Rejection-free microcanonical Monte Carlo method

1996 J. Phys. A: Math. Gen. 29 L505

(http://iopscience.iop.org/0305-4470/29/20/001)

J. Phys. A: Math. Gen. **29** (1996) L505–L509. Printed in the UK

LETTER TO THE EDITOR

Rejection-free microcanonical Monte Carlo method

C M Care

Materials Research Institute, Sheffield Hallam University, Pond Street, Sheffield S1 1WB, UK

Received 28 June 1996

Abstract. Lee has recently described a 'rejection-free' microcanonical Monte Carlo technique in which data is only collected near the centre of a microcanonical window. It is shown here that data can be efficiently collected from other energy levels in the window and that a 'rejection-free' technique can also be established for these levels. This leads to a significant improvement in efficiency over the method of Lee.

Lee [1] has described a rejection-free microcanonical Monte Carlo technique which collects data only from energy levels near the centre of a microcanonical window. In both [1] and an earlier paper [2] he asserts that it is inefficient to use data from other energy levels because of repeated rejected moves. However, it is shown below that the quality of the data collected from the majority of the other levels is essentially the same as those considered by Lee. It is further shown that it is possible to devise a rejection-free method to collect data from these other levels and that use of the method gives a significant improvement in efficiency over the rejection-free method of Lee. The rejection-free method described here, like that of Lee's, is similar in spirit to the n-fold way algorithm of Bortz *et al* [3].

There have been numerous studies of microcanonical Monte Carlo methods [4–9]. We consider here a system with discrete energy levels, E_i, and a density of states ρ_i at each level. A microcanonical algorithm is typically used to determine the density of states ratios $R_i = \rho_{i+1}/\rho_i$ with a scheme of the following general form

(i) A state is generated with energy E_i which lies within a contiguous set of energy levels, W, bounded by a minimum (maximum) allowed energy E_L (E_U) with

$$W = \{E_i : E_L \leqslant E_i \leqslant E_U\} \tag{1}$$

(ii) A new state is obtained from the current state by a Monte Carlo process with a symmetric stochastic transition matrix. In this work, and that of Lee's, this is achieved by a random spin flip.

(iii) If the new state has $E_i \in W$, the state is accepted, otherwise the old state is retained.

Steps (ii) and (iii) are repeated N_s times and a record is kept of the number of times, N_i, a state with energy E_i is observed. It is essential that rejected moves are included in this counting process. The complete algorithm generates a Markov chain with a symmetric stochastic transition matrix and such a matrix has a left eigenvector $(1, 1 \ldots 1)$ with eigenvalue unity. Provided the matrix is irreducible, this will be the limiting distribution of the Markov chain [11] and consequently

$$R_i \simeq E[N_{i+1}/N_i] \tag{2}$$

0305-4470/96/200505+05$19.50

where $E[f]$ is the expected value of the random variable, f. In practice, it is wise to run the algorithm for N_{equ} steps in the energy window, in order to allow equilibration of the system within W.

Attempted moves on states near the edges of the energy window W lead to rejected moves if the trial states generated by (ii) do not lie within W. This problem becomes particularly severe for energies near the ground state, because for low energies the R_i's are large. In this case, the algorithm mainly samples the highest energy level in the window and a large number of moves are rejected. For this reason, Lee only collects data from states at the centre of W for which all attempted moves are accepted. For a spin-$\frac{1}{2}$ Ising system in three dimensions on a simple cubic lattice this requires three 'guard' levels at the edges of W and the number of moves required to sample one state in the centre of W is $O(R^4)$ where R is a typical value for R_i in W. However, for a spin-1 system in two dimensions, the number of required 'guard' levels rises to eight at each edge of W. The number of moves required to make one observation of a level at the centre of W rises to $O(R^9)$ [8] and the use of guard levels becomes impractical.

In order to test Lee's original method [1], we undertook a simulation of an $L\times L\times L$ spin-$\frac{1}{2}$ Ising model ($E=-J\sum_{\langle i,j\rangle}\sigma_i\sigma_j$) on a simple cubic lattice using the algorithm described above, with the width of the energy window, W, set to eight and with W positioned close to the ground state. These levels were chosen because the low energy states give rise to the greatest inefficiency in the microcanonical method and also because exact results are available for the density of states [10]. The results of two simulations for $L=10$ and $L=20$ are shown in table 1 and in each case the runs consisted of $N_{\mathrm{equ}}=10^5$ equilibration moves and $N_s=4\times10^8$ data collection moves. The results were not sensitive to the choice of of N_{equ}. The energy values are given as $e_i=(E_i+qL^3/2)/(4J)$ where $q=6$ is the co-ordination number of the lattice. The fractional standard error f_i^{est} in the estimates of R_i is calculated by dividing the data into 40 blocks of attempted moves and determining the variance in the block averages. In the table we quote the quantity χ defined by

$$\chi=\left|\frac{R_i^{\mathrm{exact}}-R_i}{R_i^{\mathrm{exact}}f_i^{\mathrm{est}}}\right| \tag{3}$$

and it can be seen that all the values of χ are $O(1)$. Thus f_i^{est} represents an acceptable method of estimating the error in the simulation. If each data point were statistically independent we would expect the number of observations at each level in the window to be described by a multinomial distribution with variance for the number of observations at each level given by $N_s p_i(1-p_i)$, where p_i is the probability of observing the system in energy level E_i. We define a 'statistical inefficiency', τ, as the ratio of the measured variance to that predicted for independent data. In the calculation of the statistical inefficiency, we estimate the p_i from the total post-equilibration data set. The statistical inefficiency may equivalently be calculated as the ratio of (i) the actual number of observations at an energy level to (ii) the number of observations which would be needed to produce the same fractional error from independent data. As can be seen in table 1, the statistical inefficiency of the data with repeatedly rejected moves is very high but the fractional error is little different from the data collected near the centre of the window for which there are no rejected moves. Thus, although a large number of states are rejected, this is compensated for by the number of times these states are observed.

The method may be improved by classifying the sites in the manner shown in table 2 [3, 1]. We may then only consider sites at which a spin flip will lead to a state lying within W and hence give a successful move. In order to use this approach it is necessary to maintain appropriate tables after each spin flip [3] and this is computationally efficient only

Table 1. Typical results of standard microcanonical Monte Carlo simulation of three dimensional $S = \frac{1}{2}$ Ising model on a simple cubic lattice with $N_s = 4 \times 10^8$. e_i, N_i, R_i, f_i^{est}, R_i^{exact}, χ and τ are defined in the text.

L^3	e_i	N_i	R_i	$100 \times f_i^{\text{est}}$	R_i^{exact}	χ	τ
	12	6.026×10^3	0.2861	2.6	0.2889	0.38	1.1
	13	1.724×10^3	41.65	2.3	40.98	0.72	0.81
	14	7.180×10^4	16.39	0.68	16.56	1.5	2.8
10^3	15	1.177×10^6	0.5698	0.37	0.5673	1.2	7.7
	16	6.705×10^5	25.90	0.40	25.87	0.3	4.6
	17	1.737×10^7	11.12	0.39	11.17	1.0	1.7×10^2
	18	1.932×10^8	0.9706	0.33			2.1×10^3
	19	1.875×10^8					1.8×10^3
	12	1.57×10^2	0.03185	43	0.03603	0.27	1.5
	13	5×10^0	427.6	43	332.6	0.67	0.90
	14	2.138×10^3	123.5	3.8	133.1	1.89	3.1
20^3	15	2.641×10^5	0.07368	2.8	0.07129	1.2	4.6
	16	1.946×10^4	208.4	4.1	210.0	0.18	15
	17	4.055×10^6	86.55	3.2	88.71	0.77	4.0×10^3
	18	3.510×10^8	0.1273	2.7			4.0×10^4
	19	4.467×10^7					3.6×10^4

Table 2. Classification of sites for two dimensional $S = \frac{1}{2}$ Ising model on a square lattice. $\Delta E/4J$ is the change in energy if the spin is flipped. In general, the number of classes is $q+1$ where q is the co-ordination number of the lattice.

Class	Spin	Number of spin-up nearest neighbours	$\Delta E/4J$
1	Up	4	2
2	Up	3	1
3	Up	2	0
4	Up	1	-1
5	Up	0	-2
5	Down	4	-2
4	Down	3	-1
3	Down	2	0
2	Down	1	1
1	Down	0	2

for the relatively low energy states where there will be a high number of rejected moves. It is important to note that the result (2) can only be used directly on levels for which every move is accepted. Provided the window W is made sufficiently wide, there will be some such levels at the centre of W and this is the approach adopted by Lee. However, this wastes the move which must still be made in the boundary levels of W, and it is straightforward to develop an algorithm which makes use of these levels.

Consider a state for which it is known that N_r sites will lead to rejected moves and N_a sites will lead to accepted moves. If we run the simple spin flip algorithm there will be a number of rejected moves before a successful move is undertaken and the number of rejected moves must be counted if the result (2) is to be used. However, we can avoid undertaking the rejected moves if we consider the underlying process in the random flip

method. At each flip there is a probability $P_r = N_r(N_r + N_a)$ of the move being rejected and a probability $1 - P_r$ of the move being accepted and we therefore need to generate a random number with the same distribution as the number of flips needed to obtain a successful flip. The probability of making k flips, of which only the last is successful, is

$$p_k = P_r^{k-1}(1 - P_r) \tag{4}$$

where $1 \leqslant k < \infty$ and $\sum_{k=1}^{\infty} p_k = 1$. In order to sample from this distribution we note that the associated cumulative distribution, C_m, is given by

$$C_m = \sum_{k=1}^{m} P_r^{k-1}(1 - P_r) = 1 - P_r^m \tag{5}$$

Hence if we generate a random number, ξ, uniformly distributed in the range $0 < \xi < 1$, the number m given by

$$m = \mathrm{Int}\left[\frac{\ln(\xi)}{\ln(P_r)} + 1\right] \tag{6}$$

will have been drawn from the required distribution. Thus, if a state i includes some sites which will lead to rejected moves, a 'rejection-free' algorithm will still be maintained by

(i) randomly selecting the new site from the states which will lead to an accepted move
(ii) generating a number m as described above and adding this to N_i instead of incrementing N_i by 1.

The results of using this method are shown in table 3. The simulations consisted of $N_{\mathrm{equ}} = 10^5$ equilibration moves and $N_s = 2.4 \times 10^7$ data collection moves. It can be seen from this table that, once again, R_i data with essentially the same accuracy has been obtained for all but the lowest energies in the microcanical window, W. The error estimates from using block averaging are also seen to be reasonable when compared with the errors

Table 3. Typical results of rejection-free microcanonical Monte Carlo simulation of three dimensional $S = \frac{1}{2}$ Ising model on a simple cubic lattice with $N_s = 2.4 \times 10^7$. N_i^{corr} is the value of N_i corrected by use of equation 6. The quantities e_i, N_i, R_i, f_i^{est}, R_i^{exact}, χ, τ' and ϵ are defined in the text.

L^3	e_i	N_i	N_i^{corr}	R_i	$100 \times f_i^{\mathrm{est}}$	R_i^{exact}	χ	τ'	ϵ
	12	2.763×10^4	2.775×10^4	0.2939	1.2	0.2889	1.5	0.79	30
	13	8.107×10^3	8.154×10^3	40.35	1.1	40.98	1.5	0.87	28
	14	3.280×10^5	3.290×10^5	16.60	0.26	16.56	1.0	1.6	41
10^3	15	5.462×10^6	5.462×10^6	0.5666	0.18	0.5673	0.68	9.9	25
	16	3.095×10^6	3.095×10^6	25.86	0.20	25.87	0.23	4.8	26
	17	3.175×10^6	8.002×10^7	11.19	0.20	11.17	1.1	7.6	23
	18	5.824×10^6	8.957×10^8	0.9675	0.18			19	21
	19	6.081×10^6	8.666×10^8					16	
	12	6.485×10^3	6.489×10^3	0.03914	6.7	0.03603	1.3	1.4	260
	13	2.530×10^2	2.540×10^2	310.8	6.5	332.6	1.0	1.1	260
	14	7.891×10^4	7.894×10^4	131.9	0.55	133.1	1.6	2.4	290
20^3	15	1.041×10^7	1.041×10^7	0.07173	0.44	0.07129	1.4	4.0	250
	16	7.469×10^5	7.469×10^5	209.6	0.59	210.0	0.34	14	310
	17	7.807×10^5	1.565×10^8	88.35	0.41	88.71	1.0	13	370
	18	1.047×10^7	1.383×10^{10}	0.1242	0.43			36	250
	19	1.505×10^6	1.717×10^9					30	

with respect to the exact values (i.e. χ). The statistical efficiency, τ', in this set of data is calculated using the actual number of observations, N_i, at each level, rather than the the corrected values since this gives a more realistic measure of the true efficiency of the method.

The simplest measure of the effectiveness of the rejection-free technique is the relative efficiency, ϵ, defined by [12]

$$\epsilon = \frac{\sigma_{st}^2 t_{st}}{\sigma_{rf}^2 t_{rf}} \tag{7}$$

where σ_{st}^2 (σ_{rf}^2) is the variance and t_{st} (t_{rf}) is the computational time associated with the standard method (rejection-free method). The times were all measured on the same machine and no particular effort was made to optimise the code. The rejection-free code took approximately three times as long per attempted move than the standard code. However, the probability of a successful move in the highest energy level for the standard method is $\simeq 1/140$ ($\simeq 1/1150$) for the $L = 10$ ($L = 20$) lattices. These values combined with the time penalty for the rejection-free method are consistent with the effective time saving implied by the observed efficiencies. It was found that when $E_i = E_c$, where E_c is the expected value of the energy at the critical temperature, the efficiency, $\epsilon = 1.0$ (0.6) for $L = 10$ (20).

In conclusion, it has been shown that all the levels in the microcanonical window, except the lowest, may be used in the microcanical Monte Carlo method and further that a rejection-free technique is straightforward to devise for all the levels. In the Ising spin-$\frac{1}{2}$ models this yields a considerable improvement in efficiency over Lee's [1] algorithm. For other spin models, such as the Blume Capel [8] model, the gain is expected to be even more significant.

References

[1] Lee K-C 1995 *J. Phys. A: Math. Gen.* **28** 4835
[2] Lee K-C 1990 *J. Phys. A: Math. Gen.* **23** 2087
[3] Bortz A B, Kalos M H and Lebowitz J L 1975 *J. Comp. Phys.* **17** 10
[4] McDonald I R and Singer K 1967 *Discuss. Faraday Soc.* **43** 40
[5] Creutz M 1983 *Phys. Rev. Lett.* **50** 1411
[6] Bhanot G, Black S, Carter P and Salvador R 1987 *Phys. Lett.* **183B** 331
[7] Desai R C, Heerman D W and Binder K 1988 *J. Stat. Phys.* **53** 795
[8] Care C M 1993 *J. Phys. A: Math. Gen.* **26** 1481
[9] Promberger M and Huller A 1995 *Z. Phys.* B **97** 341
[10] Domb C and Green M S 1974 *Phase Transitions and Critical Phenomena* (New York: Academic) **3** 416
[11] Feller W 1957 *An Introduction to Probability Theory and its Applications* (2nd edn) (New York: Wiley)
[12] Hammersley J M and Handscomb D C 1964 *Monte Carlo methods* (Methuen: London)