Monte Carlo methods for magnetic systems

Zoltán Néda
Department of Theoretical and Computational Physics
Babeş-Bolyai University, Cluj-Napoca, Romania

Monte Carlo (MC) methods [1] use random numbers to compute statistical estimates for a desired parameter on a sample population.

1. Introduction and elements of statistical thermodynamics

For magnetic systems, once the Hamiltonian (H) of the system is written, the goal of theoretical calculations is to compute the relevant thermodynamic potential. In the majority of the cases we are interested in canonical ensembles, meaning that the system is considered in an environment where the temperature, T, the external magnetic field, h, and the number of elements in the system, N, is fixed. In such cases the relevant thermodynamic potential is the free-energy, $F=F(T,h,N)$. From elementary statistical physics [2] it is known, that the free-energy, which is a macroscopic quantity, can be calculated from the microscopic elements of the model, through the partition function, Z, (sometimes called as the canonical sum).

$$F = -kT\ln(Z) \quad \text{(eq. 1)} \quad \text{where} \quad Z = \sum_i \exp\left(-\frac{E_i}{kT}\right) \quad \text{(eq.2)}$$

The sum in eq. 2 extends over all possible microstates of the system which are compatible with the conditions imposed by the ensemble. E_i is the energy of the system in microstate i. and we denoted by k the Boltzmann constant. For simplicity in the following we use the $\beta = 1/kT$ notation. Once Z is known, all other thermodynamic parameters are easy to compute by computing F and using simple thermodynamic relations [2].

The primary goal of the MC type simulations will be to estimate some statistical averages for the magnetic system at various T, h and N values: the average magnetization $<M>$, the average square magnetization $<M^2>$, the average energy $<E>$ and the average square energy $<E^2>$. From the microscopic elements of the model, these averages can be computed as:

$$< X >= \frac{1}{Z}\sum_i X_i \exp(-\beta E_i)\,. \quad \text{(eq. 3)}$$

Here X stands for M, M^2, E or E^2, and X_i denotes their values in micro-state i. Once these averages are known, the specific-heat and the static susceptibility of the system can be computed using the consequences of the general fluctuation-dissipation theorem [2]:

$$C_V = \left(\frac{\partial < E >}{\partial T}\right)_{V,N} = \frac{1}{kT^2 N}\left(< E^2 > - < E >^2\right)$$
$$\chi = \left(\frac{\partial < M >}{\partial H}\right)_{H\to 0} = \frac{1}{NkT}(< M^2 > - < M >^2) \quad \text{(eqs. 4)}$$

The major difficulties in theoretically describing magnetic systems, is that the sums from eqs.2 and 3 cannot be analytically calculated. Approximations are thus needed. Due to the spectacularly increasing computational power numerical methods gain more and more popularity. A simple estimation of the sum by taking into account all microstates is impossible, since the number of elements in the sum is exponentially increasing with the system size. Exact enumerations for very small systems ($N \approx 20$) is already time-consuming on modern supercomputers. Wise computational methods are needed to estimate quickly these sums. MC methods [1] are one of these.

2. One dimensional MC Integration

Given a function *f(x)* we intend to compute it's integral, *I*, on the *[a,b]* interval:

$$I = \int_a^b f(x)dx \quad \text{(eq. 5)}$$

The simplest numerical estimation method for this integral is by considering a uniform mesh of *n* boxes on the *[a,b]* interval and estimate the value of the function in some points in these boxes. The trapezoid sum and Simpson's methods [3] are based on this deterministic approach. Another, non-deterministic possibility is to generate randomly *n* points on the *[a,b]* interval, $x_i \in [a,b]$ and calculate *f(x)* in these. When these points are generated with a uniform distribution the $\rho(x)$ probability density for having a point with coordinate *x* is: $\rho(x) = 1/(b-a)$. In such case the integral can be estimated as:

$$I = \frac{b-a}{n}\sum_{i=1}^{n} f(x_i) \quad \text{(eq. 6)}$$

The strong law of large numbers [4] guarantees us that for a sufficiently large sample one can come arbitrary close to the desired integral. A natural question that arises at this point is, how fast the convergence is….or how big *n* should be to obtain a reasonable estimate for *I* ? The nature of this convergence will tell us whether this method is useful for a numerical estimation. The converges is usually poor for $\rho(x) = Const.$ The convergence becomes better if the shape of $\rho(x)$ approximates the shape of the *f(x)* function on the *[a,b]* interval. Particularly, the convergence is infinitely rapid if $\rho(x) = |f(x)|$ (this choice however is not possible since, normalizing $\rho(x)$ is equivalent with calculating the desired *I* integral). The important sampling Monte Carlo method will calculate the *I* integral by sampling on random points on the *[a,b]* interval according to a $\rho(x)$ distribution, which approximates the shape of |*f(x)*| (points are generated with a higher probability in those regions where |*f(x)*| is large). If we generate the x_i points according to the $\rho(x)$ distribution:

$$I = \int_a^b f(x)dx = \int_a^b \frac{f(x)}{\rho(x)}\rho(x)dx = \frac{1}{n}\sum_{i=1}^{N}\frac{f(x_i)}{\rho(x_i)} \quad \text{(eq. 7)}$$

The convergence is very fast thus, if $\rho(x) \approx |f(x)|$.

3. The Metropolis MC method

In statistical thermodynamics we intend to calculate integrals (or sums) like:

$< A >= \frac{1}{Z}\int_\Omega A(x)u[H(x)]dx$ with $Z = \int_\Omega u[H(x)]dx$, where the *x* points are elements of a very high dimensional state-space (dimensionality of the order of *N*). For magnetic systems in the *{T,h,N}*-canonical ensemble we have: $u[H(x)] = \exp[-\beta H(x)]$. We will use the important sampling method to calculate these integrals. The shape of the function under the integral is dominated by the *u[H(x)]* term, we choose thus $\rho(x)$ with shape close to $u[H(x)]$:

$$< A >= \frac{\int_\Omega A(x)\rho^{-1}(x)u[H(x)]\rho(x)dx}{\int_\Omega \rho^{-1}(x)u[H(x)]\rho(x)dx} = \frac{\sum_{i=1}^{n} A(x_i)\rho^{-1}(x_i)u[H(x_i)]}{\sum_{i=1}^{n}\rho^{-1}(x_i)u[H(x_i)]} \quad \text{(eq. 8)}$$

For the particular choice: $\rho(x) = u[H(x)]/Z$ the convergence is infinitely fast, and one gets:

$$< A >= \frac{1}{n}\sum_{i=1}^{n} A(x_i) \quad \text{(eq. 9).}$$

There is however a problem with this choice, since we do not know Z (if Z would be known the problem is solved, and no further calculations would be needed). The question is thus how can one generate the x_i points in the state-space according to the $\rho(x) = u[H(x)]/Z$ (eq. 9) normalized probability distribution without the a-priori knowledge of the Z partition function? The answer was given by Metropolis et. al [5]. Their basic idea was to use a Markov process [6] for this, so that starting from an initial x_0 state, further states are ultimately distributed according to $\rho(x)$. To generate a Markov chain one needs to specify the $W(x \to x')$ transition probabilities [6] from one micro-state to another. In order that the probability density of the generated states be $\rho(x)$, it should be satisfied:

1. The Markov chain is ergodic (any state point should be reachable from any other state-point through the Markov chain)
2. For all possible x microstates: $\sum_{x' \in \Omega} W(x \to x') = 1$
3. For all possible x microstates: $\sum_{x' \in \Omega} W(x \to x')\rho(x') = \rho(x)$ (eq. 11)

Instead of condition 3 (eq. 11), a stronger but simpler condition can be used, the detailed balance: $W(x \to x')\rho(x) = W(x' \to x)\rho(x')$ (eq. 12)

By constructing such a Markov chain one will be able to generate state-space points according to the desired $\rho(x)$ probability density, without the prior knowledge of Z. For example, if one considers a system in a canonical ensemble ($u[H(x)] \sim \exp[-\beta H(x)]$) an immediate choice to satisfy (eq. 12) is the one chosen in [5], i.e. the Metropolis algorithm

$$W(x \to x') = \begin{cases} \exp(-\beta \Delta E(x,x') \cdots for \cdots \Delta E(x,x') > 0 \\ 1 \cdots\cdots\cdots\cdots\cdots\cdots for \cdots \Delta E(x,x') \le 0 \end{cases} \quad \text{(eq. 13)}$$

where, $\Delta E(x,x') = H(x') - H(x)$.

A simple Metropolis Monte Carlo method for calculating the desired statistical averages would have thus the following structure:

1. Design an ergodic Markov process on the possible microstates (each states should be reachable from the others).
2. Specify an initial x microstate for starting.
3. Choose randomly a new x' microstate (preferably so that: $W(x \to x') > 0$).
4. Compute the value of $W(x \to x')$.
5. Generate a uniformly distributed $r \in [0,1]$ random number.
6. If $r \le W$ accept as new state x' and consider the x' point in the average from eq. 9.
7. If $r > W$ the system remains in x, and count again x in the average from eq. 9.
8. Repeat steps 1→ 7 many times until the average converges.

Initially the system is „heated up" with a large number of transient steps, so that the systems approaches the desired $\rho(x)$ distribution. During these transient steps we repeat the algorithm from 1→7 without considering the generated points for the average in eq. 9.

4. The Metropolis MC algorithm for localized Ising spin systems

As an example, let us consider here the simplest two-dimensional spin system capable of a second order ferromagnetic-paramagnetic phase-transition: the 2D Ising model without an external magnetic field. The Hamiltonian of the system writes as:

$H = -J \sum_{<i,j>} \sigma_i \sigma_j$ (eq. 14), where the summation is on all nearest-neighbor spin pairs, $\sigma_i = \pm 1$ are two-state Ising spin variables and $J>0$ is the ferromagnetic coupling constant. In

case the Hamiltonian is more complicated (interaction terms with the external magnetic field, crystalline anisotropy terms or long-range dipolar magnetic interaction terms are considered) the method remains similar. We consider the problem on a square-lattice with sizes: $N = L \times L$. The following averages have to be calculated with the Metropolis sampling at various temperatures:

$$< M >= \left\langle \sum_i \sigma_i \right\rangle; \quad < M^2 >= \left\langle \left(\sum_i \sigma_i \right)^2 \right\rangle; \quad < E >= \left\langle H(\sigma_i) \right\rangle; \quad < E^2 >= \left\langle \left(H(\sigma_i) \right)^2 \right\rangle$$

The order parameter characterizing the magnetic order for $J>0$ can be defined as: $m = M / N$. The Metropolis Monte Carlo algorithm is straightforward then:

1. Fix a temperature (T).
2. Consider an initial spin configuration ($\{\sigma_i\}$, for example consider for all $i = 1, N \rightarrow \sigma_i = +1$)
3. Calculate the initial value of E and M.
4. Consider a new spin configuration by virtually "flipping" one randomly selected spin.
5. Calculate the energy E' of the new configuration, and the energy change, ΔE, due to this spin-flip.
6. Calculate the Metropolis $W(x \rightarrow x')$ probabilities for this change.
7. Generate a random number r between 0 and 1.
8. If $r \leq W$ accept the flip and update the value of the energy to E' and magnetization to M'
If $r > W$ reject the spin flip, and take again the initial E and M values in the needed averages.
9. Repeat the steps 4 - 8 many times ("heat up" the system before averages are calculated)
10. Repeat the steps 4 - 8 by collecting the values of E, E^2, M, M^2, for the needed averages.
11. Compute the averages for a large number of microstates.
12. Calculate the values of $<m>$, $<E>$, $<C_v>$ and $<\chi>$ using the given formulas at this T.
13. Change the temperature and repeat the algorithm for the new temperatures as well.
14. Construct the desired $<M(T)>$, $<E(T)>$, $<C_v(T)>$, $<\chi(T)>$ curves

5. Finite-size effects

The results of statistical thermodynamics are valid only for thermodynamic systems, meaning that one works in the $N \rightarrow \infty$ limit. Computer simulations cannot be made on infinite systems, so due to the finite size of the simulated lattices, finite-size effects are expected. The divergence in the specific heat and susceptibility (expected for infinite systems) at the critical (Curie) temperature (T_c), will not appear in MC simulations on a finite lattice (Fig. 1).

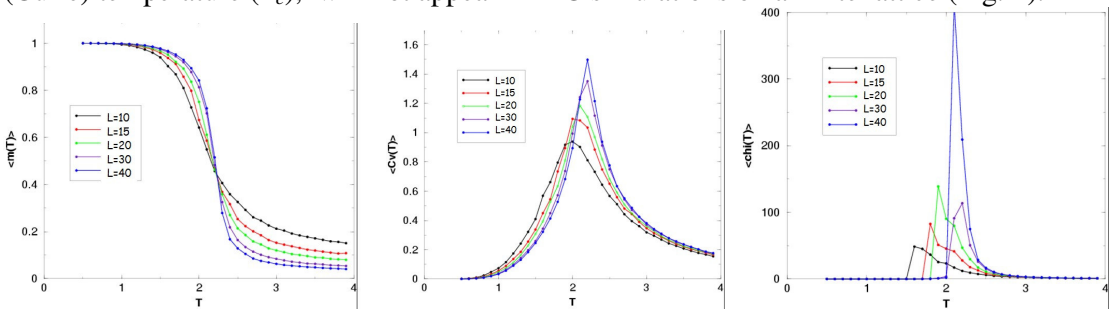

Fig. 1. *Finite-size effects for Metropolis MC simulations on a 2D Ising model Results for the order parameter (m), specific (C_V) and susceptibility (chi) are plotted as a function of temperature (T/k). (J=0.44 was chosen, so that $T_c(L = \infty) = 1$) for different lattice sizes, L.*

The reason for this is the obvious cutoff of the correlation length by the finite lattice size. Instead of a divergence at T_c, a peak is obtained in the vicinity of T_c. The height of the peak

and it's location is changing with lattice size, L (see Fig. 1), so a carefull extrapolation for $L \to \infty$ is needed [7].

6. Efficient MC techniques

The simple Metropolis algorithm as described above suffers from a series of drawbacks, solved by other MC algorithms.

- At low temperature the algorithm is inefficient due to the fact that after the equilibrium is reached (spins are ordered) most of the spin-flips are rejected. Long simulations are needed to get a reasonable estimate for the averages. This drawback is eliminated by the BKL MC algorithm [8] .
- In the neighborhood of T_c the Metropolis algorithm is again inefficient due to the critical slowing down. This critical slowing down is due to the fact that for the Metropolis algorithm the dynamical critical exponent is large ($z=2$) leading to quick divergence of the relaxation time. Many Metropolis MC steps are needed to generate statistically independent configurations. This problem is partially solved by the cluster algorithms elaborated by Swendsen and Wang [9] or Wolff [10].
- In the Metropolis MC simulations a lot of information is wasted. We use just the first and second moments of the magnetization and energy, however their distribution functions is also available. In the histogram and multi-histogram MC methods [11] one will compute averages at various temperatures from simulation results at only one fixed temperature.
- Quantum-statistical systems (Hubbard model, Stoner model, T-J model, etc..) can be also studied by MC techniques. For such problems the Quantum Monte Carlo methods are suited [12].
- Frustrated, spin-glass type systems can be studied by several MC algorithms. One of these is the simulated annealing method [13].

References:

[1] K. Binder and D.W. Hermann, *Monte Carlo Simulations in Satistical Physics. An introduction* (4th Edition) Springer (2002)

[2] C. Kittel and H. Kroemer, *Thermal Physics* (2nd, Edition), Freeman, (1980)

[3] R. Hamming, *Numerical Methods for Scientist and Engineeres* (2nd Edition), Dover, (1987)

[4] W. Feller, An *Introduction to Probability Theory and its Applications*, Vol. 1, (3rd Edition), Wiley, (1968)

[5] N. Metropolis, A.W. Roesenbluth, M.N. Rosenbluth, N. Marshall, A.H. Teller and E. Teller, *J. Chem. Phys.*, vol. 21, 1087 (1953)

[6] The Markov process is a simple stochastic process where the statistical properties of the immediate future are uniquely determined from the present, regardless of the past. Let us denote by $P(x_n \mid x_{n-1}, x_{n-1},..., x_0)$ the probability that x_n is the new state if previously the system was in states: $x_0, x_1, ..., x_{n-1}$. For a Markov process: $P(x_n \mid x_{n-1}, x_{n-1},..., x_0) = P(x_n \mid x_{n-1}) = W(x_{n-1} \to x_n)$.

[7] J.L. Cardy, *Finite-size scaling*, North-Holland, (1988)

[8] A.B. Bortz, M.H. Kalos and J.L. Lebowitz, *J. Comp. Phys.*, vol. 17, 10 (1975)

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

[10] U. Wolff, *Phys. Rev. Lett.* 62, 361–364 (1989)

[11] A.M. Ferrenberg and R.H. Swendsen; *Phys. Rev. Lett.*, vol. 61, 2635 (1988)

[12] J. Tobochnik, G. Batrouni and H. Gould, *Computers in Physics*, vol. 6, 673 (1992)

[13] S. Kirckpatrick, G.D. Gelatt and M.P. Vecchi, *Science*, vol. 220, 671 (1983)