PHYSICAL REVIEW

LETTERS

Optimized Monte Carlo Data Analysis

Alan M. Ferrenberg[a] and Robert H. Swendsen

Department of Physics, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213

We present a new method for optimizing the analysis of data from multiple Monte Carlo computer simulations over wide ranges of parameter values. Explicit error estimates allow objective planning of the lengths of runs and the parameter values to be simulated. The method is applicable to simulations in lattice gauge theories, chemistry, and biology, as well as statistical mechanics.

Recently, we showed that histograms can be used to greatly increase the amount of information obtained from a single computer simulation in the neighborhood of a critical point.[1] In particular, we demonstrated that the location and height of maxima and minima can be determined with higher accuracy and much less computer time than previously obtainable. We also noted that the region of validity of the single histogram method coincides with the finite-size scaling region, so that information about the critical region does not deteriorate with increasing system size.[2]

On the other hand, for more general problems it is often desired to investigate the behavior of the system over a wider range of parameter values. In this situation, it is necessary to perform simulations at more than one value of the parameters of interest.

This paper presents an optimized method for combining the data from an arbitrary number of simulations to obtain information over a wide range of parameter values in the form of continuous functions. The method goes beyond earlier methods[3] in that it provides an optimized combination of data from different sources, and can be applied to an arbitrary number of simulations. Errors can be calculated and provide a clear and simple guide to optimizing the length and location of additional simulations to provide maximum accuracy.

It is possible to extend the temperature range to generate the free energy and entropy from zero to infinite temperature.

Finally, our method can be used with any simulation method that provides data for a system in equilibrium,

and it requires a negligible amount of additional computer time for its implementation. The method is therefore applicable to simulations in lattice gauge theories,[4] chemistry, and biology,[5,6] as well as statistical mechanics.

We will first describe the method and then demonstrate its efficiency by calculating the properties of an $L = 16$, $d = 2$ Ising model for the full range of temperatures.

Consider the general Hamiltonian

$$H(\sigma) = H_0(\sigma) + KS(\sigma) ,$$

where $S(\sigma)$ is an operator (energy, magnetization, etc.) defined on the spins $\{\sigma_i\}$, and factors of $-1/k_B T$ have been absorbed. We are interested in the behavior of the system as a function of K. For simplicity, we will consider only one parameter K, but the generalization to an arbitrary number of parameters is straightforward. The partition function is given by

$$Z(K) = \sum_\sigma \exp[H(\sigma)] = \sum_S W(S)[KS] ,$$

where $W(S)$ is the density of states.

Consider R Monte Carlo simulations. We perform the nth simulation at K_n and store the data as histograms, $\{N_n(S)\}$, with total numbers of values $\{n_n\}$. Errors are given by

$$\delta^2 N_n(S) = g_n \overline{N_n(S)} ,$$

where we have used a bar over an expression to indicate the expectation value with respect to all Monte Carlo

(MC) simulations of length n_n. If successive MC configurations are independent, then $g_n = 1$, otherwise

$$g_n = 1 + 2\tau_n \, ,$$

where τ_n is the correlation time.

We can approximate the behavior of the partition function by

$$z_n(K) = \sum_S N_n(S) \exp[(K - K_n)S]$$

which is related to the true partition function by

$$\overline{z_n(K)} = n_n Z(K)/Z(K_n) \, .$$

The free energy is given by

$$F(K) - F(K_n) = \ln \overline{z_n(K)} - \ln n_n \, .$$

The density of states is related to the histogram by

$$W(S) = \overline{N_n(S)} n_n^{-1} \exp[f_n - K_n S] \, ,$$

where $f_n = F(K_n)$ is a parameter equal to the free energy at K_n, and will be evaluated self-consistently. If we perform simulations on a set of values $\{K_n \,|\, n = 1, R\}$, we can combine them to form a general expression which leads to an improved estimate for $W(S)$. This gives us

$$W(S) = \sum_{n=1}^{R} p_n(S) N_n(S) n_n^{-1} \exp[f_n - K_n S] \qquad (1)$$

with

$$\sum_{n=1}^{R} p_n(S) = 1 \, .$$

If we insert the actual histograms in (1) and minimize the error in the resultant estimate for $W(S)$,[7] we find

$$p_n(S) = \frac{n_n g_n^{-1} \exp[K_n S - f_n]}{\sum_{m=1}^{R} n_m g_m^{-1} \exp[K_m S - f_m]} \, . \qquad (2)$$

If we then define

$$P(S,K) = W(S) \exp[KS] \, ,$$

we obtain the essential multiple-histogram equations as

$$P(S,K) = \frac{\sum_{n=1}^{R} g_n^{-1} N_n(S) \exp[KS]}{\sum_{m=1}^{R} n_m g_m^{-1} \exp[K_m S - f_m]} \, , \qquad (3)$$

where

$$\exp[f_n] = \sum_S P(S, K_n) \, . \qquad (4)$$

The average value of any operator on S can then be evaluated as a function of K using

$$\langle A(S) \rangle(K) = \sum_S A(S) P(S,K)/z(K) \, ,$$

where

$$z(K) = \sum_S P(S,K) \, .$$

The values of f_n are found self-consistently by iterating (3) and (4). Efficient convergence is obtained by using the derivatives of the new values of f_n as functions of the old values in the iteration process. Note that from the form of Eqs. (3) and (4), an arbitrary constant can be added to each f_n without affecting the solution. This constant can be determined by evaluating the free energy at zero or infinite temperature, where it is known. Alternatively, it can be set to zero at some reference point, which is useful when applying the method to the calculation of interfacial free energies.

The statistical error in $P(S,K)$ is given by

$$\delta P(S,K) = \left[\sum_n g_n^{-1} N_n(S)\right]^{-1/2} P(S,K) \qquad (5)$$

from which it is clear that this method always reduces the statistical errors when additional MC simulations are added to the analysis. This expression also provides a clear guide for planning a series of simulations. The locations and heights of peaks in the relative error, plotted as a function of S, give direct quantitative indications of the optimum locations and lengths of additional MC simulations.

As a practical matter, it is useful to handle most of the calculations in terms of the logarithms of the various quantities in these equations. Also, if there is insufficient computer memory to save an entire two-dimensional histogram, the average of the magnetization and its square can be stored as functions of S to determine the magnetic susceptibility. It is sometimes preferable to store MC data in terms of lists rather than histograms, with obvious modifications of the equations. This will be preferable, for example, if $qn_n < r^q$, where q is the number of operators and r is the number of values each operator can assume.

If the method is restricted to two MC simulations, the calculated difference in the free energies between the simulated points

$$F(K_1) - F(K_2) = f_1 - f_2$$

is identical to that obtained by Bennett's method.[7,8]

We have tested this multiple-histogram method against the exact solution of the $d = 2$ Ising model with $L = 16$. The Hamiltonian is given by

$$H = K \sum_{\langle i,j \rangle} \sigma_i \sigma_j \, ,$$

where the spins take on the values $+1$ and -1, and the sum is over all nearest-neighbors pairs.

For temperatures close to the critical temperature, a single simulation is sufficient for high accuracy. However, for temperatures more than about 20% away from the critical temperature, additional simulations are necessary. To improve the accuracy away from K_c, we added simulations at $K = 0.3$ and 0.64 to the first simulation at $K_c = 0.4406868$. Then, using plots of the relative error from (5) as a guide, we studied the results of adding

simulations at $K = 0.0$, 0.1, 0.2, 0.375, and 0.525 for a total of eight. We simulated the model using the Swendsen-Wang algorithm,[9] taking 2×10^6 MC sweeps at each temperature, except for T_c, where we used 9.5 $\times 10^6$ sweeps. The value of g_n at T_c was estimated to be 6, with smaller values at other temperatures.

The difference between the calculated and exact values of the specific heat over the full temperature range is less than 0.3%, and would differ from the exact values[10] by less than the width of the line if we were to include a plot with this paper.

The results for the entropy show dramatic improvement as data from additional MC simulations are included in the calculation. The entropy is given by the expression

$$\mathcal{S} = \ln z(K) - K_z^{-1}(K) \sum_S SP(S,K)$$

so that entropy differences can be calculated directly. As an example, consider the $d = 2$ Ising model on an L by L lattice, for which the total differences in the entropy per site between zero and infinite temperature is $(1 - L^{-1})\ln 2$. For $L = 16$, this has the value 0.690 44. When the data are restricted to a simulation at the critical temperature, the calculated entropy difference is 0.591, which is off by 15%. Adding data from the simulations at $K = 0.3$ and 0.64 gave a value of 0.6746, which already reduced the error to 2.3%. Finally, using data from all eight simulations, we obtained 0.690 30 with a remaining error of only 0.02%. This is an order of magnitude better than would be necessary to see the L^{-2} term due to the twofold degeneracy at $T = 0$.

Another application of the method, for which we already have preliminary results, is the calculation of the free energy of a seam of bonds with coupling $K' = \alpha K$. The point $\alpha = 1$ corresponds to the usual periodic boundary conditions, $\alpha = 0$ to free boundaries, and $\alpha = -1$ to antiperiodic boundary conditions. By taking two-dimensional histograms for the seam energy and the total energy of the system, we can calculate the surface free energy and the interface free energy as functions of the temperature and α. Taking MC data for just the two points $\alpha = 1$ and -1, the results are quite good for the smallest lattices ($L = 4$ and 6) as expected from the successes of Bennett's method.[7] However, even here the weakest part of the calculation is for seam energies near zero, and the addition of data from a third simulation at $\alpha = 0$ provides improvement. For system sizes up to $L = 8$, four simulations are sufficient to reproduce the interface free energy with an error of 0.05% (and agreement with the exact result). The same data give the surface free energy with an even smaller statistical error. By extrapolating the interface free energy as a function of L we can estimate the interface free energy per unit length in the thermodynamic limit as a function of temperature. We find

that even restricting ourselves to data between $L = 4$, 6, and 8, the interfacial free energy vanishes at a temperature within 0.2% of the exact value. This extension of the calculation to temperatures other than those at which the simulations were performed does not require additional parameters beyond the set $\{f_n\}$, which were already determined.

Since this method is able to combine Monte Carlo simulation data from different sources to increase the total accuracy of the results, it could even be used to combine data from different groups working on large problems, such as those encountered in lattice gauge theories. Because there are no limitations on the method of simulation, we also expect this approach to be useful for simulations in chemistry and biology.

(a)Present address: Center for Simulational Physics, Department of Physics and Astronomy, University of Georgia, Athens, GA 30602.

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

[2]The maximum shift of the energy scales as the width of the energy distribution, which is related to the specific heat C. Since the specific heat also relates the change in energy to the corresponding change in temperature, the maximum temperature change can be easily seen to scale as $C^{-1/2}L^{-d/2}$, where L is the size, and d is the dimension of the system. At the critical point, C diverges as $L^{\alpha/\nu}$, so that the range of temperatures scales as $L^{-1/\nu}$, which coincides with the scaling region. Similar consideration of other variables shows that this scaling is a general result.

[3]Z. W. Salsburg, J. D. Jackson, W. Fickett, and W. W. Wood, J. Chem. Phys. **30**, 65 (1959); D. A. Chestnut and Z. W. Salsburg, J. Chem. Phys. **38**, 2861 (1963); I. R. McDonald and K. Singer, Discuss. Faraday Soc. **43**, 40 (1967); I. R. McDonald and K. Singer, J. Chem. Phys. **47**, 4766 (1967); J. P. Valleau and D. N. Card, J. Chem. Phys. **57**, 5457 (1972); G. Torrie and J. P. Valleau, Chem. Phys. Lett. **28**, 578 (1974); M. Falcioni, E. Marinari, M. L. Paciello, G. Parisi, and B. Taglienti, Phys. Lett. **108B**, 331 (1982); E. Marinari, Nucl. Phys. **B235** [FS11], 123 (1984); G. Bhanot, S. Black, P. Carter, and R. Salvador, Phys. Lett. B **183**, 331 (1987); G. Bhanot, K. M. Bitar, S. Black, P. Carter, and R. Salvador, Phys. Lett. B **187**, 381 (1987); G. Bhanot, K. M. Bitar, and R. Salvador, Phys. Lett. B **188**, 246 (1987); G. Bhanot, R. Salvador, S. Black, P. Carter, and R. Toral, Phys. Rev. Lett. **59**, 803 (1987).

[4]K. G. Wilson, Phys. Rev. D **10**, 2445 (1974); in *Recent Developments in Gauge Theories,* edited by G. 't Hooft et al. (Plenum, New York, 1980).

[5]J. A. McCammon and S. C. Harvey, *Dynamics of Proteins*

and Nucleic Acids (Cambridge Univ. Press, New York, 1987), and references therein.

[6]C. H. Brooks, III, M. Karplus, and B. M. Pettitt, in *Proteins: A Theoretical Perspective of Dynamics, Structure and Thermodynamics, Advances in Chemical Physics,* edited by I. Prigogine and S. Rice (Wiley, New York, 1988), Vol. 71.

[7]C. H. Bennett, J. Comput. Phys. **22**, 245 (1976).

[8]H. Park (private communication).

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

[10]A. E. Ferdinand and M. E. Fisher, Phys. Rev. **185**, 832 (1969).