JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

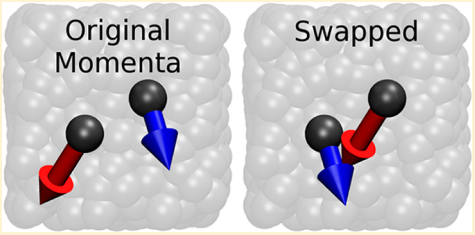
Easy Transition Path Sampling Methods: Flexible-Length Aimless Shooting and Permutation Shooting

Ryan Gotchy Mullen,[†] Joan-Emma Shea,[‡,§] and Baron Peters*[,†,‡]

[†]Department of Chemical Engineering, [‡]Department of Chemistry & Biochemistry, [§]Department of Physics, University of California, Santa Barbara, California 93106, United States

Supporting Information

ABSTRACT: We present new algorithms for conducting transition path sampling (TPS). Permutation shooting rigorously preserves the total energy and momentum of the initial trajectory and is simple to implement even for rigid water molecules. Versions of aimless shooting and permutation shooting that use flexible-length trajectories have simple acceptance criteria and are more computationally efficient than fixed-length versions. Flexible-length permutation shooting and inertial likelihood maximization are used to identify the reaction coordinate for vacancy migration in a two-dimensional trigonal crystal of Lennard-Jones particles. The optimized reaction coordinate eliminates nearly all recrossing of the transition state dividing surface.



I. INTRODUCTION

For rare events, the typical transition from one stable state (A) to another (B) is much shorter than the lifetime of A or B. Examples of rare events include chemical reactions,[1−3] molecular isomerizations[4,5] including protein folding,[6−10] protein aggregation,[7,11] nucleation in first-order phase transitions,[12,13] and the activated hops by which atoms diffuse in solids[14] and glasses.[15−17] Due to the separation of time scales inherent in rare events, a brute-force molecular dynamics (MD) simulation will include few, if any, transitions between stable states. Some methods overcome this sampling difficulty by introducing a potential energy bias along an *a priori* selected order parameter.[18−24] These methods facilitate the calculation of free energy surfaces and the exploration of configuration space. However, selecting an order parameter that preserves the barrier-crossing dynamics (i.e., the reaction coordinate) is notoriously difficult.

Transition path sampling (TPS) was developed to sample unbiased dynamical reactive trajectories and therefore enables rate constant calculations without the need for a reaction coordinate.[25] TPS has been used to investigate a wide variety of rare events, for example, protein folding,[26,27] conformational transitions,[28] enzyme catalysis,[29,30] nucleation,[31−35] amyloid fibril growth,[36] micelle budding,[37] ligand exchange,[38] reactions in solution,[39,40] rare transitions within Markov state models,[41] and glassy dynamics.[42,43]

The original TPS shooting move[44] generates trial trajectories from an existing transition path using a path based detailed balance criteria. A point on the transition path is randomly selected, the momenta are perturbed, and the trial trajectory is computed from the dynamical equations of motion. The size of the momenta perturbation can be changed to tune the acceptance probability of trial trajectories. As the perturbation decreases, the acceptance increases, with the trade off that trial trajectories diverge more slowly from the old transition path. However, even with the smallest possible change to the momenta, i.e., at the limit of machine precision, trial trajectories still diverge from the old one within picoseconds,[45] making the shooting move "best suited for the study of systems that relax...within the picosecond time scale".[46]

Several TPS methods have been developed to boost the acceptance probability of long trajectories.[46,47] Aimless shooting[48,49] maintains high acceptance and generates rapidly diverging trajectories with the highest diversity of shooting points near the stochastic separatrix. Since the momenta are drawn afresh from the Boltzmann distribution, aimless shooting is limited to the NVT ensemble and best suited for rare events with stochastic dynamics.

When the reaction coordinate is weakly coupled to other modes, transition paths are best sampled within a microcanonical framework. Chemical reactions that form or break covalent bonds have been successfully characterized by harmonic transition state theory, attesting that the bond breaking dynamics are (largely) independent of other atomic motion. Bond breaking has been simulated in TPS studies of proton hopping,[44] ligand exchange,[38] methanol coupling,[50] and peptide hydrolysis.[39] Even the breaking of a solvated ionic bond, which is part of a diffusional dissociation process, has inertial characteristics that are obscured by a stochastic thermostat.[40,51] The permutation shooting move that we present in this work rigorously preserves the kinetic energy and linear momentum, making it naturally suited for TPS with NVE dynamics.

Additionally, we formulate aimless shooting and permutation shooting for both fixed- and flexible-length TPS. In original

Received: January 14, 2015
Published: May 1, 2015

ACS Publications © 2015 American Chemical Society

DOI: 10.1021/acs.jctc.5b00032
J. Chem. Theory Comput. 2015, 11, 2421−2428

TPS, the length of all trajectories is chosen *a priori*. TPS will not give representative transition paths if the chosen length is too long or too short. The shooting move for flexible-length trajectories, introduced by van Erp et al. as part of transition interface sampling,[52] eliminated the trajectory length parameter by continuing the time propagation until a basin is entered. We will demonstrate that flexible-length versions of aimless shooting and permutation shooting are extremely simple to implement.

In Section II, we review acceptance criteria that follow detailed balance for original shooting and aimless shooting. In Section III, we develop acceptance criteria for permutation shooting and flexible-length TPS moves. In Section IV, we demonstrate flexible-length permutation shooting with a 2D model of vacancy diffusion.

II. BRIEF REVIEW OF EARLIER TPS METHODS

TPS starts from an initial trajectory that connects states A and B. Each trajectory is an ordered series of $N + 1$ time slices $\{\mathbf{x}_0, \mathbf{x}_{\Delta t}, \mathbf{x}_{2\Delta t}, ..., \mathbf{x}_{N\Delta t}\}$, where $\mathbf{x}$ is a phase space point consisting of configuration $\mathbf{r}$ and momenta $\mathbf{p}$ and the subscript indicates the time along the trajectory. Microscopically, states A and B are defined using one or more order parameters. The order parameter(s) need to distinguish basins A and B and should not overlap. Basin definitions should be small so that entering trajectories are committed to that thermodynamic state. For the original fixed-length TPS algorithms, basins must also be large enough to include typical fluctuations within a stable state; otherwise, many trajectories will terminate outside the definitions and be spuriously rejected. A population operator $h_A(\mathbf{x})$ is defined such that for $\mathbf{x}$ in A, $h_A(\mathbf{x}) = 1$; otherwise, $h_A(\mathbf{x}) = 0$. $h_B(\mathbf{x})$ is treated similarly. The population operators, in turn, are used to define a transition path indicator for trajectory i, $h_{AB}(i) = h_A(\mathbf{x}_0^i)\, h_B(\mathbf{x}_{N\Delta t}^i)$.

Next, a new trajectory n is generated from the initial (old) transition path o. As in Monte Carlo importance sampling of configuration space, n is accepted in TPS according to its statistical weight. As long as n is generated by microscopically reversible dynamics, the probability for accepting n is given by the Metropolis criterion[53]

$$P_{acc}(o \to n) = h_{AB}(n) \min\left[1, \frac{\rho_{eq}(\mathbf{x}_\tau^n)\, P_{gen}(n \to \mathbf{x}_\tau^o)}{\rho_{eq}(\mathbf{x}_\tau^o)\, P_{gen}(o \to \mathbf{x}_\tau^n)}\right] \quad (1)$$

where the shooting point is at time τ. In eq 1, all probabilities related to time propagation along the trajectories have already canceled as outlined in several earlier works.[25,54,55] Equation 1 contains only the residual factors pertaining to the probability of choosing new and old shooting points from the old and new trajectories, respectively. $\rho_{eq}(\mathbf{x})$ is the equilibrium distribution for phase space point $\mathbf{x}$. In the microcanonical ensemble

$$\rho_{eq}(\mathbf{x}) = \delta[H(\mathbf{x}) - E]/\Omega \quad (2)$$

where H is the Hamiltonian, E is the total energy, δ is the Dirac delta function, and Ω is the density of states at energy E. In the canonical ensemble, the equilibrium probability is the Boltzmann distribution

$$\rho_{eq}(\mathbf{x}) = \exp(-\beta H(\mathbf{x}))/Z \quad (3)$$

where Z is the partition function and β is the inverse temperature $1/k_B T$. $P_{gen}(o \to \mathbf{x}_\tau^n)$ is the probability of generating the new shooting point $\mathbf{x}_\tau^n$ from the old path o and depends on the details of how the time slice τ is selected on the old trajectory and how the momenta are altered.

II.A. Original Shooting. All $N + 1$ time slices are potential shooting points. The probability of choosing one is $1/(N + 1)$ and is the same for $o \to \mathbf{x}_\tau^n$ and $n \to \mathbf{x}_\tau^o$. The momenta $\mathbf{p}_\tau^o$ are perturbed by adding a random displacement vector $\delta\mathbf{p}$. The components of the displacement vector $\delta\mathbf{p}$ are sampled from a symmetric distribution with the constraint $\sum_i \delta p_i = 0$, such that the probability of choosing $\delta\mathbf{p}$ is equal to that of choosing $-\delta\mathbf{p}$.[44] The generation probability is symmetric

$$P_{gen}(n \to \mathbf{x}_\tau^o) = P_{gen}(o \to \mathbf{x}_\tau^n) \quad (4)$$

Additional constraints, e.g., to preserve angular momentum in systems without periodic boundary conditions or to enforce rigid bonds, require a complicated Gram−Schmidt procedure that, in practice, limits the utility of original shooting.

For microcanonical TPS, the sum $\mathbf{p}_\tau^o + \delta\mathbf{p}$ is rescaled such that $K(\mathbf{p}_\tau^o + \delta\mathbf{p}) = K(\mathbf{p}_\tau^o)$, giving the new momentum $\mathbf{p}_\tau^n$. From eqs 1, 2, and 4, the new trajectory is accepted if it connects basins A and B

$$P_{acc}(o \to n) = h_{AB}(n) \quad (5)$$

For constant temperature TPS, the momenta are not rescaled. Combining eqs 1, 3, and 4, the acceptance criteria is

$$P_{acc}(o \to n) = h_{AB}(n) \min[1, \exp(\beta K(\mathbf{p}_\tau^o) - \beta K(\mathbf{p}_\tau^n))] \quad (6)$$

where the potential energy Boltzmann factor $\exp(-\beta U(\mathbf{r}_\tau^o))$ is identical for the old and new shooting points.

II.B. Aimless Shooting. In aimless shooting, the time $t = 0$ is shifted to a point near the transition region along the initial trajectory. Any point with an intermediate value of the committor $(0.1 < p_B < 0.9)$ will suffice to initiate the procedure. Thereafter, new trajectories are generated from the old trajectory by randomly selecting one of only two potential shooting points, $\mathbf{r}_{-\Delta t/2}^o$ or $\mathbf{r}_{\Delta t/2}^o$. The probability of choosing a shooting point is $^1/_2$ for both $o \to \mathbf{x}_\tau^n$ and $n \to \mathbf{x}_\tau^o$. The momenta $\mathbf{p}_\tau^n$ are sampled anew from the Boltzmann distribution. The generation probability is therefore

$$P_{gen}(o \to \mathbf{x}_\tau^n) = \exp(-\beta K(\mathbf{p}_\tau^n))/2Z \quad (7)$$

which cancels the kinetic energy component of the Boltzmann distribution from eq 3. From eqs 1, 3, and 7, the acceptance criteria for new paths is

$$P_{acc}(o \to n) = h_{AB}(n) \quad (8)$$

III. NEW TPS METHODS

III.A. Permutation Shooting. Permutation shooting is an easily implemented new microcanonical TPS algorithm. It begins like aimless shooting with an initial reactive trajectory for which the $t = 0$ configurations lies near the transition state region and with $\mathbf{x}_0^o$ or $\mathbf{x}_{\Delta t}^o$ not in A or B. Thereafter, new trajectories are generated from an old trajectory as follows:

(1) Randomly select one of the two potential shooting points on the old trajectory, e.g., either $\mathbf{x}_0^o$ or $\mathbf{x}_{\Delta t}^o$, to be the shooting point $\mathbf{x}_\tau^n$ for a new trajectory.

(2) $\mathbf{x}_\tau^n$ is one of two potential shooting points on the new trajectory. Randomly select whether the other potential shooting point will be on the backward (at $t = \tau - \Delta t$) or forward (at $t = \tau + \Delta t$) half-trajectory.

DOI: 10.1021/acs.jctc.5b00032
J. Chem. Theory Comput. 2015, 11, 2421−2428

(3) Starting from $\mathbf{p}_\tau^o$, select at random two particles with the same mass and swap their momenta, as illustrated in Figure 1. Repeat i times.

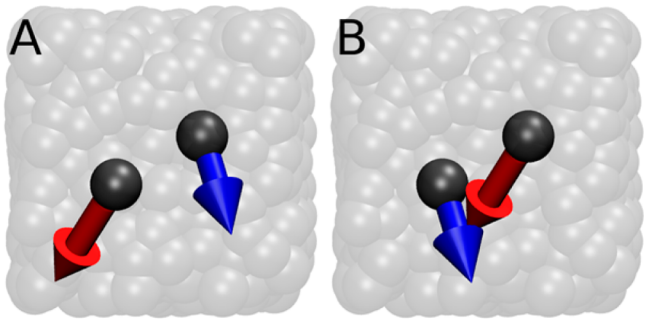


Figure 1. Snapshot of a configuration. In permutation shooting, two particles with the same mass are selected at random. In (A), the original momenta of each particle is indicated by an arrow. In (B), the momenta have been swapped.

(4) Dynamically propagate the system forward using the new momentum $\mathbf{p}_\tau^n$ and backward in time using the reversed momentum $-\mathbf{p}_\tau^n$.

(5) Accept the new trajectory if it joins the reactant and product states, A and B. Otherwise, reject the new trajectory and increase the statistical weight of the old trajectory. As done in Monte Carlo simulations, the statistical weight of the old trajectory is incremented by one observation each time that a new trajectory is rejected.

Step 5 reflects that permutation shooting obeys eqs 2 and 4 and therefore has the same acceptance criteria as microcanonical original shooting and canonical aimless shooting

$$P_{\mathrm{acc}}(o \to n) = h_{\mathrm{AB}}(n) \tag{9}$$

However, permutation shooting rigorously preserves the energy E and total momentum and therefore the corrective procedures in the original shooting move are eliminated. Additionally, the number of momenta swaps i in step 3 is an adjustable parameter that can be tuned to optimize P_{acc}. In this work, we use i equal to the number of atoms. Randomizing the time step of the alternate shooting point on the new trajectory in step 2 increases the robustness of the algorithm, as discussed in Supporting Information.

For monatomic particles and molecules with unconstrained bonds, the atomic momenta can be swapped directly in step 3. Atomic momenta cannot be swapped among molecules with rigid bonds. The constraints that maintain rigid bonds eliminate components of the momenta that would compress or stretch the bond and therefore alter the momenta. In original shooting, the momenta for molecules with rigid bonds are altered using Gram−Schmidt orthogonalization. In permutation shooting, the solution is much simpler: swap only the center-of-mass momenta. The center-of-mass momenta are easily computable, and none of the components affect bond distances.

III.B. Flexible-Length Trajectories. The length of the initial transition path $N\Delta t$ is held constant in original TPS. For long $N\Delta t$, most trajectories needlessly simulate additional time in the stable basins. If $N\Delta t$ is too short, a new trajectory will often terminate, either forward in time or backward in time, without having entered a stable basin. In this case, whether the shooting point $\mathbf{x}_\tau$ is part of a reactive or a nonreactive trajectory remains inconclusive, and the transition path ensemble (TPE) will be biased toward shorter trajectories. In a flexible-length trajectory, $\mathbf{x}_\tau$ is propagated in time until the trajectory enters either basin A or B. Each trajectory is just as long as needed to cross the barrier and no longer. Accordingly, the basin definitions should be slightly more restrictive than in fixed-length TPS to ensure the trajectories are committed to the basin upon entry.

Sampling with flexible-length trajectories is especially efficient when the reaction coordinate dynamics are stochastic because some transition paths can be significantly longer than the mean. Juraszek et al. used flexible-length original shooting to study conformational changes of the Trp-cage protein.[56] The TPE for Trp-cage folding exhibits a mean folding time $\langle N\Delta t\rangle$ = 2 ns but also includes 10 ns trajectories. If fixed-length TPS were used, all trajectories would have needed to be 10 ns long in order to accurately sample the TPE.

The method of selecting a shooting point changes the Metropolis criterion for flexible-length TPS. For original shooting, the probability of choosing time slice τ is $1/(N+1)$ because every timeslice is a potential shooting point. Since the new and old paths will not be the same length, the acceptance criterion for microcanonical original shooting with flexible-length trajectories is

$$P_{\mathrm{acc}}(o \to n) = h_{\mathrm{AB}}(n)\,\min\left[1, \frac{N^o+1}{N^n+1}\right] \tag{10}$$

For canonical original shooting with flexible-length trajectories

$$P_{\mathrm{acc}}(o \to n) = h_{\mathrm{AB}}(n)\,\min\left[1, \exp(\beta K(\mathbf{p}_\tau^o) - \beta K(\mathbf{p}_\tau^n))\frac{N^o+1}{N^n+1}\right] \tag{11}$$

For aimless shooting and permutation shooting, the probability of choosing a shooting point is $^1/_2$ on both the old and new trajectories because each only has two potential shooting points regardless of its length. The acceptance criterion for microcanonical permutation shooting with flexible-length trajectories is

$$P_{\mathrm{acc}}(o \to n) = h_{\mathrm{AB}}(n) \tag{12}$$

For canonical aimless shooting with flexible-length trajectories

$$P_{\mathrm{acc}}(o \to n) = h_{\mathrm{AB}}(n) \tag{13}$$

Thus, the acceptance rule for aimless shooting and permutation shooting is the same whether the transition paths are flexible-length or fixed-length.

IV. EXAMPLE: VACANCY HOPPING

Diffusion in single crystal domains occurs as atoms jump through interstitial sites, vacancies, or a combination of the two. Vacancy assisted diffusion is important for dopant infusion in silicon,[57−59] mass transfer in gas hydrates,[60−62] growth of nanostructured materials,[63,64] oxygen migration in TiO_2 catalysts,[65] radiation damage and creep in steels,[66,67] and many other processes.

At temperatures approaching the melting point, vacancy diffusion is a complex process that may involve chains of correlated jump events, vacancy−vacancy interactions, or collective lattice reorientation. At low temperatures, vacancy diffusion is a rare event occurring only when a nearest neighbor

DOI: 10.1021/acs.jctc.5b00032
J. Chem. Theory Comput. 2015, 11, 2421−2428

particle hops into the vacant lattice site. In this example, we restrict our work to the low-temperature regime.

IV.A. Model. We applied flexible-length permutation shooting to study vacancy migration in a two-dimensional trigonal crystal of Lennard-Jones (LJ) particles. The 2D crystal is easily visualized yet, as we will demonstrate, sufficiently complex that the mechanistically important variables are not obvious. Our system consists of 15 × 10 trigonal unit cells populated with 299 identical LJ particles of mass m, leaving one vacant lattice site. Neighboring lattice sites are spaced $2^{1/6}\sigma$ apart, which corresponds to the minimum energy $-\varepsilon$ of the Lennard-Jones potential. Dynamics were simulated using LAMMPS[68] with a $0.001\tau_{LJ}$ time step, where $\tau_{LJ} = (m\sigma^2/\varepsilon)^{1/2}$, and periodic boundary conditions.

A vacancy is surrounded by six particles, which are each equally likely to move into the vacancy. We removed this 6-fold degeneracy by tagging one particle H and studying only trajectories or configurations for which H is the hopping particle. Before a hopping event, H is located at the donor site D and is surrounded by five neighboring particles, which we numbered 1−5. After a hopping event, H is at the acceptor site A and still has five neighbors, particles 4 and 5 plus three other particles that we numbered 6−8 (Figure 2).

To locate the vacancy before and after a jump, we used a scheme similar to Geslin et al.[69] (see also ref 62). A Weeks−Chandler−Andersen (WCA)[70] probe particle with the same σ and ε that governs the LJ interactions was used to sample the potential field generated by each crystal configuration. This particle is sufficiently large that the global minima located at the unoccupied lattice site is easily distinguishable from the local minima between occupied lattice sites. At low densities or in a crystal of hard-sphere particles, the two different minima may be similar in depth. Geslin et al. locate the vacancy even in these cases by transforming the classical potential field into a pseudoquantum ground state density. For the present case, it is sufficient to locate the density at the global minima of the classical potential field. We enforce the restriction that H is the only hopping particle by rejecting any configurations for which the vacancy is surrounded by particles other than {H,4−8} or {H,1−5}.

To track the progress of the hopping event, we need to locate D and A, even when H is at neither one. In our first attempt, we define D as the minimum energy position of H considering only the interactions of H with particles 1−3 (the "D anchors") fixed in their current positions. A is similarly defined with respect to particles 6−8, the "A anchors" (Figure 2B). A hopping event therefore begins with H fluctuating around D and ends with H fluctuating around A. Our results will show that this energy minimization method for locating D and A is not sufficiently accurate for computing trial reaction coordinate velocities. However, it does distinguish prehop, hopping, and posthop configurations, which is the only way we use it here.

IV.B. Trial Order Parameters. We evaluate two order parameters that have been used to describe vacancy diffusion. The first order parameter we consider was formulated by Bennett to study vacancy diffusion in a face-centered cubic (fcc) crystal of LJ particles.[71] Bennett recognized that the hopping particle must pass through a gate composed of other nearest neighbor particles. He proposed an order parameter that tracks the position of the hopping particle relative to the center of mass of the gate particles

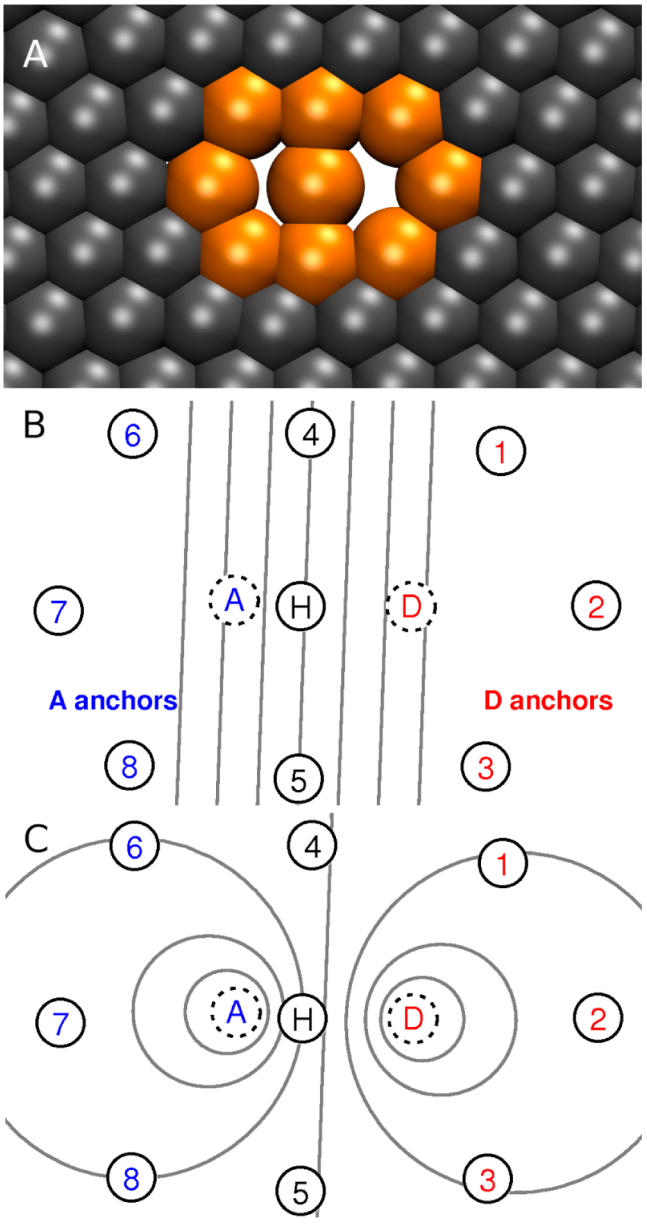


Figure 2. (A) Snapshot of a configuration near the transition state. The hopping particle and its eight neighbors are highlighted orange. (B) Diagram showing the highlighted particles, donor site D, acceptor site A, and isosurfaces of order parameter ξ (see text). Particle H lies approximately on isosurface $\xi = 0$, and isosurfaces are shown at intervals of 0.25. (C) Same configuration with isosurfaces of coordinate q (see text). H lies on isosurface $q = 0.5$, and isosurfaces are shown in intervals of 0.5.

$$\xi = \left(\mathbf{r}_H - \frac{1}{N_g}\sum_i \mathbf{r}_i\right)\cdot\hat{\mathbf{r}}_{AD} \tag{14}$$

where $\mathbf{r}_H$ is the position of the hopping particle, the N_g gate particles are at positions $\mathbf{r}_i$, and $\hat{\mathbf{r}}_{AD}$ is the unit vector in the direction from the D to A. The $\xi = 0$ isosurface is the putative dividing surface separating prehop ($\xi < 0$) and posthop ($\xi > 0$) configurations.

The second order parameter was used by Peters et al. to study methane hopping between clathrate cages in hydrates.[62] The bipolar coordinates

DOI: 10.1021/acs.jctc.5b00032
J. Chem. Theory Comput. 2015, 11, 2421−2428

$$q = \ln\left(\frac{r_{HD}}{r_{HA}}\right) \tag{15}$$

where r_{HD} and r_{HA} are the distances from H to D and from H to A, respectively, and the angle

$$\theta = \angle DHA \tag{16}$$

describes the position of the particle hopping between donor and acceptor cages. The $q = 0$ and $\xi = 0$ isosurfaces shown in Figure 2B,C are nearly, but not exactly, identical. The $\xi = 0$ isosurface depends on A, D, and separately on the gate atom positions. The $q = 0$ isosurface depends only on the A and D positions.

IV.C. Methods. We performed $N_R = 6000$ permutation shooting moves at an energy $E = -826.5\epsilon$. The two candidate shooting points were separated by $\Delta t = 25$ time steps. A new shooting point momenta $\mathbf{p}^n$ was generated by swapping $i = 299$ momenta pairs between particles. On the basis of previous applications of q, the symmetry of q, and on typical unbiased values of q, we defined the donor basin as $q < -3$ and the acceptor basin is $q > 3$. The transition path ensemble consists of 1638 unique trajectories with an average temperature $\langle T \rangle = 0.199\epsilon/k_B$ and an average trajectory length $\langle N\Delta t \rangle = 1.26\tau_{LJ}$ with a standard deviation of $0.30\tau_{LJ}$.

The order parameters ξ and q were tested by inertial likelihood maximization[72] (iLMax) to determine the best reaction coordinate. iLMax optimizes trial coordinates to predict the committor p_B and give a high transmission coefficient κ. p_B is the probability that a trajectory initiated from a configuration with Boltzmann distributed momenta will end in the product state. A high κ, approaching the classical upper limit $\kappa = 1$, indicates that recrossing of the transition state dividing surface has been effectively eliminated.

Trial coordinates have the form $q_T(z) = c_0 + c_1 z$, where z is either ξ or q and coefficients c_0 and c_1 are adjustable parameters. The trial coordinate is added to the velocity $\dot{q}_T(z)$ in the reaction probability

$$p_{RX} = \frac{1 + \mathrm{erf}[q_T(z) + c_V \dot{q}_T(z)]}{2} \tag{17}$$

z is computed for each shooting point configuration ($t = 0$). The velocities $\dot{z}$ are computed by finite difference using the shooting point configuration and the configuration $t = 1$ time step. The coefficients c_0, c_1, and c_V are adjusted to maximize the likelihood L

$$L = \prod_{\mathbf{x}\in B} p_{RX}(\mathbf{x}) \prod_{\mathbf{x}\in A} 1 - p_{RX}(\mathbf{x}) \tag{18}$$

where the first product is over shooting points that end in B and the second product is over points that end in A. The best reaction coordinate is indicated by the highest L. Adding a variable to the trial coordinate is only significant if the log-likelihood increases by several increments of a Bayesian criterion $\delta L_{min} = {}^1/_2 \ln(N_R)$, where N_R is the number of trajectories ($N_R = 6000$ in this work).

The free energy was computed for $q \geq 0$ by equilibrium path sampling (EPS,[62] also known as BOLAS[73]) at temperature $T = 0.2\epsilon/k_B$. Temperature was controlled by a Nosé−Hoover thermostat[74,75] using a damping time of 10 time steps. Since q is symmetric, the free energy $q \leq 0$ will be identical. Along the barrier, $0.0 \leq q \leq 2.3$, we used 9 EPS windows each $q_{i,max} - q_{i,min} = 0.30$ wide and overlapping by $q_{i+1,min} - q_{i,max} = 0.05$. We harvested 50 000 trajectories within each window. For the window at the barrier top, each trajectory was 200 timesteps long. On the downward slope of the barrier, $0.25 \leq q \leq 2.25$, trajectories were shortened to 100 time steps. The free energy in the acceptor basin was calculated using a 10th EPS window, $2.25 \leq q \leq 3.80$, using 100 000 trajectories 1000 time steps long. In every case, 11 equally spaced configurations were saved from each trajectory. The resulting free energy $F(q)$ was shifted vertically so that $\int_{-6}^{0} dq\, e^{-\beta F(q)} = 1$.

Likelihood maximization results were validated by computing κ from the reactive flux correlation function[76] for each trial coordinate. An ensemble of 1000 configurations from the dividing surface was randomly selected from the EPS trajectories at the barrier top. Momenta were drawn from the Boltzmann distribution at $T = 0.2\epsilon/k_B$. Forward and back trajectory pairs were computed for $1\tau_{LJ}$.

IV.D. Results. Table 1 shows the log-likelihood scores for selected trial coordinates. Log-likelihood scores are reported relative to the value of $\ln L[q]$ in increments of δL_{min}, i.e., $\Delta\ln L[q_T] = (\ln L[q_T] + 2129)/4.35$. The preliminary score $\Delta\ln L[\xi] = 178$ shows that ξ is better than q as a reaction coordinate.

Table 1. Inertial Likelihood Maximization and Transmission Coefficient Results[a]

reaction coordinate	c_V	$\Delta\ln L$	κ
D and A Located by Energy Minimization			
$0.110 + 8.4\ \xi$	0.16	178	0.80
$0.026 + 1.2\ q$	0.040	0[b]	0.45
D and A Located by Linear Combination of Surrounding Atoms			
$0.140 + 11\ \xi_{LC}$	0.18	231	0.81
$0.064 + 4.0\ q_{LC}$	0.17	310	0.93

[a]The best reaction coordinate is indicated by the highest log-likelihood, $\Delta\ln L$ (see text). [b]$\ln L[q] = -2129$.

To understand why ξ is better than q, we compare eqs 14 and 15. Both ξ and q depend on the coordinate $\mathbf{r}_H$. q is also highly sensitive to the location of D and A, whereas ξ depends on the direction from D to A. We compute velocities from configurations one time step apart. The displacement of H over one time step, averaged over the $N_R = 6000$ shooting points, is $\langle|\Delta\mathbf{r}_H|\rangle = 6.4 \times 10^{-4}\sigma$. By contrast, the average displacements of D and A are two to three times larger, $\langle|\Delta\mathbf{r}_D|\rangle = 1.5 \times 10^{-3}\sigma$ and $\langle|\Delta\mathbf{r}_A|\rangle = 1.9 \times 10^{-3}\sigma$. Large fluctuations in $\mathbf{r}_D$ and $\mathbf{r}_A$ cause large fluctuations in $\dot{q}$, which, in turn, obscure any information that would otherwise be conveyed by the velocity. One plausible explanation for the large fluctuations in $\mathbf{r}_D$ and $\mathbf{r}_A$ is that the potential energy used to locate D and A could have a broad minimum with very small second derivatives or several local minima nearly degenerate in energy. For both situations, a small perturbation of particles 1, 2, 3, 6, 7, or 8 would lead to a large shift in the optimum locations of D or A.

As an alternative, we locate D and A using lattice vectors from the positions of nearby particles. Any given particle fluctuates randomly around its lattice site, so the accuracy of estimating $\mathbf{r}_A$ and $\mathbf{r}_D$ will increase with the number of reference particles. We compute the lattice sites D and A from a linear combination of the current positions of the six anchor particles

$$\mathbf{r}_D = (3\mathbf{r}_1 + 2\mathbf{r}_2 + 3\mathbf{r}_3 + \mathbf{r}_6 + \mathbf{r}_7 + \mathbf{r}_8)/11 \tag{19}$$

and

DOI: 10.1021/acs.jctc.5b00032
J. Chem. Theory Comput. 2015, 11, 2421−2428

$$\mathbf{r}_A = (\mathbf{r}_1 + \mathbf{r}_2 + \mathbf{r}_3 + 3\mathbf{r}_6 + 2\mathbf{r}_7 + 3\mathbf{r}_8)/11 \quad (20)$$

Using eqs 19 and 20, the average displacements $\langle|\Delta\mathbf{r}_D|\rangle = 2.5 \times 10^{-4}\sigma$ and $\langle|\Delta\mathbf{r}_A|\rangle = 2.5 \times 10^{-4}\sigma$ are comparable to $\langle|\Delta\mathbf{r}_H|\rangle$. Trial coordinates computed using eqs 19 and 20 have the subscript LC. iLMax scores for ξ_{LC} and q_{LC} are both improvements over ξ and q, respectively, with the final result that q_{LC} is the top coordinate.

With the optimized reaction coordinate q_{LC}, it is now straightforward to compute the rate

$$k = \kappa[q_{LC}]\frac{\langle|\dot{q}_{LC}|\rangle_{\ddagger}}{2} e^{-\beta\Delta F(q_{LC\ddagger})} \quad (21)$$

where $\kappa[q_{LC}]$ is the transmission coefficient, $\langle|\dot{q}_{LC}|\rangle_{\ddagger}$ is the equilibrium average velocity of q_{LC} constrained to the dividing surface $q_{LC}(\mathbf{x}) = q_{LC\ddagger}$, and $\Delta F(q_{LC\ddagger})$ is the free energy difference between the ensemble of transition states and the ensemble of reactants. The free energy $F(q_{LC})$ (not shown) was estimated by reweighting the EPS data and exhibits only small deviations from $F(q)$. $F(q)$ shows a barrier at $q = 0$ that is $\Delta F(q_{\ddagger}) = 11.0k_BT$ above the free energy of the reactant state ensemble (Figure 3). The prefactor $\kappa[q_{LC}]\langle|\dot{q}_{LC}|\rangle_{\ddagger}/2$ was computed from the reactive flux correlation function[76] to be $0.660/\tau_{LJ}$. The rate constant is $k = 1.10 \times 10^{-5}/\tau_{LJ}$.

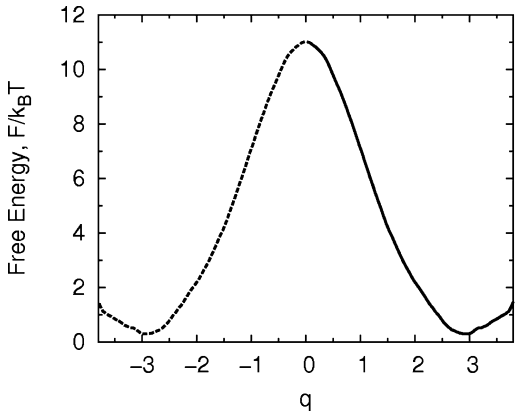


Figure 3. Free energy projected onto coordinate q. The dotted portion is the symmetric image of the computed $q > 0$ curve (solid). The integrated free energy of the reactant state $q \leq 0$ was set to zero.

Additionally, we computed κ for each trial coordinate from the normalized reactive flux correlation function (Table 1 and Figure 4). In each case, recrossing occurs during the first $0.5\tau_{LJ}$

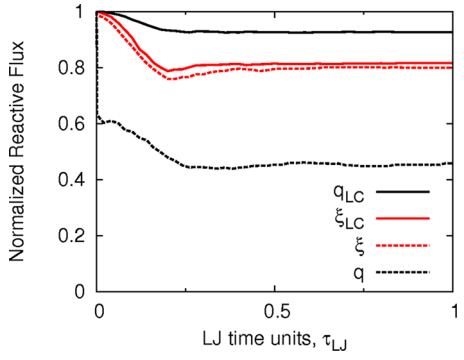


Figure 4. Normalized reactive flux correlation function for the 0-isosurface of each coordinate.

and thereafter the reactive flux plateaus. The highest κ is $\kappa[q_{LC}] = 0.93$, confirming the iLMax result that q_{LC} is the best coordinate. The lowest transmission coefficient is $\kappa[q] = 0.45$; q also scored the lowest $\Delta\ln L$. Most of the drop in $\kappa[q]$ occurs in the first $0.001\tau_{LJ}$, a consequence of the uncertainty in locating D and A by the energy minimization technique. The uncertainty in locating $q = 0$ matters most when the hopping particle H is very close to the dividing surface. The sharp drop is not physical, but rather it emerges from numerical noise in the optimization of D and A. Interestingly, $\kappa[\xi] = 0.80$ and $\kappa[\xi_{LC}] = 0.81$ are similar, highlighting that ξ is less sensitive to the way that D and A are determined. Coordinates of type ξ are less sensitive to numerical errors in the D and A optimization because the positions of the gate atoms define the origin of the ξ axis while the numerical optimization supplies the less sensitive D to A direction. Also note that the difference between $\kappa[q_{LC}]$ and $\kappa[q]$ is larger than the difference between $\kappa[q_{LC}]$ and $\kappa[\xi_{LC}]$, showing that the accuracy with which D and A are located has a greater impact than the form of the trial coordinate.

Our simulations were performed on a simple 2D Lennard-Jones crystal. However, a useful objective of our coordinate identification analyses is to discover generalizable coordinates. We anticipate that the degree to which $\kappa[q_{LC}]$ improves over $\kappa[\xi]$ (or, equivalently, $\kappa[\xi_{LC}]$) will be system-dependent. In particular, we note two studies that computed $\kappa[\xi]$ for vacancies in different 3D crystal lattices. For vacancy hopping in an fcc crystal of LJ particles, $\kappa[\xi] = 0.91$ at $k_BT/\epsilon = 0.505$.[77] The hopping particle passes through a gate composed of four nearest-neighbor particles. The largest void space between gate particles lies on the straight line path from D to A, suggesting that most paths are well-described by the linear ξ coordinate. There may be a slight advantage of using coordinate q, but, for all practical purposes, ξ is accurate and even the $\kappa[\xi]$ correction to TST could be ignored. For oxygen ion vacancy hopping in a 3D cobalt oxide rock-salt lattice, $\kappa[\xi] \approx 0.3$ over a wide range of temperatures, 995−1760 K.[78] The largest aperture between gate ions is not on the straight line D to A path described by ξ but, rather, on a curved path that may be more accurately described by the curvilinear coordinate q. These observations suggest that transition state theory models for vacancy hopping in the rock-salt system and perhaps other 3D crystals might be improved with the curvilinear coordinate q and the methods in this work.

V. CONCLUSIONS

Aimless and permutation shooting are efficient, simple algorithms to sample the transition paths in the canonical (NVT) and microcanonical (NVE) ensembles, respectively. Permutation shooting rigorously preserves the kinetic energy and total momentum of the shooting point momenta and is simple to implement even for rigid water models. Because each method only uses two potential shooting points per trajectory, the Metropolis acceptance criteria are identical regardless of whether the sampled transition paths are all the same length.

Permutation shooting was applied to vacancy migration in a two-dimensional Lennard-Jones crystal. The permutation shooting data and inertial likelihood maximization were used to examine two previously proposed reaction coordinates for vacancy hopping. Both coordinates require a procedure to identify the donor (D) and acceptor (A) lattice sites on-the-fly during the hopping event. Interestingly, transmission coefficients from the better (bipolar) coordinate are highly sensitive

DOI: 10.1021/acs.jctc.5b00032
J. Chem. Theory Comput. 2015, 11, 2421−2428

to numerical errors in a commonly used procedure for obtaining D and A. Our proposed scheme for locating D and A avoids the numerical optimization and yields transmission coefficients (in the two-dimensional Lennard-Jones crystal) of 0.93.

■ ASSOCIATED CONTENT

Supporting Information

Importance of randomly selecting potential shooting points. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.5b00032.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: baronp@engineering.ucsb.edu.

Funding

R.G.M. thanks the National Science Foundation (NSF) for a graduate research fellowship, no. DGE-1144085. B.P. thanks the NSF for support from CAREER award no. 0955502. J.E.S. thanks the NSF for support from grant no. MCB-1158577 and the David and Lucile Packard Foundation. We acknowledge support from the Center for Scientific Computing from the CNSI, MRL: NSF MRSEC (DMR-1121053) and NSF CNS-0960316.

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Klippenstein, S. J.; Pande, V. S.; Truhlar, D. G. *J. Am. Chem. Soc.* **2014**, *136*, 528−546.
(2) Broadbelt, L. J.; Pfaendtner, J. *AIChE J.* **2005**, *51*, 2112−2121.
(3) Santiso, E. E.; Gubbins, K. E. *Mol. Simul.* **2004**, *30*, 699−748.
(4) Kuharski, R. A.; Chandler, D.; Montgomery, J. A., Jr.; Rabii, F.; Singer, S. J. *J. Phys. Chem.* **1988**, *92*, 3261−3267.
(5) Bolhuis, P. G.; Dellago, C.; Chandler, D. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5877−5882.
(6) Shea, J.-E.; Brooks, C. L., III *Annu. Rev. Phys. Chem.* **2001**, *52*, 499−535.
(7) Dobson, C. M. *Nature* **2003**, *426*, 884−890.
(8) Dill, K. A.; Ozkan, S. B.; Shell, M. S.; Weikl, T. R. *Annu. Rev. Biophys.* **2008**, *37*, 289.
(9) Englander, S. W.; Mayne, L. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 15873−15880.
(10) Bryngelson, J. D.; Onuchic, J. N.; Socci, N. D.; Wolynes, P. G. *Proteins: Struct., Funct., Bioinf.* **1995**, *21*, 167−195.
(11) Straub, J. E.; Thirumalai, D. *Annu. Rev. Phys. Chem.* **2011**, *62*, 437−463.
(12) Kashchiev, D.; Van Rosmalen, G. *Cryst. Res. Technol.* **2003**, *38*, 555−574.
(13) Agarwal, V.; Peters, B. *Adv. Chem. Phys.* **2014**, *155*, 97−160.
(14) *Diffusion in Solids: Recent Developments*, 1st ed.; Nowick, A. S., Burton, J. J., Eds.; Academic Press, Inc.: New York, 1975.
(15) Angell, C. A.; Ngai, K. L.; McKenna, G. B.; McMillan, P. F.; Martin, S. W. *J. Appl. Phys.* **2000**, *88*, 3113−3157.
(16) Debenedetti, P. G.; Stillinger, F. H. *Nature* **2001**, *410*, 259−267.
(17) Berthier, L.; Biroli, G. *Rev. Mod. Phys.* **2011**, *83*, 587.
(18) Carter, E.; Ciccotti, G.; Hynes, J. T.; Kapral, R. *Chem. Phys. Lett.* **1989**, *156*, 472−477.
(19) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187−199.
(20) Laio, A.; Parrinello, M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12562−12566.
(21) Maragliano, L.; Vanden-Eijnden, E. *Chem. Phys. Lett.* **2006**, *426*, 168−175.
(22) Darve, E.; Pohorille, A. *J. Chem. Phys.* **2001**, *115*, 9169−9183.
(23) Kong, X.; Brooks, C. L., III *J. Chem. Phys.* **1996**, *105*, 2414−2423.
(24) Rosso, L.; Minary, P.; Zhu, Z.; Tuckerman, M. E. *J. Chem. Phys.* **2002**, *116*, 4389−4402.
(25) Bolhuis, P. G.; Chandler, D.; Dellago, C.; Geissler, P. L. *Annu. Rev. Phys. Chem.* **2002**, *53*, 291−318.
(26) Bolhuis, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12129−12134.
(27) Juraszek, J.; Bolhuis, P. G. *Biophys. J.* **2008**, *95*, 4246−4257.
(28) Dimelow, R. J.; Bryce, R. A.; Masters, A. J.; Hillier, I. H.; Burton, N. A. *J. Chem. Phys.* **2006**, *124*, 114113.
(29) Knott, B. C.; Haddad Momeni, M.; Crowley, M. F.; Mackenzie, L. F.; Götz, A. W.; Sandgren, M.; Withers, S. G.; Ståhlberg, J.; Beckham, G. T. *J. Am. Chem. Soc.* **2013**, *136*, 321−329.
(30) Basner, J. E.; Schwartz, S. D. *J. Am. Chem. Soc.* **2005**, *127*, 13822−13831.
(31) Moroni, D.; Ten Wolde, P. R.; Bolhuis, P. G. *Phys. Rev. Lett.* **2005**, *94*, 235703.
(32) Beckham, G. T.; Peters, B. *J. Phys. Chem. Lett.* **2011**, *2*, 1133−1138.
(33) Lechner, W.; Dellago, C.; Bolhuis, P. G. *Phys. Rev. Lett.* **2011**, *106*, 085701.
(34) Pan, A. C.; Chandler, D. *J. Phys. Chem. B* **2004**, *108*, 19681−19686.
(35) Jungblut, S.; Dellago, C. *J. Chem. Phys.* **2011**, *134*, 104501.
(36) Schor, M.; Vreede, J.; Bolhuis, P. G. *Biophys. J.* **2012**, *103*, 1296−1304.
(37) Pool, R.; Bolhuis, P. G. *J. Chem. Phys.* **2007**, *126*, 244703.
(38) Snee, P. T.; Shanoski, J.; Harris, C. B. *J. Am. Chem. Soc.* **2005**, *127*, 1286−1290.
(39) Pan, B.; Ricci, M. S.; Trout, B. L. *J. Phys. Chem. B* **2011**, *115*, 5958−5970.
(40) Mullen, R. G.; Shea, J.-E.; Peters, B. *J. Chem. Theory Comput.* **2014**, *10*, 659−667.
(41) Eidelson, N.; Peters, B. *J. Chem. Phys.* **2012**, *137*, 094106.
(42) Keys, A. S.; Hedges, L. O.; Garrahan, J. P.; Glotzer, S. C.; Chandler, D. *Phys. Rev. X* **2011**, *1*, 021013.
(43) Xi, L.; Shah, M.; Trout, B. L. *J. Phys. Chem. B* **2013**, *117*, 3634−3647.
(44) Geissler, P. L.; Dellago, C.; Chandler, D. *Phys. Chem. Chem. Phys.* **1999**, *1*, 1317−1322.
(45) Allen, R. J.; Frenkel, D.; ten Wolde, P. R. *J. Chem. Phys.* **2006**, *124*, 024102.
(46) Grünwald, M.; Dellago, C.; Geissler, P. L. *J. Chem. Phys.* **2008**, *129*, 194101.
(47) Bolhuis, P. G. *J. Phys.: Condens. Matter* **2003**, *15*, S113.
(48) Peters, B.; Trout, B. L. *J. Chem. Phys.* **2006**, *125*, 054108.
(49) Peters, B.; Beckham, G. T.; Trout, B. L. *J. Chem. Phys.* **2007**, *127*, 034109.
(50) Lo, C. S.; Radhakrishnan, R.; Trout, B. L. *Catal. Today* **2005**, *105*, 93−105.
(51) Ballard, A. J.; Dellago, C. *J. Phys. Chem. B* **2012**, *116*, 13490−13497.
(52) van Erp, T. S.; Moroni, D.; Bolhuis, P. G. *J. Chem. Phys.* **2003**, *118*, 7762−7774.
(53) Dellago, C.; Bolhuis, P. G.; Chandler, D. *J. Chem. Phys.* **1999**, *110*, 6617−6625.
(54) Dellago, C.; Bolhuis, P. G.; Geissler, P. L. Transition path sampling methods. In *Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology*; Ferrario, M., Ciccotti, G., Binder, K., Eds.; Springer-Verlag: Berlin, 2006; Vol. *1*, pp 349−391.
(55) Dellago, C.; Bolhuis, P. G. Transition path sampling and other advanced simulation techniques for rare events. In *Advanced Computer Simulation Approaches for Soft Matter Sciences*; Holm, C., Kremer, K., Eds.; Springer-Verlag: Berlin, 2009; Vol. *III*, pp 167−233.
(56) Juraszek, J.; Vreede, J.; Bolhuis, P. G. *Chem. Phys.* **2012**, *396*, 30−44.
(57) Fahey, P. M.; Griffin, P.; Plummer, J. *Rev. Mod. Phys.* **1989**, *61*, 289.
(58) Michel, A.; Rausch, W.; Ronsheim, P.; Kastl, R. *Appl. Phys. Lett.* **1987**, *50*, 416−418.

DOI: 10.1021/acs.jctc.5b00032
J. Chem. Theory Comput. 2015, 11, 2421−2428

(59) Munro, L. J.; Wales, D. J. *Phys. Rev. B* **1999**, *59*, 3969.
(60) Chatti, I.; Delahaye, A.; Fournaison, L.; Petitet, J.-P. *Energy Convers. Manage.* **2005**, *46*, 1333−1343.
(61) Lee, H.; Lee, J.-w.; Do Youn Kim, J. P.; Seo, Y.-T.; Zeng, H.; Moudrakovski, I. L.; Ratcliffe, C. I.; Ripmeester, J. A. *Nature* **2005**, *434*, 743−746.
(62) Peters, B.; Zimmermann, N. E.; Beckham, G. T.; Tester, J. W.; Trout, B. L. *J. Am. Chem. Soc.* **2008**, *130*, 17342−17350.
(63) Liu, X.; Mayer, M. T.; Wang, D. *Angew. Chem., Int. Ed.* **2010**, *49*, 3165−3168.
(64) Phillips, R. *Crystals, Defects and Microstructures: Modeling Across Scales*; Cambridge University Press: Cambridge, UK, 2001.
(65) Mars, P.; Van Krevelen, D. W. *Chem. Eng. Sci.* **1954**, *3*, 41−59.
(66) Zinkle, S.; Ghoniem, N. *Fusion Eng. Des.* **2000**, *51*, 55−71.
(67) Odette, G.; Alinger, M.; Wirth, B. *Annu. Rev. Mater. Res.* **2008**, *38*, 471−503.
(68) Plimpton, S. *J. Comput. Phys.* **1995**, *117*, 1−19.
(69) Geslin, P.-A.; Ciccotti, G.; Meloni, S. *J. Chem. Phys.* **2013**, *138*, 144103.
(70) Weeks, J. D.; Chandler, D.; Andersen, H. C. *J. Chem. Phys.* **1971**, *54*, 5237−5247.
(71) Bennett, C. H. Exact defect calculations in model substances. In *Diffusion in Solids: Recent Developments*, 1st ed.; Nowick, A. S., Burton, J. J., Eds.; Academic Press, Inc.: New York, 1975; pp 73−113.
(72) Peters, B. *Chem. Phys. Lett.* **2012**, *554*, 248−253.
(73) Radhakrishnan, R.; Schlick, T. *J. Chem. Phys.* **2004**, *121*, 2436−2444.
(74) Nosé, S. *J. Chem. Phys.* **1984**, *81*, 511−519.
(75) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695−1697.
(76) Chandler, D. *J. Chem. Phys.* **1978**, *68*, 2959−2970.
(77) Paci, E.; Ciccotti, G. *J. Phys.: Condens. Matter* **1992**, *4*, 2173.
(78) Gillan, M.; Harding, J.; Tarento, R.-J. *J. Phys. C* **1987**, *20*, 2331.

DOI: 10.1021/acs.jctc.5b00032
J. Chem. Theory Comput. 2015, 11, 2421−2428