

Markov Chain Monte Carlo Methods in Statistical Physics

K. P. N. Murthy^{*†} and V. S. S. Sastry^{††}

[†] *Materials Science Division, Indira Gandhi Centre for Atomic Research,
Kalpakkam 603 102, Tamilnadu, India
E-mail: kpn@igcar.gov.in*

^{††} *School of Physics, University of Hyderabad,
Hyderabad 500 046, Andhra Pradesh, India
E-mail: vssssp@uohyd.ernet.in*

In this paper we shall briefly review a few Markov Chain Monte Carlo methods for simulating closed systems described by canonical ensembles. We cover both Boltzmann and non-Boltzmann sampling techniques. The Metropolis algorithm is a typical example of Boltzmann Monte Carlo method. We discuss the time-symmetry of the Markov chain generated by Metropolis like algorithms that obey detailed balance. The non-Boltzmann Monte Carlo techniques reviewed include the multicanonical and Wang-Landau sampling. We list what we consider as milestones in the historical development of Monte Carlo methods in statistical physics.

We dedicate this article to Prof. Dr. G. Ananthakrishna and wish him the very best in the coming years

*International Conference on Statistical Mechanics of Plasticity and Related Instabilities,
Indian Institute of science, Bangalore,
August 29 - September 2, 2005*

*Speaker.

1. Introduction

Monte Carlo is a numerical technique that makes use of random numbers to simulate a stochastic model of a phenomenon. Historically, the first large scale Monte Carlo work carried out dates back to 1950s. It consisted of simulating neutron transport in a medium, *e.g.* a nuclear reactor core. Stanislaw Ulam, John von Neumann and Enrico Fermi were the first to propose and employ Monte Carlo technique for solving practical problems. The earliest published work on Monte Carlo is perhaps the one reported by Ulam and Metropolis in the year 1949 [1]. There were of course several isolated instances when Monte Carlo technique has been employed in some form or the other. A notable example is the Buffon's needle experiment carried out in the year 1777 for estimating the value of the irrational number π , a description of which can be found in [2].

In this paper we shall confine ourselves to discussing a few Markov chain Monte Carlo methods for simulating an equilibrium closed system described by canonical ensemble of microstates. Before we do that let us take a quick look at the random numbers that fuel the Monte Carlo machine.

2. Random Numbers

For any Monte Carlo simulation we need a sequence of real numbers distributed randomly, uniformly and independently in the range $(0, 1)$. Strictly we can call these numbers random if and only if they are generated by a random physical process like radioactive decay, thermal noise in electronic devices, cosmic ray arrival times and tossing of a coin. ¹ But for generating random numbers we invariably employ simple arithmetic operations that are fast, and that do not require much computer memory for storage.

For example consider multiplicative congruential generator [3], defined by,

$$R_{i+1} = a \times R_i \pmod{m}, \quad (2.1)$$

where a , m and $\{R_i\}$ are integers. a is called the generator or the multiplier. m is called the modulus. Start with a seed : $0 \leq R_1 \leq m - 1$. Use the above recursion and generate a sequence of integers : $\{R_1, R_2, \dots\}$. The sequence is converted to floating point numbers by dividing each by m . The multiplier a and the modulus m are to be chosen carefully. For example the choice $a = 7^5$ and $m = 2^{31} - 1$, has been shown [4] to yield good random numbers in 32-bit machines. The period of the generator is nearly 10^9 , which is not really long for several applications. The choice of $a = 13^{13}$ and $m = 2^{59}$ in the G05FAF generator of the NAG library [5] has a period of 10^{18} .

The important point is that the sequence is generated by a deterministic algorithm. These numbers are predictable and reproducible. Hence by no stretch of imagination can they be called random. We call them pseudo random. The sequence of pseudo random numbers is usually tested for randomness and then employed in Monte Carlo simulation.

¹These phenomena are known to be truly random at least according to the current day theories.

3. Canonical Ensemble

Consider a closed system of N microscopic entities (molecules; spins *etc.*), in volume V and in thermal equilibrium with the outside world (heat bath) at temperature T . Let,

$$\Omega_{CS} = \{\mathcal{C}_i : i = 1, 2, \dots, \widehat{\Omega}_{CS}\},$$

denote all possible microstates belonging to the closed system². Let $O(\mathcal{C})$ denote the value of a macroscopic variable (*e.g.* magnetization; pressure; *etc.*) when the closed system is in a microstate \mathcal{C} ($\in \Omega_{CS}$). The aim is to calculate $\langle O \rangle$.

Let $E(\mathcal{C})$ be the energy of the system when it is in microstate \mathcal{C} . Let $\beta = [k_B T]^{-1}$, where T is the temperature of the heat bath and k_B is the Boltzmann constant. The system can be found in microstate \mathcal{C} with a probability,

$$P(\mathcal{C}) = Z^{-1} \exp[-\beta E(\mathcal{C})], \quad (3.1)$$

where, $Z(T, V, N)$ is the canonical partition function given by,

$$Z(T, V, N) = \sum_{\mathcal{C} \in \Omega_{CS}} \exp[-\beta E(\mathcal{C})]. \quad (3.2)$$

Formally we have,

$$\langle O \rangle = \frac{\sum_{\mathcal{C} \in \Omega_{CS}} O(\mathcal{C}) \exp[-\beta E(\mathcal{C})]}{\sum_{\mathcal{C} \in \Omega_{CS}} \exp[-\beta E(\mathcal{C})]}. \quad (3.3)$$

Naively we can sample a large number of microstates $\{C_1, C_2, \dots, C_N\}$ with equal probability and calculate,

$$\langle O \rangle = \lim_{N \rightarrow \infty} \bar{O}_N = \frac{\sum_{i=1}^N O(C_i) \exp[-\beta E(C_i)]}{\sum_{i=1}^N \exp[-\beta E(C_i)]} \quad (3.4)$$

I must emphasize that in the above procedure all the microstates should be sampled with the same probability *i.e.* from a uniform ensemble. This is not usually viable since any finite Monte Carlo sample would contain predominantly only those microstates with large energy³. Because of the Boltzmann factor $\exp[-\beta E(\mathcal{C})]$, most of the microstates of the Monte Carlo sample will contribute negligibly to the partition sum. Hence the estimate of $\langle O \rangle$ will be statistically poor. We need a technique, like the Metropolis algorithm [8] to sample microstates from a canonical ensemble of microstates

Let $\Omega_{CS} = \{C_1, C_2, \dots\}$ denote a canonical ensemble of microstates. We take the size of the ensemble to be adequately large so that the fractional number of times a microstate $\mathcal{C} \in \Omega_{CS}$ occurs in a canonical ensemble (of microstates) equals $Z^{-1} \exp[-\beta E(\mathcal{C})]$. Let $\{C_1, C_2, \dots, C_N\}$ be

²A point in the $6N$ dimensional phase space is an example of a microstate of a classical system. Three positions and three momenta are required to specify a single molecule. There are N molecules. Hence a string of $6N$ numbers specify a microstate. Another example of a microstate is an orientational configuration of Ising spins on a lattice. Each spin can be oriented *up* or *down*. If there are N spins in the system then there are a total of 2^N microstates.

³Entropy increases with energy, rather steeply.

N microstates sampled from a canonical ensemble employing Metropolis algorithm, described in the next section. Then a simple arithmetic average provides an estimate of $\langle O \rangle$, as given by,

$$\langle O \rangle = \lim_{N \rightarrow \infty} \bar{O}_N = \frac{1}{N} \sum_{i=1}^N O(C_i). \quad (3.5)$$

4. Metropolis algorithm

In Metropolis Monte Carlo, we start with an arbitrary initial microstate C_0 and generate a Markov chain of microstates : $C_0 \rightarrow C_1 \rightarrow C_2 \rightarrow \dots \rightarrow C_n \rightarrow \dots$. The asymptotic ($n \rightarrow \infty$) part of the Markov chain contains microstates belonging to the desired canonical ensemble.

Let $C_i \in \Omega_{CS}$ be i -th entry in the Markov chain. Let $E_i = E(C_i)$ denote its energy. $P(C_i) \propto \exp[-\beta E_i]$ is its probability. Construct a trial microstate C_t by making a local change ⁴. Let $E_t = E(C_t)$ denote the energy of the trial state. $P(C_t) \propto \exp[-\beta E(C_t)]$ is its probability. Then

$$C_{i+1} = \begin{cases} C_t & \text{with probability } p \\ C_i & \text{with probability } 1 - p, \end{cases} \quad (4.1)$$

where the Metropolis acceptance probability p is given by,

$$\begin{aligned} p &= \min. \left(1, \frac{P(C_t)}{P(C_i)} \right) \\ &= \min. \left(1, \exp \left[-\beta (E_t - E_i) \right] \right). \end{aligned} \quad (4.2)$$

The heat-bath algorithm [6] or also known as the Glauber algorithm [7] consists of defining the Metropolis acceptance probability p as,

$$\begin{aligned} p &= \frac{\pi(C_t)}{\pi(C_t) + \pi(C_i)} \\ &= \frac{1}{1 + \exp[-\beta(E_t - E_i)]} \end{aligned} \quad (4.3)$$

Let $M_{i,j} \geq 0$ denote the (conditional) probability of transition from microstate \mathcal{C}_j to microstate \mathcal{C}_i . The transition matrix M is column stochastic:

$$\sum_i M_{i,j} = 1 \quad \forall j. \quad (4.4)$$

Also M is regular. This, in other words, means

$$(M^n)_{i,j} > 0 \quad (4.5)$$

⁴e.g. flip a randomly chosen Ising spin; subject a randomly chosen molecule to a random displacement in the phase space.

for some finite n and for all i, j . The Markov chain generated is irreducible and aperiodic. We have,

$$M^n |a\rangle \underset{n \rightarrow \infty}{\sim} |\pi\rangle, \quad (4.6)$$

for an arbitrary probability vector $|a\rangle$ having non-zero overlap with the invariant probability vector $|\pi\rangle$ i.e. $\langle \pi | a \rangle > 0$. $|\pi\rangle$ is the right eigenvector of M corresponding to eigenvalue unity. The largest eigenvalue of M is non-degenerate and equals unity. All other eigenvalues are all less than unity in modulus. These results follow from Peron-Frobenius theorems [9].

We have thus $M|\pi\rangle = |\pi\rangle$. This is called balance condition. It is often expressed more explicitly as

$$\sum_j [M_{i,j} \pi_j - M_{j,i} \pi_i] = 0 \quad (4.7)$$

Thus simple balance condition is adequate to ensure the existence of an invariant distribution. Often we impose a more strict detailed balance condition by demanding that each term in the sum above is zero. In other words,

$$\pi_j M_{i,j} = \pi_i M_{j,i} \quad (4.8)$$

Use of detailed balance helps us construct a transition matrix M whose invariant distribution is the desired distribution. Metropolis algorithm and the heat-bath algorithms obey detailed balance and hence asymptotic convergence of the Markov chain to the desired canonical ensemble is assured. Detailed balance also assures that the asymptotic part of the Markov chain has the time-symmetry required of an equilibrium system, see below.

Let \mathcal{F} denote a Markov chain generated by the transition matrix M with initial state belonging to an equilibrium canonical ensemble. Let us run the chain back wards and call it time-reversed Markov chain or simply reversal of \mathcal{F} . Let \mathcal{R} denote the time reversed Markov chain starting off from a microstate belonging to equilibrium canonical ensemble. Let \widehat{M} denote the transition matrix that generates \mathcal{R} . \widehat{M} is called the π -dual or time reversal of M . How is \widehat{M} related to M and $|\pi\rangle$? To answer this question let us consider a two-time joint probability matrix $W = MD$, where D is a diagonal matrix: $D_{i,j} = \pi_i \delta_{i,j}$ where $\delta_{i,j}$ is Kronecker delta. Let \widehat{W} be the joint probability matrix associated with time reversed Markov chain. It is clear that $\widehat{W} = W'$ where W' is the transpose of W . It follows then, $\widehat{M} = \widehat{W}D^{-1} = W'D^{-1} = DM'D^{-1}$. Thus we have

$$\pi_j \widehat{M}_{i,j} = \pi_i M_{j,i}. \quad (4.9)$$

A Markov chain is time-symmetric only when $W = \widehat{W} = W'$. In other words the matrix W should be symmetric. It is readily shown that the symmetry of W implies $\widehat{M} = M$, which is true if and only if M obeys detailed balance condition, see Eq. (4.8) and (4.9). Hence detailed balance is required to ensure that the generated Markov chain is time symmetric and hence can represent an equilibrium system. If M obeys only balance condition and not detailed balance condition the resulting Markov chain will not have time symmetry. There are several algorithms in vogue [10] that obey only balance condition and not detailed balance conditions. An algorithm that does not obey detailed balance will not generate time-symmetric Markov chain asymptotically. The implications of this to the study of equilibrium systems need careful investigations.

5. Critical slowing down

A problem with Monte Carlo technique is that the statistical error in the estimated average decreases with the sample size N only as $1/\sqrt{N}$. This implies that to get numbers to just one extra decimal accuracy we need to increase the sample size by a factor of hundred, which often is not possible. Much worse, the error reduction as $1/\sqrt{N}$ happens only if the microstates of the Monte Carlo sample are independent. Successive microstates generated by Metropolis algorithm are often correlated [11]. Such correlations increase the statistical error by a factor $\sqrt{1+2\tau^*}$ [11], where τ^* is the integrated correlation time. τ^* diverges when $T \rightarrow T_C$:

$$\tau^*(T) \underset{T \rightarrow T_C}{\sim} |T - T_C|^{-\Delta}. \quad (5.1)$$

This is called critical slowing down, in second order phase transition. From finite size scaling, we have $\tau^*(L) \underset{L \rightarrow \infty}{\sim} L^z$, where L is the Monte Carlo system size, $z = \Delta/\nu$ is the dynamical critical exponent and ν is the correlation length exponent. For Metropolis algorithm, $z \gg 1$. Hence Metropolis Monte Carlo simulation of large systems, close to criticality is nearly impossible. Cluster algorithms help overcome this drawback [12, 13].

In cluster algorithms, the given problem is mapped on to a bond percolation problem which helps define clusters of microscopic entities. The clusters are updated independently and randomly to generate successive entries in the Markov chain. The key point is that from interacting spins we construct non interacting clusters. Updating of clusters does not result in large energy changes. But it decorrelates successive microstates rather effectively and reduces τ^* by orders of magnitude

6. Super-critical slowing down

Near first order transition, the microstates representing the interface between ordered and disordered phases have intrinsically low probability of occurrence in a closed system. These microstates are scarcely sampled in Metropolis simulation of large Monte Carlo systems. The system takes a very long time to go from one phase to the other due to the presence of high energy barrier for large Monte Carlo system sizes. This is called super critical slowing down. The quality of local update Metropolis Monte Carlo results deteriorates exponentially with increase of system size.

7. Non-Boltzmann sampling

Metropolis and related algorithms belong to Boltzmann sampling techniques. For addressing the problems of super critical slowing down we need to go beyond Boltzmann sampling. That non-Boltzmann sampling can provide a legitimate alternative was recognized even during the early days of Monte Carlo practice [14]. However practical significance of non-Boltzmann sampling was realized only in the middle of seventies when Torrie and Valleau proposed the so called umbrella sampling [15]. Umbrella sampling is a forerunner to all the subsequent non-Boltzmann sampling techniques including multicanonical Monte Carlo [16]. Entropic sampling [17] is equivalent to multicanonical sampling. I shall take entropic sampling as a typical example of non-Boltzmann Monte Carlo and describe this technique in some what great details below. Before that let me make some general statements about non-Boltzmann sampling techniques.

The probability density for a closed system to have an energy E is given by,

$$P_B(E) \propto D(E) \exp(-\beta E), \quad (7.1)$$

where $D(E)$ is the density of states. Let us suppose we want to sample microstates in such a way that the the corresponding probability density of energy is given by,

$$P_g(E) \propto D(E) [g(E)]^{-1}, \quad (7.2)$$

where $g(E)$ is chosen suitably for obtaining the desired non-Boltzmann sampling. An ensemble, called g -ensemble, consistent with $P_g(E)$ is constructed as follows. Let C_i denote the i -th microstate in the Markov chain and C_t , the trial microstate. Let $E_i = E(C_i)$ and $E_t = E(C_t)$. The next entry C_{i+1} in the Markov chain is taken as C_t with probability p and C_i with probability $(1 - p)$ and

$$p = \min \left[1, \frac{g(E_i)}{g(E_t)} \right]. \quad (7.3)$$

The algorithm obeys detailed balance and hence the Markov chain constructed would converge asymptotically to the desired g -ensemble

When $[g(E)]^{-1} = \exp(-\beta E)$ we recover conventional Boltzmann sampling. For any other choice of $g(E)$ we get the corresponding non-Boltzmann sampling. Canonical ensemble average of a macroscopic property $O(C)$ can be obtained by un-weighting and re-weighting of $O(C)$ for each C sampled from the g -ensemble. For un-weighting we divide by $[g(E(C))]^{-1}$ and for re-weighting we multiply by $\exp[-\beta E(C)]$. The weight factor associated with a microstate C belonging to the g -ensemble is thus,

$$W(C, \beta) = g(E(C)) \exp[-\beta E(C)]. \quad (7.4)$$

We then have,

$$\langle O \rangle = \frac{\sum_C O(C) W(C, \beta)}{\sum_C W(C, \beta)} \quad (7.5)$$

The left hand side of the above is the equilibrium value of O in a closed system at β , while on the right side the summation in the numerator and in the denominator runs over microstates belonging to the non-Boltzmann g -ensemble. From a single simulation of a g -ensemble, with a temperature independent g function, we can calculate the canonical ensemble average of O at various temperatures.

8. Entropic Sampling

Entropic sampling obtains when $g(E) = D(E)$. For this choice, $P_g(E)$ is the same for all E . The system does a simple random walk on a one dimensional energy space. Hence all energy regions are visited with equal probability. As a result, in the case of first order phase transition for example, the microstates on the paths (in the configurational space) that connect ordered and disordered phases would get equally sampled.

A crucial issue that remains to be clarified pertains to the observation that we do not know $D(E)$ *a priori*. In entropic sampling, we employ a strategy to push $g(E)$ closer and closer to $D(E)$, iteratively. We divide the range of energy into a large number of bins of equal widths. We denote the discrete-energy version of $g(E)$ by the symbol $\{g_i : i = 1, M_E\}$, where M_E is the number of energy bins. We start with $\{g_i^{(0)} = 1 \forall i = 1, M_E\}$; the superscript is iteration run index and the subscript is energy bin index. The aim is to update $\{g_i\}$ from one iteration to the next: $\{g_i^{(0)}\} \rightarrow \{g_i^{(1)}\} \rightarrow \dots \{g_i^{(k)}\} \rightarrow \dots$, so that asymptotically we get $\{g_i\}$ as close to $\{D_i\}$ as desired, where $\{D_i\}$ is the discrete energy representation of $D(E)$.

The iteration is carried out as follows. In the k -th iteration, for example, we generate a large number of microstates employing acceptance probability based on $\{g_i^{(k)}\}$ and accumulate an histogram $\{h_i : i = 1, M_E\}$ of energy of visited microstates. We update $\{g_i^{(k)} : i = 1, M_E\}$ to $\{g_i^{(k+1)} : i = 1, M\}$, as given below,

$$g_i^{(k+1)} = \begin{cases} g_i^{(k)} & \text{if } h_i = 0, \\ g_i^{(k)} \times h_i & \text{if } h_i \neq 0, \end{cases}$$

for all $i = 1, 2, \dots, M_E$. The updated $\{g_i^{(k+1)}\}$ is employed in the next *i.e.* $(k+1)$ -th run, during which a fresh histogram of energy is generated. After each run, the histogram is examined for its uniformity. Flatter the histogram, closer is $\{g_i\}$ to $\{D_i\}$. Thus, the calculated histogram serves two purposes in entropic sampling. One for updating $\{g_i\}$ and the other for monitoring the convergence of $\{g_i\}$ to $\{D_i\}$. However, often, it is neither practical nor necessary to get a strictly flat histogram. An approximately flat histogram would be adequate, thanks to the unweighting followed by reweighting with the Boltzmann rule while calculating the averages. Hence the calculated macroscopic properties would come out right, even if $\{g_i\}$ does not converge strictly to $\{D_i\}$.

9. Wang-Landau algorithm

A simple and flexible variant to entropic sampling was proposed recently by Wang and Landau [18]. The distinguishing feature of this algorithm is the dynamic evolution of the acceptance probability, p . We update $\{g_i\}$ after every Monte Carlo step. Let us say the system visits a microstate in a Monte Carlo step and let the energy of the visited microstate fall in the m -th energy bin. Then g_m is updated to $f \times g_m$, where f is the Wang-Landau factor. The updated $\{g_i\}$ becomes operative immediately for determining the acceptance/rejection criteria from the very next trial move. We set $f = f_0$ for the zero-th run. f_0 can be any number greater than unity. The choice of $f_0 = e$ is recommended by Wang and Landau. We generate a large number of microstates employing the dynamically evolving p . At the end of a run we calculate the histogram of energy of microstates visited by the system during the run. Because of the continuous updating of p , the energy span of the density of states increases significantly and the energy histogram serves to monitor the convergence of $\{g_i\}$ to $\{D_i\}$. A run should be long enough to facilitate the system to span the energy over the desired range and to render the histogram of energy approximately flat. At the end of, say, the v -th run, the Wang-Landau factor for the next run is set as $f = f_{v+1} = \sqrt{f_v}$.

After several runs, this factor would be very close to unity. This implies that there would occur no significant change in $\{g_i\}$ during later runs. For example with the square-root rule and $f_0 = e$, we have $f_{25} = \exp(2^{-25}) \lesssim 1 + 10^{-7}$. It is clear that f decreases monotonically with increase of the run index and reaches unity asymptotically. Wang and Landau recommend the square-root rule; any other consistent rule should do equally well.

From the converged g , the desired macroscopic properties of the system can be calculated. To this end we invoke the connection between the density of states and microcanonical entropy, $S(E) = k_B \log[D(E)]$. Thus the Monte Carlo estimate of microcanonical entropy is $\alpha(E) = k_B \log[g(E)]$. For implementing such a scheme we need to normalize $g(E)$. The normalization constant should be obtained from known properties of the system. For example in Ising model, the ground state is doubly degenerate: $D(E_{\min}) = 2$. The total number of microstates equals 2^V where V is the number of Ising spins in the Monte Carlo model. $\int_{E_{\min}}^{E_{\max}} D(E)dE = 2^V$. The normalized $g(E)$ provides a good approximation to $D(E)$.

Alternately, we can take the output $\{g_i\}$ from the above and carry out a single long non-Boltzmann sampling run which generates microstates belonging to the g -ensemble. Note that during the production run we do not update $g(E)$. By un-weighting and re-weighting of the microstates generated in the production run, we can calculate the desired macroscopic properties of the system as a function of β . More importantly, it is adequate if the system visits the energy region of interest and not necessarily the entire range and the histogram of energy in the region of interest is approximately flat.

The usefulness of the Wang-Landau algorithm has been unambiguously demonstrated for systems with discrete energy spectrum. However, when we try to apply this technique to systems with continuous energy, there are serious difficulties. Liquid crystalline materials with continuous energy spectrum provide such an example. Hence we modify the Wang-Landau algorithm for simulating phase transition in liquid crystalline materials details of which can be found in [19]

10. Epilogue

We have presented a quick and brief review of a few Markov Chain Monte Carlo techniques for applications in statistical physics. We have discussed the Metropolis algorithm which launched the Monte-Carlo-for-statistical-physics business in the first place. The Metropolis algorithm remains to date the best algorithm in this field. We also saw that the Metropolis algorithm which obeys detailed balance condition ensures that the Markov chain generated has time symmetry which is an important characteristic of an equilibrium system. We also saw of critical slowing down near second order phase transition and cluster algorithms. The super critical slowing near first order phase transitions can be tackled by non-Boltzmann Monte Carlo techniques. We discussed in detail a few non-Boltzmann sampling techniques which include entropic sampling and the recent Wang-Landau algorithm. Tables (1) lists what we consider as milestones in the historical development of Monte Carlo methods in statistical physics. Of course only a few of the milestone topics were discussed here. For discussions on the other topics see [20–23].

Table 1: Milestones in Monte Carlo Statistical Physics

When?	What?	Who?
1951	Linear congruential generator	Lehmer [3]
1953	Metropolis algorithm	Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller [8]
1964	Hand Book on Monte Carlo Method	Hammersley and Handscomb [20]
1969	Random clusters	Kasteleyn, Fortuin [24]
1975	n-fold way	Bortz, Kalos Lebowitz [25]
1976	Cluster counting algorithm	Hoshen and Kopelman [26]
1977	Umbrella Sampling	Torrie and Valleau [15]
1980	Ising critical droplets	Coniglio and Klein [27]
1987	Swendsen-Wang cluster algorithm	Swendsen and Wang [12]
1988	Histogram reweighting	Ferrenberg and Swendsen [28]
1989	Wolff cluster algorithm	Wolff [13]
1991	Multicanonical and Entropic sampling	Berg and Neuhaus [16]; Lee [17]
1995	Absorbing Markov Chain	Novotny [29]
2000	A Guide to Monte Carlo Simulations in Statistical Physics	Landau and Binder [22]
2001	Wang-Landau algorithm	Wang and Landau [18]

References

- [1] N. Metropolis and S. Ulam, *J. Statistical Assoc.* **44**, 335 (1949).
- [2] A. Hall, *On an experimental determination of π* , *Messeng. Math.* **2**, 113 (1873).
- [3] D. H. Lehmer, *Ann. Comp. Lab. Harvard Univ.* **26** 141 (1951).
- [4] W. H. Press, S. Teukolsky, W. T. Vetterling and B. P. Flannery, *Numerical Recipes*, Cambridge (1992).
- [5] NAG Fortran Library Manual Mark 14, 7 Numerical Algorithm Groups Inc. (1990).
- [6] M. Cruetz, *Phys. Rev. Lett.* **43**, 553 (1979).
- [7] R. J. Glauber, *J. Math. Phys.* **4** 294 (1963).
- [8] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).
- [9] F. R. Gantmacher, *Applications of the theory of matrices*, Interscience (1959);
J. R. Norris, *Markov Chains*, Cambridge University press (1997);
J. G. Kennedy, J. L. Snell, and A. W. Knapp, *Denumerable Markov chains*, Springer-Verlag, New York (second edition) (1976).
- [10] C. J. Geyer and E. A. Thompson, *J. Am. Stat. Assoc.* **90**, 909 (1995);
D. Frenkel and B. Smit, *Understanding molecular simulations*, Academic, San Diego (1996);
M. C. Tesi, E. J. Janse van Rensburg, E. Orlandini and S. G. Whittington, *J. Stat. Phys.* **82**, 155 (1996);
M. C. Tesi, E. J. Janse van Rensburg, E. Orlandini and S. G. Whittington, *J. Phys. A* **29**, 2451 (1996);
U. E. Hansmann, *Chem. Phys. Lett.* **281**, 140 (1997);
G. Thorleifsson and M. Falcioni, *Comp. Phys. Commun.* **109**, 161 (1998);
V. I. Manousiouthakis and M. W. Deem, *J. Chem Phys.* **110**, 2753 (1999);
Onuttom Narayan and A. P. Young, *Phys. Rev. E* **64**, 021104 (2001).
- [11] H. Müller Krumbhaar and K. Binder, *J. Stat. Phys.* **8**, 1 (1973).
- [12] R. H. Swendsen and J.-S. Wang, *Phys. Rev. Lett.* **58**, 86 (1987)
- [13] U. Wolff, *Phys. Rev. Lett.* **62**, 361 (1989).
- [14] L. D. Fosdick, *Methods Comp. Phys.* **1**, 245 (1963)
- [15] G. M. Torrie and J. P. Valleau, *J. Comp. Phys.* **23**, 187 (1977)
- [16] B. A. Berg and T. Neuhaus, *Phys. Rev. Lett.* **68**, 9 (1992).
- [17] J. Lee, *Phys. Rev. Lett.* **71**, 211 (1993); Erratum: **71**, 2353 (1993).
- [18] F. Wang, and D. P. Landau, *Phys. Rev. Lett.* **86**, 2050 (2001).
- [19] D. Jayasri, V. S. S. Sastry and K. P. N. Murthy, *Phys. Rev. E* **72**, 036702 (2005).
- [20] J. M. Hammersley and D. C. Handscomb, *Monte Carlo Methods*, Chapman and Hall, London (1964)
- [21] K. Binder and D. W. Heermann, *Monte Carlo Simulation in Statistical Physics: An Introduction*, Springer (1988)
- [22] D. P. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics*, Cambridge University Press (2000)

- [23] M. E. J. Newman and G. T. Barkema, *Monte Carlo Methods in Statistical Physics*, Oxford University Press, Oxford (2001);
B. A. Berg, *Markov Chain Monte Carlo Simulations and Their Statistical Analysis*, World Scientific Publishers, Singapore, (2004);
M. H. Kalos and P. A. Whitlock, *Monte Carlo Methods, Volume I: Basics*, Wiley-Interscience Publications, John Wiley and Sons, New York (1986);
R. Y. Rubenstein, *Simulation and the Monte Carlo Method*, Wiley Series in Probability and Mathematical Statistics, John Wiley and Sons, New York (1981);
K. P. N. Murthy, *Monte Carlo: Basics*, Monograph ISRP/TD-3, Indian Society for Radiation Physics (2000). (eprint: arXiv: cond-mat/0104215 v1 12 April 2001);
K. P. N. Murthy, *Monte Carlo Methods in Statistical Physics*, Universities Press (India) Pvt. Ltd., distributed by Orient Longmann (2004).
- [24] P. W. Kasteleyn, C. M. Fortuin, *J. Phys. Soc. Japan Suppl.* **26**, 11 (1969);
C. M. Fortuin and P. W. Kasteleyn, *Physica(Utrecht)* **57**, 536 (1972).
- [25] A. B. Bortz, M. H. Kalos and J. L. Lebowitz, *J. Comp. Phys.* **17**, 10 (1975).
- [26] J. Hoshen and R. Kopelman, *Phys. Rev. B* **14**, 3438 (1976).
- [27] A. Coniglio and W. Klein, *J. Phys. A* **13**, 2775 (1980).
- [28] A. M. Ferrenberg and R. H. Swendsen, *Phys. Rev. Lett.* **61**, 2635 (1988)
- [29] M. A. Novotny, *Phys. Rev. Lett.* **74**, 1 (1995); *Erratum*, **75**, 1424 (1995); eprint: arXiv: cond-mat/9411081 (1994);
M. A. Novotny, *Computers in Physics* **9**, 46 (1995).