

Non-reversible Monte Carlo simulations of spin models

Heitor C. M. Fernandes, Martin Weigel

Institut für Physik, KOMET 331, Johannes Gutenberg-Universität Mainz, Staudinger Weg 7, 55128 Mainz, Germany

Abstract

Monte Carlo simulations are used to study simple systems where the underlying Markov chain satisfies the necessary condition of global balance but does not obey the more restrictive condition of detailed balance. Here, we show that non-reversible Markov chains can be set up that generate correct stationary distributions, but reduce or eliminate the diffusive motion in phase space typical of the usual Monte Carlo dynamics. Our approach is based on splitting the dynamics into a set of replicas with each replica representing a biased movement in reaction-coordinate space. This introduction of an additional bias in a given replica is compensated for by choosing an appropriate dynamics on the other replicas such as to ensure the validity of global balance. First, we apply this method to a mean field Ising model, splitting the system into two replicas: one trying to increase magnetization and the other trying to decrease it. For this simple test system, our results show that the altered dynamics is able to reduce the dynamical critical exponent. Generalizations of this scheme to simulations of the Ising model in two dimensions are discussed.

Keywords: Monte Carlo simulations, Markov chains, random-walk suppression

1. Introduction

In the realm of Monte Carlo simulations, techniques based on Markov chains are by far the most common methods [1]. Based on the principle of detailed balance, the Metropolis-Hastings algorithm [2] is the workhorse of the Monte Carlo technique, unsurpassed in its most general applicability and technical simplicity. By construction, under its guidance the system performs a (biased) random walk in configuration space. While it samples states directly according to the equilibrium distribution, as desired, the diffusive motion in phase space makes convergence relatively slow. This is not much of a problem for the strongly localized distributions typical of, say, the canonical ensemble off criticality. The broadened distributions close to continuous phase transitions and, even more so, the asymptotically flat distributions typically encountered in generalized-ensemble simulations such as the multicanonical method [3], are much more problematic, however, when sampling with a random-walk like exploration of configuration space. At this point, the diffusive behavior of the standard Metropolis based approaches regularly turns into the dominant bottleneck of the simulation since equally sampling an extensive phase space volume becomes prohibitively slow. Here, we investigate, in how far the diffusive dynamics of conventional local Monte Carlo simulations can be altered by relaxing the condition of detailed balance, while still ensuring convergence to the exact equilibrium distribution.

For a Monte Carlo Markov chain, the conservation of probability requires the transition probabilities $T(x_i \rightarrow x_j)$ to satisfy the global balance equation

$$\sum_j T(x_i \rightarrow x_j) \pi_i = \sum_j T(x_j \rightarrow x_i) \pi_j \quad (1)$$

where π is the stationary probability distribution, and x_i and x_j denote representatives of the configuration space under consideration. The typically chosen detailed balance condition, $T(x_i \rightarrow x_j) \pi_i = T(x_j \rightarrow x_i) \pi_j$, is obviously sufficient, but not necessary for (1) to hold. While Monte Carlo algorithms satisfying global, but violating detailed balance have been regularly used before, e.g., in updating spins on a lattice in sequential instead of random order [1] or for the checkerboard decomposition used in parallel implementations [4], the systematic engineering of non-reversible dynamics for the purpose of speeding up convergence has not been regularly attempted in statistical physics [5] (an exception is the hybrid Monte Carlo method [6]). In a mathematical context, the authors of Ref. [7] (see Refs. [8, 5] for related ideas) suggest to construct a general non-reversible Monte Carlo dynamics for a Markov chain on a *one-dimensional* configuration space $\{k = 1, 2, \dots\}$ by replicating the whole chain of states into + and - copies and performing two-step transitions as indicated in Fig. 1:

- (a) a traditional, reversible Metropolis update between states $(k, \pm)$ and $(k \pm 1, \mp)$ according to the desired stationary distribution $\pi(k, +) = \pi(k, -) = \pi(k)$,

Email addresses: heitor@uni-mainz.de (Heitor C. M. Fernandes), weigel@uni-mainz.de (Martin Weigel)

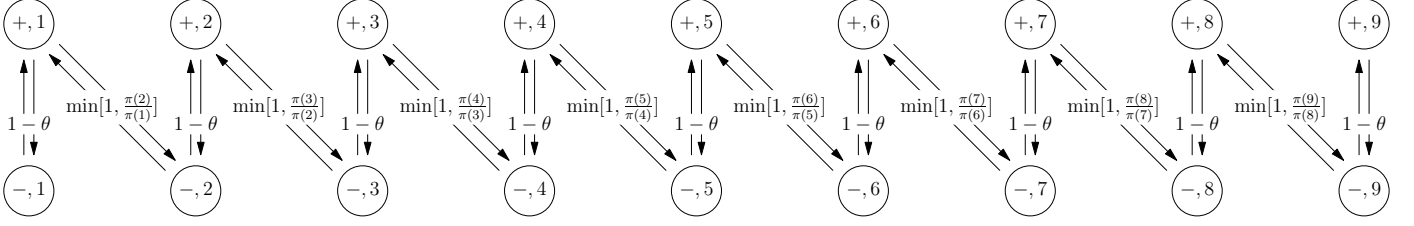


Figure 1: Non-reversible dynamics on a one-dimensional chain of states with stationary distribution π and a fixed probability $1 - \theta$ of replica change.

i.e., with acceptance rate

$$p_{\text{acc}} = \min \left[1, \frac{\pi(k \pm 1)}{\pi(k)} \right] \quad (2)$$

and

- (b) a transition $(k, \pm) \mapsto (k, \mp)$ with a fixed probability $1 - \theta$.

Both transitions obviously leave π invariant, such that π is stationary, but the combination of the two steps is not reversible. Moreover, it is clear that if θ is small, the chain will continue moving in the same direction as long as the Metropolis steps are accepted.

2. Mean Field Ising model

In the spirit of the above construction, we considered a ferromagnetic, mean-field Ising model with N spins and Hamiltonian

$$\mathcal{H} = -\frac{J}{2N} \sum_{i,j} s_i s_j = -\frac{J}{2N} M^2, \quad (3)$$

which has the peculiarity that, due to the indistinguishability of aligned spins, its state is completely described by the total magnetization M . Therefore, all microstates belonging to fixed magnetization M (or fixed energy $E = -JM^2/2N$) are completely equivalent. Under the usual single-spin flip dynamics with $M \mapsto M \pm 2$ the system is, therefore, exactly described by the one-dimensional chain depicted in Fig. 1 with $k = (M+N)/2$. The system undergoes a paramagnetic to ferromagnetic ordering transition at the critical coupling $T_c = 1$, and with a regular, reversible single-spin flip Metropolis dynamics, critical slowing down is observed. To update the configuration directly in terms of the magnetization M (and not in terms of the individual spins s_i), the degeneracy of microstates needs to be taken into account. The canonical distribution of magnetizations therefore is given by

$$P(M) = \frac{N!}{N_+!N_-!} \exp \left(\frac{\beta J M^2}{2N} \right), \quad (4)$$

where N_+ and N_- denote the number of up and down spins, respectively. Since $M = N_+ - N_-$, this is equivalent to

$$P(M) = \frac{N!}{\left(\frac{N+M}{2}\right)! \left(\frac{N-M}{2}\right)!} \exp \left(\frac{\beta J M^2}{2N} \right). \quad (5)$$

Hence, the acceptance probabilities (2) for the moves $M \mapsto M \pm 2$ become

$$p_{\text{acc}} = \min \left[1, \frac{N \mp M}{N \pm M + 2} \exp \left\{ \frac{2\beta J}{N} (\pm M + 1) \right\} \right]. \quad (6)$$

The form of these probabilities as a function of M for $T = T_c = 1$ is illustrated in Fig. 2. Figure 3 shows time series of non-reversible simulations of this type for the mean-field model for different values of the replica-change probability θ . For large values of θ , the algorithm behaves very similarly to the case of regular, reversible Metropolis simulations, and the temporal evolution of the magnetization M resembles a random walk. As θ is gradually reduced, however, the system evolves for an increasing number of consecutive steps in the same direction before a change of replica and, therefore, a reversal of the direction of evolution of M occurs. Inspection of the acceptance probabilities in Fig. 2 reveals that, in fact, moves decreasing the magnetization $M \mapsto M - 2$ are *always* accepted for $M > 0$ and, vice versa, moves increasing $M \mapsto M + 2$ are always accepted if $M < 0$. Hence, the system must cross the symmetric point $M = 0$ before a reversal of the evolution direction can occur. This condition leads to the characteristic zig-zag pattern of the magnetization time series displayed in Fig. 3.

The effect of this changeover from diffusive to quasi ballistic evolution of the magnetization is corroborated by an analysis of the integrated autocorrelation times τ_{int} . From a binning analysis [9], we determine $\tau_{\text{int}}(M)$ from the original time series data from the limit of large block lengths, where some care has to be taken due to the fact that the autocorrelation function does not simply decay exponentially, but features a superimposed oscillatory behavior resulting from the zig-zag pattern of the time series shown in Fig. 3. Estimating the autocorrelation times for a range of different system sizes results in the dynamical scaling data presented in Fig. 4. Fits of the expected functional form

$$\tau_{\text{int}} = AN^z \quad (7)$$

to the data for a range of successively smaller values of the replica-change parameter θ show a gradual reduction of the effective dynamical critical exponent z from its value $z = 1.4390(33)$ for the reversible dynamics down to $z = 0.833(13)$ for the irreversible dynamics in the limit $\theta \rightarrow 0$. This is consistent with the finding $z \approx 0.85$ for a related approach discussed in Ref. [5].

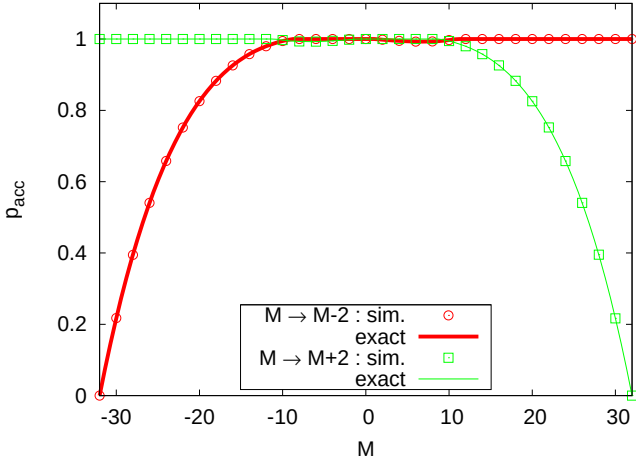


Figure 2: Acceptance probabilities according to Eq. (6) for non-reversible simulations of the mean-field Ising model (3) at the asymptotic critical temperature $T_c = 1$ for a system of $N = 32$ spins.

A certain reduction of autocorrelation times is also observed in the high-temperature paramagnetic and low-temperature ordered phases with this approach. The irreversible dynamics does not lead to a significant increase in tunneling events between the symmetric peaks in the ordered phase, however. These results will be discussed elsewhere [10].

3. Two-dimensional Ising model

For general short-range models and appropriate reaction coordinates such as, e.g., energy and magnetization for spin systems, the equivalence of microstates at fixed E or M seen in the mean-field model is lost. Consequently, representing phase space in analogy to Fig. 1 corresponds to a projection, not an identity. Here, we consider non-reversible simulations of the ferromagnetic, nearest-neighbor Ising model with Hamiltonian

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} s_i s_j, \quad (8)$$

where spins are located on the sites of an $L \times L$ square lattice. We studied two types of non-reversible simulations, using either the magnetization M or the energy E as the projection coordinate.

For the case of the magnetization, again two chains, one increasing $M \mapsto M+2$ and one decreasing $M \mapsto M-2$, are sufficient. Depending on which of the chains the system currently is in, we either pick one of the up or one of the down spins at random and propose a spin flip. Due to the different *a priori* probabilities for choosing a spin before and after the flip, we again get additional symmetry factors in the acceptance probabilities for the transition $M \mapsto M \pm 2$,

$$p_{\text{acc}} = \min \left[1, \frac{N_{\mp}}{N_{\pm} + 1} e^{-\beta \Delta E} \right], \quad (9)$$

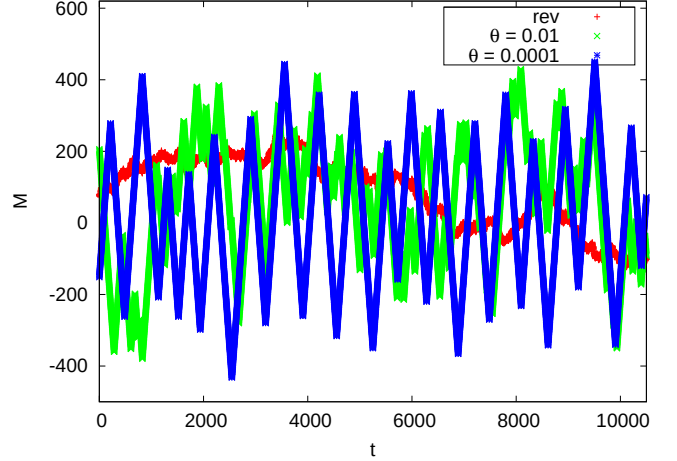


Figure 3: Time series of the magnetization M in a non-reversible simulation of the mean-field Ising model at $T = T_c$ for different choices of the replica-change probability θ .

which is analogous to the expression (6). Employing this type of dynamics at the asymptotic critical point of the short-range model, $T_c = 2/\ln(1 + \sqrt{2}) \approx 2.269$, we do not see any improvement of convergence over the reversible dynamics and, in fact, for large value of the replica-change probability θ , autocorrelation times are even larger for the non-reversible than for the regular, reversible approach [10].

When taking the energy as the projection coordinate, on a square lattice a single spin flip can lead to changes $\Delta E = 0, \pm 4J$ or $\pm 8J$. It is rather clear, therefore, that a two-chain description is not very appropriate. If we insist on only distinguishing between $\Delta E > 0$ and $\Delta E < 0$ with two chains, possibly throwing in some micro-canonical $\Delta E = 0$ moves to ensure ergodicity, we indeed do not find any reduction of random-walk behavior and consequently of autocorrelation times [10]. Generalizations for more than two chains can be worked out along the lines of the so-called “fiber algorithm” [7], see also Ref. [8]. Here, we use a setup with three chains, labeling the current move type by $|\Delta E|$ and y , such that there are altogether three chains with magnitudes $|\Delta E| = 0, 4J$ and $8J$, and $y = \pm 1$ indicates the direction of chain traversal, i.e., $\Delta E = y|\Delta E|$. Transitions are then performed in three steps:

- reversibly update $(E, y) \mapsto (E + y|\Delta E|, -y)$ with the Metropolis acceptance probability,

$$p_{\text{acc}} = \min \left[1, \frac{N_{\Delta E=y|\Delta E|}}{N'_{\Delta E=-y|\Delta E|}} e^{-\beta \Delta E} \right], \quad (10)$$

- unconditionally negate $y \mapsto -y$,
- with probability θ , randomly choose a new step size $|\Delta E|$.

Some extra bookkeeping is involved in keeping track of the numbers $N_{\Delta E=y|\Delta E|}$ of possible moves with $\Delta E =$

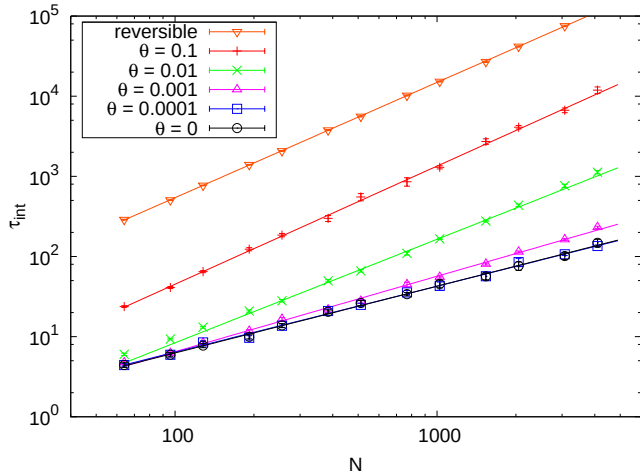


Figure 4: Integrated autocorrelation times τ_{int} of the magnetization from reversible and non-reversible simulations of the mean field Ising model of Eq. (3) at the asymptotic critical temperature $T_c = 1$ as a function of system size. The lines are fits of the functional form (7) to the data.

$y|\Delta E|$ before and after performing the considered transition. Since the moves with $\Delta E = 0$ do not interfere with the global direction of energy change, one might as well restrict attention to the two chains with $|\Delta E| = 4J$ and $8J$, and simply perform one or more $\Delta E = 0$ moves in between the updates changing the energy. Studying the short-range model (8) at criticality, the three-chain non-reversible algorithm walking in energy space indeed achieves some reduction of the random-walk behavior as is apparent from the data for the integrated autocorrelation times $\tau_{\text{int}}(E)$ and $\tau_{\text{int}}(M)$ of the internal energy and magnetization for different system sizes collected in Fig. 5. In contrast to the case of the mean-field model, however, no clear-cut reduction of the dynamical critical exponent z is achieved but, in any case, at least a constant speedup-factor around $\tau_{\text{int}}^{\text{irrev}}/\tau_{\text{int}}^{\text{rev}} \approx 5$ is observed. Some reduction of autocorrelation times is also seen off criticality [10].

A general drawback of the considered approaches is that movement in one chosen direction of the reaction coordinates E or M only continues until a rejection of one of the reversible Metropolis steps occurs. It is due to the extra factors depending on $N_{\Delta E=y|\Delta E|}$ in Eq. (10), which depend on the microstate under consideration and are not the same for all states with the same energy, that such rejections occur even in regions where the corresponding generalized Boltzmann factor would indicate that an attempt to change the energy by ΔE via a standard Metropolis update should always be accepted. In mathematical terms, complete random-walk suppression with the outlined techniques is preempted by the lack of lumpability [11] of the underlying Markov chain on the microscopic states with respect to the chosen state-space partition, e.g., according to the energy E [10].

Generalizations of the outlined approach for the case of multicanonical simulations are possible with minor adap-

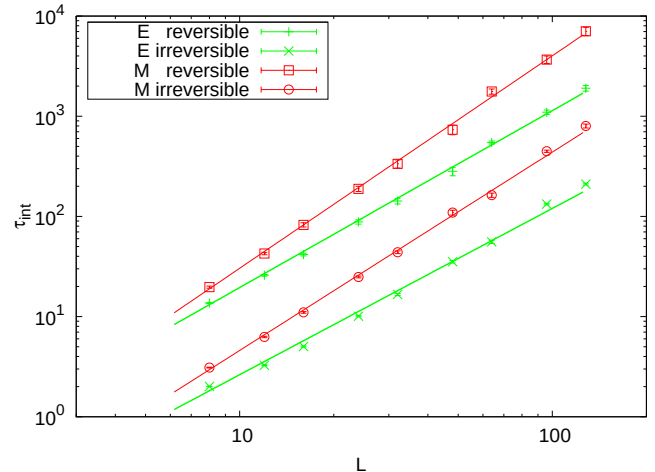


Figure 5: Finite-size scaling of the integrated autocorrelation times $\tau_{\text{int}}(M)$ and $\tau_{\text{int}}(E)$ of the internal energy and magnetization of non-reversible simulations of the two-dimensional Ising model at criticality using a state-space partition according to the energy E . Fits of the functional form $\tau_{\text{int}} = AL^z$ to the data yield the estimates $\tau_{\text{int}}(E) = 1.670(52)$ and $\tau_{\text{int}}(M) = 1.980(29)$ for the non-reversible dynamics, to be compared to the estimates $\tau_{\text{int}}(E) = 1.766(24)$ and $\tau_{\text{int}}(M) = 2.121(27)$ for the regular, reversible simulation.

tations [10]. Although for the ideal case of a perfectly flat distribution one might expect rejections of the Metropolis steps to be rare, the microstate factors counter-balancing the change in the number of possible moves of a given type appear to be rather wildly varying between different pairs of microstates connected by the move, such that random-walk suppression is not perfect from this approach. In this respect, the mean-field model considered in Sec. 2 is a rather special case, since all microstates are completely equivalent.

The authors acknowledge support by the “Center for Computational Sciences in Mainz” (SRFN). M.W. acknowledges computer time provided by NIC Jülich under grant No. hmz18 and funding by the DFG under contract No. WE4425/1-1.

- [1] B. A. Berg, Markov Chain Monte Carlo Simulations and Their Statistical Analysis, World Scientific, Singapore, 2004.
- [2] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, E. Teller, Equation of state calculations by fast computing machines, J. Chem. Phys. 21 (1953) 1087.
- [3] B. A. Berg, T. Neuhaus, Multicanonical ensemble: A new approach to simulate first-order phase transitions, Phys. Rev. Lett. 68 (1992) 9.
- [4] K. Binder, D. P. Landau, A Guide to Monte Carlo Simulations in Statistical Physics, Cambridge University Press, Cambridge, second edition, 2005.
- [5] K. S. Turitsyn, M. Chertkov, M. Vucelja, Irreversible Monte Carlo Algorithms for Efficient Sampling, Preprint arXiv:0809.0916.
- [6] S. Duane, A. D. Kennedy, B. J. Pendleton, D. Roweth, Hybrid Monte Carlo, Phys. Lett. B 195 (1987) 216–222.
- [7] P. Diaconis, S. Holmes, R. M. Neal, Analysis of a nonreversible Markov chain sampler, Ann. Appl. Prob. 10 (2000) 726–752.
- [8] P. Gustafson, Y. C. MacNab, S. Wen, On the value of derivative evaluations and random walk suppression in Markov chain Monte Carlo algorithms, Stat. Comput. 14 (2004) 23–38.

- [9] H. Flyvbjerg, H. G. Petersen, Error estimates on averages of correlated data, *J. Chem. Phys.* 91 (1989) 461.
- [10] H. C. M. Fernandes, M. Weigel, in preparation.
- [11] P. Buchholz, Exact and ordinary lumpability in finite Markov chains, *J. Appl. Prob.* 31 (1994) 59–75.