PROCEEDINGS OF THE 6TH
JOINT EPS-APS INTERNATIONAL
CONFERENCE ON PHYSICS COMPUTING

PC '94

Physics Computing '94

Palazzo dei Congressi
Lugano, Switzerland

22-26 August 1994

EDITORS:

Ralf Gruber and Marco Tomassini

Swiss Scientific Computing Centre
Manno, Switzerland

AMERICAN PHYSICAL SOCIETY

EUROPEAN PHYSICAL SOCIETY

A General Class of Hybrid Monte-Carlo Methods

Raúl Toral
Departament de Física, Universitat de les Illes Balears, 07071-Palma de Mallorca, Spain.
A.L. Ferreira
Departamento de Física, Universidade de Aveiro, 3800 Aveiro, Portugal

Abstract

We propose a simple modification of the Hybrid Monte-Carlo method to sample equilibrium distributions of continuous field models. Applications include the simulation of systems with a conserved order parameter and an efficient implementation of the Fourier acceleration scheme.

A large class of problems can be reduced to computing averages according to a statistical distribution $\exp(-\mathcal{H})$, where the Hamiltonian function $\mathcal{H}([\phi])$ depends on the $N = L^d$ scalar variables $[\phi] \equiv (\phi_1, \ldots, \phi_N)$. Many numerical methods, such as Molecular Dynamics, Langevin integration and Monte–Carlo (MC), have been used to study these equilibrium properties of field–theoretical models[1]. Monte–Carlo methods have the advantage that their only errors are of statistical origin and can be consequently decreased by increasing the number of samplings, whereas Molecular Dynamics and Langevin suffer from systematic errors coming from the finite time–step used in the numerical integration of the equations of motion. The main problem of Monte–Carlo simulations is that in some cases, and specially near a second order phase transition, the number of configurations necessary to achieve a given small error is very large (usually, grows as some power of the system size N) thus requiring too much computer time. This is due to the phenomenon of critical slowing down (CSD) that dramatically increases correlations near a critical point. More specifically, CSD theory[2] tells us that near second–order phase transitions the correlation time, $\tau_{\hat{O}}$, of a measured observable $\hat{O}$, increases as $\tau_{\hat{O}} \sim \xi^z$, being ξ the correlation length and z the dynamical critical exponent. $\tau_{\hat{O}}$ can be defined as some measure of the relaxation time of the correlation function of the observable $\hat{O}$:

$$C_{\hat{O}}(t) = \frac{< \hat{O}(t)\hat{O}(0) > - < \hat{O}(t) >< \hat{O}(0) >}{< \hat{O}^2(0) > - < \hat{O}(0) >^2} \tag{1}$$

For finite systems close enough to the critical point, the correlation time increases with system side as $\tau_{\hat{O}} \sim L^z$, according to finite size scaling theory.

For the local updating schemes such as heat-bath or Metropolis the exponent z being near 2 strongly demands on computer time. For spin models the collective updating scheme of Swendsen and Wang [3] has proven quite successful in reducing the dynamical critical exponent and overcoming CSD. For continuous field models, several methods such as Multigrid Monte-Carlo[4], time-step matrix Langevin method[5] or overrelaxation (see [6] and references therein), have been proposed, although is not still clear if some reducing can be achieved in non–trivial models such as the ϕ^4 model.

In its simplest form, MC introduces a stochastic dynamics which involves the proposal of a new field configuration plus an acceptance/rejection step. Ideally, every proposal should be completely uncorrelated from the previous one. However, it is a major problem how to generate essentially independent configurations and still keep the acceptance probability within some reasonable limits. The Hybrid Monte Carlo (HMC) algorithm [7], uses the global update given by Molecular Dynamics to generate a new configuration which is then accepted or rejected according to the usual rules of Monte–Carlo. We now describe a generalisation of the standard hybrid Monte–Carlo method. The evolution is performed by an appropriate numerical integration

of the corresponding system dynamics. To this end, a set of conjugate momenta variables $[p] \equiv (p_1, \ldots, p_N)$ associated with a kinetic energy $\mathcal{H}_K = \sum_{i=1}^{N} p_i^2/2$ is introduced. The variables p_i can be in general a vector variable with D components, $p_i = (p_i^1, p_i^2, \ldots, p_i^D)$. The total Hamiltonian is $\hat{\mathcal{H}} = \mathcal{H} + \mathcal{H}_K$. We consider the following dynamics written in compact vector notation:

$$\begin{aligned} \frac{d\phi}{dt} &= \sum_{s=1}^{D} \mathcal{A}^s p^s \\ \frac{dp^s}{dt} &= (\mathcal{A}^s)^T F \qquad , s = 1, \ldots, D \end{aligned} \tag{2}$$

here the $\mathcal{A}^s$ are some linear operators which can be represented as a matrix, and F_j represents the force as computed from the Hamiltonian $-\frac{\partial}{\partial \phi_j}\mathcal{H}$. The standard HMC can be obtained simply by setting $D = 1$ and $\mathcal{A}$ equal to the identity operator. It is easily verified that, for arbitrary matrices $\mathcal{A}$, the evolution equations (2) exactly conserve energy, i.e., $d\hat{\mathcal{H}}/dt = 0$. For the numerical integration of the previous equations of motion the "leap-frog" scheme can be used:

$$\begin{aligned} \phi' &= \phi + \delta t \sum_{s=1}^{D} \mathcal{A}^s p^s + \frac{(\delta t)^2}{2} \sum_{s=1}^{D} \mathcal{A}^s (\mathcal{A}^s)^T F([\phi]) \\ p'^s &= p^s + \frac{\delta t}{2} (\mathcal{A}^s)^T \left(F([\phi]) + F([\phi']) \right) \end{aligned} \tag{3}$$

The total energy, as a result of the time discretization used in the leap–frog scheme, is no longer conserved and its variation $\Delta\hat{\mathcal{H}}$ can be controlled by varying δt. Detailed balance is satisfied if the configuration obtained after evolving n time steps is accepted with probability $\min[1, \exp(-\Delta\hat{\mathcal{H}})]$. As in the standard HMC the momenta variables are refreshed after every acceptance/rejection step according to the Gaussian distribution of independent variables $\exp(-\mathcal{H}_K)$. The evolution given by n leap–frog steps and the acceptance/rejection step constitute what is called 1 MC trial.

The presence of arbitrary matrices $\mathcal{A}$ is closely related to the matrix–time–step integration techniques for the Langevin equation introduced in reference [5]. In fact, it can be easily shown that our evolution equations for $n = 1$ coincide exactly with the Euler scheme for solving the Langevin evolution equations. The important part, however, is that the presence of an acceptance/rejection step makes the hybrid Monte–Carlo approach "exact", in the sense that averages are not biased by the choice of an arbitrarily large time step.

We have introduced a class of HMC methods characterized by the matrices $\mathcal{A}^s$. One can then choose the matrices $\mathcal{A}^s$ that better suit a particular problem. We will show two applications: simulation of conserved order parameter systems[8] and implementation of Fourier acceleration.

The case of a system whose order parameter $\Phi = \sum_i \phi_i$ is conserved can be simulated by introducing this new conservation law via the matrices $\mathcal{A}^k$. For that purpose, we choose $D = d$ (the system dimensionality) and the operator $\mathcal{A}^k$ equal to the $k - th$ component of the lattice gradient operator $\vec{\nabla}_L$. The condition $D = d$ is necessary in order not to introduce any spurious conservation laws in the system. It can be shown that the corresponding leap–frog evolution equations:

$$\begin{aligned} \phi' &= \phi - \frac{(\delta t)^2}{2} \vec{\nabla}_L^2 F([\phi]) + \delta t \; \vec{\nabla}_L \cdot \vec{p} \\ \vec{p}' &= \vec{p} - \frac{\delta t}{2} \vec{\nabla}_L^T \cdot \left(F([\phi]) + F([\phi']) \right) \end{aligned} \tag{4}$$

exactly conserve the order parameter. Moreover, the previous equation for the field ϕ is exactly the Euler scheme (apart from a trivial time rescaling) for the solution of the celebrated Cahn–Hilliard-Cook equation[9] for the study of the dynamics of systems with a conserved order parameter.

Fourier acceleration turns out to be a particular case of our general equations in which the matrix $\mathcal{A}$ is chosen to be diagonal in Fourier space. A simple example is given by the Gaussian model, defined by the

following Hamiltonian:

$$\mathcal{H} = \sum_{i=1}^{N} \left[\frac{\mu}{2} \phi_i^2 + \frac{1}{2} \mid \vec{\nabla}_L \phi_i \mid^2 \right] \tag{5}$$

index i runs over the $N = L^2$ sites of a 2-dimensional square lattice, with periodic boundary conditions (a similar analysis can be carried out in any spatial dimension but we refer to the case d=2 for simplicity). Introducing the associated momenta (scalar, $D = 1$) variables, the total Hamiltonian $\hat{\mathcal{H}}$ in terms of the Fourier transform of fields and momenta, $\hat{\phi}_k$ and $\hat{p}_k$ reads:

$$\hat{\mathcal{H}} = \sum_{k=1}^{N} \left[\frac{\omega_k^2}{2} |\hat{\phi}_k|^2 + \frac{1}{2} |\hat{p}_k|^2 \right] \tag{6}$$

where $\omega_k^2 = \mu + 4\big(\sin^2(k_x/2) + \sin^2(k_y/2)\big)$.

If we choose for $\mathcal{A}$ a diagonal matrix in Fourier space of diagonal elements $\hat{A}_k$, the leap–frog algorithm in Fourier space reads:

$$\begin{aligned} \omega_k \hat{\phi}_k(\delta t) &= (1 - \frac{c_k^2}{2})\, \omega_k \hat{\phi}_k(0) + c_k\, \hat{p}_k(0) \\ \hat{p}_k(\delta t) &= -\frac{c_k}{2}(2 - \frac{c_k^2}{2})\, \omega_k \hat{\phi}_k(0) + (1 - \frac{c_k^2}{2})\, \hat{p}_k(0) \end{aligned} \tag{7}$$

where $c_k = \hat{A}_k \omega_k \delta t$.

We now turn to the question of the optimal choice for the matrix $\mathcal{A}$. From equation (7) it is immediately seen that if we choose the matrix $\mathcal{A}$ such that $\hat{A}_k = 1/\omega_k$ the iteration equations get independent of the mass μ and all the modes are equally updated. This is, in effect, an exact implementation of the method of Fourier acceleration. This choice of the matrix clearly reduces completely CSD ($z = 0$) in the sense that correlation times are independent of the mass even when the mass goes to zero and the model becomes critical.

By using the evolution equations together with the assumption that the field variables $\hat{\phi}_k(0)$ are in thermal equilibrium and, therefore, follow the distribution $\exp(-\hat{\mathcal{H}})$, one can compute the equilibrium average discretization error as:

$$< \hat{\mathcal{H}}(n\delta t) - \hat{\mathcal{H}}(0) > \equiv < \Delta \hat{\mathcal{H}} > = \frac{N \delta t^4}{32 - 8\delta t^2}\, \sin^2(n\theta) \tag{8}$$

with $\theta = \cos^{-1}(1 - \frac{\delta t^2}{2})$. This is related to the average acceptance probability p by the relation[10]: $p = \mathrm{erfc}(\frac{1}{2}\sqrt{< \Delta \hat{\mathcal{H}} >})$

By making the assumption that the probability to accept or reject the whole configuration at a given step is equal to p and independent of the previous time-evolution of the system, one obtains the correlation functions for the magnetization:

$$C_M(m) \equiv C_M(m\ n\delta t) = \left[1 - 2p \sin^2(\frac{n\theta}{2}) \right]^m \tag{9}$$

and the energy:

$$C_{\mathcal{H}}(m) \equiv C_{\mathcal{H}}(m\ n\delta t) = [1 - p \sin^2(n\theta)]^m \tag{10}$$

Both correlation functions decay exponentially and a measure of the correlation time is given simply by $\tau_M = -1/\log(|C_M(1)|)$ and $\tau_{\mathcal{H}} = -1/\log(C_{\mathcal{H}}(1))$. In figure 1 we compare both correlation times given by the above expressions with simulation results for the case $L = 32$ and $n = 4$ as a function of δt. We observe a rather good agreement between the analytical expression and the simulation results. We have used the shown minimum in the correlation time of the energy as a function of δt to find the optimal n for a given system size that minimizes the computer effort $(2\tau_{\mathcal{H}} + 1)n$. It can be shown numerically that the optimal values scale as $n \sim L^{1/2}$ and $\delta t \sim L^{-1/2}$. For large L, the corresponding correlation times τ_M and $\tau_{\mathcal{H}}$ turn out to be L–independent and are given by $\tau_M = 2.5$, $\tau_{\mathcal{H}} = 1.5$ approximately. The optimal acceptance

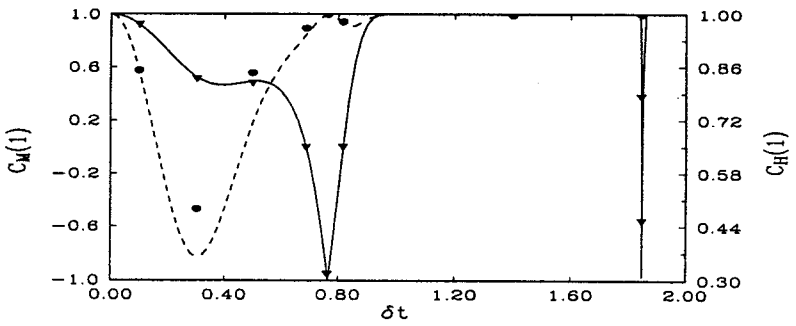

Figure 1: Comparison of $C_M(1)$ (continuos line) and $C_{\mathcal{H}}(1)$ (dashed line) obtained from eqs. (9) and (10) for a system of size $L = 32$ and $n = 4$ with simulation results.

probability $p \approx 0.67$ and the product $n\delta t \approx 1$ are also independent of the system size. So, the time needed to get an independent configuration is constant in the natural units of the dynamics, but the computer effort measured by the number n of updates needed to get 1 Monte Carlo step increases as $L^{1/2}$.

Financial support from the Dirección General de Investigación Científica y Técnica (Grant No. PB-92-0046, Spain) and JNICT Portugal, are acknowledged. We thank B. Mehlig for fruitful discussions.

References

[1] See, for example, K. Binder (ed) *The Monte-Carlo Method in Condensed Matter Physics*, Topics App. Phys. Vol 71 (Springer, Berlin, Heidelberg 1992). A D Sokal Nuclear Physics B (Proc. Suppl.) **20** (1991) 55.

[2] P C Hohenberg and B Halperin, Rev. Mod. Phys. **49**, 435, (1977)

[3] R H Swendsen and J S Wang, Phys. Rev. Lett. **58** (1987) 86.

[4] J Goodman and A D Sokal, Phys. Rev. **D 40** (1989) 2035.

[5] G.G. Batrouni, G. R. Katz, A.S. Kronfled, G.P. Lepage, B. Svetitsky and K.G. Wilson, Phys. Rev. **D 32** (1985) 2736.

[6] D Kennedy, Nuclear Physics B (Proc. Suppl.) **30** (1993) 96.

[7] S. Duane, A. D. Kennedy, B. J. Pendleton and D. Roweth, Phys. Lett. **B 195** (1987) 216.

[8] A L Ferreira and R Toral, Phys. Rev. **E 47** (1993) R3848.

[9] J.D. Gunton. M. San Miguel and P. S. Sahni, in *Phase Transitions and Critical Phenomena*, edited by C. Doamb and J.L. Lebowits (Academic, London, 1983), Vol. 8.

[10] S. Gupta, A. Irback, F. Karsch and B. Petersson, Phys. Lett. **B 242** (1990) 437.