Flat-Histogram Dynamics and Optimization in Density of States Simulations of Fluids†

M. Scott Shell,* Pablo G. Debenedetti,‡ and Athanassios Z. Panagiotopoulos§

Department of Chemical Engineering, Princeton University, Princeton, New Jersey 08544

Received: May 28, 2004; In Final Form: July 25, 2004

We analyze the dynamics of flat-histogram Monte Carlo algorithms that determine the density of states as a function of potential energy. In contrast to conventional Boltzmann sampling, these methods aid equilibration by enforcing a flat energy distribution, which enables a system to escape local energy minima. For two model fluids, we perform an extensive characterization of flat-histogram dynamics, that is, the dynamical behavior when the sampling is such that the system makes a random walk in energy. We show that the tunneling time, defined as the average number of steps required for the system to move between its low- and high-energy bounds, correlates uniquely with the entropy range of the energy window sampled, independent of system size and of the particular system investigated. We also demonstrate that the rate of statistical improvement of flat-histogram calculations scales with the tunneling time, and hence we propose that an optimal workload distribution in parallel implementations of these methods should divide the total entropy range equally among each processor.

I. Introduction

Flat-histogram Monte Carlo (MC) methods are a growing class of techniques designed to enhance configurational sampling and directly obtain free energies of complex model systems. Such methods include the multicanonical,[1,2] transition-matrix,[3−5] and Wang−Landau algorithms.[6,7] In off-lattice systems, these approaches have been used extensively to study protein folding,[8−11] vapor−liquid equilibria,[12−14] polymer phase behavior,[15−17] simple glasses,[18,19] and chemical reactions,[20] to name just a few examples. Generally speaking, flat-histogram methods aim to produce a nearly uniform distribution in one or more macroscopic observables, such as potential energy or number of particles, within a predetermined range. To achieve a flat distribution in an MC simulation, one must sample microstates according to an inverse partition function, since this will cancel up to a constant when integrated to determine the corresponding macroscopic probabilities. The main objective of these algorithms, therefore, is to iteratively determine the weights, free energies, or entropies (corresponding to the partition function) which will ensure an equiprobable distribution.

Much effort has focused on algorithms that generate a flat distribution in potential energies,[3,6,7,18,21−23] which often fall in the class of so-called "density of states" (DOS) flat-histogram methods. These techniques facilitate equilibration of complex systems by permitting enhanced flexibility in sampling energy space in that the system is given a greater probability of escaping low-lying energy minima as compared to traditional Boltzmann sampling at low temperatures. Parallel tempering methods also aim to enhance low-temperature equilibration by exchanging configurations between two or more realizations of the system spanning range of temperatures.[24,25] In this latter case, the multiple canonical MC simulations are augmented with the addition of periodic configuration swaps between adjacent temperatures. When adjacent temperatures are not close enough to produce overlap in the sampled energies, the majority of configuration swaps are rejected; this becomes especially problematic as the system size increases since the fluctuations in energy in each canonical instance become increasingly narrow. DOS methods can be similarly implemented in parallel fashion (with or without configuration swaps)[5−7,26] but have the additional advantage that the amount of energy space overlap between adjacent realizations is exactly specified: the energy range of interest is divided into a series of overlapping windows, each assigned to a unique processor which samples its subrange uniformly. After each processor finishes, data from each in the form of calculated weights or free energies are then patched together.

A key question in parallelized DOS algorithms is how to construct the energy windows for optimal performance. The optimal strategy should distribute the workload evenly among the available processors to minimize the simulation time necessary to achieve a given statistical accuracy in the overall calculation. To our knowledge, this issue has not yet been systematically addressed in off-lattice systems. Furthermore, while several studies have examined flat-histogram dynamics in lattice models,[5,27−30] they have mainly addressed the effects of system size and frustration on the dynamics and not the optimization of a parallel processing scheme. In the present work, we propose an empirically based procedure for the design of parallel DOS implementations. Our approach is based on a detailed characterization of the random-walk dynamics in two model systems, the Lennard-Jones and Dzugutov[31] potentials. We measure the "tunneling time" in DOS simulations—the number of MC steps τ required to sample between the minimum and maximum energy bounds—and demonstrate that simulation demand correlates with tunneling time, which in turn strongly depends on the entropy range of the energy window of interest.

† Part of the special issue "Frank H. Stillinger Festschrift".

* Address correspondence to this author. E-mail: Shell@princeton.edu.
‡ E-mail: pdebene@princeton.edu.
§ E-mail: azp@princeton.edu.

10.1021/jp047677j CCC: $27.50
Published on Web 09/10/2004

For a single DOS run at constant N and V, the probability of observing a specific configuration is specified as

$$\pi(\mathbf{q}^N) \propto \frac{1}{\Omega[U(\mathbf{q}^N)]} = \exp[-S(U(\mathbf{q}^N))] \quad \text{for } U_{\min} < U(\mathbf{q}^N) < U_{\max}$$

$$\pi(\mathbf{q}^N) = 0 \quad \text{otherwise} \qquad (1)$$

where $\mathbf{q}^N$ is the $3N$ coordinates of the particles, U is the potential energy, Ω is the density of states, and $S = \ln \Omega$ is the corresponding dimensionless entropy. The simulation parameters $U_{\min}$ and $U_{\max}$ indicate the bounds on the energy range sampled, and the constant of proportionality is given by the probability normalization condition. It is straightforward to see that when eq 1 is summed over all configurations of given energy U, the macroscopic distribution becomes constant: $\rho(U) \propto \Omega(U)/\Omega(U) = \text{const}$. The task of the simulation is to determine the density of states directly (hence the naming of these methods) and to use Ω post-simulation to determine conventional thermodynamic properties via standard reweighting techniques. Since Ω can span numerous orders of magnitude which exceed standard double precision calculations, it is customary to determine the entropy S (i.e., $\ln \Omega$) instead. For off-lattice systems, a small energy bin width must be chosen and the discretized function $S(U)$ is stored as an array in the computer. During the simulation, the system can be propagated by the usual single-particle displacement moves, whose acceptance criterion is

$$\mathrm{P}_{\mathrm{acc}}(\mathrm{o} \rightarrow \mathrm{n}) = \min\{1, \exp[S(U_{\mathrm{o}}) - S(U_{\mathrm{n}})]\} \qquad (2)$$

where the abbreviations "o" and "n" indicate the original and new configurations. Proposed moves which would take the system outside of its energy bounds are always rejected. It is important to recognize that additive shifts in the entropy have no effect on the simulation, since these are always removed from the normalization of eq 1. Therefore, the simulation can only determine S up to an additive constant. Conventionally, this constant is simply adjusted such that the minimum value of the entropy is zero. Regardless, classical thermodynamic properties derived from S remain unaffected by the additive uncertainty.

As a DOS simulation progresses, it has an important feedback mechanism for the calculation of S. When energies are sampled uniformly, that is, when a measured histogram is statistically flat, S has converged upon the true entropy. The popular Wang−Landau algorithm makes direct use of this feedback mechanism.[6,7] After each MC step, the entropy at the ending energy is increased by an amount g, called the modification factor, which is greater than zero.[32] This update occurs regardless of whether the ending state results from an acceptance or rejection. In this way, when energies are sampled with equal probability, on average the entropy function is not modified, since the updates simply serve to change the arbitrary additive constant up to which it is known. The modification factor has a key role in this process: it starts off around one, which allows rapid generation of an initial estimate for S, and it is decreased in stages until it is nearly zero.[6,7] The intermediate values of g resolve the entropy with increasing precision, and microscopic detailed balance is achieved asymptotically as $g \rightarrow 0$. During the progression of stages, the value of the modification factor is halved ($g_{\mathrm{new}} = (1/2)g_{\mathrm{old}}$) whenever the energy histogram for the current g is determined to be sufficiently flat.

The amount of MC steps required to converge a DOS simulation can vary significantly with the model system and the specific energy range studied, and this is particularly true of the Wang−Landau method. A Wang−Landau simulation at high energies can require orders of magnitude less time than one at very low energies, where very few states are available and where the dynamics of the system in energy space are much slower. Consider a simulation which uses a converged S and is propagated according to eq 2, which we term true flat-histogram sampling. In this case, the system will perform a random walk in energy which eventually produces a flat histogram. This is a pseudo rather than a pure random walk since the structure and connectivity of configuration space pose constraints on the paths the system can take. There is a time scale associated with this random walk, measured in number of MC steps, which is characteristic of true flat-histogram sampling. It is reasonable to suggest that this time scale is also relevant to the "speed" of a Wang−Landau calculation, even though the entropy does not converge until its very last stages of the modification factor. In other words, the limiting speed at which a system can perform a random walk in energy space is also the limiting rate at which statistics can be gathered for computing the entropy.

The relevant time scale in true flat-histogram sampling of a particular energy range is the tunneling time, τ, rigorously defined as the average number of steps needed to move from the low- to the high-energy bound and vice versa. The tunneling time has been studied in detail for lattice systems,[5,27−30] where τ has been identified with the time necessary to move between the ground and anti-ground-state configurations. For moves involving single-spin flips, ref 29 has found the dependence on system size in 2D models to be $\tau \sim N^{2.4}$ for the Ising ferromagnet, $\tau \sim N^{2.9}$ for the fully frustrated Ising model, and $\tau \sim \exp(\sqrt{N}/4.21)$ for the Ising spin glass. Other authors have proposed $\tau \sim N^{2.2}$ and $\tau \sim N^{2.6}$ for the 2D[28] and 3D[5] spin glass models, respectively, although ref 29 has questioned these results, citing the need for special treatment of the long-tailed tunneling time distributions which occur in spin glass systems.

In the current work, we study the flat-histogram tunneling time in two model continuum systems. In contrast to lattice models, one cannot explore the complete range of energy in these systems, which is of course unbounded to the right. Therefore, when using continuum models, it becomes necessary to prespecify the energy range of interest and to measure a tunneling time specific to that range. One might anticipate some degree of similarity in our results with those of lattice systems, given the short-ranged interactions and locality of the MC moves; however, the increased dimension and complexity of configurational space in the continuum models should have an effect on at least the magnitude of the tunneling times. We examine the familiar Lennard-Jones system in our study and choose the Dzugutov potential as an instructive but computationally inexpensive contrast. Similar to ref 29, we measure the distribution of tunneling times rather than the averages alone, but unlike previous investigations, we make our measurements for a collection of representative energy ranges at each system size. In this way, we not only determine the dependence of τ on system size, but also on the properties of the specific energy window as well. Our analysis of tunneling times allows us to propose an optimal distribution of workload in parallel DOS implementations.

The remainder of this paper is organized as follows. Section II reviews the computational details of our study and Section

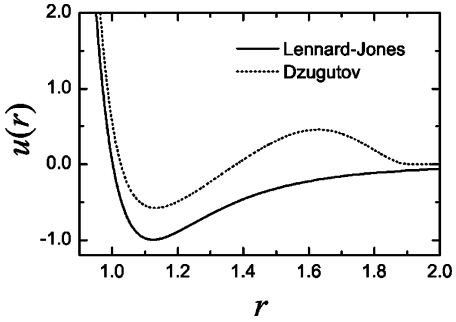

Figure 1. Potential energy functions for the Lennard-Jones (solid) and Dzugutov (dotted) systems.

III presents the tunneling time results. In Section IV, we discuss the optimization scheme and comment on the relationship between our findings and previous work on lattice systems.

II. Methods

We study the Lennard-Jones (LJ) and Dzugutov (DZ) models in our investigation of tunneling times. Each system is cast in a cubic simulation cell to which periodic boundary conditions are applied. The pairwise LJ potential has the following form when expressed in dimensionless units:

$$u(r) = 4(r^{-12} - r^{-6}) \tag{3}$$

where u has a minimum at $u(r = 2^{1/6}) = -1$. We examine this system at reduced density $\rho = 0.88$, which is well within the liquid phase and corresponds with a previous Wang−Landau study by us.[23] The potential is truncated at $r_c = 2.5$ and the long range correction is applied.[33] We examine three system sizes for the LJ potential, $N = 110$, 250, and 500. In comparison, the form of the pairwise Dzugutov potential is given by

$$u(r) = A(r^{-m} - B)\exp\left(\frac{c}{r-a}\right)\theta(a-r) + B\exp\left(\frac{d}{r-b}\right)\theta(b-r) \tag{4}$$

where $m = 16$, $A = 5.82$, $B = 1.28$, $a = 1.87$, $b = 1.94$, $c = 1.1$, and $d = 0.27$, and the energy minimum occurs at $u(r \approx 1.13) \approx -0.581$. The switching function $\theta(x)$ equals 1 for $x \geq 0$ and 0 for all other x. At short distances, the DZ potential resembles the shifted LJ model, although at longer distances it possesses a repulsive "bump" and then tapers continuously to zero at $r = b$. Figure 1 compares the two potentials. We study systems of $N = 110$ and 250 DZ particles, for which no potential truncation scheme is necessary given its finite range. We have chosen the DZ model because it provides a useful single-component contrast to the LJ system but is still computationally inexpensive enough to permit the long runs necessary to extract adequate tunneling time statistics. The DZ potential was designed to favor icosahedral ordering in a simple monatomic liquid;[31] at low temperatures in the liquid state, it exhibits pronounced glassy kinetics[34] and can enter a dodecagonal quasicrystalline phase.[35] Compared to the LJ system, the DZ model has substantially different short-range ordering,[31] possesses a different ground-state crystalline configuration,[36] and has distinct melting behavior.[37] It has also been suggested that the DZ potential may not exhibit a liquid−vapor transition.[38]

To generate the function $S(U)$ required to perform the flat-histogram simulations in energy, we use a modified Wang−Landau procedure which has improved convergence and statistical accuracy in off-lattice systems.[23] In this version, the modification factor is decreased more rapidly ($g_{new} = 0.1g_{old}$) than in the original scheme, which quickly generates a rough estimate of the entropy. Once the condition $g < 10^{-5}$ is met, the move proposal statistics are recorded for the remainder of the simulation, that is, the frequency of proposed moves between each pair of energies is measured. The calculated move proposal probabilities are then used to extract an additional, more accurate estimate of the entropy, and this improved $S(U)$ replaces the Wang−Landau version at each change of the modification factor. Additional details are provided in ref 23. We perform these simulations for 2.5×10^8, 5×10^8, and 10×10^8 MC steps for the $N = 110$, 250, and 500 LJ systems, respectively. In the Dzugutov case, we perform the simulations for 5×10^7 and 10×10^7 steps for $N = 110$ and 250. In all cases, the discretization of S into energy bins is done by setting the bin width to $\Delta U = 1$ (dimensionless units), with the exception of the 500-particle LJ system, for which memory requirements necessitated $\Delta U = 2$. The calculated entropy curves for the 110-particle cases are shown in Figure 2.

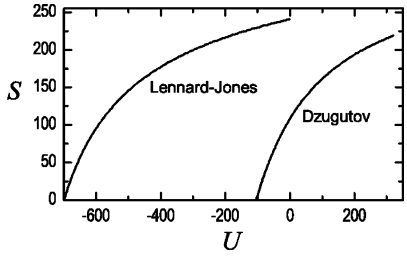

Figure 2. Entropy curves used for flat-histogram sampling at the smallest system size, $N = 110$. Both curves are discretized into bins of width 1 and adjusted to zero entropy at the minimum energy.

Using the calculations for $S(U)$, we study the flat-histogram dynamics of simulations over a variety of energy ranges. Table 1 lists the specifics of these ranges for each model and system size. For the smallest LJ system, for example, we examine 14 windows of various widths within the energy range −700 to 0, which roughly corresponds to $T = 0.7-10.0$. Similarly for the 110-particle DZ system, the windows are within the energy range −105 to 320, which corresponds to $T = 0.7-6.0$. At larger system sizes, the energy ranges are scaled accordingly. The tunneling time distribution is measured in the following manner, which is similar to that used by ref 5 for lattice systems. Starting at a random energy, we wait until the system reaches the lowest energy bin (of width ΔU), and we record the step number t_1. Then, we wait until the system has reached the highest energy bin and record t_2. Subsequently, for t_3 we search for the lowest bin again and the process continues, alternating high- and low-energy bins, until the dataset t_i contains several tens of thousands of entries. We then extract the distribution of tunneling times from the individual values given by $(t_{i+1} - t_i)$. Very long runs are required to resolve the tunneling time distributions to high accuracy; each energy window run is performed for 1×10^{10} MC steps for the 110-particle systems and 2.5×10^9 steps for the larger ones.

For comparison, we also examine the distribution of tunneling times in a pure random walk, often called the first passage time distribution. Let $P(x, t)$ denote the probability of observing the random walker at position x and time t. P is initially concentrated at the location $x = 0$ and we are interested in the probability flux of P at $x = L$ as the random walk progresses in time. This pure, one-dimensional random walk obeys the

TABLE 1: Summary of Energy Ranges Studied for the Lennard-Jones (LJ) and Dzugutov (DZ) Systems[a]

run ID	model	N	U_{min}	U_{max}	ΔU	ΔS	$\langle\tau\rangle$
1a	LJ	110	−700	−500	200	145.2	276 960
1b	LJ	110	−600	−400	200	81.3	54 625
1c	LJ	110	−500	−300	200	53.8	20 613
1d	LJ	110	−400	−200	200	39.1	10 371
1e	LJ	110	−300	−100	200	30.1	6201
1f	LJ	110	−200	0	200	24.1	4143
1g	LJ	110	−600	−300	300	103.7	105 480
1h	LJ	110	−300	0	300	40.9	13 019
1i	LJ	110	−600	−200	400	120.6	157 050
1j	LJ	110	−400	0	400	63.3	34 945
1k	LJ	110	−600	−100	500	134.0	212 320
1l	LJ	110	−500	−400	100	31.8	5245
1m	LJ	110	−300	−200	100	16.9	1676
1n	LJ	110	−100	0	100	10.7	898
2a	LJ	250	−1600	−1143	457	346.4	1 696 900
2b	LJ	250	−1371	−914	457	191.1	315 320
2c	LJ	250	−1143	−686	457	125.8	120 450
2d	LJ	250	−914	−457	457	90.7	57 977
2e	LJ	250	−686	−229	457	69.7	34 060
2f	LJ	250	−457	0	457	55.4	21 444
2g	LJ	250	−1371	−686	685	243.6	564 040
2h	LJ	250	−914	−229	685	121.7	121 700
2i	LJ	250	−1371	−457	914	282.8	823 380
2j	LJ	250	−914	0	914	146.6	203 620
2k	LJ	250	−1371	−229	1142	313.5	1 117 200
2l	LJ	250	−1143	−914	229	74.4	31 508
2m	LJ	250	−686	−457	229	39.1	8306
2n	LJ	250	−229	0	229	24.8	3913
3a	LJ	500	−3200	−2286	914	697.2	7 353 600
3b	LJ	500	−2742	−1828	914	384.6	1 512 100
3c	LJ	500	−2286	−1372	914	253.0	564 970
3d	LJ	500	−1828	−914	914	183.0	267 430
3e	LJ	500	−1372	−458	914	140.6	151 100
3f	LJ	500	−914	0	914	112.2	94 357
4a	DZ	110	−105	16	121	118.6	205 550
4b	DZ	110	−44	77	121	79.8	59 475
4c	DZ	110	16	138	122	56.7	23 313
4d	DZ	110	77	199	122	41.9	11 069
4e	DZ	110	138	259	121	32.1	6188
4f	DZ	110	199	320	121	25.3	3917
4g	DZ	110	−44	138	182	103.7	120 460
4h	DZ	110	138	320	182	43.3	13 374
4i	DZ	110	−44	199	243	121.7	180 980
4j	DZ	110	77	320	243	67.2	37 954
4k	DZ	110	−44	259	303	135.8	23 7490
4l	DZ	110	16	77	61	32.8	5420
4m	DZ	110	138	199	61	18.0	1606
4n	DZ	110	259	320	61	11.2	781
5a	DZ	250	−240	34	274	266.4	1 288 100
5b	DZ	250	−103	171	274	181.9	286 040
5c	DZ	250	34	309	275	129.3	114 220
5d	DZ	250	171	446	275	95.8	55 072
5e	DZ	250	309	583	274	73.3	30 997
5f	DZ	250	446	720	274	58.1	19 870
5g	DZ	250	−103	309	412	236.6	541 520
5h	DZ	250	171	583	412	127.9	66 262
5i	DZ	250	−103	446	549	277.7	732 080
5j	DZ	250	171	720	549	153.8	178 210
5k	DZ	250	−103	583	686	309.8	985 670
5l	DZ	250	34	171	137	74.6	30 874
5m	DZ	250	309	446	137	41.1	7889
5n	DZ	250	583	720	137	25.9	3321

[a] ΔS is the change in entropy over a given energy range, and $\langle\tau\rangle$ is the corresponding average tunneling time measured.

diffusion equation, which in dimensionless form is

$$\frac{\partial^2 P}{\partial \tilde{x}^2} = \frac{\partial P}{\partial \tilde{t}} \tag{5}$$

where $\tilde{x} = x/L$, $\tilde{t} = Dt/L^2$, and D is the diffusivity of the walker.

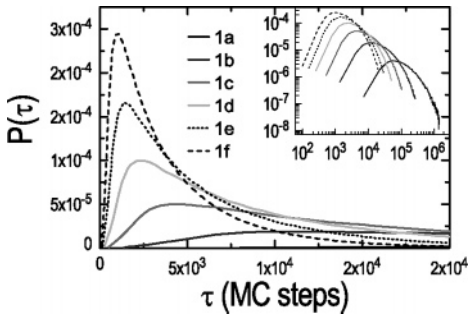

Figure 3. Tunneling time distributions for six runs in the $N = 110$ Lennard-Jones fluid. The inset shows the same runs in a log−log scale. The energy range in each case has a width of 200 energy units, and the minimum energies are −700, −600, −500, −400, −300, and −200 for runs 1a−1f, respectively. Details of the runs can be found in Table 1.

The boundary conditions for examining the first passage distribution are reflecting at $\tilde{x} = 0$ and absorbing at $\tilde{x} = 1$. Mathematically, these and the initial condition are expressed as

$$\frac{\partial P}{\partial \tilde{x}}(\tilde{x} = 0) = 0 \tag{6a}$$

$$P(\tilde{x} = 1) = 0 \tag{6b}$$

$$P(\tilde{t} = 0) = \delta(\tilde{x}) \tag{6c}$$

where δ is the Dirac delta function. This problem is readily solved by Fourier series. The final result is

$$P(\tilde{x}, \tilde{t}) = 2\sum_{n=0}^{\infty} \cos[(n + 1/2)\pi \tilde{x}]\exp[-(n + 1/2)^2\pi^2\tilde{t}] \tag{7}$$

The first passage distribution is given by $j(\tilde{t}) = -\partial P/\partial \tilde{x}$ $(\tilde{x} = 1, \tilde{t})$. Applying this expression to eq 7, one finds in the long-time limit that the probability decays exponentially as $j(\tilde{t} \to \infty) \approx \pi \exp(-\pi^2\tilde{t}/4)$. It proves instructive to look for similar exponential decay in the long-time behavior of the measured tunneling time distributions. The scaling of real time depends on the length squared, implying that the average tunneling time in a pure random walk in energy will exhibit a quadratic dependence on the energy range, scaling as $(U_{max} - U_{min})^2$.

III. Results

The average tunneling times measured in each simulation run are also reported in Table 1. Figure 3 shows a small sample of the measured tunneling time distributions; it contains $P(\tau)$ for the first six runs of the LJ system, each of which was assigned an energy range of width 200. The figure shows that each distribution peaks relatively rapidly before a subsequent slow decay and that the scale of the average tunneling time varies significantly between the very low- and high-energy windows. The inset displays the same $P(\tau)$ on a log−log scale. Remarkably, each distribution appears to have a similar underlying functional form, which is simply scaled in τ. Motivated by this observation, we plot in Figure 4 the same functions rescaled by their average tunneling time. That is, we plot $P \times \langle\tau\rangle$ as a function of $\tau/\langle\tau\rangle$. All of these rescaled results appear to fall onto the same master curve within the statistical uncertainty of the data. The same is true for the remainder of the energy windows studied, even for different system sizes and including the DZ potential. For comparison, we have plotted the analogous

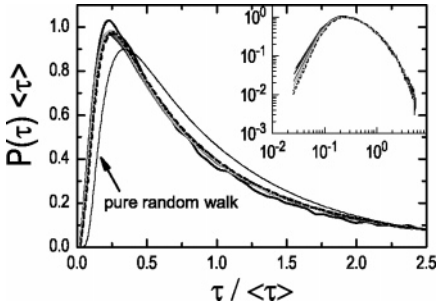

Figure 4. Scaled tunneling time distributions for the same runs as in Figure 3. In each case, the *x*-axis has been scaled by the average tunneling time, and the *y*-axis has been normalized accordingly. The thin solid line is the scaled distribution for a pure random walk in energy (not shown in the inset). The other line types correspond to those of Figure 3.

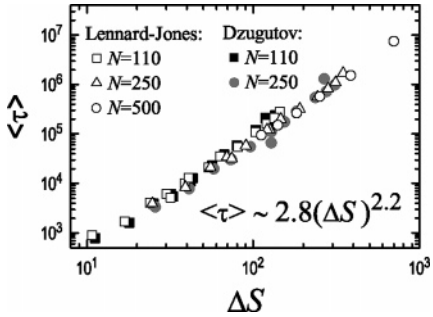

Figure 5. Correlation between the average tunneling times in each run $\langle\tau\rangle$ and the entropy range sampled ΔS. The equation is a least-squares fit to all the data displayed.

prediction for a pure random walk in energy. The pure random walk does not coincide with the master curve; its peak is shifted to later times and its long-time tail decays at a faster rate. It therefore appears that these systems possess a common feature borne out in their configuration space organization which gives rise to the master functional form observed but which biases the system from performing a pure random walk. This bias is further evidenced by the fact that the average upward and downward tunneling times are distinct, that is, we find that the time for passing from the minimum to maximum energy is consistently less than that of the opposite direction (results not shown). Consequently, the dynamics of the system are not simply captured by an effective diffusivity which depends on the energy range of interest. Presumably, there exists a unique dynamical equation which describes the measured random walk distributions, analogous to the diffusion equation discussed in Section. II. This is the subject of future research.

For all of the measured tunneling time distributions in both systems, the long-time tails decay exponentially. This contrasts with results for a lattice-spin glass model studied by Dayal et al.,[29] in which the long-time distribution possessed more slowly decaying power-law behavior. As these authors noted, the power-law trend can be problematic for the calculation of certain moments of the distribution, including the average, which become ill-defined if the exponent is too small in absolute value. To extract the tunneling time dependence on system size, Dayal et al. fitted their measured distributions to a "fat-tailed" model and examined the fit parameters rather than the moments.[29] We find no need to use such a procedure in the current work, as the long-time exponential decay ensures that the average tunneling time is always well-defined.

Our main interest is in the dependence of the average tunneling time on the properties of the system and energy range sampled. Remarkably, we find that all tunneling times correlate strongly with the range of *entropy* sampled. That is, $\langle\tau\rangle$ is a strong function of the quantity $\Delta S = S(U_{max}) - S(U_{min})$. To our knowledge, this observation is new for continuum models and the equivalent relationship has also not been found for lattice systems. Figure 5 shows the average tunneling times as a function of ΔS. All of the data appear to collapse onto a single master curve, which can be fit to the power-law relationship:

$$\langle\tau\rangle = 2.8(\Delta S)^{2.2} \tag{8}$$

The most interesting feature of this correlation is that it appears to hold for all system sizes and for both the LJ and DZ systems. We have also fit the results for each system size separately, which gives only slightly better fits in which the power-law exponents obtained are in the range 2.3−2.4. The existence of a strong master correlation therefore implies that the system size and model dependence of the tunneling time are largely contained within the entropy itself. Intuitively, the role of the entropy in determining flat-histogram dynamics is not completely unexpected. Reference 29 found that the tunneling times in spin glasses were strongly correlated with the ratio of the ground-state degeneracy to that of the first excited energy level. Similarly, the quantity ΔS is related to the relative configurational space volumes at the high- and low-energy bounds, since $\exp(\Delta S) = \ln[\Omega(U_{max})/\ \Omega(U_{min})]$. That is, ΔS stems from the increased number of configurations available at U_{max} as compared to U_{min}. Thus, one can think of the random walk in energy as the process of navigating this path of changing configuration space volumes.

Previous work on lattice models has focused on the correlation between tunneling time and system size, with the intensive energy bounds fixed. Our work also provides a system size prediction, captured by the *N*-dependence of ΔS. Assuming that the system sizes observed are large enough to exhibit extensive behavior, the relationship $\Delta S = N\Delta s$ should hold when the energy bounds are scaled by *N*, where *s* is the intensive entropy. Substituting this expression into eq 8, one obtains $\langle\tau\rangle \propto N^{2.2}$ for a fixed set of intensive energy bounds u_{min} and u_{max}. The scaling exponent in this prediction is almost identical to that which has been found for Ising models employing single-spin flips[5,28,29](see also Introduction). In our simulations, each step in the random walk typically makes a change in energy of order one, such that transitions only occur between neighboring energy levels. This is the result of tuning the maximum particle displacement for 50% acceptance, which produces "local" changes in the system energy, similar to the effect of single-spin flips in lattice models. Thus, we hypothesize that the exponents found in this and the previous works might be a common feature of simulations of short-ranged, unfrustrated systems involving local MC moves.

Finally, we have looked at the relationship between the average tunneling time and the simulation demands of flat-histogram algorithms. For this study, we started with $S = 0$ for all energies and then measured the time dependence of the statistical error in *S* as it was calculated by the Wang−Landau method. The error was measured as the standard deviation in *S* predictions among 10 independent runs for each window. More procedural details are given in ref 23. Results for the first six runs in the 110-particle LJ systems are displayed in the top panel of Figure 6. Despite a moderate amount of statistical noise, these results demonstrate that each range requires a distinct number of MC steps to reach the same statistical accuracy. In the bottom panel of this figure, we show the same curves rescaled such

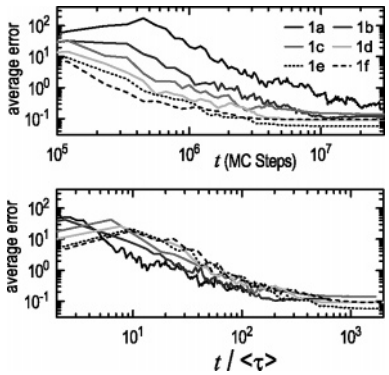

Figure 6. Progression of statistical error in Wang−Landau simulations for the energy ranges of 1a−1f in Table 1. The error is measured from the standard deviation in S among 10 independent runs for each window. The top panel shows the raw data in terms of absolute number of MC steps, and the bottom shows the same data with each curve time-rescaled by the appropriate tunneling time.

that the time axis is normalized by the average tunneling time for each window. The time-adjusted results nearly collapse onto each other within the noise of the data, although this rescaling does not completely remove all of their dependence on the energy range. For the most part, however, the tunneling time appears to be a dominant factor in the rate at which Wang−Landau simulations can resolve their calculations. At a basic level, this behavior can be rationalized by the fact that the system must perform at least one complete sweep of the energy range at each modification factor, which necessitates one tunneling time at a minimum. Large values of the modification factor enhance the ability of the system to tunnel initially,[29] but it is likely the later stages at small g, which dominate the simulation, are more strongly influenced by time scale of the true flat-histogram dynamics.

IV. Discussion and Conclusion

For the two model systems investigated, the computed dependence of the tunneling time on the entropy suggests an optimal strategy for parallel implementations of density of states methods. The most efficient distribution of energy windows occurs when the statistical error in the calculated results decreases at the same rate in each processor. In this manner, each processor reaches the same accuracy simultaneously and no idle time is spent waiting for a subset to catch-up to the desired precision. Since the rate of statistical improvement appears to depend largely on the tunneling time and hence the entropy, the optimal strategy should allocate the equivalent range of entropy to each energy window. That is, the complete energy range should be divided among the available processors such that each U_{min} and U_{max} gives the same ΔS. This is demonstrated graphically in Figure 7, in which the total span of entropy is divided equally among four processors. This energy window setup would then optimize the calculations for $S(U)$ in a parallel Wang−Landau simulation, for example.

At first, the proceeding prescription for energy window design may seem problematic since the entropy function required for the optimization is none other than the unknown function to be determined. This is not a significant issue, however, because an approximate entropy function can be used quite effectively to determine the energy window allocation. Given two pairs of data (T_1, U_1) and (T_2, U_2), an estimate for the temperature dependence of the potential energy is readily extracted by fitting to standard thermodynamic models. For example, in the constant

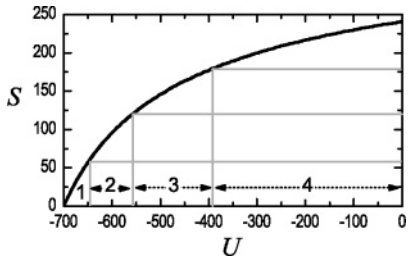

Figure 7. Schematic of the optimal design of energy windows for the 110-particle Lennard-Jones system when four processors are available. The total range of entropy is divided equally among the four, and the corresponding energies are extracted. Although not shown, each window would also be adjusted to create overlap with its neighbors.

heat capacity approximation, we have

$$U(T) \approx C_V T + U_0 \tag{9}$$

where $C_V = (U_2 - U_1)/(T_2 - T_1)$ is the constant heat capacity and $U_0 = (T_2U_1 - T_1U_2)/(T_2 - T_1)$ is the zero-temperature energy. After standard thermodynamic calculus, the corresponding entropy function is

$$S(U) = C_V \ln\left(\frac{U - U_0}{C_V}\right) \tag{10}$$

where the arbitrary additive constant has been omitted. A slightly more accurate description of the potential energy in liquids is available from an expression derived by Rosenfeld and Tarazona:[39]

$$U(T) \approx aT^{3/5} + U_0 \tag{11}$$

whereupon fitting to the two data points yields $a = (U_2 - U_1)/(T_2^{3/5} - T_1^{3/5})$ and $U_0 = (T_2^{3/5}U_1 - T_1^{3/5}U_2)/(T_2^{3/5} - T_1^{3/5})$. After transformation of this expression for the potential energy, the entropy function takes the form

$$S(U) = -\frac{3a}{2}\left(\frac{U - U_0}{a}\right)^{-2/3} \tag{12}$$

With either eq 10 or 12, it is possible to construct a rough estimate of the entropy to aid in the determination of energy window ranges. In our experience, this estimate does not need to be exceedingly accurate for range design. For example, we applied eq 12 to the 110-particle Lennard-Jones system, using the two-temperature-energy pairs (0.8, − 686) and (2.0, − 583), and predicted the energy ranges within $U_{min} = -700$ and $U_{max} = 0$ which would split the entropy evenly across four processors. The predicted ranges had actual ΔS values within 3% of the target average value, evaluated from the simulation-determined entropy. This low deviation is particularly impressive given the significant extrapolation to higher energies relative to the original two data points.

The generality of our recipe for parallel DOS implementations merits some discussion. The equal-entropy allotment rule is based upon the assumption that the tunneling time increases superlinearly with the range of entropy. To rationalize this assumption, let us define t_{tot}* as the total MC time of the overall simulation, equivalent to the sum of the times t_i* required for each processor to reach a given statistical accuracy. Our results have established that t* is a function of τ and hence of ΔS. In mathematical terms, then, $t_{tot}* = \sum t*(\Delta S_i)$ with the constraint that the individual ΔS_i values sum to the total entropy range. Therefore, if t* is *any* convex function of ΔS, then the minimum

in t_{tot}* occurs when each entropy range is the same. This is equivalent to the statement that the average of a function is always greater than the function of the average, when the second derivative of that function is positive. If we accept that t* always correlates linearly with the τ, then the equal-entropy rule applies whenever the tunneling time is a convex function of the entropy range. Equation 8 clearly meets this requirement in the present study, but any convex function will suffice.

The findings in this work raise an important question: what is the equivalent behavior of the tunneling time when alternate move types are employed or when the flat histogram occurs in variables other than the energy? Clearly, more advanced particle moves can accelerate the system's path through configuration space and thereby decrease the average tunneling time, but it is unknown whether such improvements will alter both fit parameters in eq 8 or the prefactor alone. The issue becomes even more uncertain when the moves modify different macroscopic variables. For example, flat-histogram techniques are frequently used to enhance sampling along the N coordinate when performing grand canonical simulations at near-coexistence conditions.[12,14] In this case, one generates a flat distribution in particle number, and the Helmholtz free energy $A(N;T)$ rather than the entropy is the unknown weighting function. One might hypothesize that tunneling time for a range of N might also correlate with the entropy, which in this case is $\Delta S = \beta\Delta\langle U\rangle - \beta\Delta A$ where $\beta = 1/k_BT$. These are all nontrivial issues for future study. The computational time requirements for measuring the tunneling time statistics in these cases might be significantly increased because of either expensive advanced moves or inherently longer tunneling times.

Recently, Earl and Deem have proposed a useful strategy for distributing processor workload in traditional parallel tempering simulations.[40] Given the number of MC move attempts to be performed at each temperature and their computational expense, Earl and Deem's approach distributes the workload for n replicas evenly across the available (possibly heterogeneous) processors. The method accomplishes this task by allowing each processor to operate on more than one replica, such that a processor that finishes work on one replica can begin work on another immediately. Our algorithm differs in that the workload is distributed not by dividing each replica's duty among several processors, but by altering the parameters characterizing the replicas themselves, namely, the potential energy range of interest. Thus, we take advantage of the flexibility of DOS simulations, which permit greater freedom in designing each replica than the usual parallel tempering situation. Our methodology is also readily extended to a heterogeneous set of processors; in this case, one aims to equate the "real-time" tunneling time among each processor, which is the tunneling time in number of MC steps times the computational time for a move.

Kofke and co-workers also identified the entropy as the key factor in optimizing staged free-energy calculations.[41,42] These authors focused on free-energy perturbation (FEP) methods, in which free-energy differences are computed by sampling a reference system alone, as in the particle-insertion method for determining the chemical potential. Their approach was substantially different from the present work; they derived an analytical expression for the expected ensemble-averaged error in FEP techniques and demonstrated that the statistical quality of the results depended on the entropy difference between the reference and target states. This enabled them to propose an optimal choice of reference state. Although that work is of a different nature than the dynamical issues we have addressed, it is worth considering the broad role that the entropy plays in free-energy calculations.

In the present work, we have examined the random walk dynamics for flat-histogram Monte Carlo simulations in potential energy. We have studied two models, the Lennard-Jones and Dzugutov systems, and have measured the tunneling behavior for the time required to move between low- and high-energy bounds. We have found that the average tunneling time for all energy ranges, system sizes, and models studied correlates with the range of entropy sampled according to the same power-law expression and that this result is in agreement with previous studies on the Ising model. We have also shown that the simulation time required to calculate the density of states to a specified accuracy scales with the tunneling time. These considerations enable one to optimize the energy window setup in parallel flat-histogram simulations by allocating each processor an equal entropy workload. This recipe should prove useful in aiding the determination of the density of states in low-temperature, complex, dense, or glassy systems.

Acknowledgment. We gratefully acknowledge the support of the Fannie and John Hertz Foundation and of the Department of Energy, Division of Chemical Sciences, Geosciences, and Biosciences, Office of Basic Energy Science (grants DE-FG02-87ER13714 to P.G.D. and DE-FG02-01ER15121 to A.Z.P.).

References and Notes

(1) Berg, B. A.; Neuhaus, T. *Phys. Rev. Lett.* **1992**, *68*, 9.
(2) Lee, J. *Phys. Rev. Lett.* **1993**, *71*, 211.
(3) Smith, G. R.; Bruce, A. D. *J. Phys. A* **1995**, *28*, 6623.
(4) Fitzgerald, M.; Picard, R. R.; Silver, R. N. *J. Stat. Phys.* **2000**, *98*, 321.
(5) Wang, J. S.; Swendsen, R. H. *J. Stat. Phys.* **2002**, *106*, 245.
(6) Wang, F. G.; Landau, D. P. *Phys. Rev. E* **2001**, *64*, 056101.
(7) Wang, F. G.; Landau, D. P. *Phys. Rev. Lett.* **2001**, *86*, 2050.
(8) Hansmann, U. H. E.; Okamoto, Y. *Curr. Opin. Struct. Biol.* **1999**, *9*, 177.
(9) Hansmann, U. H. E. *Int. J. Mod. Phys. C* **1999**, *10*, 1521.
(10) Mitsutake, A.; Sugita, Y.; Okamoto, Y. *Biopolymers* **2001**, *60*, 96.
(11) Rathore, N.; Knotts, T. A.; de Pablo, J. J. *J. Chem. Phys.* **2003**, *118*, 4285.
(12) Wilding, N. B. *Am. J. Phys.* **2001**, *69*, 1147.
(13) Shell, M. S.; Debenedetti, P. G.; Panagiotopoulos, A. Z. *Phys. Rev. E* **2002**, *66*, 056703.
(14) Errington, J. R. *J. Chem. Phys.* **2003**, *118*, 9915.
(15) Noguchi, H.; Yoshikawa, K. *J. Chem. Phys.* **1998**, *109*, 5070.
(16) Huang, L.; He, X. H.; Wang, Y. Y.; Chen, H. N.; Liang, H. J. *J. Chem. Phys.* **2003**, *119*, 2432.
(17) Vorontsov-Velyaminov, P. N.; Volkov, N. A.; Yurchenko, A. A. *J. Phys. A* **2004**, *37*, 1573.
(18) Bhattacharya, K. K.; Sethna, J. P. *Phys. Rev. E* **1998**, *57*, 2553.
(19) Faller, R.; de Pablo, J. J. *J. Chem. Phys.* **2003**, *119*, 4405.
(20) Calvo, F. *Mol. Phys.* **2002**, *100*, 3421.
(21) Wang, J. S.; Tay, T. K.; Swendsen, R. H. *Phys. Rev. Lett.* **1999**, *82*, 476.
(22) Yan, Q. L.; de Pablo, J. J. *Phys. Rev. Lett.* **2003**, *90*,
(23) Shell, M. S.; Debenedetti, P. G.; Panagiotopoulos, A. Z. *J. Chem. Phys.* **2003**, *119*, 9406.
(24) Lyubartsev, A. P.; Martsinovski, A. A.; Shevkunov, S. V.; Vorontsovvelyaminov, P. N. *J. Chem. Phys.* **1992**, *96*, 1776.
(25) Marinari, E.; Parisi, G. *Europhys. Lett.* **1992**, *19*, 451.
(26) Faller, R.; Yan, Q. L.; de Pablo, J. J. *J. Chem. Phys.* **2002**, *116*, 5419.
(27) Berg, B. A.; Celik, T. *Phys. Rev. Lett.* **1992**, *69*, 2292.
(28) Zhan, Z. F.; Lee, L. W.; Wang, J. S. *Physica A* **2000**, *285*, 239.
(29) Dayal, P.; Trebst, S.; Wessel, S.; Wurtz, D.; Troyer, M.; Sabhapandit, S.; Coppersmith, S. N. *Phys. Rev. Lett.* **2004**, *92*, 097201.
(30) Huse, D. A.; Trebst, S.; Troyer, M. *e-print cond-mat/0401195*, 2004.
(31) Dzugutov, M. *Phys. Rev. A* **1992**, *46*, R2984.
(32) Our discussion of the Wang−Landau method entails a slightly different presentation of the algorithm in that we refer directly to the entropy rather than the density of states itself as in the original papers.
(33) Frenkel, D.; Smit, B. *Understanding molecular simulation: from algorithms to applications*; Academic: San Diego, CA, 2002.

(34) Dzugutov, M.; Simdyankin, S. I.; Zetterling, F. H. M. *Phys. Rev. Lett.* **2002**, *89*, 195701.
(35) Dzugutov, M. *Phys. Rev. Lett.* **1993**, *70*, 2924.
(36) Roth, J.; Denton, A. R. *Phys. Rev. E* **2000**, *61*, 6845.
(37) Shell, M. S.; Debenedetti, P. G.; Stillinger, F. H. *J. Phys. Chem. B* **2004**, *108*, 6772.
(38) Roth, J. *Eur. Phys. J. B* **2000**, *14*, 449.
(39) Rosenfeld, Y.; Tarazona, P. *Mol. Phys.* **1998**, *95*, 141.
(40) Earl, D. J.; Deem, M. W. *J. Phys. Chem. B* **2004**, *108*, 6844.
(41) Lu, N. D.; Kofke, D. A. *J. Chem. Phys.* **1999**, *111*, 4414.
(42) Kofke, D. A. *J. Chem. Phys.* **2002**, *117*, 6911.