

ORNL/TM-2005/1

MULTIPLE WALKERS IN THE WANG-LANDAU ALGORITHM

January 2005

Prepared by
G. Brown
T. C. Schulthess
W. R. Elwasif
D. E. Bernoldt

DOCUMENT AVAILABILITY

Reports produced after January 1, 1996, are generally available free via the U.S. Department of Energy (DOE) Information Bridge:

Web Site: <http://www.osti.gov/bridge>

Reports produced before January 1, 1996, may be purchased by members of the public from the following source:

National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
Telephone: 703-605-6000 (1-800-553-6847)
TDD: 703-487-4639
Fax: 703-605-6900
E-mail: info@ntis.fedworld.gov
Web site: <http://www.ntis.gov/support/ordernowabout.htm>

Reports are available to DOE employees, DOE contractors, Energy Technology Data Exchange (ETDE), and International Nuclear Information System (INIS) representatives from the following sources:

Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831
Telephone: 865-576-8401
Fax: 865-576-5728
E-mail: reports@adonis.osti.gov
Web site: <http://www.osti.gov/contact.html>

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

ORNL/TM-2005/1

**MULTIPLE WALKERS IN THE WANG-LANDAU
ALGORITHM**

G. Brown, T. C. Schulthess, W. R. Elwasif, and D. E. Bernoldt

Date Published: January 2005

Prepared by
OAK RIDGE NATIONAL LABORATORY
P. O. Box 2008
Oak Ridge, Tennessee 37831-6285
managed by
UT-Battelle, LLC
for the
U. S. DEPARTMENT OF ENERGY
under contract DE-AC05-00OR22725

Abstract

The mean cost for converging an estimated density of states using the Wang-Landau algorithm is measured for the Ising and Heisenberg models. The cost increases in a power-law fashion with the number of spins, with an exponent near 3 for one-dimensional models, and closer to 2.4 for two-dimensional models. The effect of multiple, simultaneous walkers on the cost is also measured. For the one-dimensional Ising model the cost can increase with the number of walkers for large systems. For both the Ising and Heisenberg models in two-dimensions, no adverse impact on the cost is observed. Thus multiple walkers is a strategy that should scale well in a parallel computing environment for many models of magnetic materials.

Understanding the statistical mechanics of coupled magnetic moments is essential to understanding the properties of magnetic materials. On a fundamental level this involves evaluating sums over all possible configurations, such as that involved in the partition function

$$Z = \sum_{\mathbf{S}_i} \exp(-\beta E[\mathbf{S}_i]), \quad (1)$$

where \mathbf{S}_i is one of the configurations of all the spins in the system, $\beta = 1/k_B T$ is the inverse temperature in energy units, and $E[\mathbf{S}_i]$ is the energy of a particular configuration of spins. Other quantities of interest, such as the internal energy

$$\langle E \rangle = (1/Z) \sum_{\mathbf{S}_i} E[\mathbf{S}_i] \exp(-\beta E[\mathbf{S}_i]) \quad (2)$$

are related to the partition function. Semiclassical models of magnetic materials, e.g. the Ising and Heisenberg models, have proven very useful in studying the statistical mechanics of magnetic materials, but direct evaluation of Eqs. (1) or (2) is rarely possible given the number of configurations.

Monte Carlo methods which randomly sample the configurations are a particularly useful way of estimating these sums. It is possible for each configuration to be drawn completely at random. However, for a particular value of the inverse temperature β , only a small subset of configurations make a significant contribution. Thus the estimate converges much more rapidly if the choices are made using an importance sampling approach, which involves a random walk through configuration space. The random walk consists of a chain of configurations, with each step of the random walk corresponding to a change in the previous configuration. Whether to accept the new configuration, or retain the previous configuration, depends on the configurations importance. One common is to use the Metropolis [1] formula $\min[1, \exp(-\beta \delta E)]$, to determine the probability of accepting the new configuration. Here $\delta E = E'_{i+1} - E_i$ is the energy difference between the proposed and original configurations, respectively, and the prime reinforces the provisional nature of the proposed configuration. Importance sampling of this type can be used to estimate the property of a system for temperatures not much different from β^{-1} .

Flat-histogram methods use the density of states, $g(E)$, for importance sampling. The analog of the Metropolis formula in this case is to accept the new configuration with probability $\min[1, g(E_i)/g(E'_{i+1})]$. The effect is to create an equal probability of visiting each energy level in the system. In

other words, a histogram of where the random walk is at the end of each Monte Carlo move would be essentially flat (save for statistical $1/\sqrt{N}$ noise). Even for situations with energy a continuous variable the histogram will be flat, and this observation is important for properly handling such systems.

The main obstacle to flat-histogram methods is that $g(E)$ is not known, otherwise one would directly evaluate formulas such as Eq. (1). Instead, an estimate of the density of states $\hat{g}(E)$ must be constructed self-consistently as the Monte Carlo estimate is generated. Recently, Wang and Landau [4] developed an algorithm for finding $\hat{g}(E)$ directly. With this estimate the partition function can then be calculated as

$$Z = \sum_E g(E) \exp(-\beta E) \quad (3)$$

by noting that the density of states is defined as

$$g(E) = \sum_{\mathbf{S}_i} \delta(E - E[\mathbf{S}_i]) , \quad (4)$$

and $\delta(X)$ is the Dirac δ -function. Quantities such as the internal energy are calculated in an analogous fashion. In the Wang-Landau approach, the majority of the computational effort comes in evaluating $\hat{g}(E)$, while Eq. (3) can be rapidly evaluated for many temperatures once $\hat{g}(E)$ is known.

An important quantity in the Wang-Landau algorithm is visit histogram, $H(E)$, which records where the random walk has been. At the end of each step in importance-sampling chain, the $H(E)$ is updated based on the accepted configuration and the estimated density of states is changed by $\ln[\hat{g}(E_i)] \leftarrow \ln[\hat{g}(E_i)] + \ln f$, where f is the modification factor which is initially $\ln f=1$. (Note that we refer to the \ln of many fundamental quantities since that is how they are encountered in actual implementations of the Wang-Landau algorithm.) The estimate $\hat{g}(E)$ is considered to be converged when

$$\min [H(E)] \leq A \text{ mean}[H(E)] , \quad (5)$$

where both the mean and minimum value are reëvaluated after each Monte Carlo step. The the flatness parameter $0 < A < 1$ controls the accuracy of the estimated $\hat{g}(E)$, with increasing accuracy as A approaches unity. Once the estimate has been converged for a particular value of $\ln f$, the value of f is reduced such that $\ln f' = \ln(f)/2$, while $\hat{g}(E)$ is retained unchanged. For each iteration in f , a new visit histogram $H(E)$ is calculated, and the process is repeated until $\ln f \leq 1 \times 10^{-6}$, and the estimated $\hat{g}(E)$ is considered final.

It is possible for multiple, simultaneous walkers to be employed. The Wang-Landau algorithm discussed above is unchanged, except that both

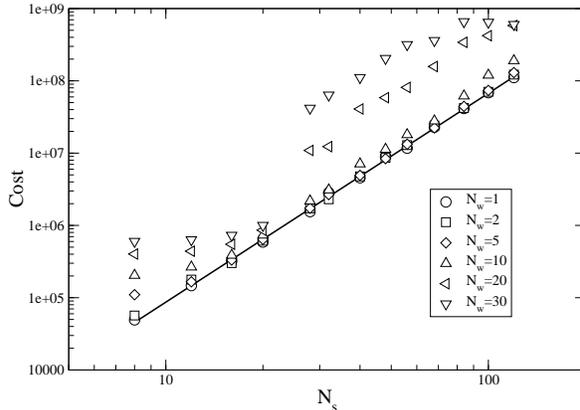


Figure 1: Mean cost for final estimate of density of states for the one-dimensional Ising Model. The solid line is a least-squares fit to the data for one walker ($N_w=1$), and it corresponds to a cost $S \sim N_s^{2.89}$. For this system, more than 10 walkers usually increases the cost compared to one walker.

$H(E)$ and $\hat{g}(E)$ are updated for each of the N_w walkers after each Monte Carlo step. All walkers share the same $H(E)$ and $\hat{g}(E)$. Traditionally, the random walks for the Wang-Landau algorithm begin from a ground state. However, when $N_w > 1$ it is important that all walkers start from different configurations. To randomize these starting configurations, all but one of the walks underwent five configuration updates that were automatically accepted.

The scaling of computational cost of the Wang-Landau algorithm is reported as the number of classical spins, N_s , and walkers, N_w , are changed. The cost is taken to be the number of energy evaluations S , required to converge to a final $\hat{g}(E)$. One- and two-dimensional arrangements of Ising and Heisenberg spins are considered. For all models, the only contribution to the Hamiltonian is from nearest-neighbor exchange

$$E = -(1/2) \sum_i \sum_{j \in \text{NN}_d(i)} J \hat{\mathbf{s}}_i \cdot \hat{\mathbf{s}}_j, \quad (6)$$

where $\text{NN}_d(i)$ is the set of nearest neighbors to point i on a square lattice of d dimensions. For the Ising model $\hat{\mathbf{s}}_i \in \pm 1$, while for the Heisenberg model $\hat{\mathbf{s}}_i$

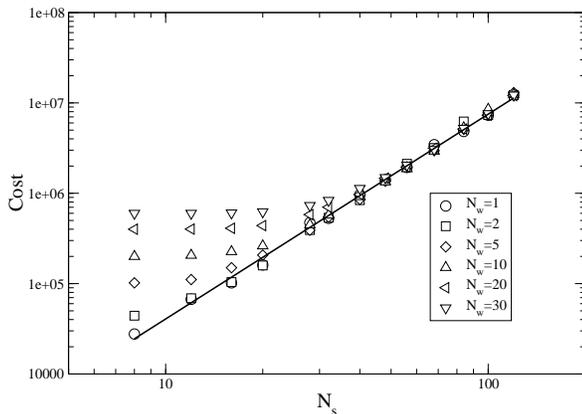


Figure 2: Mean cost for final estimate of the density of states for the two-dimensional Ising Model. The solid line is a least-squares fit to the data for one walker ($N_w=1$), and it corresponds to a cost $S \sim N_s^{2.27}$. For this system, multiple walkers only increase the cost for small systems.

are three-dimensional unit vectors. For convenience, \mathbf{J} is taken to be unity, and the possible energies allowed for a system lie between $-N_s$ and $+N_s$.

The mean cost for the Ising model on a one-dimensional lattice with N_s spins is shown in Fig. 1. Unless stated otherwise, all results in this report are for an accuracy $A=0.75$ and with the mean cost estimated from 100 trials. For $N_s < 20$, the increasing cost with N_w is largely an artifact from the minimum number of steps required before convergence. For larger systems, the cost is only significantly higher for $N_w > 10$. The solid line in Fig. 1 is a least-squares fit to the $N_w=1$ data, $S=112N_s^{2.89}$. The quality of the fit suggests that a power-law dependence of the cost on system size is a good assumption.

The mean cost for the Ising model on a two-dimensional lattice is shown in Fig. 2. Here a $L \times 4$ lattice with $N_s=4L$ spins was employed to allow a large number of different N_s while keeping N_s small. The finite size effects for quantities calculated from these lattices would be large, but here only the cost of converging is studied. The increasing cost with N_w for small systems is an artifact due to the minimum number of energy evaluations required before an estimate $\hat{g}(E)$ could be considered converged at each iteration of

In f . Otherwise, there cost is independent of N_w . The dependence of the cost on N_s is estimated from a least-squares fit to the results for $N_w=1$, the solid line in Fig. 2, or $S=220N_s^{2.27}$. From the figure, it is also apparent that this estimate of the cost works quite well for all N_w for lattices large enough.

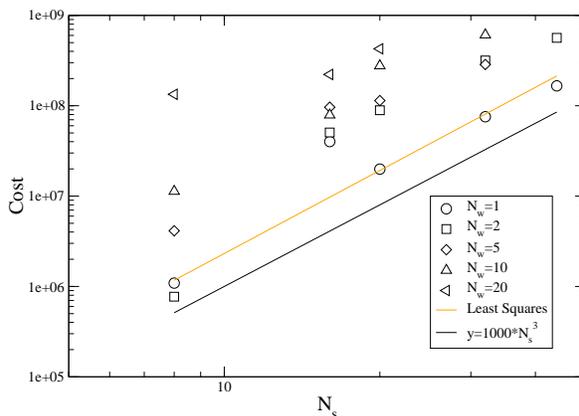


Figure 3: Mean cost for final estimate of the density of states for a one-dimensional arrangement of Heisenberg spins constrained to the unit sphere. The lines indicate that the approximate best estimate for the cost is $S \sim N_s^3$.

In addition to the Ising model, the Heisenberg model with three-vector spins constrained to lie on the unit sphere have also been considered. One issue that must be considered that there are not discrete energy levels for the Heisenberg model as there are in the Ising model. One solution to this problem is to divide the continuous energy values into a finite set of energy bins. We have taken bins with of width $\Delta E=1$ [5, 6]. In addition, it is common to increase the convergence by considering only a subset of energies [4], and here we have chosen to restrict the random walkers to the energy window $-7N_s/4 \leq E \leq +N_s/5$. The omission of the energies $-2N_s \leq E < -7N_s/4$ from $\hat{g}(E)$ prevents accurate estimates of the sums for extremely low temperatures, but still allows accurate estimates for temperatures well below the Curie temperature [7].

The mean cost for a one-dimensional lattice of these Heisenberg spins is shown in Fig. 3. The results clearly are not as well behaved as those for the one- or two-dimensional Ising model. There is a general trend for increasing

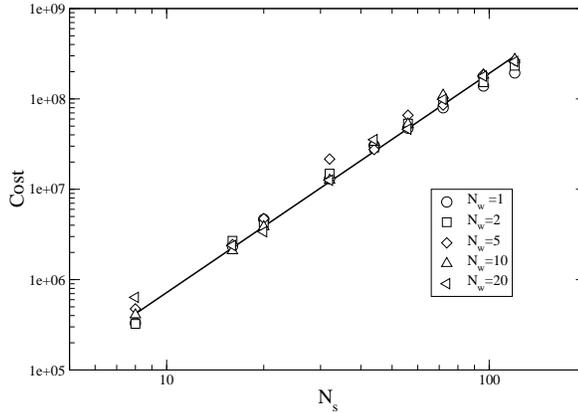


Figure 4: Mean cost for final estimate of the density of states for a two-dimensional arrangement of Heisenberg spins constrained to the unit sphere. The lines is a best fit to all data, and corresponds to a cost $S \sim N_s^{2.43}$. The number of walkers does not affect the cost of estimating $g(E)$ for the two-dimensional Heisenberg model.

cost with increasing N_w . The cost increases as $S \sim N_s^3$ as indicated by the least-squares fit to the $N_w=1$ data (orange line), and the line corresponding to $S=1000N_s^3$ provided as a guide to the eye.

Much better results are obtained for a two-dimensional arrangement of Heisenberg spins, again on a $4 \times L$ lattice. The results of the mean cost for this system are shown in Fig. 4. For this data, the artifact seen for small N_s in the two-dimensional Ising model does not occur since a large number of evaluations are required for convergence. The data are quite closely coupled for all N_s , indicating that multiple walkers do not interfere with convergence. The solid line is a least-squares fit to all data, and it corresponds to $S=2700N_s^{2.43}$. For this model, there seems to be a systematic trend for the cost to be less than the power-law estimate for large N_s . Further study would be required to distinguish whether this is an oscillation around the fit power-law form, or perhaps a power-law with smaller exponent is valid for larger systems.

In conclusion, the computational cost of obtaining an estimated density of states $\hat{g}(E)$ with the Wang-Landau algorithm as been measured for

Ising and Heisenberg models in one and two dimensions. A power-law dependence with respect to the numbers of spins, with exponents of about 3 for the one-dimensional systems and 2.4 for the two-dimensional systems. These results are consistent with the previous studies of the “escape-time” for a flat-histogram random walker to cross the entire range of energies [8, 9]. This power-law dependence of the cost suggests that the Wang-Landau algorithm, and other flat-histogram methods, are only competitive for small- and medium-sized systems. If the effects of critical slowing down near the Curie temperature can be ignored, then the cost of traditional importance-sampling Monte Carlo increases linearly with N_s . In such a case, Wang-Landau will be more efficient for only for small systems. The cross-over value of N_s will depend largely on the desired accuracy and the number of temperatures statistical mechanics properties will be evaluated at. This report also demonstrates that multiple, simultaneous walkers do not adversely affect the cost to converge estimates for the two-dimensional Ising and Heisenberg models. Thus implementing multiple Wang-Landau walkers should scale well in a parallel computing environment.

References

- [1] N.K Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).
- [2] D. P. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics* (Cambridge University Press, Cambridge, 2000).
- [3] M. H. Kalos and P. A. Whitlock: *Monte Carlo Methods* (Wiley, New York, 1986).
- [4] F. Wang and D. P. Landau, *Phys. Rev. E* **64**, 056101 (2001).
- [5] M. S. Shell, P. G. Debenedetti, and A. Z. Panagiotopoulos, *Phys. Rev. E* **66**, 056703 (2002).
- [6] S. J. Mitchell, M. A. Novotny, and J. D. Muñoz, *Int. J. Mod. Phys. C*, **10**, 1503-1512 (2000).
- [7] G. Brown and T. C. Schulthess, *J. Appl. Phys.* (in press).
- [8] B. J. Schulz, K. Binder, and M. Mueller, *Int. J. Mod. Phys. C* **13**, 477 (2002).
- [9] P. Dayal, S. Trebst, S. Wessel, D. Würtz, M. Troyer, S. Sabhapandit, and S. N. Coppersmith, *Phys. Rev. Lett.* **92**, 097201 (2004).

INTERNAL DISTRIBUTION

1. D. E. Bernoldt
2. G. Brown
3. W. R. Elwasif
4. J. A. Nichols
5. T. C. Schulthess
6. T. Zacharia
7. Central Research Library
- 8 - 9. ORNL Laboratory Records - OSTI
10. ORNL Laboratory Records - RC