Low-Temperature Structural Transitions: Circumventing the Broken-Ergodicity Problem

Vladimir A. Sharapov, Dario Meluzzi, and Vladimir A. Mandelshtam

Chemistry Department, University of California at Irvine, Irvine, California 92697, USA
(Received 1 December 2006; published 8 March 2007)

We consider systems undergoing very-low-temperature solid-solid transitions, exhibiting the well-known ''broken-ergodicity'' problem that is often so severe that even the replica exchange method converges too slowly. We propose an improvement of the latter, which consists of coupling the lower-temperature random walks to analytically generated random walks corresponding to an auxiliary harmonic superposition system. Numerically accurate results are obtained for several Lennard-Jones clusters, which have so far been treated only by the harmonic superposition approximation.

In the present Letter we are concerned with the so-called ''solid-solid'' structural transformations that generally occur in clusters, such as van der Waals clusters. The latter are often modeled by the Lennard-Jones (LJ) pair potential. Consequently, we will limit our discussion to the LJ clusters. However, both the methodology that we develop and our conclusions can possibly be extended to a broad range of problems, from protein folding to condensed-phase structural transformations.

A systematic analysis of LJ clusters reveals that generic properties, such as the melting temperature or temperature of surface reconstruction (see, e.g., Refs. [1–3]), change monotonically with size. However, the rich size-specific behavior of LJ clusters arises at low temperatures because the global energy minimum is highly sensitive to cluster size. Well-known examples of this behavior include clusters with sizes $n = 38$, 75–77, 98, and 102–104 [2,4–11]. In each of these cases, the highly symmetric nonicosahedral global minimum belongs to a narrow funnel of the potential energy surface (see the overview [12] describing a ''potential energy funnel''). A high potential barrier separates this narrow funnel from an icosahedral funnel. Albeit energetically less favorable, the low-energy icosahedral structures have higher vibrational entropy and, consequently, become thermodynamically more favorable at some finite, but usually low temperatures [2,8]. The corresponding structural transformation is then characterized as a ''solid-solid'' transition. Accurate numerical simulations of such phenomena are extremely challenging, and have become possible only recently due both to the efficiency of the replica exchange method (REM) [13,14] and to gradually increasing computer power, which permits extremely long Monte Carlo (MC) runs. Selected examples of such calculations can be found in Refs. [1,3,9,10,15]. However, despite some recent successes, it is not hard to find a case that causes the standard version of the REM to display the ''broken-ergodicity'' behavior.

In the standard REM framework, the sampling efficiency at the lowest temperature, T_{min}, depends on exchanges with higher-temperature random walks, which can eventually exchange with the one that can sample the whole configuration space. On the other hand, when the temperature is too high the volume of the configuration space may become too large implying that the system may still need very long times to find the relevant funnel(s) that could be very narrow. Furthermore, to maintain equal exchange rates (e.g., 50%) between adjacent replicas, the required total number of replicas grows approximately as $K \sim T_{min}^{-1}$ [16]. This factor too can affect the correlation times significantly, since a very small value of T_{min} results in a large number of replicas.

Figure 1 gives an example of a system (the LJ_{31} cluster) that undergoes a low-temperature structural transformation from the Mackay global energy minimum to the lowest anti-Mackay minimum [3,17]. The heat capacity was calculated with a standard version of the REM. Although a rigorous correlation time analysis has never been performed for this case, the lower-temperature peak shows noticeable differences between the two independent calcu-
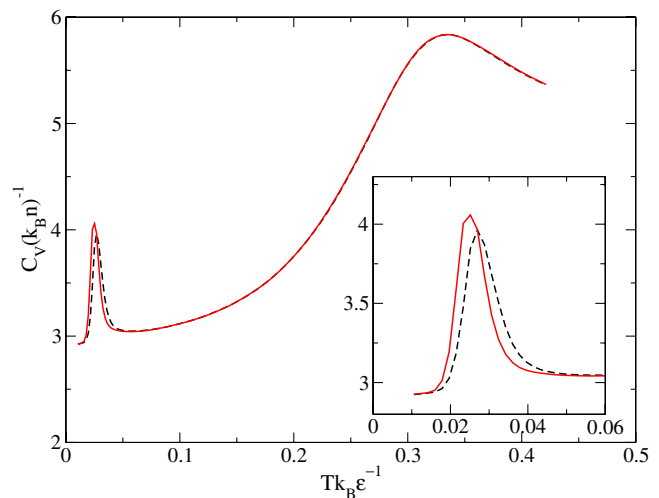

FIG. 1 (color online). Low-temperature heat capacity for the LJ_{31} cluster from two independent REM calculations, in the absence of exchanges with the proposed auxiliary harmonic superposition system. In each calculation, averaging was over 1.5×10^9 MC steps, but the convergence of the low-temperature peak is still not perfect.

105701-1

lations, each using as many as 1.5×10^9 MC steps, where we define an MC step as one attempt to move a single particle in all the Metropolis random walks. (In order to reduce the statistical errors to the thickness of the curve at least 10^{10} MC steps are needed.) We note also that the case of LJ_{31} is relatively easy, because both low-temperature structures are of the icosahedral type, thus the energy barrier separating those structures is not too high. The situation becomes much worse for LJ_{38} [9], where the random walks have to be an order of magnitude longer. We now believe that the results reported in Ref. [10] on the decahedral-icosahedral low-temperature transition in LJ clusters with $n = 75–77$ are not well converged even after about 10^{11} MC steps, which is around the practical limit of currently available computers. For even more pathological cases ($n = 98, 102–104$), it is unlikely that the standard version of the REM can ever converge in the region of the solid-solid transitions.

Since the REM was introduced, numerous modifications and improvements of this method have been suggested. However, in the present Letter we cannot afford an adequate review of the relevant literature. Our proposal to overcome the broken-ergodicity problem, still within the REM framework, is to add to the set of conventional Metropolis random walks an auxiliary random walk that can switch often between two or more funnels. For a given set of energy minima, we consider an effective potential arising from the superposition of harmonic potentials approximating the true potential in the vicinity of each minimum. Use of such potential allows us to generate independent, canonically distributed random points analytically, thus minimizing the correlation times in the auxiliary system. To the best of our knowledge, this latter feature of our method is new, while the idea of introducing an auxiliary system with better sampling properties and able to exchange its configurations with the system of interest is well known.

We consider a n-atom system described by the potential $V(r)$ using the mass-scaled Cartesian coordinates. Without loss of generality, for all configurations we assume that the center of mass is located at the origin of the coordinate system.

Let $\{r_\alpha\}$ be configurations of different energy minima that are separated by high barriers and are assumed to contribute to the properties of the system at low temperatures. Formulation of the present method would be straightforward if the energy minima were points in the configuration space. Unfortunately, for a cluster, the situation is complicated because $V(r)$ is invariant not only under translations, inversion, and permutations of atoms, but also under rotations of the cluster around its center of mass. Factoring out the translations, every minimum gives rise to $2\prod_k n_k!/h_\alpha$ of three-dimensional manifolds, where h_α is the order of the point group for the αth minimum, and n_k defines the number of atoms of the kth type. For each minimum, the Hessian matrix K_α has exactly six zero eigenvalues, so, strictly speaking, the $3n$-dimensional

Gaussian distribution arising from the harmonic approximation is not a meaningful approximation to the true low-temperature canonical distribution and is not even defined uniquely. In what follows, we give an unambiguous definition of the harmonic superposition system.

For the αth minimum, let Ω_α be a disk of sufficiently small radius δ that belongs to the purely vibrational hyperplane, i.e., the $(3n - 6)$-dimensional hyperplane orthogonal to all the translational and rotational eigenvectors of K_α:

$$r \in \Omega_\alpha, \quad \text{iff} \quad \| r - r_\alpha \| < \delta \quad \text{and} \quad r \perp \text{Ker} K_\alpha. \quad (1)$$

An unambiguous harmonic potential in the vicinity of the αth minimum can be defined for points $r \in \Omega_\alpha$,

$$V_\alpha(r) = E_\alpha + \tfrac{1}{2}(r - r_\alpha)^T K_\alpha(r - r_\alpha), \quad (2)$$

together with the normalized Gaussian distribution,

$$G_\alpha(r; \beta) = \frac{1}{Z_\alpha(\beta)} e^{-\beta V_\alpha(r)}. \quad (3)$$

Here the vibrational partition function for the αth minimum is

$$Z_\alpha(\beta) = e^{-\beta E_\alpha} \left(\frac{\beta}{2\pi} \prod_i \lambda_{\alpha i} \right)^{-1/2}$$

with the product taken over $(3n - 6)$ nonzero eigenvalues, $\lambda_{\alpha i}$, of K_α.

For a point r in the vicinity of one of the minima define its projection

$$r_\Omega := \mathcal{R} \mathcal{P} r \in \Omega_\alpha, \quad (4)$$

where $\mathcal{P}$ is an appropriate permutation of the atoms and $\mathcal{R}$, is a rotation (possibly including the inversion) of the cluster around the origin. (An efficient algorithm of finding such $\mathcal{P}$ and $\mathcal{R}$ will be described elsewhere.) We can now extend the definition of the harmonic potential to the subset that can be generated from Ω_α by the inversion and all possible rotations and permutations:

$$V_{\text{harm}}(r) := V_\alpha(r_\Omega). \quad (5)$$

Consequently, we also define the canonical distribution function for our harmonic superposition system [18],

$$G(r; \beta) \sim e^{-\beta V_{\text{harm}}(r)}. \quad (6)$$

At sufficiently large values of the inverse temperature $\beta = 1/k_B T$ Eq. (6) mimics the true Boltzmann distribution,

$$W(r; \beta) \sim e^{-\beta V(r)}. \quad (7)$$

Most importantly, random points distributed according to $G(r; \beta)$ can be generated analytically as the latter can be represented as a superposition of normalized Gaussian distributions. In order to correctly account for permutation/inversion and rotational symmetries we generate the random points only within the set $\Omega := \cup_\alpha \Omega_\alpha$ with the reduced distribution function

$$G_\Omega(r;\beta) = \sum_\alpha P_\alpha(\beta)G_\alpha(r;\beta), \qquad r \in \Omega, \qquad (8)$$

where the weights are

$$P_\alpha(\beta) = \frac{1}{Z(\beta)}\frac{Z_\alpha(\beta)}{h_\alpha}, \qquad Z(\beta) = \sum_\alpha \frac{Z_\alpha(\beta)}{h_\alpha}.$$

For sufficiently big β the Gaussians practically vanish outside the disks Ω_α, so we have the normalization conditions:

$$\int_{\Omega_\alpha} G_\alpha(r;\beta)dr = 1; \qquad \sum_\alpha P_\alpha(\beta) = 1. \qquad (9)$$

Moreover, the two probability distributions [Eqs. (6) and (8)] are related as

$$\frac{\int_{\mathbb{R}^{3n}} F(r)G(r;\beta)dr}{\int_{\mathbb{R}^{3n}} G(r;\beta)dr} = \frac{\int_\Omega F(r)G_\Omega(r;\beta)dr}{\int_\Omega G_\Omega(r;\beta)dr} \qquad (10)$$

for any function $F(r)$ that is invariant to translations, rotations, inversion, and permutations of identical atoms.

We now consider a set of coupled random walks running at different temperatures according to the standard REM scheme [13,14], which maintains the Boltzmann probability distribution (7) in each random walk. Such a scheme includes at least two types of moves, namely, a standard Metropolis move involving a single random walk, and a move involving exchanges of configurations between two random walks with similar temperatures. In addition to these two moves we introduce a *new move* that involves configuration exchange between a particular random walk and the auxiliary harmonic superposition system. That is, a random trial point $r' \in \Omega$ is generated analytically according to the probability distribution (8). The trial point is then accepted with probability

$$P(r \to r') = \min\left\{1, \frac{W(r';\beta)G(r;\beta)}{W(r;\beta)G(r';\beta)}\right\}$$
$$= \min\{1, e^{\beta[\Delta V(r')-\Delta V(r)]}\}, \qquad (11)$$

where $\Delta V(r) = V_{\text{harm}}(r) - V(r)$. Clearly, the number of minima in the auxiliary system can be greater than two and the new move can be utilized independently for different random walks running at different temperatures.

The new move (as well as the standard REM moves) satisfies the detailed balance condition, which ensures that the points in the random walks generated by such moves will be distributed according to $W(r;\beta)$. Furthermore, the new move is to be implemented only for particular random walk(s), for which the switching rate is significant. Exchanges of configurations between the coupled random walks maintain ergodicity of the system at all temperatures.

The sampling efficiency of the new move depends on how frequently the random walk switches from one minimum to another. The switching rate in turn depends on the following two factors. (i) The probabilities, $P_\alpha(\beta)$, for the trial point r' to appear in different minima. This factor can be optimized by adjusting the value of β. For example, for

equal statistical weights [$P_1(\beta) = P_2(\beta)$] for two selected minima correspond to the value [2]

$$\beta_{\text{harm}} = \frac{\frac{1}{2}\sum_i \ln(\lambda_{1i}/\lambda_{2i}) + \ln(h_1/h_2)}{E_2 - E_1}. \qquad (12)$$

(ii) How well the harmonic potential approximates the true potential for points sampled by the canonical distribution at $\beta = \beta_{\text{harm}}$. The acceptance probability (11) will have an appreciable value when $V(r)$ is not very different from $V_{\text{harm}}(r)$. However, when the harmonic approximation is poor, regardless of the other factors, $P(r \to r')$ will be nearly zero. For example, the harmonic approximation is poor for the LJ$_{38}$ cluster, where the temperature of the octahedral-to-icosahedral solid-solid transition appears to be too high. To use the present method for such cases, the procedure must be modified, for example, by introducing additional swapping parameters. (The corresponding extensions of the method will be explored elsewhere.)

Expression (12) is known in the context of the harmonic superposition approximation (HSA) [2,18]. As long as the potential energy is well represented by its harmonic approximation for the corresponding energy range, Eq. (12) provides a good estimate for the temperature of the structural transformation from one minimum to the other. Consequently, the HSA becomes more accurate for cases with lower transition temperatures.

We carried out numerical tests on the LJ$_{31}$ cluster to verify that the heat capacity computed by the new method is indistinguishable, within the statistical errors, from that computed by the standard version of the REM (the comparison is not shown here). Although more than two energy minima contribute to the low-temperature peak, our auxiliary harmonic superposition system included only the Mackay global minimum and the lowest anti-Mackay minimum. However, the latter circumstance does not make our method approximate, while having the two minima still removes the numerical bottleneck associated with the slow switching rate between the two major funnels. Notably, for the case of LJ$_{31}$, the numerical gain due to the exchanges with the harmonic superposition system turns out to be between 2 and 3 orders of magnitude compared to the regular REM calculation. The present method is expected to yield even higher gains for more difficult cases.

Figure 3 shows heat capacities for several LJ clusters with nonicosahedral global minima in the low-temperature regime. The global minimum of LJ$_{98}$ has tetrahedral symmetry, while the global minima of LJ$_{102-104}$ have decahedral symmetry. The next-in-energy local minimum in each case is a Mackay icosahedron. So far the corresponding solid-solid structural transitions for these systems have been characterized only using the HSA [2]. In the present study, the replica temperatures were chosen to cover only the ranges around the low-temperature peak. This choice is possible because at sufficiently low temperatures only the two lowest energy minima make a noticeable contribution to the properties of the system. For each case we performed
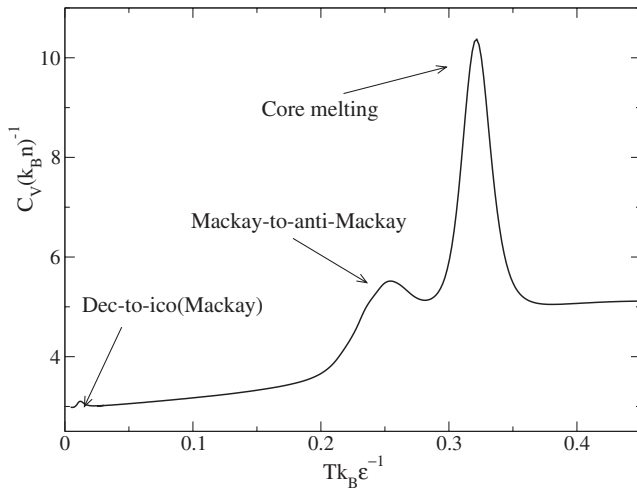
FIG. 2. The heat capacity of LJ_{102} obtained by pasting together the results from two independent calculations: one using the present method for temperatures $T < 0.03$, the other using the conventional REM for $T > 0.03$ [3]. The three peaks correspond to three distinct structural transformations.

about 10^9 MC steps per temperature. Most interestingly, as in the case of LJ_{31} [17], the HSA gives an excellent estimate of the heat capacity at low temperatures, which is now confirmed numerically.

At higher temperatures all clusters in this size range undergo two more structural transformations [3]: the Mackay-to-anti-Mackay surface reconstruction and core melting. This is shown in Fig. 2 for LJ_{102}, for which the heat capacity has two other peaks at higher temperatures.

In conclusion, we presented a substantial improvement of the REM to accurately describe very-low-temperature
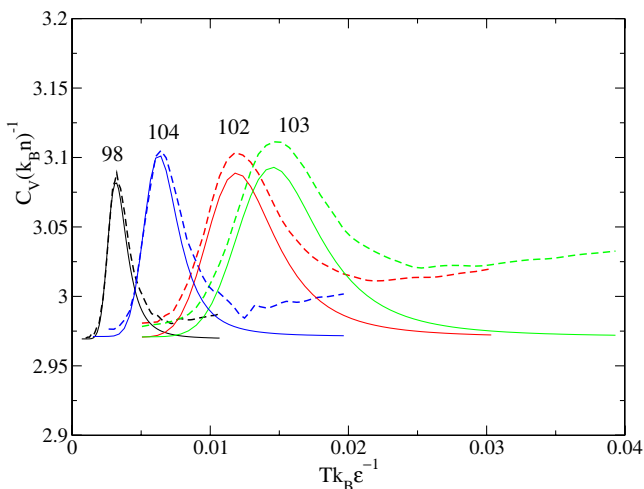

FIG. 3 (color online). The heat capacities of LJ_n clusters with $n = 98$, $102-104$, computed by the present method (dashed lines) and by the harmonic superposition method (solid lines). In each case the peak is due to the solid-solid structural transition from the global nonicosahedral energy minimum to the lowest icosahedral local minimum.

structural transformations in many-body systems and demonstrated its applicability to several cases of LJ clusters that seemed so far to be numerically intractable. A quantum version of our method can be implemented in the context of the variational Gaussian wave packet method [19]. In principle, the method can also be applied in a similar context to other systems, e.g., to describe solid-solid phase transitions in bulk materials or equilibrium properties of proteins involving different conformations. However, the requirement of an *a priori* knowledge of the minima representing the relevant funnels may make the method difficult to apply, especially in the cases corresponding to excessively large numbers of local minima involved.

[1] D. D. Frantz, J. Chem. Phys. **115**, 6136 (2001).
[2] J. P. K. Doye and F. Calvo, Phys. Rev. Lett. **86**, 3570 (2001); J. Chem. Phys. **116**, 8307 (2002).
[3] V. A. Mandelshtam and P. A. Frantsuzov, J. Chem. Phys. **124**, 204511 (2006).
[4] J. P. K. Doye, D. J. Wales, and R. S. Berry, J. Chem. Phys. **103**, 4234 (1995).
[5] J. Pillardy and L. Piela, J. Phys. Chem. **99**, 11805 (1995).
[6] J. P. K. Doye and D. J. Wales, Chem. Phys. Lett. **247**, 339 (1995).
[7] R. H. Leary and J. P. K. Doye, Phys. Rev. E **60**, R6320 (1999).
[8] J. P. K. Doye, M. A. Miller, and D. J. Wales, J. Chem. Phys. **110**, 6896 (1999).
[9] F. Calvo, J. P. Neirotti, D. L. Freeman, and J. D. Doll, J. Chem. Phys. **112**, 10340 (2000); **112**, 10350 (2000).
[10] V. A. Mandelshtam, P. A. Frantsuzov, and F. Calvo, J. Phys. Chem. A **110**, 5326 (2006).
[11] T. V. Bogdan, D. J. Wales, and F. Calvo, J. Chem. Phys. **124**, 044102 (2006).
[12] C. L. Brooks, III, J. N. Onuchic, and D. J. Wales, Science **293**, 612 (2001).
[13] C. J. Geyer, in *Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface*, edited by E. K. Keramidas (Interface Foundation, Fairfax Station, VA, 1991), p. 156.
[14] K. Hukushima and K. Nemoto, J. Phys. Soc. Jpn. **65**, 1604 (1996).
[15] E. G. Noya and J. P. K. Doye, J. Chem. Phys. **124**, 104503 (2006).
[16] C. Predescu, M. Predescu, and C. V. Ciobanu, J. Chem. Phys. **120**, 4119 (2004).
[17] F. Calvo and J. P. K. Doye, Phys. Rev. E **63**, 010902 (2000).
[18] D. Wales, Mol. Phys. **78**, 151 (1993).
[19] P. Frantsuzov, D. Meluzzi, and V. A. Mandelshtam, Phys. Rev. Lett. **96**, 113401 (2006).