PHYSICAL REVIEW
LETTERS

VOLUME 71 | 12 JULY 1993 | NUMBER 2

New Monte Carlo Algorithm: Entropic Sampling

Jooyoung Lee
Supercomputer Computations Research Institute B-186, Florida State University, Tallahassee, Florida 32306-4052
(Received 28 April 1993)

We present a new Monte Carlo sampling algorithm, with which one can obtain any desired distribution of the sampling in one Monte Carlo simulation. The free energy and the entropy of a system can thus be obtained from a simple exercise of this algorithm. The main idea is to sample directly the entropy of a system at infinite temperature. Importance sampling is shown to be a particular case of the new algorithm. The algorithm is tested against the exact partition function of the $L = 4$ simple cubic Ising model. A comparison with the multicanonical ensemble for the $L = 12$, $q = 10$ Potts model shows that the new algorithm is more general and more efficient.

PACS numbers: 05.50.+q, 11.15.Ha, 64.60.Fr, 75.10.Hk

With fast-growing computer technology, Monte Carlo (MC) simulations have been very successful in studying various statistical systems including neural networks, problems in biology and chemistry, lattice-gauge theories, and optimization problems in various areas, not to mention the statistical physics in the study of phase transitions and critical phenomena. Calculation of entropy from simulation methods [1, 2] has been a very difficult task, since entropy is a function of the probability with which a typical equilibrium configuration of the system is sampled. Several methods [3–8] have been suggested, and they have some advantages and disadvantages.

Almost all MC sampling algorithms are based on the idea of importance sampling, introduced by Metropolis *et al.* [1]. The thermodynamic average $\langle \mathcal{O} \rangle$ of an observable $\mathcal{O}(x)$ can be estimated [9] by

$$\bar{O} = \frac{\sum_{i=1}^{n} \mathcal{O}(x_i) P^{-1}(x_i) \exp[-\beta \mathcal{H}(x_i)]}{\sum_{i=1}^{n} P^{-1}(x_i) \exp[-\beta \mathcal{H}(x_i)]}, \quad (1)$$

where x_i represents a configuration at time i of a given system with a Hamiltonian $\mathcal{H}$. β is the inverse temperature with $k_B = 1$, and $P(x)$ is a sampling probability. Equation (1) becomes exact in the infinite time limit, $n \to \infty$. If $P(x)$ is chosen to be constant, very few samples contribute significantly to the sum in Eq. (1), and an enormously large value of n is required to get a reasonable estimate of $\langle \mathcal{O} \rangle$.

Importance sampling comes in if one chooses $P(x)$ as the Boltzmann weight $\exp[-\beta \mathcal{H}(x)]$. It is generally a good sampling algorithm for Eq. (1), but it can fail to access all the possible equilibrium distributions of the partition function if there exists a large barrier between them (i.e., the required sample size n becomes too large). One confronts similar problems if one needs to use the histogram method [10] for a wide range of the coupling constant, say temperature T, due to the small number of samples away from the equilibrium position of the internal energy E.

Recently three approaches have been suggested to overcome these problems. The first one is the multihistogram method [11, 12], in which many histograms with overlapping distributions serve as connecting information bridges. This approach can widen the range of T [13, 14], but it is not suited for sampling systems with a large barrier. The second approach is to use many microcanonical ensembles [15–17]. This approach is used mostly to obtain the whole partition function numerically for relatively small systems. These two methods suffer from the following two common difficulties: Errors propagate through the neighboring sets of data, requiring a systematic and elaborate study of error propagation [11], and many simulations at different temperatures are necessary.

Recently the multicanonical ensemble (MCE) was introduced by Berg and Neuhaus [18] as an approach to simulate strong first-order phase transitions. It has been very successful in sampling over a large barrier. The sam-

0031-9007/93/71(2)/211(4)$06.00

pling is based on a Boltzmann-like weight, but it uses different sets of temperatures for small intervals of internal energy, in such a way as to drag a system across the barrier in either direction. The set of parameters needs to be tuned to get the desired uniform distribution of the observable. The tunneling time was shown [18] to increase only as a power of the system size, which is an enormous improvement compared with exponential increase with conventional importance sampling.

In this Letter we present a new MC sampling algorithm, with which one can obtain any distribution of the sampling in one MC simulation. The conventional importance sampling is shown to be a particular case of the new algorithm. We first describe the algorithm and then consider as an example the $L = 4$ simple cubic Ising model by numerically obtaining all the coefficients of the partition funtion. The results of the new algorithm are also compared with existing MCE [18] data for the $L = 12$, $q = 10$ Potts model [19].

The main idea of the new algorithm lies in the sampling of the entropy of a given system at infinite temperature. For this work, configurations will be projected onto internal energy E, but the method can be generalized for any quantity such as magnetization, or for combinations of several order parameters.

A partition function can be written as

$$Z(\beta) = \sum_x \exp[-\beta E(x)] = \sum_E \Omega(E) \exp(-\beta E) = \sum_E \exp[S(E) - \beta E], \quad (2)$$

where $\Omega(E)$ and $S(E)$ are the density of states and the bulk entropy for a given E. $\sum_x$ is the sum over all possible configurations. In the case of importance sampling, to obtain a Boltzmann distribution $P(x) \propto \exp[-\beta E(x)]$ or $P(E) \propto \exp[S(E) - \beta E]$, it is sufficient to impose, along with ergodicity, the detailed balance condition,

$$\frac{W(x \to x')}{W(x' \to x)} = \exp\{-\beta[E(x') - E(x)]\}, \quad (3)$$

since the ratio of the transition probabilities W depends only on the change in E.

At infinite temperature, we have $Z(\beta = 0) = \sum_x 1 = \sum_E \exp S(E)$. By exact analogy, if one wants to obtain an arbitrary distribution

$$P(E) \propto \exp[A(E)] = \exp[S(E) - J(E)], \quad (4)$$

it is sufficient to impose, along with ergodicity, a similar detailed balance condition,

$$\frac{W(x \to x')}{W(x' \to x)} = \exp\{-[J(E(x')) - J(E(x))]\}. \quad (5)$$

Equations (4) and (5) are the main equations upon which the new algorithm is based. It is trivial to observe that importance sampling is a particular case of the new algorithm, with $J(E) = \beta E$.

Therefore to obtain the partition function, one need only obtain $S(E)$. Our goal is then to obtain $S(E)$ from a chosen $J(E)$, and to obtain $P(E)$ from MC simulations. For example, $J(E) = S(E)$ with the exact $S(E)$ will produce, within the statistical errors, a completely uniform distribution of internal energy over the entire energy space. With a reasonably rough estimate of $S(E)$ one is able to sample the entire energy space in one MC simulation to obtain a better estimate of $S(E)$, provided that the MC time is long compared with the ergodicity time [18].

In practice, a rough estimate of $S(E)$ is necessary before a long MC simulation can be undertaken. It can be obtained quite easily for small system sizes, without any background information about the model, by iterating in the following fashion: (1) Choose a small system and start with a random configuration. Initially $J(E) = S(E)$ is set to zero for all E. (2) Obtain the histogram $H(E)$ of E for a short MC run. A trial move is accepted according to Eq. (5). (For the first iteration, this is just a random sampling.) (3) New estimate of $S(E)$ is given now as follows:

$$S(E) = \begin{cases} J(E) & \text{for } H(E) = 0, \\ J(E) + \ln H(E) & \text{otherwise.} \end{cases} \quad (6)$$

This defines one iteration. With the new estimate of $J(E) = S(E)$, one repeats the iteration until one reaches a desired range (or the whole range) of the internal energy sampled. Sometimes it is a good idea to increase the number of MC sweeps as the covered energy range becomes large. Since the bulk entropy scales with volume, $S(E)$ estimated from a given system size serves as an excellent guess for $S(E)$ for a larger system. We find that a rough estimate of the entropy $S(E)$ can be obtained with minimal usage of computer time compared to the time required for the final run.

We have tested the new algorithm on the $L = 4$ simple cubic Ising model. The Hamiltonian is given by

$$-\beta\mathcal{H} = \beta \sum_{\langle i,j \rangle} \sigma_i \sigma_j, \quad (7)$$

where $\sigma_i = \pm 1$, and the sum is over all nearest-neighbor pairs. The partition function can be written as a sum of polynomials in $u = \exp(-4\beta)$,

$$Z(u) = \sum_{n=0}^{96} N(n) u^n, \quad (8)$$

where the sum over $n \equiv E/4$ runs over the interval $[0, 3L^3/2]$ for even L. For $L = 4$, $n = 0(96)$ corresponds to complete (anti)ferromagnetic order. It takes only 3000 MC sweeps to obtain a rough estimate of $S(E)$. We performed about 10^6 MC sweeps to get a better estimate of $S(E)$ and 56 successive 10^6 MC sweeps to get the final results. Errors have been obtained by measuring quantities at each 10^6 MC sweeps. One MC sweep here corresponds to one trial update of the whole system. One can easily obtain $N(n)$ for small n by exact enumeration. Using exact value of $N(9) = N(87) = 71\,680$ [15, 16], we restricted $9 \le n \le 87$ (which was not necessary

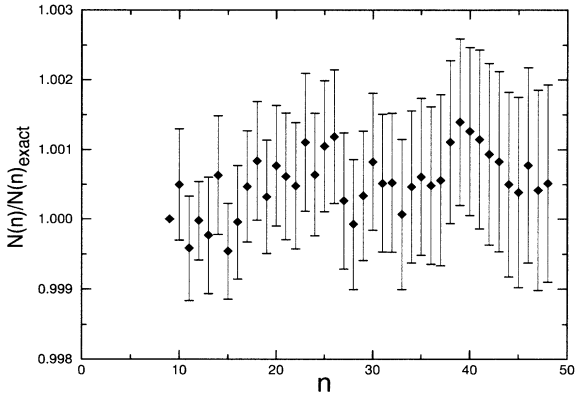

FIG. 1. Numerically obtained coefficients of the partition function of the $L = 4$ simple cubic Ising model. The values are normalized to 1 using the exact results from Ref. [20]. Values for $n \geq 49$ are not shown since they are mirror images of $n \leq 47$. $N(n)$ varies over an order of 10^{14} between $n = 9$ and $n = 48$. $N(9) = N(87) = 71\,680$ is used as input.

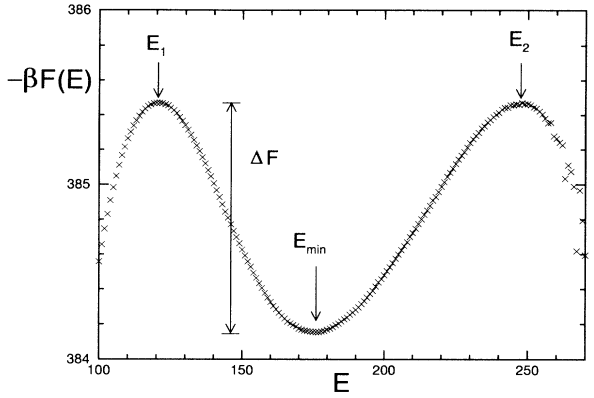

FIG. 2. Restricted free energy of $L = 12$, $q = 10$ Potts model, $-\beta F = \beta E + S(E)$, at $\beta_c(L) = 1.407\,77$ from the numerically obtained $S(E)$ for $100 \leq E \leq 273$. It is proportional to the logarithm of the Boltzmann equilibrium distribution.

though). In Fig. 1, we show $N(n)$ from our algorithm as compared with the exact values [20]. The overall agreement is quite impressive. We can estimate the degeneracy of the ground state as 1.999 9(3) by normalizing the sum of the density of states as 2^{64}. We find the zero of the partition function closest to the real axis to be $0.384\,279(21) + i0.087\,755(21)$ which is in excellent agreement with the exact value $0.384\,282 + i0.087\,739$ [15, 17, 20]. The estimate is improved compared to the result in Ref. [17], where roughly 10 times as many equivalent MC sweeps were performed in many microcanonical ensembles. We also find that for an accuracy up to 10^{-6} for the zero of the partition function closest to the real axis, we need only $0 \leq n \leq 49$. Currently we are applying the algorithm to larger lattice sizes. For $L = 14$, we find [21] that we need less than 700 intervals in n out of 4117, which is a very encouraging result for the future calculations for even bigger systems.

The MCE [18, 22] has been very successful in sampling over a large free-energy barrier, such as in the case of strong first-order transitions. To see how and why it works from the viewpoint of our algorithm, we applied our algorithm on the $L = 12$, $q = 10$ Potts model, with $100 \leq E \leq 273$, where E, the negative of the internal energy, is defined as (S in Ref. [18])

$$E = \sum_{\langle i,j \rangle} \delta_{q_i,q_j}, \tag{9}$$

where δ is a Kronecker delta, and $q_i = 0, \ldots, 9$. We performed 4×10^7 MC sweeps. For each 10^7 MC sweep we estimated new $S(E)$. In Fig. 2 we show the restricted free energy [23] $-\beta F(E) = \beta E + S(E)$ at $\beta_c = 1.407\,77$ where the two peaks are equal in height. In Fig. 3, we show $-dS(E)/dE \approx -[S(E+1) - S(E-1)]/2$ along with $\beta_c(L) = 1.407\,77$. This derivative is the solution of $dF(E)/dE = 0$, and only if $\beta(E) = -dS(E)/dE$ is the distribution of energy at E completely uniform in conventional importance sampling. A trial move was updated according to Eq. (5), which is equivalent to using $\beta(E)$ for each individual E, consequently creating a uniform distribution over the entire range of E in MC sampling. MCE can be viewed as a set of step functions to approximate Fig. 3 between E_1 and E_2. With a finite range of intervals as in Ref. [18], MCE cannot give a completely uniform distribution, even in infinitely long sampling time. Therefore it is not surprising to observe that the tunneling time from our algorithm is slightly less than half for $L = 12$ as compared to a Metropolis algorithm [21], whereas for the MCE, it reduced by a factor of 1.45 [18] compared to a heat-bath algorithm. While this exercise gave us a better understanding of the MCE, we find some discrepancy with Ref. [18] for various quan-

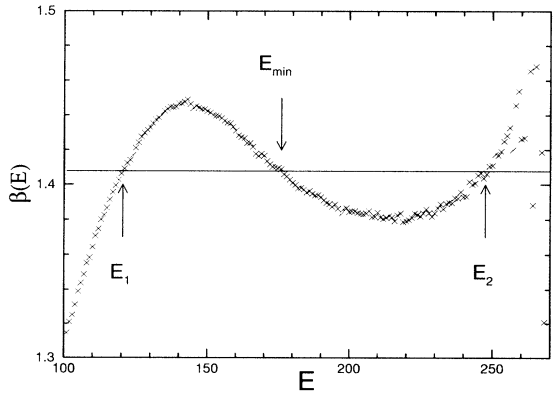

FIG. 3. The function $\beta(E) = -dS(E)/dE$, as a numerical derivative of the entropy. In conventional importance sampling, the distribution of energy is flat only at E if the simulation is performed at $\beta(E)$.

TABLE I. Comparison of the numerical estimates of various quantities for the $q = 10$ Potts model between this work and Ref. [18].

	L	$\beta_c(L)$	$E_1(L)$	$E_{\min}(L)$	$E_2(L)$	$F_L^s = \Delta F/L$
This work	12	1.407 77(11)	120.5(3)	176.0(5)	247.5(3)	0.1100(4)
Ref. [18]	12	1.407 38(09)	116	169	243	0.1071(6)
This work	16	1.415 14(06)	223.0(6)	315.7(10)	436.5(5)	0.1091(7)
Ref. [18]	16	1.415 34(12)	216	309	429	0.1086(7)
This work	24	1.421 08(04)	522.2(15)	718.5(23)	975.8(5)	0.1065(6)
Ref. [18]	24	1.42100(08)	523	723	978	0.1058(8)

tities as listed in Table I, for $L = 12, 16, 24$. Some of our results are checked also by a Metropolis algorithm. For E_1 and E_2 our data seem to be more consistent with L^{-1} scaling [23] than corresponding data from Ref. [18]. We are currently checking on larger sizes using heat-bath type updates in our algorithm.

In summary, we present a powerful and general Monte Carlo sampling algorithm, based on the sampling of the entropy at infinite temperature, from which one can sample any distribution of the observables. One can then obtain the whole partition function from a single Monte Carlo simulation. The conventional importance sampling is shown to be a particular case of the new algorithm. An application to the $L = 4$ simple cubic Ising model reproduced all the coefficients of the exact partition function within 0.1%. A comparison with the multicanonical ensemble for the $L = 12$, $q = 10$ Potts model shows that the new algorithm is more general and more efficient. The algorithm should be of wide applicability, since it can be applied to any systems where conventional Monte Carlo simulation methods are applicable.

We would like to thank M. A. Novotny, P. A. Rikvold, H. Meirovitch, H. Park, M. Y. Choi, and J. M. Kosterlitz for valuable dicussions. We also thank Bryan M. Gorman for proofreading. This work was supported by the Florida State University Supercomputer Computations Research Institute (U.S. DOE through Contract No. DE-FC05-85ER25000).

[1] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
[2] B. J. Alder and T. E. Wainwright, J. Chem. Phys. **31**, 459 (1959).
[3] J. P. Hansen and L. Verlet, Phys. Rev. **184**, 151 (1969).
[4] Z. W. Salsburg, J. D. Jacobson, W. Fickett, and W. W. Wood, J. Chem. Phys. **30**, 65 (1959).
[5] J. P. Valleau and D. N. Card, J. Chem. Phys. **57**, 5457 (1972).
[6] H. Meirovitch, Chem. Phys. Lett. **45**, 389 (1977).
[7] K. K. Mon, Phys. Rev. Lett. **54**, 2671 (1985).
[8] K. Binder, in *Monte Carlo Methods in Statistical Physics,* edited by K. Binder (Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1986).
[9] L. D. Fosdick, Methods Comput. Phys. **1**, 245 (1963).
[10] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **61**, 2635 (1988).
[11] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **63**, 1195 (1989).
[12] C. H. Bennett, J. Comput. Phys. **22**, 245 (1976).
[13] E. Marinari, Nucl. Phys. **B235 [FS11]**, 123 (1984).
[14] N. A. Alves, B. A. Berg, and R. Villanova, Phys. Rev. B **41**, 383 (1990).
[15] B. Bhanot, S. Black, P. Carter, and R. Salvador, Phys. Lett. B **183**, 331 (1987).
[16] K.-C. Lee, J. Phys. A **23**, 2087 (1990).
[17] B. Bhanot, R. Salvador, S. Black, P. Carter, and R. Toral, Phys. Rev. Lett. **59**, 803 (1987).
[18] B. A. Berg and T. Neuhaus, Phys. Rev. Lett. **68**, 9 (1992).
[19] R. B. Potts, Proc. Cambridge Philos. Soc. **48**, 106 (1952); F. Y. Wu, Rev. Mod. Phys. **54**, 235 (1982).
[20] R. B. Pearson, Phys. Rev. B **26**, 6285 (1982).
[21] J. Lee (unpublished).
[22] B. A. Berg, U. Hansmann, and T. Neuhaus, Phys. Rev. B **47**, 497 (1993).
[23] J. Lee and J. M. Kosterlitz, Phys. Rev. Lett. **65**, 137 (1990); Phys. Rev. B **43**, 3265 (1991).