqrt{\sin^2(2\pi k/L) + (m_{acc} + 2\sin^2(\pi k/L))^2}$$

$$\delta t_k^F = \Delta t\sqrt{\sin^2(2\pi k/L) + (m_{acc} + 2\sin^2(\pi k/L))^2}/(m_{acc}+2)$$

Autocorrelation time

Suppose that Q is some observable and Q_t is a measurement of Q in configuration corresponding to Monte-Carlo time t. Then the autocorrelation function is defined as:

$$c(t) = \frac{< Q_0 Q_t > - < Q >^2}{< Q^2 > - < Q >^2}$$

Typically the function $c(t)$ is approximately exponential $c(t) = e^{-t/\tau}$ where τ is the autocorrelation time - the time between decorrelated configurations. Clearly we can define τ also from the relation

$$\tau = \int_0^\infty e^{-t/\tau} dt = \sum_t c(t)$$

and this yields a robust way to estimate τ even when $c(t)$ is not exactly an exponential - measured this way it is sometimes referred to as the integrated autocorrelation time. It also provides a convenient way to find the autocorrelation time from a sequence of N Monte Carlo measurements

$$\tau = \sum_{i=1}^{N} \sum_{j=i}^{j<M} \frac{Q(i)Q(j) - < Q >^2}{< Q^2 > - < Q >^2} \tag{7}$$

where we measure the autocorrelation function only out to M steps (we must choose $M >> \tau$).

The efficiency of different simulation algorithms can be described in terms of the dynamic critical exponent z:

$$\tau \sim (1/M_{\text{lattice}})^z = (L/M_{\text{physical}})^z \sim L^z$$

An algorithm which resulted in $z = 0$ is said to suffer no critical slowing down while most local algorithms such as conventional HMC typically yield $z \sim 2$.

Fig. 1 represents the autocorrelation time τ as a function of m_{acc} for lattice SUSY QM with $m = 10$, $g = 100$ for a lattice with $L = 64$ sites. Notice that the dimensionless parameter characterizing the interaction strength g/m^2 is unity in this case - we are in a strong coupling regime. The massgap of the theory for these values of the bare parameters can be calculated via Hamiltonian methods and yields a value $m_g \sim 16.0$.

For large $m_{\text{acc}} \to \infty$ the update can be shown to reduce to the usual local HMC algorithm and, as the plot makes clear, leads to a maximal value

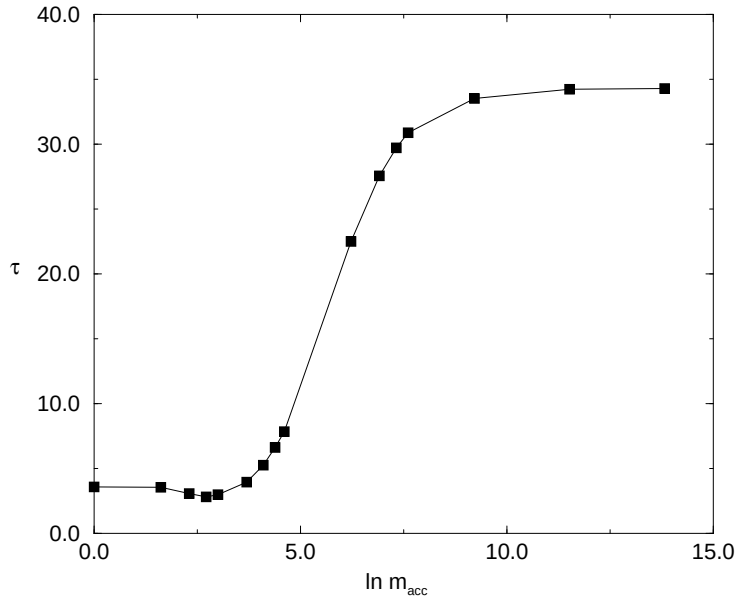

Figure 1: $\ln \tau$ vs $\ln m_{acc}$ for Supersymmetric Quantum Mechanics with $L = 64$, $m = 10$, $g = 100$.

for τ. Indeed the minimal value of τ is can be seen to be achieved by setting $m_{\rm acc}$ approximately equal to the massgap of the model $\ln m_{acc} \sim 2.7$. In the vicinity of this point, one can see that the minimum autocorrelation $\tau_{\rm min}$ is more than an order of magnitude smaller than the τ achieved by the usual local algorithm for this lattice size.

To derive a dynamic critical exponent we have also compared the ddependence of the autocorrelation time on lattice size L both for the usual local HMC and the Fourier accelerated algorithm with $m$acc $= 15.0$. For the case $m_{\rm acc} = \infty$ fig 2 shows a plot of $\ln \tau$ against $\ln L$ together with a straight line fit yielding $z \sim 1.7$. The plot makes it clear that this is a lower bound for z so that in all likelihood z achieves a value of $z = 2$ for very large lattices. This is to be contrasted to fig 3 which shows the same plot for $m_{\rm acc} = 15$. A fit in this case yields a value of z consistent with zero. For the largest lattice size $L = 256$ we see that more than two orders of magnitude decrease in τ are gained by use of the accelerated algorithm.

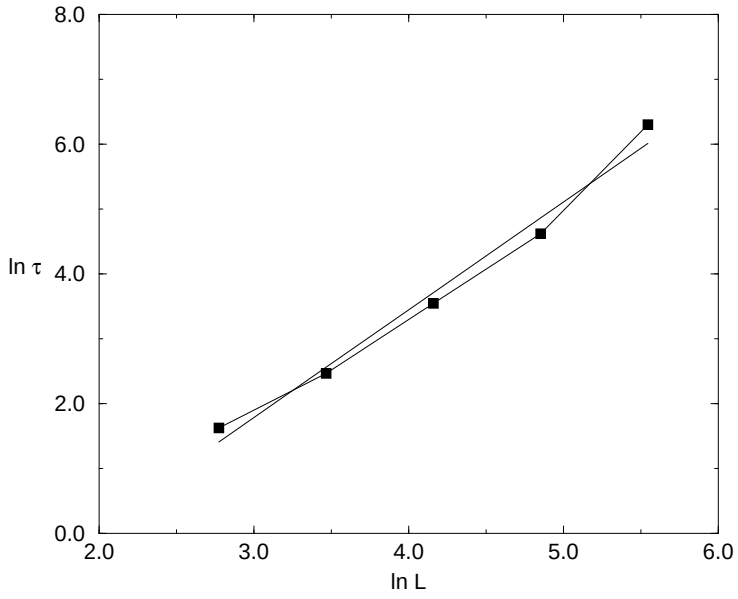

Figure 2: $\ln \tau$ vs $\ln L$ for Supersymmetric Quantum Mechanics with $m_{acc} \sim \infty$, $m = 10$, $g = 100$. Straight line fit yields $z = 1.7$

Discussion

We have shown that the use of Fourier acceleration can yield large payoffs in the simulation of non-gauge field theories with dynamical fermions. We have presented results which support this conclusion for supersymmetric quantum mechanics in which the HMC algorithms can be pushed to large ($L = 256$) lattice size and correlation length ($\xi \sim 16$). We have obtained similar, though less quantitative, results for two-dimensional Wess-Zumino models [6].

Acknowledgements

Simon Catterall was supported in part by DOE grant DE-FG02-85ER40237

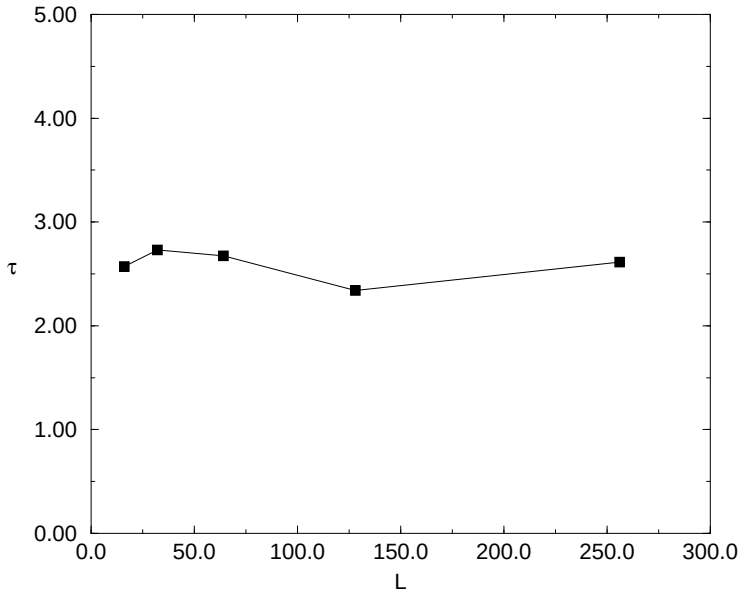

Figure 3: $\ln \tau$ vs $\ln L$ for Supersymmetric Quantum Mechanics with $m_{acc} = 15$, $m = 10$, $g = 100$.

References

[1] S. Duane, A. Kennedy, B. Pendleton and D. Roweth, Phys. Lett. B195B (1987) 216.

[2] D. Espriu, V. Koulovassilopoulos and A. Travesset, Phys. Rev. D56 (1997) 6885.

[3] G. Katz, G. Batrouni, C. Davies, A. Kronfeld, P. Lepage, P. Rossi, B. Svetitsky and K. Wilson, Phys. Rev. D37:1589, 1988.

[4] E. Witten, Nucl. Phys. B185 (1981) 513.

[5] S. Catterall and E. Gregory, Phys. Lett. B487 (2000) 349.

[6] S. Catterall and S. Karamov, hep-lat/0108024, submitted to Phys. Rev. D.