

Enhanced configurational sampling with hybrid non-equilibrium molecular dynamics–Monte Carlo propagator ^F

Cite as: J. Chem. Phys. **148**, 014101 (2018); <https://doi.org/10.1063/1.5004154>

Submitted: 11 September 2017 . Accepted: 27 November 2017 . Published Online: 02 January 2018

Donghyuk Suh , Brian K. Radak , Christophe Chipot, and Benoît Roux

COLLECTIONS

Paper published as part of the special topic on [Enhanced Sampling for Molecular Systems](#)

^F This paper was selected as Featured



View Online



Export Citation



CrossMark

ARTICLES YOU MAY BE INTERESTED IN

[Molecular dynamics based enhanced sampling of collective variables with very large time steps](#)

The Journal of Chemical Physics **148**, 024106 (2018); <https://doi.org/10.1063/1.4999447>

[SSAGES: Software Suite for Advanced General Ensemble Simulations](#)

The Journal of Chemical Physics **148**, 044104 (2018); <https://doi.org/10.1063/1.5008853>

[Perspective: Maximum caliber is a general variational principle for dynamical systems](#)

The Journal of Chemical Physics **148**, 010901 (2018); <https://doi.org/10.1063/1.5012990>

Lock-in Amplifiers
Find out more today



 Zurich
Instruments



Enhanced configurational sampling with hybrid non-equilibrium molecular dynamics–Monte Carlo propagator

Donghyuk Suh,^{1,a)} Brian K. Radak,^{2,b)} Christophe Chipot,^{3,4,c)} and Benoît Roux^{5,6,d)}

¹*Department of Chemistry, University of Chicago, Chicago, Illinois 60637, USA*

²*Leadership Computing Facility, Argonne National Laboratory, Argonne, Illinois 60439-8643, USA*

³*Laboratoire International Associé Centre National de la Recherche Scientifique and University of Illinois at Urbana-Champaign, UMR 7565, Université de Lorraine, BP 70239, 54506 Vandœuvre-lès-Nancy, France*

⁴*Department of Physics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA*

and Beckman Institute, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA

⁵*Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, Illinois 60637-1454, USA*

⁶*Center for Nanoscale Materials, Argonne National Laboratory, Argonne, Illinois 60439-8643, USA*

(Received 11 September 2017; accepted 27 November 2017; published online 2 January 2018)

Molecular dynamics (MD) trajectories based on classical equations of motion can be used to sample the configurational space of complex molecular systems. However, brute-force MD often converges slowly due to the ruggedness of the underlying potential energy surface. Several schemes have been proposed to address this problem by effectively smoothing the potential energy surface. However, in order to recover the proper Boltzmann equilibrium probability distribution, these approaches must then rely on statistical reweighting techniques or generate the simulations within a Hamiltonian tempering replica-exchange scheme. The present work puts forth a novel hybrid sampling propagator combining Metropolis-Hastings Monte Carlo (MC) with proposed moves generated by non-equilibrium MD (neMD). This hybrid neMD-MC propagator comprises three elementary elements: (i) an atomic system is dynamically propagated for some period of time using standard equilibrium MD on the correct potential energy surface; (ii) the system is then propagated for a brief period of time during what is referred to as a “boosting phase,” via a time-dependent Hamiltonian that is evolved toward the perturbed potential energy surface and then back to the correct potential energy surface; (iii) the resulting configuration at the end of the neMD trajectory is then accepted or rejected according to a Metropolis criterion before returning to step 1. A symmetric two-end momentum reversal prescription is used at the end of the neMD trajectories to guarantee that the hybrid neMD-MC sampling propagator obeys microscopic detailed balance and rigorously yields the equilibrium Boltzmann distribution. The hybrid neMD-MC sampling propagator is designed and implemented to enhance the sampling by relying on the accelerated MD and solute tempering schemes. It is also combined with the adaptive biased force sampling algorithm to examine. Illustrative tests with specific biomolecular systems indicate that the method can yield a significant speedup. *Published by AIP Publishing.* <https://doi.org/10.1063/1.5004154>

I. INTRODUCTION

Classical molecular dynamics (MD) and Metropolis-Hasting Monte Carlo (MC) simulations based on detailed atomic models are powerful tools to study the properties of complex biomolecular systems.^{1–3} While simulations based on realistic all-atom (AA) models arguably offer the most detailed information, such models evolve on a complex and rugged energy surface and their dynamics are often burdened by a host of slow processes. For this reason, achieving an adequate sampling of all the relevant configurations

of a system from straight MD or MC simulation is often challenging.

A number of schemes have been proposed to enhance sampling and accelerate the exploration of configurational space by smoothing the underlying potential energy surface of a system.^{4–19} Among those schemes, two simple and attractive ideas are the hyperdynamics accelerated MD (aMD)^{9,10} and the replica-exchange with solute tempering (REST2) algorithms.¹⁶ However, to recover the Boltzmann equilibrium distribution of the proper Hamiltonian, one must either carry out a post-hoc reweighting analysis or generate the simulation within the context of a Hamiltonian tempering replica-exchange scheme involving multiple copies of the system.²⁰ While the former suffers from most configurations being statistically meaningless after reweighting, the latter substantially increases the computational cost of the simulation.

^{a)}Electronic mail: dsuh@uchicago.edu

^{b)}Electronic mail: brian.radak@anl.gov

^{c)}Electronic mail: chipot@ks.uiuc.edu

^{d)}Electronic mail: roux@uchicago.edu. URL: <http://thallium.bsd.uchicago.edu/RouxLab/>.

One avenue to address these issues is to carry out the propagation in such a way that the resulting configurational sampling reflects the proper Hamiltonian. For example, it is possible to reformulate the self-guided Langevin dynamics (SGLD) in such a way as to restore microscopic detailed balance.¹⁹ In the present effort, we wish to explore hybrid dynamical propagation schemes that combine the strength of non-equilibrium molecular dynamics (neMD) and Metropolis Monte Carlo (MC) to achieve enhanced sampling.^{21–26} In hybrid neMD-MC, the value of some chosen variable or coupling parameter is altered gradually in a time-dependent and controlled fashion, while the remaining degrees of freedom are allowed to evolve freely according to the normal equations of motion. The configuration generated by such a non-equilibrium “switching” trajectory is then treated as a candidate that must be either accepted or rejected via a Metropolis criterion to generate the equilibrium Boltzmann distribution.

The hybrid neMD-MC propagator designed here comprises an “equilibrium phase,” a “boosting phase,” and a “Metropolis MC step”:

- (i) Equilibrium phase: An atomic system is dynamically propagated for some period of time using standard equilibrium MD on the correct potential energy surface.
- (ii) Boosting phase: The system is then propagated for a brief period of time τ via a time-dependent Hamiltonian that is evolved toward the perturbed potential energy surface and then back to the correct potential energy surface.
- (iii) Metropolis MC step: The resulting configuration at the end of the neMD trajectory is then accepted or rejected according to a Metropolis criterion before returning to step 1. Using a symmetric switching schedule for ramping the Hamiltonian up and down, as well as keeping detailed balance with symmetric two-end momentum reversal, ensures that the algorithm strictly produces a Boltzmann equilibrium distribution.²⁷

In contrast to the hybrid neMD-MC simulations guided by a coarse-grained (CG) model introduced previously,²⁵ the algorithm described above relies on a perturbed Hamiltonian, but does not require the construction of a CG model to generate the target configuration. The hybrid neMD-MC sampling propagator implemented here rests on two schemes during the boosting phase, the so-called hyperdynamics accelerated MD (aMD)^{9,10} and the replica-exchange with solute tempering (REST2).¹⁶ Nevertheless, the strategy allows virtually any number of variations. Furthermore, the hybrid propagator may be naturally combined with a number of enhanced sampling strategies and free energy techniques.¹⁷ For example, preliminary results are also shown using a time-dependent bias along a collective variable determined via the adaptive biasing force (ABF) approach.²⁸ In Sec. II, we formulate the theoretical basis of the hybrid neMD-MC sampling propagator. The performance of the method is then illustrated with specific biomolecular systems. Our results indicate that the method can yield a significant speedup for biomolecular systems.

II. THEORY AND COMPUTATIONAL METHODS

A. Theory

Let us consider a classical system with time-dependent Hamiltonian $H[\mathbf{x}; \lambda(t)]$, where $\mathbf{x}$ represents all coordinates and momenta and $\lambda(t)$ is a time-dependent coupling parameter. Here, $\lambda = 0$ denotes the unperturbed Hamiltonian, H_0 , while $\lambda = 1$ corresponds to the perturbed Hamiltonian, H_1 , offering the maximum boost. In the conventional hybrid neMD-MC scheme for alchemical changes to the Hamiltonian, one varies λ from 0 to 1 in a neMD trajectory in order to generate a proposed move from $\mathbf{x}$ to $\mathbf{x}'$. The final configuration at the end of the non-equilibrium switching period corresponds to a different Hamiltonian. For example, in constant-pH simulations based on a hybrid neMD-MC algorithm, this could be the unprotonated or protonated state of a residue.^{21,22} In contrast, the boosting phase of the hybrid neMD-MC propagator starts and ends with the same unperturbed Hamiltonian, H_0 . In other words, the system is first brought to a perturbed Hamiltonian, H_1 , associated with a potential energy surface that is smoother and easier to sample ($\lambda = 0 \rightarrow \lambda = 1$) and then brought back to the unperturbed Hamiltonian H_0 ($\lambda = 1 \rightarrow \lambda = 0$). A random walk in phase space for a system at thermodynamic equilibrium must obey microscopic detailed balance,

$$\pi(\mathbf{x})T_p(\mathbf{x} \rightarrow \mathbf{x}')T_a(\mathbf{x} \rightarrow \mathbf{x}') = \pi(\mathbf{x}')T_p(\mathbf{x}' \rightarrow \mathbf{x})T_a(\mathbf{x}' \rightarrow \mathbf{x}), \quad (1)$$

where $\pi(\mathbf{x})$ is the equilibrium probability of being in state $\mathbf{x}$, $T_p(\mathbf{x} \rightarrow \mathbf{x}')$ is the transition probability for generating a proposed move from $\mathbf{x}$ to $\mathbf{x}'$, and $T_a(\mathbf{x} \rightarrow \mathbf{x}')$ is the transition probability for accepting a proposed move from $\mathbf{x}$ to $\mathbf{x}'$. Here, $\pi(\mathbf{x}) = Q^{-1} \exp[-H(\mathbf{x})/k_B T]$, where Q is the canonical partition function. After the boosting phase is completed, the energy difference $\Delta H = H[\mathbf{x}_B, \lambda = 0] - H[\mathbf{x}_A, \lambda = 0]$ is calculated, and a standard Metropolis acceptance criterion

$$T_a = \min[1, e^{-\beta \Delta H}] \quad (2)$$

is used to enforce microscopic detailed balance ($\beta \equiv 1/k_B T$, where k_B is Boltzmann’s constant and T is the absolute temperature). The hybrid propagator with the different stages is depicted schematically in Fig. 1.

A “symmetric two-ends momentum reversal” prescription is used in association with the non-equilibrium switches,^{22,25,27} by which one randomly chooses with equal

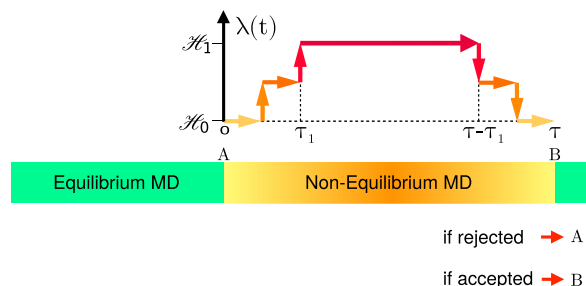


FIG. 1. Hybrid neMD-MC propagator scheme comprising an equilibrium MD phase, a non-equilibrium MD boosting phase, and an acceptance or rejection via a Metropolis MC step.

probability to carry out these trajectories as a “forward” or “backward” MD propagation. Forward propagation means that the non-equilibrium switch trajectory is simply continued using current positions and momenta, whereas backward propagation means that the momenta of the system are flipped at the beginning and at the end of the non-equilibrium switch trajectory (i.e., it is equivalent to a propagation going back in time). It was shown previously that the symmetric two-ends momentum reversal prescription increases the sampling rate by reducing the likelihood that different regions of configurational space remain isolated from one another.^{25,27} The energy difference ΔH in Eq. (2) may be further decomposed as a sum of two terms, q and w , where q is the heat exchange between the system and an external heat bath during propagation and w is the non-equilibrium work done during the perturbation. Since it is not trivial to keep track of heat-exchange during stochastic integration (such as Langevin dynamics), we instead opt to use a deterministic integrator such that the heat exchange is formally zero. The energy difference between states A and B is equal to the non-equilibrium work applied during the trajectory augmented by the shadow work done. The shadow work corresponds to the non-equilibrium work caused by the integrator error associated with the finite time step.²⁹ During the boosting phase, we used a schedule symmetric in time for perturbing the Hamiltonian by the coupling parameter, $\lambda(t)$,

$$\lambda(t; \tau, \tau_1) = \begin{cases} t/\tau_1 & (0 \leq t < \tau_1), \\ 1 & (\tau_1 \leq t < \tau - \tau_1), \\ (\tau - t)/\tau_1 & (\tau - \tau_1 \leq t < \tau), \end{cases} \quad (3)$$

where τ is the total switching time and τ_1 is the ramping time (see Fig. 1). The middle region, during which λ is a constant equal to 1, is the boosting phase that lasts for $\tau - 2\tau_1$.

In practice, the time-dependent coupling parameter $\lambda(t)$ was implemented via stepwise variations with a sequence of n_s discrete steps,

$$\lambda(t; \tau, \tau_1) = \begin{cases} 0, & 0 \leq t < \tau_1/n_s, \\ 1/n_s, & \tau_1/n_s \leq t < 2\tau_1/n_s, \\ \dots & \\ (n_s - 1)/n_s, & \tau_1(n_s - 1)/n_s \leq t < \tau_1, \\ 1, & \tau_1 \leq t < \tau - \tau_1, \\ 1 - 1/n_s, & \tau - \tau_1 + \tau_1/n_s \leq t < \tau - \tau_1 + 2\tau_1/n_s, \\ \dots & \\ 1/n_s, & \tau - 2\tau_1/n_s \leq t < \tau - \tau_1/n_s, \\ 0, & \tau - \tau_1/n_s \leq t < \tau \end{cases} \quad (4)$$

as depicted in Fig. 1 (see also the Tcl pseudocode in the Appendix).

The hybrid neMD-MC propagator algorithm was implemented according to the following steps:

- (i) The system is dynamically propagated for some period of time using unbiased equilibrium MD with the unperturbed Hamiltonian.

- (ii) The forward/backward direction for the switch is randomly chosen with equal probability for the neMD switching trajectory; the momenta are flipped if a backward propagation is chosen.

- The system is propagated via the time-dependent Hamiltonian $H[x, \lambda(t)]$ toward the perturbed potential energy surface ($\lambda = 0 \rightarrow \lambda = 1$) in a time τ_1 .
- The system is propagated with the perturbed Hamiltonian for some time to enhance barrier crossings ($\lambda = 1$) for a time $\tau - 2\tau_1$.
- The system is propagated via a time-dependent Hamiltonian $H(t)$ evolved back toward the unperturbed Hamiltonian ($\lambda = 1 \rightarrow \lambda = 0$) in a time τ_1 .
- The momenta are flipped again if the switch involved a backward propagation.

- (iii) The resulting configuration is accepted or rejected according to the Metropolis-Hasting criterion: If the move is accepted, we repeat the cycle; if the move is rejected, we return to the conformation at the end of step (i).

The performance of the hybrid neMD-MC propagator depends on our ability to generate a proposition likely to help overcome the barriers in a rugged potential energy surface. For this purpose, the parameters controlling the schedule, τ , τ_1 , and the choice of the perturbed Hamiltonian, H_1 , are critical. Two popular perturbation schemes designed to enhance sampling by deforming the potential energy surface were considered here for the boosting phase. The first one is the accelerated MD (aMD),^{9,10}

$$U_{\text{aMD}}[\mathbf{x}, \alpha, E] = \begin{cases} U[\mathbf{x}], & U[\mathbf{x}] \geq E, \\ U[\mathbf{x}] + \frac{(E - U[\mathbf{x}])^2}{\alpha + (E - U[\mathbf{x}])}, & U[\mathbf{x}] < E, \end{cases} \quad (5)$$

where α is a tuning parameter that determines the depth of the modified potential energy basins lying below the minimum threshold energy E (the aMD potential is flat when $U[\mathbf{x}]$ is equal or smaller than E if α is zero and increasingly unperturbed when α becomes very large).^{9,10} The aMD prescription may be applied to the total potential energy function or to various contributions such as the torsional potentials. In the present study, we have used Eq. (5) to reduce the energy barriers in the torsional potentials. To construct a time-dependent Hamiltonian for the boosting phase of the hybrid-aMD propagator, the threshold energy E was kept fixed while the tuning parameter α in Eq. (5) was effectively replaced by $\alpha/\lambda(t)$, with $\lambda(t)$ following the symmetric schedule given in Eq. (3). The singularity at the first step is avoided by setting E equal to zero until $\lambda(t)$ reaches its first non-zero value at time $t \geq \tau_1/n_s$ following Eq. (4).

The second perturbation scheme considered here is the replica-exchange with solute tempering (REST2),¹⁶

$$U_{\text{REST2}}[\mathbf{x}, \gamma] = U_v + \gamma U_u + \gamma^{1/2} U_{uv}, \quad (6)$$

where U_v is the solvent potential energy, U_u is the solute potential energy, and U_{uv} is the solute-solvent interaction

potential energy. It should be noted that we use the acronym REST2 here to indicate solute interaction tempering, even in the absence of multiple replicas. Let the coupling parameter $\gamma_{\max}$ correspond to the maximum solute tempering allowed, traditionally expressed as T/T_u in terms of an effective temperature T_u ascribed to the solute.¹⁶ To construct a time-dependent Hamiltonian for the boosting phase of the hybrid-REST2 propagator, the coupling parameter γ in Eq. (6) was effectively replaced by the time-dependent form $\gamma(t) = \lambda(t)(\gamma_{\max} - 1) + 1$, with $\lambda(t)$ following the symmetric schedule given in Eq. (3). As a result, the Hamiltonian is not perturbed when $\lambda = 0$, and the coupling parameter of the solute-tempering boost reaches its maximum of $\gamma_{\max}$ when $\lambda = 1$.

B. Computational methods

The hybrid neMD-MC propagator was tested on several biologically relevant molecular systems. The first objective of these tests is to show that the propagator does produce the correct Boltzmann equilibrium distribution. The second objective of these tests is to show that the propagator achieves convergence of the equilibrium properties more rapidly than simple unbiased MD. The hybrid neMD-MC propagator was implemented as a general Tcl script for the simulation program NAMD.³⁰ The Tcl script is given in the [Appendix](#).

The first test case is deca-alanine in vacuum with a dielectric constant $\epsilon = 20$. Eight individual 200-ns trajectories were generated using (i) equilibrium MD, (ii) two hybrid propagators using either REST2 or aMD as a boost, (iii) two hybrid propagators randomly accepting proposed moves by means of mean-acceptance ratio from the corresponding propagators, and (iv) two accelerating schemes, namely, REST2 and aMD. The equations of motion were integrated with a time step of 1 fs. The temperature was maintained at 300 K using Langevin dynamics with a damping coefficient of 1.0 ps^{-1} . Nonbonded short range interactions were truncated at 9 Å with a switching function effective from 8 Å. The parameters of the switching schedule were set to $t_{\text{eqMD}} = 5 \text{ ps}$, $\tau = 5 \text{ ps}$, $\tau_1 = 2 \text{ ps}$, number of switches $n_s = 39$, maximum REST2 boost $\gamma = 0.8$, and maximum aMD boost $(E, \alpha) = (50, 10)$. The CHARMM 22 protein force field³¹ was employed to model deca-alanine, allowing a direct comparison with a previous theoretical study.²⁸ A second test case considers again deca-alanine in vacuum, but this time with a dielectric constant, ϵ , of 1. This system is normally difficult to sample using simple unbiased equilibrium MD trajectories because of the strong internal backbone-backbone hydrogen bonds, making the usage of some form of importance sampling strategy a necessity. Here, the adaptive biasing force (ABF) method was considered.^{32,33} Results from ABF with standard MD and from ABF with hybrid neMD-MC boosted by aMD and REST2 were examined. Eight individual 40-ns trajectories were generated for each case, using α -helix and C_5 -extended conformation as initial coordinates. After optimization, the switching schedule was set to $t_{\text{eqMD}} = 7 \text{ ps}$, $\tau = 2 \text{ ps}$, $\tau_1 = 1 \text{ ps}$, number of switches $n_s = 18$, maximum REST2 boost $\gamma = 0.5$, and maximum aMD boost $(E, \alpha) = (40, 1)$. Extensive computations based on the multiple-walker ABF method^{32,33} were performed to provide a reference potential of mean force (PMF) that can be used to assess the converge of the

different methods. The computation was carried out with eight independent walkers, corresponding to an aggregated simulation time of 8 μs .

A third test case considers the folding/unfolding of the acetyl and NH_2 terminally capped $(\text{AAQAA})_3$ peptide solvated in explicit water. The 15 residue peptide was simulated using the hybrid-REST2 neMD-MC propagator. As a basis of comparison, the system was also simulated with simple unbiased MD, as well as multiple-copy REST2 and temperature replica-exchange MD (TREMMD) simulations.

For the optimization, different switching schedules were tested while varying τ , τ_1 , the maximum boost (REST2 or aMD), and number of switches. After optimization, the parameters of the switching schedule were set to $t_{\text{eqMD}} = 2 \text{ ps}$, $\tau = 24 \text{ ps}$, $\tau_1 = 8 \text{ ps}$, number of switches $n_s = 790$, and maximum REST2 boost $\gamma = 0.7$. The system was equilibrated at 300 K with initial structure of complete α -helical conformation solvated with 3964 water molecules in a cubic cell of initial dimensions equal to $55 \times 55 \times 55 \text{ Å}^3$. The equations of motion were integrated with a time step of 2 fs. The temperature was maintained at 300 K using Langevin dynamics with a damping coefficient of 1 ps^{-1} . Nonbonded short-range interactions were truncated at 12 Å with a switching function effective from 10 Å. The CHARMM 36 force field was used to describe the peptide, with the TIP3P water model. The REST2 calculation was done with 10 replicas ranging up to the same maximum boost of $\gamma = 0.7$ as in the hybrid neMD-MC simulations and exchange attempts every 0.2 ps. The TREMMD simulation was done with 32 replicas ranging from 278 K to 416 K and exchange attempts every 10 ps, matching exactly the protocol of Best *et al.*³⁴ However, the total length of the present TREMMD simulation is much shorter (1.875 ns per replica and 60 ns in total) than that of Best *et al.*³⁴ (150 ns per replica and 4.8 μs in total). To provide an objective comparison of the actual computational cost of these single- and multiple-copy approaches, the results are reported in terms of MD steps per replica times the number of replica involved.

III. RESULTS AND DISCUSSION

A. Validation

We first examine the results for deca-alanine in vacuum with a dielectric constant $\epsilon = 20$. The main purpose of this test is to validate the hybrid neMD-MC propagator. Because of the high dielectric constant, this is a convenient toy model, easy to sample directly by brute-force MD. The PMF as a function of the end-to-end distance extracted from unbiased MD is shown in Fig. 2. It is a double-well potential, with basins at 6 Å and 12 Å, separated by a small barrier of about 0.5 kcal/mol. The PMF calculated using the hybrid propagator is also shown in Fig. 2. The comparison shows that the hybrid method correctly reproduces the PMF extracted from the unbiased MD trajectory. The small error bars corresponding to a 99% confidence interval are indicative of the high convergence. As expected, the PMFs generated by pure REST2 and aMD without any reweighting do not reproduce the correct result. The biased PMFs nonetheless serve to display the overall free energy surface when the system is propagated at the highest boosting level using these two perturbation schemes designed to

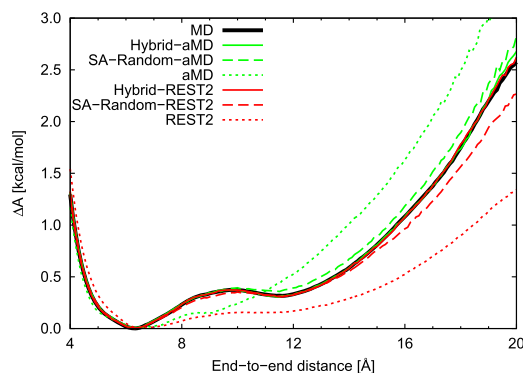


FIG. 2. Validation test using deca-alanine with dielectric constant of 20. SA-Random stands for hybrid scheme that randomly accepts proposed moves using the same mean acceptance ratio achieved from hybrid simulations.

enhance sampling by deforming the potential energy surface. Finally, the random acceptance simulations with the same mean acceptance ratio from the hybrid-aMD (0.1522 ± 0.002) and hybrid-REST2 simulations (0.716 ± 0.003) deviate from the correct PMF. This shows that the acceptance criterion with the symmetric two-ends momentum reversal is critical to obey detailed balance and obtain correct results.

B. Optimization

Generally, the efficiency of neMD-MC depends on maximizing the production of uncorrelated configurations while trying to minimize the effort spent to generate these.²⁶ The main drawback of neMD-MC simulations is that each new attempted MC move requires a non-equilibrium MD simulation; a low acceptance rate necessarily implies that a large fraction of computer time is discarded by the algorithm. Whereas long switches might yield a high acceptance probability but are computationally prohibitive, short switches are computationally inexpensive, but expected to yield vanishingly low acceptance probabilities. The most efficient algorithm is obviously a compromise balancing between these two opposing factors. It is generally unclear as to how to systematically achieve this balance.

In conventional neMD-MC schemes, the switch is meant to allow transitions between two different discrete states, thus longer switch times lead to higher acceptance rates as the transformation approaches the adiabatic limit.²⁶ In a previous study of a constant-pH algorithm, an optimal switching time for changing the protonation state of propionic acid in an explicit solvent was determined to be about 15 ps.²⁶ However, the hybrid propagator described here is not designed to produce transitions between discrete states. One implication is that a high acceptance probability, i.e., the number of accepted moves over the total number of attempted moves, does not guarantee that one is using the most effective perturbation scheme and boosting schedule. For example, the accepted candidate configuration may be too close to the starting configuration if the perturbation is too mild. Within the constraint of a fixed budget of computer time, increasing indefinitely the switching time also implies a concomitant decrease in the time available to perform equilibrium MD. Under these premises, simply maximizing the acceptance probability is not a useful criterion for optimizing efficiency.

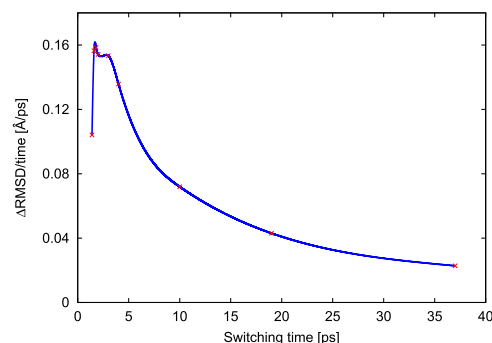


FIG. 3. Conformational evolution speed for deca-alanine in vacuum with a dielectric constant of 5 as a function of the boosting schedule. The conformational evolution speed is root-mean-square deviation (RMSD) of the peptide (calculated from all heavy atoms) at a time t relative to its configuration at a previous time $t - \Delta T$. For each case, the time interval ΔT was set to its switching time. The blue line is a fitted function for visual guide.

Several factors could affect the efficiency of the hybrid neMD-MC scheme described here. As a simple measure of efficiency of the hybrid neMD-MC propagation, we define a “conformational evolution speed” given by the net root-mean-square deviation (RMSD) of the peptide at a time t relative to the peptide configuration at time $t - \Delta T$. Monitoring the conformational evolution speed, while including the total cost to generate the non-equilibrium switches, allows us to assess the net efficiency of the hybrid neMD-MC propagator in terms of the perturbation scheme and boosting schedule. As illustrated in Fig. 3, it is indeed possible to identify a maximum efficiency in the case of the deca-alanine system. Based on this metric, a switching time of about 1.8 ps yields the most efficient hybrid neMD-MC propagator for the deca-alanine system.

For the (AAQAA)₃ system in an explicit solvent, we monitored the fraction of secondary-structure elements in the peptide chain as a function of switching time. As illustrated in Fig. 4, the efficiency is fairly poor when the switching time is too short. The simulation started with the peptide in an α -helical conformation, which persisted for an extended

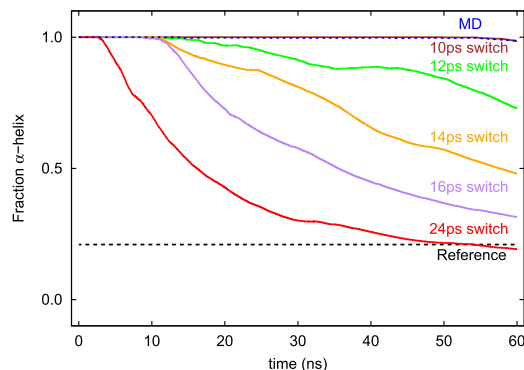


FIG. 4. Optimization test for the (AAQAA)₃ system in an explicit solvent by monitoring time cumulative average for the α -helicity. 4 independent trajectories were simulated for all cases, and 4 additional independent simulations were operated for MD and 24 ps switch. The peptide is defined to be in the α -helical conformation when the (ϕ, ψ) backbone dihedral angles of three consecutive residues fulfill the conditions $|-65^\circ - \phi| < 35^\circ$ and $|-37^\circ - \psi| < 30^\circ$.

TABLE I. Acceptance probability for the (AAQAA)₃ peptide.^a

τ (ps)	τ_1 (ps)	P_a
10	1	0.42
12	2	0.60
14	3	0.72
16	4	0.79
24	8	0.92

^a τ is the total switching time, τ_1 is the ramping time, and the boosting phase lasts $\tau - 2\tau_1$; 4 neMD-MC simulations were carried out for a switching time of 10, 12, 14, and 16 ps, and 8 neMD-MC simulations were carried out for a switching time of 24 ps.

period of time. In contrast, the efficiency is considerably improved by employing a switching time of 24 ps. While it is possible that this could lead to further improvement, the behavior of the system was not examined for switching times longer than 24 ps. In practice, the efficiency of the (AAQAA)₃ system in an explicit solvent is already very good with 24 ps according to Fig. 4. As shown in Table I, the acceptance probability P_a is about 0.92 with a switching time of 24 ps. This is approaching the limit afforded by the shadow work, which would be present with normal equilibrium MD propagation.

C. Performance

Deca-alanine is an extremely well-characterized benchmark system for testing the ability of new sampling schemes to capture subtle conformational equilibria.^{28,35} Its global free-energy minimum corresponds to a fully α -helical conformation resulting from the formation of robust $i \rightarrow i + 4$ intra-molecular hydrogen bonds. When generating a PMF along its end-to-end distance, the metastable states that exist along the orthogonal degrees of freedom in the “rugged” region between 4 and 12 Å can dramatically slow the convergence of conformational sampling. Furthermore, there is a very large free energy difference of about 30 kcal/mol between the completely folded form (14 Å) and the family of unfolded structures (32 Å). As a result, biased simulations must be used to overcome the considerable imbalance in the equilibrium probability along the end-to-end distance. Here, we use the adaptive biasing force (ABF) method.^{32,33} It should be emphasized that the necessity to adopt some form of importance sampling method to explore the end-to-end distance of the deca-alanine system remains true, whether a hybrid neMD-MC propagator is used or not.

The main results are displayed in Fig. 5. Comparison of the various methods reveals that the hybrid propagator with ABF clearly converged faster toward the reference PMF. While the standard deviations of the PMF based on eight independent simulations, either MD or the hybrid, are comparable in regions of the reaction pathway that are easily sampled (>12 Å), those for the hybrid method are certainly lower in the rugged region (<12 Å). This is indicative that the hybrid method facilitates the sampling of the numerous metastable states lying along orthogonal degrees of freedom in this region. Interestingly, when simulating multiple independent trajectories with pure ABF, we observed that some trajectories occasionally remained “trapped” in metastable regions for an

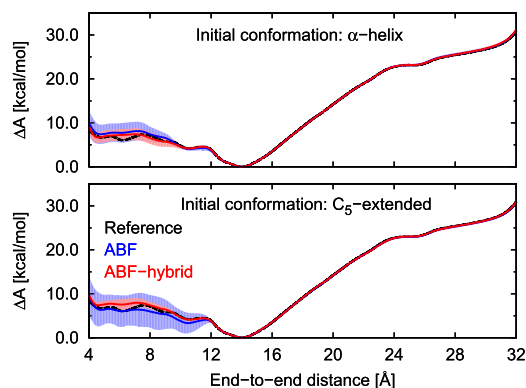


FIG. 5. PMF of the end-to-end distance of tests on deca-alanine in vacuum computed using ABF and ABF with hybrid-(REST2, aMD). Results using α -helix (top) and C_5 -extended (bottom) conformation as the starting structure are shown. 8 independent trajectories of 40 ns were simulated using ABF and ABF-hybrid.

extended fraction of the simulation time. This occurred for ABF with both starting conformations, leading to an over-stabilization of the metastable states and resulting in slow convergence. If one consider these trapped simulations as outliers and discard them, then ABF-MD performs overall better than ABF-hybrid. Importantly, this highly undesirable trapping phenomenon was not observed for the ABF-hybrid simulations. In assessing the overall efficiency of the ABF-hybrid simulations, it is important to note that no data are collected about the gradient during the neMD switching trajectories. Thus, the total amount of data accumulated to estimate the mean gradient is necessarily smaller than with ABF-MD. The hybrid scheme with ABF having only 70% of equivalent MD data points, yet performing better, underscores its power.

As a final illustrative test, we consider the folding/unfolding of the 15 residue peptide (AAQAA)₃ in an explicit solvent. This system has been the object of extensive simulation studies and is very well characterized.^{25,34} To highlight the conformational sampling challenge, the simulations were all started in the long-lived α -helix metastable state. Let us first compare the results from simple brute-force equilibrium MD and a hybrid neMD-MC propagator with REST2 boosting. Figure 6 shows the efficiency of the hybrid-REST2 propagator through time-cumulative averages of secondary structures (α -helix and β -strand). It is observed that the simulation converges towards the reference equilibrium average much faster with the hybrid propagator. This trend is systematic, as indicated by the 99% confidence interval estimated from the 8 independent trajectories. Rapid conformational fluctuations of the (AAQAA)₃ peptide are visibly occurring with the hybrid-REST2 propagator, as shown by the time-series of the individual simulations (Figs. S1–S4 of the [supplementary material](#)). Conversely, the systems simulated via simple brute-force equilibrium MD remain essentially trapped in the starting α -helical conformation. While the brute-force equilibrium MD generated 13 times more data points than the hybrid propagator, the conformational space of the (AAQAA)₃ peptide was explored much more efficiently with the hybrid propagator. Remarkably, REST2 (10 replicas) and TREMD (32 replicas) result in a weaker performance than

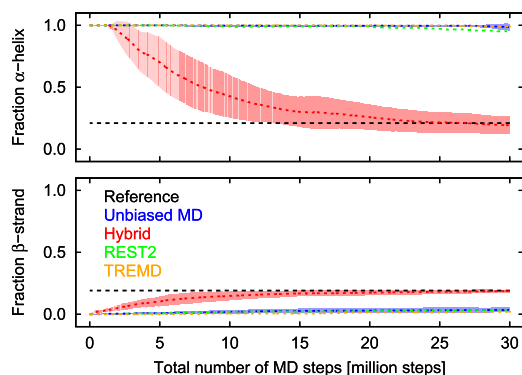


FIG. 6. Cumulative time-average for the secondary structure elements of a (AAQAA)₃ peptide in an explicit solvent using unbiased MD (single-copy), hybrid-REST2 neMD-MC (single-copy), REST2 with 10 replicas, and TREMD with 32 replicas. 8 independent trajectories of 60 ns were simulated using unbiased MD and the hybrid neMD-MC propagator, 6 ns each of 10 replicas were simulated using REST2, and 1.875 ns each of 32 replicas were simulated using TREMD. The total number of MD steps on the x-axis represents the actual computational costs of each approach as the number of steps per replica times the number of replica. The peptide is defined to be in the α -helical conformation when the (ϕ, ψ) backbone dihedral angles of three consecutive residues fulfill the conditions $|-65^\circ - \phi| < 35^\circ$ and $|-37^\circ - \psi| < 30^\circ$. The peptide is in the β -strand conformation when (ϕ, ψ) fulfill the conditions $|-140^\circ - \phi| < 40^\circ$ and $|150^\circ - \psi| < 30^\circ$.

the single-copy hybrid-REST2 neMD-MC propagator. The conformational sampling of (AAQAA)₃ in solution remains inefficient with these multiple-copy approaches, despite the fact that there were frequent exchanges between the replicas (Figs. S5 and S6 of the [supplementary material](#)). The difference in performance is particularly glaring when comparing in terms of total computational cost of all approaches (Fig. 6). It is of interest to compare these results with the ABF calculations for the deca-alanine toy model discussed above. In the case of the ABF-hybrid simulations, nearly 70% of the data produced correspond to equilibrium sampling and were accumulated to estimate the mean gradient. Increasing the amount of non-equilibrium switches is not advantageous because it reduces the amount of data used to estimate the mean gradient needed in ABF. This points to the limitation associated with the inherent loss of data during the non-equilibrium switches in the context of ABF. When the data collected from simple brute-force equilibrium MD are sufficiently uncorrelated, as in the case of deca-alanine, then increasing the proportion of the sampling carried out via the hybrid propagator becomes statistically less efficient. Even if the hybrid propagator actually moves the system around conformational space rapidly, enough data points must still be accumulated to accurately determine the average gradient that is needed for the ABF method. This stands in contrast with the (AAQAA)₃ simulation, for which the fraction of helicity could be determined accurately even though only 7.7% of the data produced corresponds to equilibrium sampling. The hybrid propagator can efficiently explore the conformational space accessible to the peptide in solution and the result without relying on accumulated equilibrium data as in the case of ABF. Even though the brute-force equilibrium MD simulations include about 13 times more data points than the hybrid simulation, it is extremely inefficient to explore the conformational space accessible to the (AAQAA)₃ peptide in solution.

IV. CONCLUSION

Powerful hybrid neMD-MC algorithms can be designed to boost and accelerate the conformational sampling of complex molecular systems. At the heart of the hybrid neMD-MC algorithms is a transient non-equilibrium perturbation of the system, which is aimed at overcoming the barriers in the rugged energy landscapes. Importantly, the neMD-MC algorithms robustly generate the correct Boltzmann equilibrium distribution. A vast range of time-dependent perturbations are possible to construct a family of hybrid neMD-MC propagators adapted to different situations. In the present test, we boosted the conformational sampling by relying on Hamiltonian perturbations based on aMD and REST2, but the strategy allows virtually any accelerated methods to be introduced during the neMD steps to efficiently lower the free-energy barriers. The benchmark systems tested demonstrate the correctness and effectiveness of the hybrid propagator, emphasizing its faster convergence compared to equilibrium MD. Although the switching schedule was optimized heuristically in the present study, it is possible to infer it analytically when the Hamiltonian is perturbed linearly.²⁶ Depending on the problem at hand, hybrid neMD-MC propagators may advantageously be combined with a number of established strategies and free energy methods,¹⁷ including umbrella sampling (US),³⁶ adaptive biasing force (ABF),²⁸ Hamiltonian tempering replica-exchange,³⁷ and alchemical free energy perturbation.¹² Future work will explore how the present hybrid neMD-MC propagators behave in the case of very complex biological objects and can be possibly tailored to address the rugged free-energy landscape underlying intricate processes therein.

SUPPLEMENTARY MATERIAL

See [supplementary material](#) for details (time-series, time-cumulative average, and accepted exchanges) about the simulations of the (AAQAA)₃ peptide in an explicit solvent (MD, hybrid neMD-MC, REST2, and TREMD).

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation (NSF) through Grant No. MCB-1517221. Additional support was provided by the France and Chicago Collaborating in The Sciences (FACCTS) program (to B.R. and C.C.). An award of computer time was provided by the Innovative and Novel Computational Impact on Theory and Experiment (INCITE) program; this research used resources of the Argonne Leadership Computing Facility (ALCF), which is a DOE Office of Science User Facility supported under Contract No. DE-AC02-06CH11357. Additional computational resources were provided by the University of Chicago Research Computing Center and the Grand Equipement National de Calcul Intensif-Centre Informatique National de l'Enseignement Supérieur (GENCI-CINES) at Montpellier, France (to C.C.).

APPENDIX: TCL PSEUDOCODE FOR NEMD-MC PROPAGATOR

Schematic Tcl pseudocode for executing the hybrid neMD-MC propagator with the NAMD³⁰ program:


```

proc save_callback {labels values} {
  global saved_labels saved_values
  set saved_labels $labels
  set saved_values $values
}
proc save_array {} {
  global saved_labels saved_values saved_array
  foreach label $saved_labels value $saved_values {
    set saved_array($label) $value
  }
}
set eqMD 1000
#eqMD steps per cycle
set Relax 100
#lowest propagation without perturbation
set Ramping 7900
#total steps during ramping except for relaxation
set Highest_boost 4000
#total steps at the highest boost
set num_runs 2308
#number of cycles
set num_switches 790
#number of switches per each direction (up, down)
set final 0.7
#highest boost for REST2
set diff [expr (1.0-$final)/$num_switches]
#perturbed amount per switch
set i 0
while {$i < $num_runs} {
  run [expr $eqMD / 2]
  save_array
  set TOT_E_A $saved_array(TOTAL)
  checkpoint
  if { rand() > 0.5 } {
    set flipped_vel 1
    rescalelevels -1.0
  } else {
    set flipped_vel 0
  }
  Langevin off
  set step_up 1
  set current 1.0
  run [expr $Relax / 2]
  set current [expr $current - $diff]
  while {$step_up <= $num_switch} {
    sptScaleFactor $current
    run [expr int($Ramping/($num_switch*2))]
    set current [expr $current - $diff]
    incr step_up
  }
  run $Highest_boost
  set step_down [expr $num_switch - 1]
  while {$step_down >= 0} {
    set current [expr $current + $diff]
    sptScaleFactor $current
    run [expr int($Ramping/($num_switches*2))]
    incr step_down -1
  }
  set current [expr $current + $diff]
  sptScaleFactor $current
  run [expr $Relax / 2]
  save_array
  set TOT_E_B $saved_array(TOTAL)
  if { $flipped_vel } {
    rescalelevels -1.0
  }
  set BOLTZMAN 0.001987191
  set beta [expr 1. / ($temp * $BOLTZMAN)]
  set delta [expr $beta * ($TOT_E_B - $TOT_E_A)]
  set MC [expr exp(-1. * $delta) > rand()]
  if { $MC } {
    puts "EXCHANGE_ACCEPT RUN $i"
    incr exchange_accepted
  } else {
    puts "EXCHANGE_REJECT RUN $i"
    revert
  }
  incr exchange_attempted
  Langevin on

```

```

run [expr $steps_for_MD / 2]
incr i
}

```

- ¹M. Karplus and J. A. McCammon, "Molecular dynamics simulations of biomolecules," *Nat. Struct. Mol. Biol.* **9**, 646–652 (2002).
- ²N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, and A. H. Teller, "Equation of state calculations by fast computing machines," *J. Chem. Phys.* **21**, 1087–1092 (1953).
- ³W. K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," *Biometrika* **57**, 97–109 (1970).
- ⁴T. Huber, A. E. Torda, and W. F. van Gunsteren, "Local elevation: A method for improving the searching properties of molecular dynamics simulation," *J. Comput.-Aided Mol. Des.* **8**, 695–708 (1994).
- ⁵V. Leone, F. Marinelli, P. Carloni, and M. Parrinello, "Targeting biomolecular flexibility with metadynamics," *Curr. Opin. Struct. Biol.* **20**, 148–154 (2010).
- ⁶X. Wu and B. R. Brooks, "Self-guided Langevin dynamics simulation method," *Chem. Phys. Lett.* **381**, 512–518 (2003).
- ⁷A. Damjanović, X. Wu, B. R. Brooks *et al.*, "Backbone relaxation coupled to the ionization of internal groups in proteins: A self-guided Langevin dynamics study," *Biophys. J.* **95**, 4091–4101 (2008).
- ⁸X. Wu and B. R. Brooks, "Toward canonical ensemble distribution from self-guided Langevin dynamics simulation," *J. Chem. Phys.* **134**(13), 134108 (2011).
- ⁹A. F. Voter, "Hyperdynamics: Accelerated molecular dynamics of infrequent events," *Phys. Rev. Lett.* **78**, 3908 (1997).
- ¹⁰D. Hamelberg, J. Mongan, and J. A. McCammon, "Accelerated molecular dynamics: A promising and efficient simulation method for biomolecules," *J. Chem. Phys.* **120**, 11919 (2004).
- ¹¹M. Fajer, D. Hamelberg, and J. A. McCammon, "Replica-exchange accelerated molecular dynamics (REXAMD) applied to thermodynamic integration," *J. Chem. Theory Comput.* **4**, 1565–1569 (2008).
- ¹²W. Jiang and B. Roux, "Free energy perturbation Hamiltonian replica-exchange molecular dynamics (FEP/H-REMD) for absolute ligand binding free energy calculations," *J. Chem. Theory Comput.* **6**, 2559–2565 (2010).
- ¹³L. Maragliano and E. Vanden-Eijnden, "A temperature accelerated method for sampling free energy and determining reaction pathways in rare events simulations," *Chem. Phys. Lett.* **426**, 168–175 (2006).
- ¹⁴H. Lei and Y. Duan, "Improved sampling methods for molecular simulation," *Curr. Opin. Struct. Biol.* **17**, 187–191 (2007).
- ¹⁵M. Christen and W. F. van Gunsteren, "On searching in, sampling of, and dynamically moving through conformational space of biomolecular systems: A review," *J. Comput. Chem.* **29**, 157–166 (2008).
- ¹⁶L. Wang, R. Friesner, and B. J. Berne, "Replica exchange with solute scaling: A more efficient version of replica exchange with solute tempering (REST2)," *J. Phys. Chem. B* **115**, 9431–9438 (2011).
- ¹⁷A. Mitsutake, Y. Mori, and Y. Okamoto, "Enhanced sampling algorithms," *Methods Mol. Biol.* **924**, 153–195 (2013).
- ¹⁸R. C. Bernardi, M. C. Melo, and K. Schulten, "Enhanced sampling techniques in molecular dynamics simulations of biological systems," *Biochim. Biophys. Acta* **1850**, 872–877 (2015).
- ¹⁹X. Wu, B. R. Brooks, and E. Vanden-Eijnden, "Self-guided Langevin dynamics via generalized Langevin equation," *J. Comput. Chem.* **37**, 595–601 (2016).
- ²⁰M. Arrar, C. A. F. de Oliveira, M. Fajer, W. Sinko, and J. A. McCammon, "w-REXAMD: A Hamiltonian replica exchange approach to improve free energy calculations for systems with kinetically trapped conformations," *J. Chem. Theory Comput.* **9**, 18–23 (2013).
- ²¹H. A. Stern, "Molecular simulation with variable protonation states at constant pH," *J. Chem. Phys.* **126**, 164112 (2007).
- ²²H. A. Stern, "Erratum: Molecular simulation with variable protonation states at constant pH," *J. Chem. Phys.* **127**, 079901 (2007).
- ²³A. J. Ballard and C. Jarzynski, "Replica exchange with nonequilibrium switches," *Proc. Natl. Acad. Sci. U. S. A.* **106**, 12224–12229 (2009).
- ²⁴J. P. Nilmeier, G. E. Crooks, D. D. L. Minh, and J. D. Chodera, "Nonequilibrium candidate Monte Carlo is an efficient tool for equilibrium simulation," *Proc. Natl. Acad. Sci. U. S. A.* **108**, E1009–E1018 (2011).
- ²⁵Y. Chen and B. Roux, "Enhanced sampling of an atomic model with hybrid nonequilibrium molecular dynamics–Monte Carlo simulations guided by a coarse-grained model," *J. Chem. Theory Comput.* **11**, 3572–3583 (2015).

- ²⁶B. K. Radak and B. Roux, "Efficiency in nonequilibrium molecular dynamics Monte Carlo simulations," *J. Chem. Phys.* **145**, 134109 (2016).
- ²⁷Y. Chen and B. Roux, "Efficient hybrid non-equilibrium molecular dynamics–Monte Carlo simulations with symmetric momentum reversal," *J. Chem. Phys.* **141**, 114107 (2014).
- ²⁸J. Comer, J. C. Gumbart, J. Hénin, T. Lelièvre, A. Pohorille, and C. Chipot, "The adaptive biasing force method: Everything you always wanted to know but were afraid to ask," *J. Phys. Chem. B* **119**, 1129–1151 (2015).
- ²⁹D. A. Sivak, J. D. Chodera, and G. E. Crooks, "Using nonequilibrium fluctuation theorems to understand and correct errors in equilibrium and nonequilibrium simulations of discrete Langevin dynamics," *Phys. Rev. X* **3**, 011007 (2013).
- ³⁰J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, and K. Schulten, "Scalable molecular dynamics with NAMD," *J. Comput. Chem.* **26**, 1781–1802 (2005).
- ³¹A. D. MacKerell, Jr., D. Bashford, M. Bellott, R. L. Dunbrack, Jr., J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher III, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, and M. Karplus, "All-atom empirical potential for molecular modeling and dynamics studies of proteins," *J. Phys. Chem. B* **102**, 3586–3616 (1998).
- ³²J. Comer, J. Phillips, K. Schulten, and C. Chipot, "Multiple-replica strategies for free-energy calculations in NAMD: Multiple-walker adaptive biasing force and walker selection rules," *J. Chem. Theory Comput.* **10**, 5276–5285 (2014).
- ³³K. Minoukadeh, T. Lelièvre, and C. Chipot, "Potential of mean force calculations: A multiple-walker adaptive biasing force approach," *J. Chem. Theory Comput.* **6**, 1008–1017 (2010).
- ³⁴R. B. Best, X. Zhu, J. Shim, P. E. M. Lopes, J. Mittal, M. Feig, and A. D. MacKerell, Jr., "Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ , ψ and side-chain χ_1 and χ_2 dihedral angles," *J. Chem. Theory Comput.* **8**, 3257–3273 (2012).
- ³⁵S. Park, F. Khalili-Araghi, E. Tajkhorshid, and K. Schulten, "Free energy calculation from steered molecular dynamics simulations using Jarzynski's equality," *J. Chem. Phys.* **119**, 3559–3566 (2003).
- ³⁶G. M. Torrie and J. P. Valleau, "Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling," *J. Comput. Phys.* **23**, 187–199 (1977).
- ³⁷W. Jiang, Y. Luo, L. Maragliano, and B. Roux, "Calculation of free energy landscape in multi-dimensions with Hamiltonian-exchange umbrella sampling on petascale supercomputer," *J. Chem. Theory Comput.* **8**, 4672–4680 (2012).