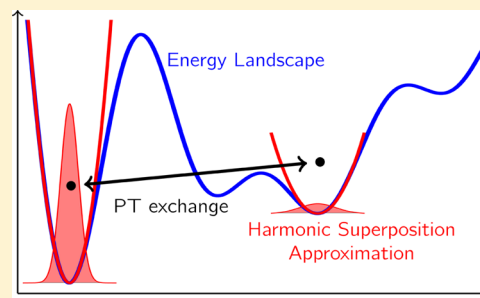


# Superposition-Enhanced Estimation of Optimal Temperature Spacings for Parallel Tempering Simulations

Andrew J. Ballard\* and David J. Wales

University Chemical Laboratories, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom

**ABSTRACT:** Effective parallel tempering simulations rely crucially on a properly chosen sequence of temperatures. While it is desirable to achieve a uniform exchange acceptance rate across neighboring replicas, finding a set of temperatures that achieves this end is often a difficult task, in particular for systems undergoing phase transitions. Here we present a method for determination of optimal replica spacings, which is based upon knowledge of local minima in the potential energy landscape. Working within the harmonic superposition approximation, we derive an analytic expression for the parallel tempering acceptance rate as a function of the replica temperatures. For a particular system and a given database of minima, we show how this expression can be used to determine optimal temperatures that achieve a desired uniform acceptance rate. We test our strategy for two atomic clusters that exhibit broken ergodicity, demonstrating that our method achieves uniform acceptance as well as significant efficiency gains.



## 1. INTRODUCTION

Effective equilibrium sampling of complex systems remains a major challenge in computational thermodynamics. Systems exhibiting broken ergodicity, reliant on rare event fluctuations, or otherwise locally “trapped”, are characterized by long equilibration times due to large (free) energy barriers separating metastable states. This problem is particularly acute for systems undergoing strong first-order-like phase transitions, such as transitions corresponding to protein crystallization, crystal polymorphism, and other changes in morphology for condensed-matter systems.<sup>1–3</sup>

Parallel tempering (PT)<sup>3–5</sup> has emerged as a powerful method to analyze such complex systems. Here, several copies of a given system are simulated at different temperatures, with occasional moves that attempt to swap configurations of neighboring replicas. By virtue of these exchange moves, the low-temperature replicas gain access to larger thermal fluctuations of the high-temperature replicas, facilitating quicker transitions between metastable states.

The effectiveness of a given PT simulation depends crucially upon the sequence of temperatures that define the replicas. The temperature spacing should be small enough to produce a reasonable acceptance rate between neighboring replicas, yet large enough so that computational resources are not wasted on an unnecessarily large number of replicas. Generally, a uniform acceptance rate is desirable across all neighboring replica pairs.<sup>6–11</sup> This condition ensures that the progression of replica trajectories among a ladder of temperatures is uniform, providing an equilibration of the combined replica system that is not hindered by bottlenecks along the temperature axis.

In general it is a difficult problem to find a temperature progression that achieves a uniform acceptance rate, often requiring an iterative or adaptive approach before production runs.<sup>12–14</sup> However, theoretical progress has also been made

under certain assumptions about the thermodynamic behavior of the system of interest. Kofke has shown that a geometric progression of temperatures

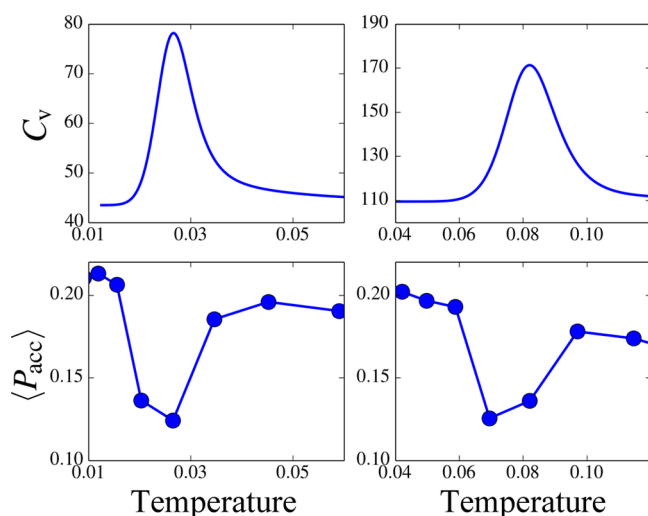
$$\frac{T_{i+1}}{T_i} = \text{const} \quad (1)$$

leads to a uniform acceptance rate for systems with constant heat capacity.<sup>7,8</sup> Today the geometric spacing is a standard rule of thumb, considered a best guess if nothing else is known about the system of interest. Yet it is not always optimal. Indeed, systems undergoing phase transitions will generally have a significant heat capacity peak, in which case a constant heat capacity approximation is obviously inaccurate, and hence a temperature progression given by eq 1 is certainly not optimal (see Figure 1). Other methods for determination of optimal PT temperatures have focused on uniform acceptance for solvated protein systems,<sup>15</sup> optimizing replica diffusion through iterative approaches,<sup>16</sup> or have made temperature a dynamical parameter.<sup>17–20</sup> In the following we will be concerned with obtaining a uniform acceptance profile, although the method we describe can be used to optimize according to other criteria.

In the present contribution we present a method for determining optimal PT temperatures, which can be especially useful for systems undergoing phase transitions. The method relies upon knowledge of the competing phases, represented as configurational basins on the potential energy landscape. Using the harmonic superposition approximation (HSA),<sup>2</sup> in which the landscape is approximated by harmonic wells at the various minima, we approximate thermodynamic quantities, such as the energy density of states, heat capacities, and the various free

Received: September 2, 2014

Published: November 4, 2014



**Figure 1.** Bottleneck problem in parallel tempering: While a geometric spacing of temperatures (eq 1) provides a uniform acceptance rate when the heat capacity is roughly constant, in the vicinity of a phase transition it leads to a bottleneck. These effects are clearly illustrated for two atomic clusters of 31 (left panels) and 75 atoms (right panels) that exhibit broken ergodicity.<sup>2</sup>

energies of the competing phases. With the approximate distributions obtained by the HSA we can estimate the parallel tempering acceptance probability, using information contained solely within a database of configurational local minima. To arrive at our result, we subdivide the configuration space into separate basins of attraction on the potential energy landscape and express the PT acceptance rate of a given replica pair as a weighted sum over contributions from individual pairs of minima (eq 2). In this way, the replica pair acceptance rates are determined explicitly from the relative weights of the various competing phases, along with their associated conditional acceptance probabilities. Using the known analytic energy distributions within the HSA, we derive an analytic expression, eq 10, for the conditional acceptance rates. The overall PT rate is then obtained by weighting each of these terms by the HSA local free energies of the individual wells. With this expression at our disposal, we find an optimized temperature progression via a simple algorithm, described in Section 3, which matches each neighboring pair of temperatures to a target acceptance rate.

Our method exploits the HSA to represent the landscape. Because the HSA becomes exact in the limit of zero temperature, we expect our method to be ideal for transitions at low temperatures compared to the barrier heights between phases. Indeed, low-temperature solid–solid phase transitions of many atomic clusters are very accurately described by the HSA.<sup>2</sup> While the HSA is in practice never an exact approximation, it is clear why this method could more accurately predict thermodynamics of systems undergoing phase transitions: While a constant heat capacity assumption is certainly unjustifiable for an arbitrary system, at a relatively low temperature it is reasonable to assume that, within a certain basin of attraction, the system samples an energy landscape that is roughly harmonic, to a first approximation.

While our method relies upon knowledge of competing minima, finding these structures is computationally inexpensive compared to characterizing the thermodynamics.<sup>21,22</sup> Because the end result of such techniques is not thermodynamics but a

list of minima, the dynamics does not require detailed balance to be satisfied. Software such as GMIN<sup>23</sup> and its python-based counterpart *pele*<sup>24</sup> employ global optimization tools of energy landscape exploration and provide rapid, efficient, and accurate determination of low-lying minima. We also note that recent developments in thermodynamic sampling algorithms (which do satisfy detailed balance) have exploited the HSA to obtain more accurate thermodynamic results using information contained in a database of minima. Examples include parallel tempering with an auxiliary “reservoir” replica, which samples the HSA distribution,<sup>25,26</sup> superposition-enhanced nested sampling, which draws samples from the HSA,<sup>27</sup> and the hybrid basin-sampling method.<sup>21</sup> Similar methods such as smart-darting,<sup>28</sup> based upon smart-walking,<sup>29</sup> employ similar ideas with collective movesets that hop between minima.

The remainder of the paper is organized as follows. In Section 2 we derive an expression for the PT acceptance probability within the HSA approximation. Our central result, eq 10, allows us to estimate PT acceptance rates solely from a database of minima. Following this estimation, in Section 3 we describe how this expression can be exploited to determine a set of temperatures which achieve a uniform acceptance rate. In Section 4 we then test this prescription on two systems exhibiting low-temperature solid–solid phase transitions, namely Lennard-Jones clusters with 31 and 75 particles, LJ<sub>31</sub> and LJ<sub>75</sub>, showing that our method can achieve uniform acceptance rates across phase transitions and perform more efficiently than simulations using standard temperature progressions. Finally, in Section 5, we discuss the conditions under which the current strategy can be most effective and useful.

## 2. DERIVATION OF ACCEPTANCE PROBABILITY EXPRESSION

Consider a PT simulation for a system described by a potential energy  $V(x)$ , where  $x$  is a point in the system’s configuration space. Let  $A$  and  $B$  be two replicas in this simulation, at temperatures  $T_A$  and  $T_B$ , respectively, which sample the system canonically, so that  $\rho_A(x) \propto e^{-V/kT_A}$  etc. In the following we assume  $T_B > T_A$ . We desire an expression for the average PT acceptance rate,  $\langle P_{acc} \rangle$ , between our pair of replicas ( $A, B$ ). To begin our analysis we divide the configuration space into a finite number of wells, indexed by  $w$ . Here a point  $x$  in configuration space belongs to well  $w$  if it lies in the corresponding basin of attraction for a given local minimum, defined by steepest-descent pathways which converge to the minimum in question.<sup>2,30</sup> Without loss of generality we can write any configuration space average as a weighted sum over conditional averages. Hence we can express the average PT acceptance probability of this pair as

$$\langle P_{acc} \rangle = \sum_{w,w'} p_w^A p_{w'}^B \bar{P}_{acc}(w, w') \quad (2)$$

where  $p_w^i = Z_w^i/Z^i$  is the equilibrium occupation probability for well  $w$  in replica  $i$  and  $\bar{P}_{acc}$  is the average acceptance probability, conditioned upon  $A$  and  $B$  residing in wells  $w$  and  $w'$ , respectively.

We desire an explicit expression for  $\langle P_{acc} \rangle$  in terms of  $T_A$  and  $T_B$ , via eq 2. While the well weights  $p_w^i$  can be approximated by the corresponding HSA well free energies (see Section 3), more challenging is the conditional acceptance,  $\bar{P}_{acc}(w, w')$ . The

central result of this section is an analytic expression for  $\bar{P}_{\text{acc}}(w, w')$ , eq 10, which we now derive.

We begin by writing the conditional acceptance rate as a configuration-space average:

$$\bar{P}_{\text{acc}}(w, w') = \int dx \rho_A(x|w) \int dy \rho_B(y|w') \min\{1, e^{\Delta\beta\Delta V}\} \quad (3)$$

where  $\Delta\beta = \beta_B - \beta_A < 0$ ,  $\Delta V(x, y) = V(y) - V(x)$ , and  $\beta_i = 1/kT_i$ . The conditional configurational distributions  $\rho_i(x|w) \propto e^{-\beta_i V}$  for  $x \in w$  and 0 otherwise. Using properties of the min function it can easily be shown<sup>31</sup> that eq 3 can be rewritten as

$$\bar{P}_{\text{acc}}(w, w') = 2 \int \int_{\Delta\beta\Delta V(x,y)>0} dx dy \rho_A(x|w) \rho_B(y|w') \quad (4)$$

or alternatively, in terms of energy distributions  $p_A$  and  $p_B$ :

$$\bar{P}_{\text{acc}}(w, w') = 2 \int_{-\infty}^{\infty} d\epsilon p_A(\epsilon|w) \int_{-\infty}^{\epsilon} d\epsilon' p_B(\epsilon'|w') \quad (5)$$

The distribution  $p_i(\epsilon|w)$  corresponds to the canonical energy distribution of replica  $i$  (at temperature  $T_i$ ) when restricted to well  $w$ .

Our expression for the conditional acceptance probability, eq 5, is exact, but the energy distributions  $p_i(\epsilon|w)$  are unknown. In the following analysis we approximate these quantities by the (known) HSA distributions, described by a gamma distribution. For replica  $A$ , for instance, we have

$$p_A(\epsilon|w) = \frac{\beta_A^\kappa}{\Gamma(\kappa)} (\epsilon - \epsilon_w)^{\kappa-1} e^{-\beta_A(\epsilon - \epsilon_w)} \quad (6)$$

where  $\epsilon_w$  is the potential energy of the local minimum associated with well  $w$ ,  $\Gamma$  is the gamma function, and  $2\kappa$  is the number of configurational degrees of freedom. Note that this distribution only depends on the energy,  $\epsilon_w$ , of the minimum in question and not the shape (curvature) of the well. In the limit  $\kappa \rightarrow \infty$  this distribution approaches a Gaussian, described by cumulants:

$$\mu_A(w) = \kappa T_A + \epsilon_w; \quad \sigma_A^2(w) = \kappa T_A^2 \quad (7)$$

$$\mu_B(w') = \kappa T_B + \epsilon_{w'}; \quad \sigma_B^2(w') = \kappa T_B^2$$

Assuming this Gaussian limit has been reached (since  $2\kappa$  is the number of degrees of freedom, this is not a severe assumption), we can evaluate eq 5 as

$$\begin{aligned} \bar{P}_{\text{acc}}(w, w') &= 2 \int_{-\infty}^{\infty} du_1 p_A(u_1|w) \int_{-\infty}^{u_1} du_2 p_B(u_2|w') \\ &= \int_{-\infty}^{\infty} du_1 p_A(u_1|w) \text{erfc}\left(\frac{\mu_B - u_1}{\sqrt{2}\sigma_B}\right) \end{aligned} \quad (8)$$

In going from the first to second line, we used the definition of the error function and exploited the fact that it is an odd function. Equation 8 is an integral of a Gaussian convoluted with an error function, which is

$$\bar{P}_{\text{acc}}(w, w') = \text{erfc}\left(\frac{\Delta\mu_{ww'}}{\sqrt{2}r_{ww'}}\right) \quad (9)$$

where  $\Delta\mu_{ww'} = \mu_B(w') - \mu_A(w)$  and  $r_{ww'} = (\sigma_A(w)^2 + \sigma_B(w')^2)^{1/2}$ .

Substituting in the cumulants for our replicas, eq 7, we arrive at our final expression:

$$\bar{P}_{\text{acc}}(w, w') = \text{erfc}\left[\chi_0 \left(1 + \frac{\Delta\epsilon_{ww'}}{\kappa\Delta T}\right)\right] \quad (10)$$

where

$$\chi_0 = \sqrt{\frac{\kappa}{2}} \frac{(\gamma - 1)}{\sqrt{1 + \gamma^2}} \quad (11)$$

Here  $\gamma = T_B/T_A$ ,  $\Delta T = T_B - T_A$ , and  $\Delta\epsilon_{ww'} = \epsilon_{w'} - \epsilon_w$  is the potential energy difference between the local minima corresponding to wells  $w$  and  $w'$ .

It is instructive to consider limiting cases of eq 10. For  $w = w'$ , we have  $\Delta\epsilon_{ww'} = 0$ , and our expression reduces to the previous result derived by Kofke for constant  $C_V$  systems with Gaussian energy distributions (see eq 13 of refs 8 and 7). Here we find that  $\gamma = T_B/T_A$  alone controls the acceptance rate, validating a geometric temperature progression. However, if  $w \neq w'$ , then the acceptance deviates from “geometric behavior” by a factor  $\Delta\epsilon_{ww'}/\kappa\Delta T$ , which becomes significant when the energy splitting between minima is comparable to the difference in average internal energy of the two replicas. We expect such deviations to appear near a phase transition, when multiple wells contribute to the thermodynamics.

### 3. OPTIMIZING THE REPLICA SPACING

With eq 10 in hand, we can now estimate the full acceptance probability  $\langle P_{\text{acc}} \rangle$ , in terms of  $T_A$  and  $T_B$ . The occupation probabilities  $p_w^i$ , which weight the individual contributions  $\bar{P}_{\text{acc}}(w, w')$ , are conveniently expressed in terms of their free energies  $f_w$ . For a replica at inverse temperature  $\beta$ , we have

$$p_w = \frac{e^{-\beta f_w}}{\sum_{w'} e^{-\beta f_{w'}}} \quad (12)$$

Within the HSA, the free energy is given by

$$f_w = \epsilon_w - \beta^{-1}s_w \quad (13)$$

with local minimum potential energy  $\epsilon_w$  and well entropy

$$s_w = -\kappa \ln \bar{\nu}_w - \ln n_w \quad (14)$$

Here the entropy is specified by two standard features of the well: (1) the geometric mean vibrational frequency,  $\bar{\nu}_w$ , which describes the curvature of the minimum; and (2) the number of distinguishable permutation-inversion isomers,  $n_w$ , determined by the point group associated with the given minimum.<sup>2</sup>

Equations 2, 10, and 12–14 provide an explicit expression for our acceptance probability, which we write as

$$\langle P_{\text{acc}} \rangle \equiv g(T_A, T_B) \quad (15)$$

and depends upon the database of minima. Using this expression, we now optimize a sequence of  $M$  temperatures, as follows. First, a target acceptance rate  $p_{\text{acc}}^*$  is chosen, and an initial temperature  $T_1$  is specified. A sequence of temperatures  $T_2, \dots, T_M$  is then built from the bottom up: given a temperature  $T_i$ ,  $T_{i+1}$  is determined from the condition:

$$g(T_i, T_{i+1}) = p_{\text{acc}}^* \quad (16)$$

This is a standard optimization problem, which can be solved in various ways.<sup>2,32</sup> Hence we generate a sequence of  $M$  temperatures which, within the HSA approximation, each have a uniform acceptance of  $p_{\text{acc}}^*$  and which relies only on the database of minima.

#### 4. PERFORMANCE ON TEST SYSTEMS

In this section we evaluate the effectiveness of our method by applying it to two model systems, clusters of Lennard-Jones particles, interacting via the pairwise potential:

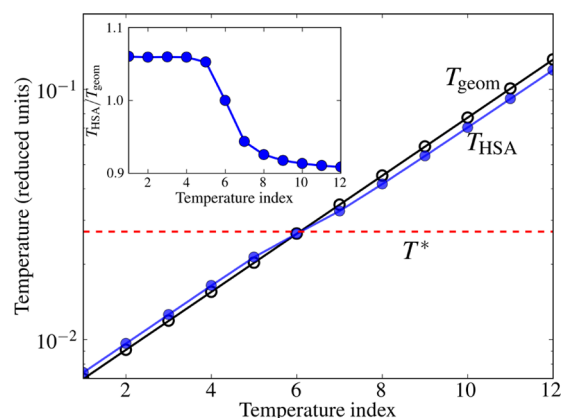
$$u(r) = 4\epsilon \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^6 \right] \quad (17)$$

These clusters can exhibit strong first-order transitions at low temperatures<sup>2,25,26,33–35</sup> and as such are prototypical examples of complex systems exhibiting broken ergodicity. Hence they have been used extensively for benchmark studies of thermodynamic sampling methods.<sup>2,21,25–27</sup> In what follows we will focus on LJ<sub>31</sub> and LJ<sub>75</sub>, clusters of 31 and 75 particles, respectively. Both of these systems exhibit low-temperature solid–solid transitions indicated by pronounced peaks in the heat capacity profile (Figure 1), making them ideal candidates for testing our method.

To evaluate the effectiveness of our strategy we perform two PT simulations for each of our LJ systems: one simulation employing an HSA-optimized set of temperatures, and the other using a standard geometric temperature progression. Because this approach is intended to be useful near low-temperature phase transitions, we consider a temperature range that spans the peak in the heat capacity profile (Figure 1). For each of our test systems this range was achieved using  $M = 12$  replicas. The precise HSA-optimized sequence was determined via the optimization scheme described in Section 3, using a relatively cautious<sup>14,36</sup> target acceptance rate of  $p_{\text{acc}}^* = 0.22$ . The corresponding reference simulations employed a temperature sequence defined by the recursion  $T_i = \gamma T_{i-1}$ , where the constant  $\gamma$  was chosen to achieve the acceptance rate  $p_{\text{acc}}^* = 0.22$  for the lowest temperature pair.

In order to uniquely specify the temperature sequence, we must also specify initial conditions from which the optimization scheme and geometric recursion begin. Rather than (arbitrarily) choosing an initial temperature  $T_1$ , we instead require an intermediate replica  $R \approx M/2$  to be located at the peak of the  $C_V$  curve, i.e.,  $T_R = T^*$ . In this way, the temperature progression of the HSA and reference simulations coincide exactly at replica  $R$ , located at the  $C_V$  peak. For our simulations we chose  $R = 5$  and  $6$  for LJ<sub>31</sub> and LJ<sub>75</sub>, respectively. This choice of initial condition is convenient for the purposes of the advanced simulation strategy which we describe below, in addition to being less arbitrary than a value of  $T_1$  chosen ad hoc. In Figure 2 we compare the temperature set obtained from HSA optimization to the reference geometric progression for the case of LJ<sub>31</sub>. The corresponding plot for LJ<sub>75</sub> is qualitatively similar.

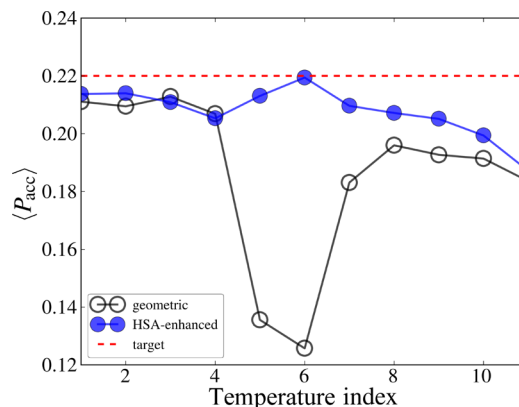
To ameliorate poor PT convergence due to broken ergodicity, we employed an advanced simulation technique, reservoir replica exchange,<sup>25,26</sup> whereby the PT simulation employs an additional “reservoir replica,” which samples the HSA approximation to the underlying landscape. Because the reservoir replica is an efficient sampler, it provides an ersatz high-temperature replica: its frequent transitions between metastable states provide a channel for rapid PT equilibration when coupled to the sequence of  $M$  “physical” replicas. By employing this method, we were able to obtain converged thermodynamics without the need for additional replicas with temperatures approaching melting. For our simulations, the reservoir replica was coupled to replica  $R$  at the  $C_V$  peak ( $T_R = T^*$ ), shown elsewhere to be a good choice for these clusters.<sup>26</sup>



**Figure 2.** HSA-optimized temperature progression (blue), compared to reference geometric progression (black) on a log scale, and their ratio (inset) on a linear scale. The progression of the HSA-optimized replicas mimics the reference sequence at high and low temperatures, where a geometric sequence is known to be optimal. At the phase transition (temperature index 6), there is a crossover.

Since replica  $R$  coincides for our HSA and reference simulations, the reservoir method couples to the same distribution in both the HSA and reference simulations, reducing any systematic discrepancies due to this auxiliary sampling.

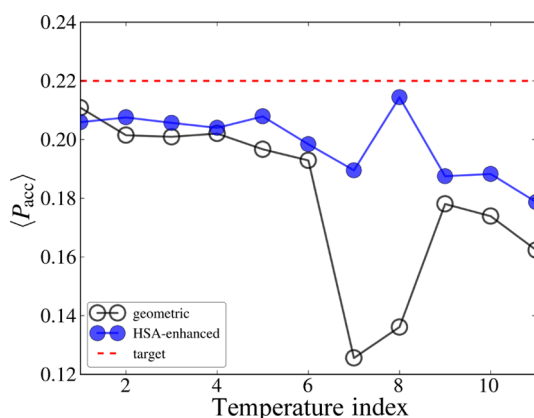
We now present the results of our PT simulations, focusing first on the acceptance rate. Figure 3 displays the acceptance



**Figure 3.** Acceptance profile for LJ<sub>31</sub>: The average replica exchange acceptance probability for pair  $(T_i, T_{i+1})$  is plotted vs temperature index  $i$ , for temperature spacings chosen geometrically (empty black circles) and chosen by HSA-optimization (filled blue circles). The dip in the geometric case coincides with the LJ<sub>31</sub> heat capacity peak (Figure 1). At higher temperatures both profiles deviate from uniform behavior, as the HSA becomes less accurate.

profile of LJ<sub>31</sub> for both the reference and HSA-enhanced simulations. The reference simulation exhibits a roughly constant profile near the terminal temperatures but displays a noticeable dip in the vicinity of the phase transition. The HSA-enhanced simulations, however, display a profile that is much more uniform across the phase transition. Our simulations of LJ<sub>75</sub> show similar behavior: Figure 4 indicates that the reference simulations of LJ<sub>75</sub> fail to achieve a uniform profile near the phase transition, whereas with the HSA-optimized set this dip is absent. From these results it is clear that HSA-optimized temperature set can achieve a more uniform acceptance profile than standard geometric progression.





**Figure 4.** Average replica exchange acceptance probability for  $\text{LJ}_{75}$ , for temperature spacings chosen geometrically (empty black circles) and chosen by HSA optimization (filled blue circles).

We have seen above that our strategy provides a more uniform acceptance profile near a  $C_V$  peak. We now investigate whether it can also lead to gains in simulation efficiency. To characterize convergence of our PT simulation, we need a measure that captures the dynamics of our entire  $M$  replica system. This measure can be achieved by following the replica trajectories as they evolve between the set of  $M$  temperatures. By “replica trajectory” we refer to a time series trace of the temperature index  $k$ , where discrete jumps  $k \rightarrow k \pm 1$  occur by successful replica exchanges with neighboring temperatures. By projecting onto  $k$ , the dynamics of a given replica trajectory  $r$  are represented as a random walker in this discrete space, with transitions in  $k$  related to the PT acceptance profile. The entire  $M$  replica system will be converged once each of our  $M$  trajectories has exchanged many times between the terminal temperatures ( $k = 1$  and  $k = M$ ), and the time scale under which these transitions occur will give an indication of the convergence condition for our PT simulation.

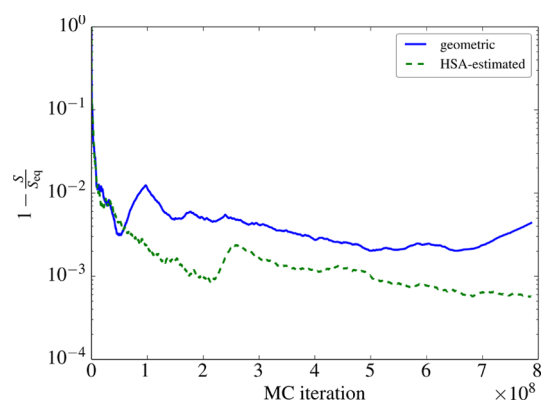
To characterize convergence from our replica trajectories, we use a technique developed by Doll et al.<sup>37</sup> For each trajectory  $r$ , let  $\mathbf{p}^r(t) = (p_1^r, \dots, p_M^r)$  denote the vector of observed occupation probabilities up to time  $t$ . That is, element  $p_k^r(t)$  is the fraction of simulation time trajectory  $r$  was observed in temperature index  $k$  up to time  $t$ . Since we have  $M$  trajectories, the dynamics of our simulation are represented by  $M$  such vectors,  $\mathbf{p}^1, \dots, \mathbf{p}^M$ . Importantly, we know the converged limit of  $\mathbf{p}^r$ : As  $t \rightarrow \infty$  each temperature is equally occupied,<sup>37</sup> so  $\mathbf{p}^r \rightarrow \mathbf{p}_{\text{eq}} = (1/M, \dots, 1/M)$  for each  $r$ . The approach of  $\mathbf{p}^r(t)$  to  $\mathbf{p}_{\text{eq}}$  will determine the overall simulation convergence. To analyze this convergence we define an occupation-based entropy:<sup>37</sup>

$$S^{(r)}(t) = - \sum_{i=1}^M p_i^r(t) \log p_i^r(t) \quad (18)$$

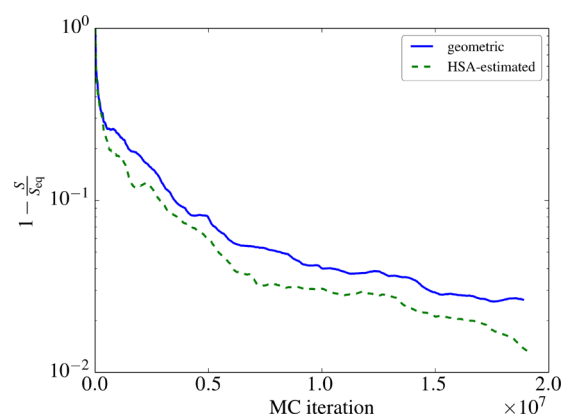
By following the approach of  $S^{(r)}$  to its limiting value,  $S_{\text{eq}} = \ln(M)$ , we can straightforwardly probe convergence of our  $M$  replica system. Below, we report values averaged over the  $M$  replica trajectories:  $S(t) = (1/M) \sum_r S^{(r)}(t)$ .

In Figures 5 and 6 we display  $S(t)$  for  $\text{LJ}_{31}$  and  $\text{LJ}_{75}$ , respectively. For both systems, we clearly see that  $S$  approaches  $S_{\text{eq}}$  quicker with HSA-optimized temperatures compared to the reference simulations, indicating that HSA-optimized spacings provide more rapid equilibration.

As a second test of convergence, we directly calculate a time scale associated with the temperature index  $k$  for our replica



**Figure 5.** Convergence of  $\text{LJ}_{31}$  parallel tempering simulations: The convergence of the replica occupation-based entropy  $S$  (eq 18) to its equilibrated value,  $S_{\text{eq}}$  is plotted on a log scale as a function of simulation time. The HSA-enhanced temperature choice performs significantly better than the standard geometric temperature progression.



**Figure 6.** Convergence of  $\text{LJ}_{75}$  simulations in terms of replica occupation entropy  $S$ : In the long run, the HSA-enhanced temperature choice shows an improvement over a standard geometric temperature progression.

trajectories. For a given replica trajectory, we estimate a characteristic relaxation time associated with  $k$  as

$$\tau = \frac{1}{\sigma_k^2} \int_0^\infty dt C_k(t) \quad (19)$$

where

$$C_k(t) = \langle k(t)k(0) \rangle - \langle k \rangle^2 \quad (20)$$

is the autocorrelation function of  $k$  and  $\sigma_k^2$  is its variance. In Table 1 we display our estimates of  $\tau$  for simulations with HSA-optimized and geometric temperature sequences. The values are reported as averages over the  $M$  replica trajectories, with individual  $\tau$  values obtained from block averaging.<sup>38</sup> As shown

**Table 1.** Autocorrelation Time of PT Temperature Index  $k$  for Reference Simulations ( $\tau_{\text{geom}}$ ), HSA-Enhanced ( $\tau_{\text{HSA}}$ ), and Their Ratio<sup>a</sup>

| system           | $\tau_{\text{geom}}$ | $\tau_{\text{HSA}}$ | $\tau_{\text{geom}}/\tau_{\text{HSA}}$ |
|------------------|----------------------|---------------------|--|
| $\text{LJ}_{31}$ | 4.97                 | 0.640               | 7.76                                   |
| $\text{LJ}_{75}$ | 1.42                 | 0.90                | 1.59                                   |

<sup>a</sup>Times are reported per  $10^6$  MC steps.

in Table 1, the relative speedup due to HSA optimization ranges from 1.59 for LJ<sub>75</sub> to 7.76 for LJ<sub>31</sub>, providing an improvement.

Based upon this time scale analysis, it is clear that LJ<sub>31</sub> benefits from HSA-optimized temperatures. The results from LJ<sub>75</sub>, on the other hand, are more difficult to discern: with HSA-optimization, LJ<sub>75</sub> only receives a marginal benefit as judged by the convergence time scales. The fact that  $\tau_{\text{HSA}} \approx \tau_{\text{geom}}$  despite a more uniform acceptance profile (Figure 4) may indicate that the bottleneck for equilibration is not captured simply by a dip in the acceptance profile, but rather is determined by rare structural fluctuations within a given replica acting on much longer time scales. Indeed, it is known that low-temperature simulations of LJ<sub>75</sub> are extremely difficult to converge, even when using parallel tempering.<sup>26</sup> In this case, the convergence time scales reported for LJ<sub>75</sub> can only be considered lower limits to equilibration.

## 5. DISCUSSION AND CONCLUSION

We have applied our optimization technique to systems undergoing first-order-like phase transitions, for which there is currently no successful analytic theory for PT temperature spacings. The results above clearly show that our optimized temperature sequence can provide PT results with (1) a more uniform acceptance profile and (2) faster convergence. A standard geometric progression of temperatures applied to such a situation leads to a bottleneck in the acceptance profile, which ultimately slows system equilibration and hinders replica mixing. This failure is due to the breakdown of the assumption of constant heat capacity, resulting from a shift of metastability from one phase to another. Our approach, on the other hand, can be very effective in this regime. The method is based upon the HSA, an approximation of the underlying canonical distribution that is able to predict changes of state and other nontrivial thermodynamic behavior. By utilizing the HSA, our temperature optimization scheme is sensitive to changes in populations of various phases that occur near such transitions and hence can tolerate a nonuniform heat capacity.

Although our method requires no particular assumptions about the heat capacity profile, it does, importantly, rely on the HSA to faithfully represent the underlying canonical distribution. This representation is necessary in order to accurately estimate the thermodynamic behavior of the system, in particular its acceptance probability profile. Because the HSA is accurate only for low temperatures (and only exact at  $T = 0$ ), our method will generically be most reliable for systems in or near the solid phase. We therefore intend to use it for suitable applications, such as crystal polymorphs, crystal–quasicrystal coexistence, and other systems undergoing structural rearrangements.

In a similar vein, our method also requires a database of configurational minima of the potential energy landscape, which must first be found before the HSA can be constructed and utilized. Although the total number of minima can often be large, fortunately our method only requires those minima that are important for the thermodynamics. For the systems we are considering (solids at reasonably low temperatures), this number is relatively small. Furthermore, finding these relevant minima can be computationally inexpensive compared to equilibrium thermodynamic sampling as a whole;<sup>22,21</sup> primarily because minima searching techniques need not satisfy detailed balance and hence permit a wider range of step-taking routines. By investing a relatively minor amount of time in global

optimization (minutes in the case of our test systems), we obtain quicker convergence of our (much) longer PT simulations thanks to the HSA-optimized temperatures. Basin-hopping<sup>33,39,40</sup> is one efficient strategy for minima searching and exploration of the potential energy landscape and is currently available in our group software.<sup>23,24</sup>

Finally, we mention that our method can be extended to optimize temperatures according to other criteria as well. Here temperatures were selected to achieve a uniform acceptance profile. Others, however, argue instead that temperature spacings should minimize replica round-trip times.<sup>16</sup> To achieve this, Katzgraber et al. use an iterative procedure that optimizes the flow of replica trajectories between terminal temperatures, achieving sizable efficiency gains over a uniform acceptance strategy.<sup>16</sup> The tools developed here, in particular the estimates for acceptance rates, could be used to optimize in a similar way. This analysis is in progress.

## AUTHOR INFORMATION

### Corresponding Author

\*E-mail: ab2111@cam.ac.uk.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

It is a pleasure to acknowledge stimulating conversations with Dr. Jacob Stevenson. This project was funded by EPSRC Grant EP/I001352/1 and the ERC.

## REFERENCES

- (1) Doye, J. P. K.; Miller, M. A.; Wales, D. J. *J. Chem. Phys.* **1999**, *110*, 6896–6906.
- (2) Wales, D. J. *Energy Landscapes*; Cambridge University Press: Cambridge, 2003.
- (3) Frenkel, D.; Smit, B. *Understanding Molecular Simulation*; Academic Press: San Diego, CA, 2002.
- (4) Swendsen, R.; Wang, J.-S. *Phys. Rev. Lett.* **1986**, *57*, 2607–2609.
- (5) Earl, D.; Deem, M. W. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3910–3916.
- (6) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **1999**, *314*, 141–151.
- (7) Kofke, D. A. *J. Chem. Phys.* **2002**, *117*, 6911–6914.
- (8) Kofke, D. A. *J. Chem. Phys.* **2004**, *120*, 10852.
- (9) Nadler, W.; Hansmann, U. H. E. *Phys. Rev. E* **2007**, *75*, 026109.
- (10) Nadler, W.; Hansmann, U. H. E. *J. Phys. Chem. B* **2008**, *112*, 10386–10387.
- (11) Denschlag, R.; Lingenehl, M.; Tavan, P. *Chem. Phys. Lett.* **2009**, *473*, 193–195.
- (12) Sanbonmatsu, K.; García, A. *Proteins: Struct., Funct., Bioinf.* **2002**, *46*, 225–234.
- (13) Schug, A.; Herges, T.; Wenzel, W. *Proteins: Struct., Funct., Bioinf.* **2004**, *57*, 792–798.
- (14) Rathore, N.; Chopra, M.; de Pablo, J. J. *J. Chem. Phys.* **2005**, *122*, 024111.
- (15) Patriksson, A.; van der Spoel, D. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2073–2077.
- (16) Katzgraber, H. G.; Trebst, S.; Huse, D. A.; Troyer, M. *J. Stat. Mech.: Theory Exp.* **2006**, P03018.
- (17) Rick, S. W. *J. Chem. Phys.* **2007**, *16*, 054102.
- (18) Kim, J.; Straub, J. E.; Keyes, T. *Phys. Rev. Lett.* **2006**, *97*, 50601.
- (19) Kim, J.; Keyes, T.; Straub, J. E. *J. Chem. Phys.* **2009**, *130*, 124112.
- (20) Kim, J.; Straub, J. E.; Keyes, T. *J. Phys. Chem. B* **2012**, *116*, 8646–8653.
- (21) Wales, D. J. *Chem. Phys. Lett.* **2013**, *584*, 1–9.
- (22) Oakley, M. T.; Johnston, R. L.; Wales, D. J. *Phys. Chem. Chem. Phys.* **2013**, *15*, 3965–3976.

(23) Wales, D. J. *GMIN: A program for finding global minima and calculating thermodynamic properties from basin-sampling*; <http://www-wales.ch.cam.ac.uk/software.html>.

(24) Stevenson, J. D.; Rühle, V. R.; Wales, D. J. *pele: Python Energy Landscape Explorer*; <http://github.com/pele-python/pele> (Accessed 2013).

(25) Sharapov, V. A.; Meluzzi, D.; Mandelshtam, V. A. *Phys. Rev. Lett.* **2007**, *98*, 105701.

(26) Sharapov, V. A.; Mandelshtam, V. A. *J. Phys. Chem. A* **2007**, *111*, 10284–10291.

(27) Martiniani, S.; Stevenson, J. D.; Wales, D. J.; Frenkel, D. *Phys. Rev. X* **2014**, *4*, 031034.

(28) Andricioaei, I.; Straub, J. E.; Voter, A. F. *J. Chem. Phys.* **2001**, *114*, 6994–7000.

(29) Zhou, R.; Berne, B. J. *J. Chem. Phys.* **1997**, *107*, 9185–9196.

(30) Mezey, P. G. *Potential Energy Hypersurfaces*; Elsevier: Amsterdam, 1987.

(31) Park, S. *Phys. Rev. E* **2008**, *77*, 016709.

(32) Bonnans, J. F.; Gilbert, J. C.; Lemaréchal, C.; Sagastizábal, C. A. *Numerical Optimization: Theoretical and Practical Aspects*; Springer: Berlin, 2006.

(33) Wales, D. J.; Doye, J. P. K. *J. Phys. Chem. A* **1997**, *101*, 5111–5116.

(34) Doye, J. P. K.; Wales, D. J.; Miller, M. A. *J. Chem. Phys.* **1998**, *109*, 8143–8153.

(35) Frantsuzov, P. A.; Mandelshtam, V. A. *Phys. Rev. E* **2005**, *72*, 037102.

(36) Kone, A.; Kofke, D. A. *J. Chem. Phys.* **2005**, *122*, 206101.

(37) Doll, J. D.; Plattner, N.; Freeman, D. L.; Liu, Y.; Dupuis, P. *J. Chem. Phys.* **2012**, *137*, 204112.

(38) Flyvbjerg, H.; Petersen, H. *J. Chem. Phys.* **1989**, *91*, 461–466.

(39) Li, Z.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 6611–6615.

(40) Wales, D. J.; Scheraga, H. A. *Science* **1999**, *285*, 1368–1372.