DYNAMICAL CRITICAL PROPERTIES OF THE HYBRID MONTE CARLO ALGORITHM

Sourendu Gupta

Theory Division, CERN, CH-1211, Geneva 23, Switzerland.

ABSTRACT

We study the dynamical critical behaviour of the Hybrid Monte Carlo algorithm for the two-dimensional XY model. We observe that randomising trajectory lengths decreases computational effort dramatically in the spin-wave phase, but increases it in the vortex phase. We also find that the dynamical critical exponent is close to 2 in the Langevin limit. When sufficiently long trajectories are used, this exponent drops to 1 in the vortex phase, and to zero in the spin-wave phase. The computational effort in molecular dynamics units, on the other hand, scales with an exponent 1. We present an argument for the generality of this observation.

CERN-TH.6178/91
August 1991

1. Introduction.

In this paper we investigate autocorrelations in the Hybrid Monte Carlo (HMC) algorithm [1]. This, and the related Hybrid Stochastic (HS) algorithm [2], have become the methods of choice in simulations of QCD with dynamical fermions. Till now the dynamical properties of HMC have been studied in detail only for free field theory [3,4], and generalisations to other models have been based on some early studies of the HS algorithm [5].

Here we report simulations of an interacting theory—the two-dimensional XY model. Although this model contains no fermions, some of the systematics observed in QCD have their analogues here. We study them in detail, with particular attention paid to methods of removing some of the more undesirable features. We test criteria for optimising runs with respect to CPU time spent in decorrelating long-distance observables. In addition we perform a very high statistics measurement of dynamical critical exponents.

The HMC algorithm is used to simulate a lattice theory defined by an action $S[\{\phi\}]$ for fields ϕ . Extra fields π are introduced quadratically in the action, and interpreted as ‘momenta’ conjugate to the original fields ϕ . The joint evolution in a fictitious time is determined by the equations of motion arising from the Hamiltonian

$$H = \frac{1}{2} \sum_i \pi_i^2 + S[\{\phi_i\}]. \quad (1)$$

The Monte Carlo procedure consists of choosing $\{\pi_i\}$ randomly from a Gaussian distribution, and evolving the system through a finite step approximation of the equations of motion. The change of energy in one such trajectory, ΔH , is then used in a Metropolis question either to accept or reject the configuration so generated. There are two algorithmic parameters— Δt is the step size in time, and N_{md} is the number of steps per trajectory. Instead of N_{md} one could choose to quote the trajectory length $T = N_{md}\Delta t$. We restrict our attention to this minimal algorithm, since it is the one most widely used in modern simulations of QCD.

The natural unit for the autocorrelation time is a trajectory, since a new configuration is generated at the end of each. The CPU time involved in generating such a trajectory is proportional to N_{md} . In this paper we shall report both the integrated autocorrelation times τ for different operators and the effort $E = N_{md}\tau$. Since the amount of CPU time spent on a trajectory is proportional to N_{md} , E is a direct measure of the time spent in obtaining one statistically uncorrelated

configuration. We shall concentrate only on operators which transform as singlets under global symmetries of the action. Other operators can be completely decorrelated by taking any valid algorithm and adding a random global symmetry transformation at each step.

We summarise here the main results known about the dynamics of the HMC for the free field theory (FFT) in d dimensions, with action

$$S = \frac{m^2}{2} \sum_i \phi_i^2 + \frac{1}{2} \sum_{\langle ij \rangle} (\phi_i - \phi_j)^2, \quad (2)$$

where $\langle ij \rangle$ denotes a sum over nearest neighbours, and ϕ_i are real fields. The equations of motion are completely decoupled into independent harmonic oscillators by making the canonical transformation generated by

$$G = \pi_k M_{k,l} \phi_l, \quad \text{where} \quad M_{k,l} = \frac{1}{L^{d/2}} e^{ik \cdot l}. \quad (3)$$

This takes the site variables ϕ_l into the Fourier modes ϕ_k . Each mode has a frequency $\omega_k^2 = m^2 + 4 \sum_\mu \sin^2 k_\mu/2$. It was observed in Ref. 3 that if the trajectory length T is constant, then those modes for which $T\omega_k/\pi$ is close to an integer develop very large correlation lengths. It was shown that this problem could be avoided by randomising trajectory lengths. In Refs. 2,4 it was shown that, if trajectory lengths are randomly drawn from a distribution with mean length $\bar{T}$, then the autocorrelation time $\tau(k)$ for a mode k is given by

$$\tau(k) = \begin{cases} 2 \left[1 - \sqrt{1 - (2\omega_k \bar{T})^2} \right]^{-1} & (\bar{T} < 1/2\omega_k) \\ 2 & (\bar{T} \geq 1/2\omega_k). \end{cases} \quad (4)$$

$\tau(k)$ is minimised at $\bar{T} = 1/2\omega_k$. An explicit check shows that $\tau(0)$ is irrelevant for all operators which transform as singlets under the global symmetry of the action in Eq. (2). Hence the slowest modes on an L^d lattice are those with one of the momentum components equal to $2\pi/L$ and the others 0. These modes are optimised by taking $\bar{T} = L/2\sqrt{m^2 L^2 + 4\pi^2}$. Then E scales linearly with L at fixed mL , or as $1/m$ at fixed L when $mL \gg 1$. For much smaller $\bar{T}$, both E and τ scale as L^2 or $1/m^2$ in these two regimes.

For the 2d XY model, we find that τ has a shallow minimum when the acceptance is between 60 and 80%. On keeping the acceptance fixed and varying the

trajectory length, we find that autocorrelation times and the computational effort decrease to a minimum before beginning on a nearly linear rise. However, unlike for FFT, the minima of τ and E need not occur at the same trajectory length. Since optimal trajectory lengths for long-distance operators are significantly larger than those for short-distance operators, we concentrate on the former. For very short trajectories τ scales roughly as ξ^2 (where ξ is the correlation length). When the trajectory length is chosen to minimise τ , then it scales roughly as ξ in the vortex phase and is independent of L , at fixed β , in the spin-wave phase. On the other hand, E scales with exponent 1. These observations suggest that Δt can be fixed by a choice of acceptance, and N_{md} independently chosen to optimise autocorrelations. The scale of the optimal trajectory length is set by the time required for a local disturbance to traverse the lattice. This provides a natural explanation for the linear scaling of effort, and is, conversely, a testable argument for the generality of this observation. In the vortex phase, and for short trajectories in the spin-wave phase, we find that it is advisable not to randomise trajectory lengths; whereas for long trajectories in the spin-wave phase it seems necessary to do so.

The plan of this paper is as follows. In section 2 we present the model and the details of the simulation. Autocorrelations and the tuning of parameters are dealt with in section 3. Measurements of the dynamical critical exponent are discussed in section 4, and conclusions presented in section 5.

2. The XY Model and the Simulations.

We have applied the HMC algorithm to the two-dimensional XY model. The fields U_i at sites i of an L^2 lattice are elements of $U(1)$, and the action is

$$S = \beta \sum_{\langle ij \rangle} (1 - \text{Re} U_i U_j^\dagger), \quad (5)$$

The model is known to have a Kosterlitz-Thouless transition between a vortex phase (at low β) and a spin-wave or perturbative phase (at high β). Several high-statistics studies [6] have located the transition at $\beta_{KT} \approx 1.1$. The model is critical at every $\beta > \beta_{KT}$, with continuously varying critical exponents. Hence for each β one obtains a different continuum theory.

Using the group derivatives defined in Ref. 7, the equations of motion are found to be

$$\partial_t U_i = i\pi_i U_i, \quad \text{and} \quad \partial_t \pi_i = -\beta \sum_{\mu} \text{Im} U_i (U_{i+\mu}^\dagger + U_{i-\mu}^\dagger). \quad (6)$$

These equations are solved by a leap-frog algorithm with a time step Δt . It has been shown [8,9] that these integration schemes conserve phase space volume at each step.

We have measured the average action, $\langle s \rangle$, the specific heat, c_v , and the magnetic susceptibility, χ . In terms of the action density in each configuration,

$$s = \frac{1}{V} \sum_{\langle ij \rangle} \text{Re} U_i U_j^\dagger, \quad (7)$$

the specific heat can be written as

$$c_v = V \langle (s - \langle s \rangle)^2 \rangle. \quad (8)$$

Similarly, in terms of the non-singlet magnetisation density

$$m = \frac{1}{V} \sum_i U_i, \quad (9)$$

the magnetic susceptibility χ can be written as

$$\chi = V \langle |m|^2 \rangle. \quad (10)$$

Using the fluctuation-dissipation theorem, χ can be expressed as a sum over the spin-spin correlation function at all distances. In the spin-wave phase χ is expected to diverge in the thermodynamic limit.

The runs have been made on lattices of size 16^2 , 32^2 and 64^2 . The program performs well—it takes 0.8 microseconds of CPU time per leap-frog step per spin on a Cray-XMP, and achieves about 170 million vector operations per second. Nevertheless, the actual update can be fairly slow. For example, with a trajectory length $N_{md} = 400$ on a 64^2 lattice, one new configuration is obtained every 1.3 seconds. For reasons given in the next section, we have performed runs with fixed trajectory lengths in the vortex phase, and randomised them in the spin-wave phase.

On all lattices up to size 64^2 we have statistics of between 10^4 and 10^5 autocorrelation lengths. All the averages have been obtained after cutting at least 50 autocorrelation lengths of data for thermalisation. The details of the runs in the vortex phase of the model are shown in Tables 1 and 2. There is good agreement between these results and those in Ref. 6. For the runs in the spin-wave phase, the corresponding details are shown in Table 3.

In Fig. 1 we compare data from our simulations to the relation [10] between the observed acceptance probability, P_{acc} , and the average energy violation, $\langle \Delta H \rangle$, given by the formula

$$P_{acc} = \text{erfc} \left(\frac{1}{2} \sqrt{\langle \Delta H \rangle} \right). \quad (11)$$

From Tables 1–3, it can be checked that, at fixed β , when the trajectory length becomes of order 1, keeping constant P_{acc} for changing lattice size L requires that Δt be scaled as $L^{-d/4}$. We also show in Fig. 2 a phenomenon already noticed in QCD—the saturation of P_{acc} as a function of N_{md} , when Δt , β and L are fixed. This behaviour follows from very general arguments when the Hamiltonian and all its derivatives are bounded.

It is clear from the data that finite size effects are most easily seen in χ . In common with earlier measurements [6] we find that these effects are small when $\xi \leq L/4$. In our measurements of the dynamical critical exponent in the vortex phase, we avoid doing a finite-size scaling analysis by restricting ourselves to this region of the physics.

3. Systematics of Autocorrelations

The autocorrelation function for a measurement of A is defined by

$$C(t) = N \langle (A_i - \langle A \rangle)(A_{i+t} - \langle A \rangle) \rangle, \quad (12)$$

where $\langle A \rangle$ is the average over the set of measurements, and the constant N is fixed by the requirement $C(0) = 1$. At large t , $C(t) \sim \exp[-t/\tau_{exp}]$, and a measurement of the decay rate gives the exponential autocorrelation time τ_{exp} . This is the autocorrelation time for the slowest mode that couples to A . When τ_{exp} is short, then its measurement is prone to errors. For this reason we prefer to measure the integrated autocorrelation time, defined through the relation

$$\tau_{int} = \sum_{t>0} C(t). \quad (13)$$

A further advantage of this measure of the autocorrelation time is that it is directly related to error estimates on the expectation value of A , since an unbiased estimate of the variance in the measurement, σ_A^2 , is given in terms of the naïve estimator σ_0^2 by

$$\sigma_A^2 = \sigma_0^2(1 + 2\tau_{int}). \quad (14)$$

In the rest of this paper τ refers to the integrated autocorrelation time.

We have measured τ for $\langle s \rangle$, c_v and χ . In each case, we have followed the correlation function self-consistently to 10τ , and obtained a stable value for the measurement using a jack-knife estimator with up to 10 blocks. We find that the measured values of τ are stable against the cut on thermalisation and the number of jack-knife blocks. Halving the data sample, in all cases, increases the error estimates by 50%, but does not lead to a significant drift in the estimates of τ . In Fig. 3 we show how the average value and the error estimate of τ move as the amount of statistics is changed. The near constancy of the average, and the fact that error estimates scale inversely as the square root of the statistics, strongly suggests that it is possible to measure such small autocorrelation times rather reliably.

3.1. Tuning of parameters

For fixed N_{md} the dependence of τ on P_{acc} is shown in Fig. 4. There is a shallow minimum in τ for $0.6 \leq P_{acc} \leq 0.8$. The existence of such a minimum is, of course, expected. The pleasant aspect of this curve is that the minimum is not very sharp, implying that P_{acc} does not have to be fine-tuned to yield low autocorrelations.

The dependence of E_s and E_χ on N_{md} at fixed Δt is shown in Fig. 5 in both phases of the model. There is a fairly sharp minimum in E . The initial decrease of E with N_{md} implies that the exact Langevin algorithm demands larger computational effort than HMC, as has been noticed before [11]. At the optimum, one gains more than a factor of 10 in effort compared to $N_{md}=1$. On the other hand, the eventual increase in E implies that directed motion in phase space grows progressively less efficient in killing autocorrelations. A comparison of figures 5 and 6 shows that the minima of E and τ do not necessarily occur at the same trajectory length. Our results on the scaling of these two quantities, presented later, nevertheless indicate that, asymptotically, the minima of E and τ must occur at the same value of N_{md} .

We observe that E is minimised for short-ranged operators (like the action), at trajectory lengths much smaller than required for long-distance measurements (like χ). For the latter, the optimal trajectory length T_{opt} , is clearly L -dependent. It varies from about 1 to 10 on the lattices studied and, at fixed β , increases linearly with L . We show this in Fig. 7 for measurements made in the spin-wave

phase. The same behaviour is seen in the vortex phase.

3.2. Randomising trajectory lengths.

In order to study phase locking, we have measured the autocorrelations $\tau(k)$ of the Fourier modes k of the angle variables ϕ_l , where $U_l = \exp(i\phi_l)$. Runs in the vortex phase with a fixed trajectory length do not give rise to very long autocorrelation times in isolated Fourier modes. Nor is this behaviour seen for simulations in the spin-wave phase when the trajectory length is small. In Fig. 8, we show $\tau(k)$ as a function of k for simulations in these two regimes. For convenience we have shown only the autocorrelations of modes with one of the components of k equal to zero. Narrow peaks in $\tau(k)$ are completely absent. Nevertheless, if one randomises trajectory lengths by drawing them from an exponential distribution, then the effort at any value of $\langle N_{md} \rangle$ increases. For example, at the optimal trajectory length for $\beta = 0.78$ on a 32^2 lattice, we find

$$E_X = \begin{cases} 54 \pm 3 & (N_{md} = 50) \\ 100 \pm 9 & (\langle N_{md} \rangle = 50) \end{cases} \quad (15)$$

The reason for this is clear. When drawing trajectory lengths at random, they are chosen off-optimal sufficiently often to increase the effort significantly.

The lack of phase locking at small β can be understood through a strong coupling expansion. It is clear that as $\beta \rightarrow 0$, the equations of motion decouple, and each site-variable becomes a rigid rotor. For such a system, the frequency of motion depends on the energy, and at every refreshing of momenta this frequency is changed. Non-zero values of β can be handled by going to the action-angle variables at $\beta = 0$, and setting up a perturbation theory in β . This expansion assures us that there is no need to refresh momenta at any $\beta < 50$, provided, of course, that there is no phase transition at a coupling smaller than this which invalidates the approximations.

In fact the situation changes dramatically in the spin-wave phase on using long trajectory lengths. There are clear peaks in $\tau(k)$ at several values of k , making it necessary to randomise trajectory lengths. The effort is reduced drastically when this is done. As a typical example, for runs on a 64^2 lattice at $\beta = 1.50$, we obtain

$$E_X = \begin{cases} (2 \pm 1) \times 10^4 & (N_{md} = 400) \\ 620 \pm 45 & (\langle N_{md} \rangle = 400) \end{cases} \quad (16)$$

This difference of behaviour between long and short trajectories is consistent with the observations made in Ref. 12.

Slow fluctuations are clearly visible in the run time history of χ (Fig. 9), reminiscent of similar structures seen in QCD simulations. They completely disappear when the trajectory length is randomised. Another signal of phase locking comes from the evolution of χ and its error estimate with increasing statistics. We find that the average moves very quickly to the true expectation value, but the errors begin to scale with the square root of the statistics only very slowly. This is a clear signal that a small number of modes begin to dominate the long-time behaviour of the measurement. On a d -dimensional lattice the total number of modes is L^d , whereas for a fixed trajectory length the number of phase locked modes is asymptotically constant. Hence these are thermodynamically insignificant, at least for the expectation value of a fluctuation measure.

4. The Dynamical Critical Exponent.

In this section we report our measurements of the dynamical critical exponent, z . When ξ is much less than the lattice size L , this is defined by the scaling of operators of fixed physical size, through the relation

$$\tau(\tau/\xi) = \xi^z f(\tau/\xi). \quad (17)$$

Since χ can be written as a sum over the spin-spin correlation length at all separations, its autocorrelations can be used to extract z . On the other hand, since the action is the correlation function at unit lattice spacing, it cannot provide a measure of z when the correlation length is finite. In both phases of the model, we have extracted two values for z ; one for $N_{md} = 1$, and the other for N_{md} large enough that τ reaches its minimum value.

In the vortex phase of the 2d XY model, we restrict ourselves to $L/\xi < 4$, in order to avoid finite-size effects. We extract the critical exponent by fitting data for fixed L to the form

$$\ln \tau_X(\xi) = \ln \tau_1 + z \ln \xi, \quad (18)$$

using the measurements of ξ reported in Ref. 6. The data and the fitted curves are shown in Fig. 10, and the values of the parameters extracted from the fits are listed in Table 4. We find that, for $N_{md} = 1$, the value of z is close to 2, and drops to 1 when very long trajectories are used. It is interesting to note that effort is smaller for $N_{md} = 300$ than for $N_{md} = 1$ at all values of ξ on the 32^2 lattice.

In the spin-wave phase of the 2d XY model, we have extracted z using data from simulations with $L=16, 32$ and 64 at $\beta=1.3$ and 1.5 . The critical exponent is determined through finite size scaling fits of the form

$$\ln r(L) = \ln r_1 + z \ln L. \quad (19)$$

As in the vortex phase, we have evaluated z for two different trajectory lengths. The results are summarised in Table 5. The dynamical critical exponent drops from 2 to 0 as N_{md} is increased from 1 to a sufficiently large value.

Since the value of N_{md} at which $z = 0$ increases linearly with L , we have attempted to determine the scaling properties of E . This is intrinsically more difficult, since one has to scan a range of trajectory lengths, determine the optimal length, and then try to analyse the scaling. A large part of the errors can be ascribed to the inability to determine T_{opt} to sufficient accuracy. The results of the fits in the spin-wave phase of the model are

$$E_X(L) = \begin{cases} (0.6 \pm 0.2)L^{1.38 \pm 0.09} & (\beta = 1.30) \\ (0.6 \pm 0.1)L^{1.36 \pm 0.07} & (\beta = 1.50) \end{cases} \quad (20)$$

Since this scaling relation has been obtained at fixed P_{acc} , a part of the exponent is due to the known dependence of Δt on L . In order to correct for this, we measure the scaling for the effort expressed in molecular dynamics time units. We find

$$(\Delta t)E_X(L) = \begin{cases} (0.5 \pm 0.1)L^{0.88 \pm 0.09} & (\beta = 1.30) \\ (0.4 \pm 0.1)L^{0.88 \pm 0.07} & (\beta = 1.50) \end{cases} \quad (21)$$

Clearly, the variation $\Delta t \sim L^{-d/4}$ is taken care of by this means, and we obtain an exponent close to that found in the vortex phase.

In the spin-wave phase, autocorrelations and computational effort for the action scale properly. For $N_{md} = 400$, we find that τ_s scales with exponent 0. At the optimal trajectory length determined by E_X we obtain

$$E_s(L) = \begin{cases} (0.47 \pm 0.09)L^{1.43 \pm 0.06} & (\beta = 1.30) \\ (0.4 \pm 0.1)L^{1.40 \pm 0.08} & (\beta = 1.50) \end{cases} \quad (22)$$

Again, correcting for the L -dependence of the step size, we obtain

$$(\Delta t)E_s(L) = \begin{cases} (0.40 \pm 0.07)L^{0.93 \pm 0.06} & (\beta = 1.30) \\ (0.31 \pm 0.09)L^{0.91 \pm 0.08} & (\beta = 1.50) \end{cases} \quad (23)$$

There is complete consistency between measurements of the dynamical critical exponent from these two different operators.

An attempt to find the scaling of effort in the vortex phase yields

$$(\Delta t)E_X(\xi) = (1.8 \pm 0.2)\xi^{1.54 \pm 0.07} \quad (L = 32).$$

The mismatch between this exponent and the one for τ , taken at face value, would lead to the absurd conclusion that for sufficiently large ξ , the effort at its minimum is larger than that for very long trajectories. Clearly, the implication is that at sufficiently large ξ the minimum of E and τ must occur at the same trajectory length, and both would then scale with exponent 1.

We present an argument which makes plausible the idea that $z \approx 1$ in general. Since local updates must yield $z \approx 2$, we look for a mechanism that could convert HMC into a global scheme. It is clear that the state of any spin at a given time depends on the states of its nearest neighbours at the previous leap-frog step, the next-nearest at the preceding step, and so on. Thus, information on spins far removed can be conveyed to any spin, and influence its update, provided only that the evolution is carried out long enough. This flow of information is disrupted at the refreshing of momenta. If there is an identifiable time, T_{fp} , in which signals travel across a lattice of size L , then this must set the scale for T_{opt} .

We start with free field theory, perturbing the field at a point, arbitrarily labelled as the origin, by an amount Δ . Expanding this perturbation in Fourier modes, we obtain $\delta\phi_k(0) = \Delta/L^{d/2}$. The induced perturbation at point r and time T is given by

$$\delta\phi_r(T) = \frac{\Delta}{2L^d} \sum_k \left\{ e^{i(k \cdot r + \omega_k T)} + e^{i(k \cdot r - \omega_k T)} \right\}. \quad (24)$$

This describes a lattice approximation to a spherical wave of 'sound'. Here T_{fp} is the time required for the peak of the excitation in the slowest mode ($k = 2\pi/L$) to reach a distance $L/2$, and is equal to π/ω_k . Not surprisingly, this sets the scale for T_{opt} . When the trajectory length is very small, the suggested update depends only on the nearest neighbours; whereas for trajectories longer than T_{fp} , each spin can feel the influence of the furthest spin in the lattice, and the update becomes global.

For the XY model we measured T_{fp} directly. This can be done by taking two sets of initial conditions $\{U_0, \pi_0\}$ and $\{U'_0, \pi_0\}$, differing only in the spin at a

given site l , and evolving both using the MD equations. A measure of the local perturbation at site i is given by $Re(U_i - U_i')$. Given an initial perturbation Δ , an estimate of T_{fp} is obtained by measuring the time at which the perturbation first reaches a value of $\bar{\Delta}$ at a distance $L/2$ from l . Since the theory is non-linear, $\bar{\Delta}$ can be larger than Δ . Furthermore, T_{fp} can depend on both these quantities.

We find that the time t_i at which the perturbation at a site i first exceeds $\bar{\Delta}$ is linear in $|i - l|$, showing that there is a well-defined speed with which perturbations travel through the lattice. For $\bar{\Delta}/\Delta$ small, this speed is independent of the magnitude of the perturbation. At the same time, measurements on lattices of size 16^2 and 32^2 yield nearly the same speed, and hence a linear dependence of T_{fp} on L . For example, at $\beta = 1.30$ and for $\Delta = 0.1$, we find $T_{fp} = 0.18 \pm 0.03$ on a 16^2 lattice, and 0.33 ± 0.02 on a 32^2 lattice. As in the case of FFT, the ratio between T_{fp} and T_{opt} is large. Although this prevents us from using this measurement to tune the trajectory length, it does demonstrate that the linear scaling law arises from the same mechanism as in FFT.

A possible objection to this argument is that long-ranged couplings may invalidate the idea of a slow outward propagation of perturbations. This is not necessarily true. Consider the toy model of scalar fields ϕ with an extended quadratic action

$$S = \beta \sum_l \gamma_l \sum_{|j-l|=1} \phi_l \phi_j, \quad (25)$$

where γ_l measures the strength of the coupling between spins at distance l . A necessary condition for the theory to make sense (S is extensive) is that the series in γ_l is absolutely convergent. With the canonical transformation given in Eq. (3), the molecular dynamics Hamiltonian decomposes into a sum over harmonic oscillators with frequencies given by

$$\omega_k^2 = \beta \sum_{l,\mu} \gamma_l \cos lk_\mu. \quad (26)$$

Since the series in γ_l is convergent, ω_k exists, and the arguments in the previous paragraphs follow. It seems that in a quite general class of models it might be possible to tune HMC parameters to yield a linear scaling of E with β or L .

5. Conclusions.

In summary, we find that autocorrelation times for HMC are minimised when the acceptance is in the range $0.6 \leq P_{acc} \leq 0.8$, when all other parameters are

kept fixed. We find a good scaling behaviour of autocorrelation times at fixed P_{acc} . This implies that the two parameters of the algorithm (Δt and N_{md}) can be tuned separately. Tuning of Δt corresponds to fixing the acceptance, and N_{md} can then be varied to achieve the best possible autocorrelation properties.

In the spin-wave phase of the model, it seems necessary to randomise trajectory lengths, specially when very long trajectories are used. Phase locking shows up in simulations of the two-dimensional XY model as slow oscillations in run time histories of measurements. Since the number of phase locked modes is thermodynamically insignificant, expectation values can converge fairly well to the correct value, but error estimates may take a long time to refine. In the vortex phase of the model it is advisable not to randomise trajectory lengths. This conclusion can be generalised through strong-coupling arguments.

Another set of results pertains to the choice of N_{md} . The computational effort for decorrelating measurements can be optimised by a choice of the trajectory length. The optimal trajectory length increases linearly with L at fixed β , in both phases of the XY model. Since T_{opt} can be meaningfully measured on small lattices, we obtain an operational definition of optimal trajectory lengths. It is highly recommended that this measurement be performed, since there can be more than an order of magnitude gain in computational effort by doing this tuning.

The dynamical critical exponent is 2 when short trajectories are used for the simulation. It falls to 1 on using long trajectories. Working at the optimal trajectory length at fixed acceptance, the computational effort switches from a scaling behaviour with exponent $(2 + d/4)$ to $(1 + d/4)$ as the correlation length increases.

Although all the results here have been obtained with Hybrid Monte Carlo simulations, they have obvious implications for simulations performed with the Hybrid Stochastic algorithm. In the latter, measurements can be performed at each MD step; hence every statement about the computational effort for HMC translates into a statement for autocorrelations in HS. It is of some practical importance to check whether the dynamical phenomena seen in purely bosonic models undergo major changes when fermions are introduced. We leave this for the future.

I would like to thank Anders Irbäck, Sudeshna Sinha and Ulli Wolff for discussions.

REFERENCES

- 1) S. Duane, A. D. Kennedy, B. J. Pendleton and D. Roweth, *Phys. Lett. B* **195** (1987) 216.
- 2) S. Duane, *Nucl. Phys. B* **257** [FS14] (1985) 652.
- 3) P. McKenzie, *Phys. Lett. B* **226** (1989) 369.
- 4) A. D. Kennedy and B. J. Pendleton, Talk presented at the International Symposium Lattice '90, Tallahassee, USA, October 8-12th, 1990.
- 5) E. Dagotto and J. B. Kogut, *Phys. Rev. Lett.* **58** (1987) 299.
- 6) R. Gupta, J. DeLapp, G. G. Batrouni, G. C. Fox, C. F. Baillie and J. Apatolakis, *Phys. Rev. Lett.* **61** (1988) 1996; U. Wolff, *Nucl. Phys. B* **322** (1989) 759; R. G. Edwards, J. Goodman and A. D. Sokal, *Nucl. Phys. B* **354** (1991) 289.
- 7) G. G. Batrouni, G. R. Katz, G. S. Kronfeld, G. P. Lepage, B. Svetitsky and K. G. Wilson, *Phys. Rev. D* **32** (1985) 2736.
- 8) M. Creutz, *Phys. Rev. D* **38** (1988) 1228.
- 9) R. Gupta, G. W. Kilcup and S. R. Sharpe, *Phys. Rev. D* **38** (1988) 1278.
- 10) S. Gupta, A. Irback, F. Karsch and B. Petersson, *Phys. Lett. B* **242** (1990) 437.
- 11) A. Vaccaro, Columbia preprint CU-TP-513, March 1991.
- 12) M. Fukugita and A. Ukawa, *Nucl. Phys. B* **300** [FS22] (1988) 433.

FIGURE CAPTIONS

- 1) The acceptance P_{acc} as a function of ΔH for the 2-d XY model for simulations in the vortex (open symbols) and spin-wave (filled symbols) phase on 16^2 (squares) and 32^2 (circles) lattices. The line shows the prediction of Eq. (11).
- 2) P_{acc} as a function of N_{md} on 32^2 lattices. The data are taken at $\beta = 1.30$ with fixed Δt . The lines are meant to guide the eye.
- 3) The dependence of τ_χ and its error on the number of trajectories used in its determination. The circles denote data taken at $\beta = 0.86$, and the squares for $\beta = 0.90$, on 64^2 lattices for $N_{md} = 750$ and $P_{acc} = 0.75$.
- 4) The dependence of τ_χ on P_{acc} for $N_{md}=10$ on a 32^2 lattice at $\beta=0.82$ (filled circles) and $\beta=1.30$ (open circles).
- 5) The variation of (a) E_χ and (b) E_s with T on 32^2 lattices for $P_{acc}=0.75$. The filled circles denote data taken at $\beta = 0.78$ and the open circles at $\beta = 0.90$.
- 6) The variation of (a) τ_χ and (b) τ_s with T on 32^2 lattice. For an explanation of the symbols, see the caption to Fig. 5.
- 7) The variation of E_χ with T at $\beta = 1.30$. Filled circles denote data taken on 16^2 lattices, open circles on 32^2 , and filled squares on 64^2 . Note the linear dependence of T_{opt} on L .
- 8) The autocorrelation time $\tau(k)$ of Fourier modes k of the fields. Filled circles show data taken on a 64^2 lattice at $\beta = 2$, $\Delta t = 0.10$ and $N_{md} = 10$, open circles that on a 32^2 lattice at $\beta = 0.82$, $\Delta t = 0.212$ and $N_{md} = 300$.
- 9) A representative portion of the run time history for χ taken on 64^2 lattices at $\beta = 1.50$ (a) with fixed and (b) randomised trajectory lengths. In both cases $\langle N_{md} \rangle = 400$, and $P_{acc} = 75\%$. N_t denotes the number of trajectories.
- 10) The scaling of τ_χ with ξ in the vortex phase. The circles are for $L = 32$ and the squares for $L = 64$. The lines show the fits reported in Table 4.

TABLES

TABLE 1

The values of Δt , N_{md} and P_{acc} , the number of trajectories N_t (in thousands), $\langle s \rangle$, c_v and χ at different values of β for runs in the vortex phase on 32^2 lattices. All errors mentioned in this paper are 1σ confidence limits.

β	N_t	Δt	N_{md}	P_{acc}	$\langle s \rangle$	c_v	χ
0.78	42.0	0.223	1	0.84	0.9577(4)	1.007(9)	22.9(6)
	30.0	0.223	300	0.73	0.9574(7)	0.97(2)	23.8(2)
0.82	42.0	0.212	1	0.85	1.0235(5)	1.14(1)	36.(1)
	10.5	0.212	300	0.73	1.025(1)	1.09(3)	36.7(5)
0.86	44.0	0.205	1	0.85	1.0936(7)	1.31(2)	64.(2)
	14.0	0.205	300	0.73	1.093(1)	1.33(4)	60.4(7)
0.88	42.0	0.205	1	0.84	1.1275(8)	1.37(2)	76.(3)
	15.0	0.205	300	0.72	1.128(1)	1.32(4)	78.2(9)
0.90	42.0	0.198	1	0.85	1.163(1)	1.45(2)	104.(5)
	20.0	0.198	300	0.74	1.161(1)	1.43(4)	102.(1)

TABLE 2

The values of Δt , N_{md} and P_{acc} , the number of trajectories N_t (in thousands), $\langle s \rangle$, c_v and χ at different values of β for runs in the vortex phase on 64^2 lattices.

β	N_t	Δt	N_{md}	P_{acc}	$\langle s \rangle$	c_v	χ
0.82	17.5	0.150	750	0.74	1.0241(5)	1.12(3)	36.1(4)
0.86	13.8	0.145	750	0.73	1.0928(6)	1.29(5)	60.0(7)
0.88	17.5	0.145	750	0.72	1.1262(5)	1.31(3)	80.1(8)
0.90	13.9	0.140	750	0.73	1.1602(6)	1.39(5)	110.(1)
0.95	15.3	0.130	750	0.75	1.2441(5)	1.49(4)	307.(4)

TABLE 3

The values of Δt , N_{md} and P_{acc} , the number of trajectories N_t (in thousands), $\langle s \rangle$, c_v and χ for different β and lattice sizes L for runs in the spin-wave phase.

β	L	N_t	Δt	N_{md}	P_{acc}	$\langle s \rangle$	c_v	χ
1.30	16	80.0	0.210	1	0.82	1.5534(9)	0.76(2)	155.2(9)
		17.0	0.210	400	0.75	1.552(1)	0.75(2)	155.1(2)
	32	80.0	0.149	1	0.87	1.5517(7)	0.75(2)	560.(6)
		19.5	0.149	400	0.75	1.5521(6)	0.75(2)	550.9(8)
	64	80.0	0.105	1	0.91	1.5522(4)	0.75(3)	1998.(46)
		10.0	0.105	400	0.75	1.5508(3)	0.76(2)	1972.(4)
1.50	16	10.0	0.190	400	0.76	1.6244(9)	0.68(2)	169.7(3)
	32	10.0	0.134	400	0.77	1.6246(4)	0.68(2)	619.(1)
	64	10.0	0.095	400	0.77	1.6243(2)	0.66(2)	2257.(38)

TABLE 4

Results of fits to the dynamical critical exponent for the 2d XY model in the vortex phase, using data collected in the runs shown in Tables 1 and 2.

L	N_{md}	τ_1	z	χ^2/dof
32	1	60.(10)	2.0(1)	1.2/3
	300	0.14(2)	1.1(1)	2.0/3
64	750	0.08(1)	1.01(7)	4.4/3

TABLE 5

Results of fits to the dynamical critical exponent for the 2d XY model in the spin-wave phase, using data collected in the runs shown in Table 3.

β	N_{md}	τ_1	z	χ^2/dof
1.30	1	0.3(1)	2.2(1)	0.11/1
	400	2.1(4)	-0.06(5)	0.23/1
1.50	400	1.7(6)	0.01(9)	1.06/1

Figure 1

Figure 3

Figure 2

Figure 4

Figure 5

Figure 6

Figure 7

Figure 8

Figure 9

Figure 10