LLNL-JRNL-492635

LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

Nonequilibrium candidate Monte Carlo: A new tool for equilibrium simulation

J. P. Nilmeier, G. E. Crooks, D. L. Minh, J. D. Chodera

August 10, 2011

Proceedings of the National Academy of Science

Disclaimer

This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.

Nonequilibrium candidate Monte Carlo is an efficient tool for equilibrium simulation

Jerome P. Nilmeier[a], Gavin E. Crooks[b], David D. L. Minh[c], and John D. Chodera[d,1]

[a]Biosciences and Biotechnology Division, Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, CA 94550; [b]Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720; [c]Biosciences Division, Argonne National Laboratory, Argonne, IL 60439; and [d]California Institute for Quantitative Biosciences (QB3), University of California, Berkeley, CA 94720

Edited by Michele Parrinello, ETH Zurich, Lugano, Switzerland, and approved August 30, 2011 (received for review April 21, 2011)

Metropolis Monte Carlo simulation is a powerful tool for studying the equilibrium properties of matter. In complex condensed-phase systems, however, it is difficult to design Monte Carlo moves with high acceptance probabilities that also rapidly sample uncorrelated configurations. Here, we introduce a new class of moves based on nonequilibrium dynamics: Candidate configurations are generated through a finite-time process in which a system is actively driven out of equilibrium, and accepted with criteria that preserve the equilibrium distribution. The acceptance rule is similar to the Metropolis acceptance probability, but related to the nonequilibrium work rather than the instantaneous energy difference. Our method is applicable to sampling from both a single thermodynamic state or a mixture of thermodynamic states, and allows both coordinates and thermodynamic parameters to be driven in nonequilibrium proposals. Whereas generating finite-time switching trajectories incurs an additional cost, driving some degrees of freedom while allowing others to evolve naturally can lead to large enhancements in acceptance probabilities, greatly reducing structural correlation times. Using nonequilibrium driven processes vastly expands the repertoire of useful Monte Carlo proposals in simulations of dense solvated systems.

expanded ensembles | Markov chain Monte Carlo | Metropolis–Hastings | molecular dynamics

In this paper, we describe a new technique for constructing efficient Markov chain Monte Carlo (MCMC) (1) moves that both have high acceptance rates and also allow rapid transit through configuration space, greatly enhancing convergence rates in simulations of dense solvated systems. The Metropolis Monte Carlo (MC) (2, 3) sampling procedure is generalized by using nonequilibrium processes to generate candidates for equilibrium simulations. Within this framework, moves that are efficient for an isolated part of a system but lead to near-universal rejection in standard Monte Carlo simulations of dense mixtures can be converted to nonequilibrium processes that generate candidates with higher acceptance probabilities. In this new procedure, the acceptance criteria is related to the nonequilibrium work, rather than the potential energy difference used in traditional Monte Carlo moves.

Since their introduction in the mid-twentieth century, MC (2, 3) and molecular dynamics (MD) (4) simulations have become favored tools for sampling from complex multidimensional distributions, such as configurations of microscopic physical systems in thermodynamic ensembles. However, these methods can produce highly correlated samples, leading to slow convergence of estimated expectations. Whereas MD requires the use of small timesteps for numerical stability and to approximate sampling from the desired distribution, MC simulations can, in principle, make use of nonlocal moves that accelerate mixing of the Markov chain. Indeed, vast improvements in efficiency have been obtained by applying cleverly constructed move sets that exploit physical intuition about the system under study, such as cluster moves in Potts and Ising model simulations (5, 6).

Designing efficient moves requires striking a balance between rapid traversal of phase space and ensuring reasonable acceptance probabilities. For complex heterogeneous systems such as solvated biomolecules, achieving this balance remains challenging. Typically, efficient moves exploit physical insight into kinetically slow processes and energetically favorable configurations. Often, the experimenter may possess physical insight about one component in the system (e.g., a biomolecule) that permits the design of moves that would be efficient in the absence of other components (e.g., solvent), but encounter energetically unfavorable interactions in their presence, reducing acceptance rates to levels where standard MC provides no benefit. As an illustrative example, consider a bistable dimer—a pair of particles interacting with a potential with minima in compact or extended configurations, separated by a high barrier (see Fig. 1). For simulations of this system in a vacuum, a simple and effective standard MC move is to instantaneously increase the interparticle distance from a compact to extended configuration (or conversely, to decrease the distance from an extended to compact configuration). When the dimer is immersed in a dense solvent, however, this move is met with near-universal rejection because solvent molecules overlap with proposed configurations.

One approach that can allow unperturbed degrees of freedom to relax, and hence maintain a reasonable acceptance rate, is to use a nonequilibrium process to generate candidate configurations. Using the appropriate acceptance criterion for the final configuration will preserve the equilibrium distribution. In the case of the bistable dimer immersed in dense solvent, the extension (or contraction) may be carried out over a finite number of increments interleaved with standard Metropolis Monte Carlo or molecular dynamics steps that allow the solvent to reorganize to avoid overlap with the dimer particles.

The basic idea of using nonequilibrium driven processes as Monte Carlo moves has precedents in both the statistical (7, 8) and chemical (9–12) literature. Among the latter, Athènes developed "work-bias Monte Carlo" to enhance acceptance rates in grand canonical Monte Carlo simulations (9), Stern presented a scheme to sample an equilibrium mixture of protonation states at constant pH in explicit solvent (11) [although an inexact variant was proposed earlier (10)], and Nilmeier et al. (12) proposed the driving of a subset of degrees of freedom to enhance acceptance rates (using an approximate acceptance criterion). Non-

Author contributions: J.P.N., G.E.C., D.D.L.M., and J.D.C. designed research; D.D.L.M. and J.D.C. performed research; J.P.N., G.E.C., D.D.L.M., and J.D.C. contributed new analytic tools; J.D.C. analyzed data; and J.P.N., G.E.C., D.D.L.M., and J.D.C. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: Simulation and analysis scripts, as well as simulation data generated for this paper, are available online at http://simtk.org/home/ncmc/; simtk.org is an online repository for exchanging scientific data produced by SimBios, a National Institutes of Health Center for Biomedical Computing.

[1]To whom correspondence should be addressed. E-mail: jchodera@berkeley.edu.

See Author Summary on page 18199.

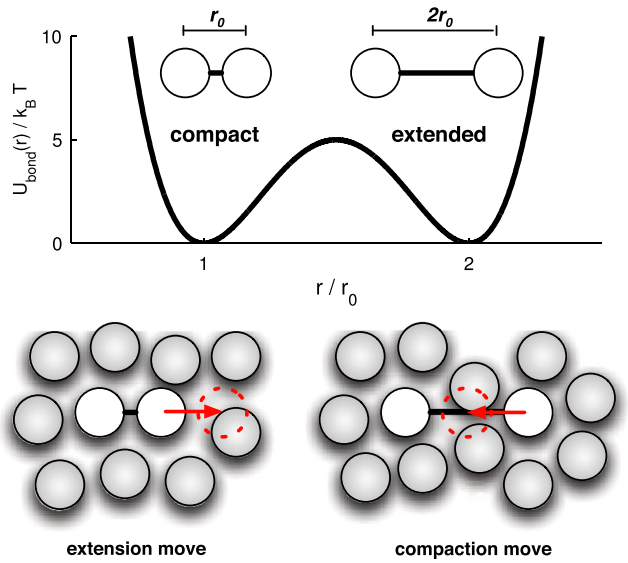

Fig. 1. Bistable dimer potential and instantaneous MC moves in WCA fluid. An extension move increases the dimer extension by $\Delta r = +r_0$, whereas a compaction move decreases the dimer extension by $\Delta r = -r_0$. Both move types meet with near-universal rejection when implemented as instantaneous MC moves in a dense WCA fluid. Note that the lower panel is only a cartoon—the actual described simulation is of a 3D system.

equilibrium processes have also been used to generate configurations for parallel tempering simulations (13–15).

Here, we unify these ideas and significantly extend the application of nonequilibrium moves in physical simulations. We present a theoretical framework, nonequilibrium candidate Monte Carlo (NCMC), that is applicable to both single thermodynamic states (e.g., NVT, NpT, μVT ensembles) as well as mixtures of thermodynamic states [e.g., expanded ensemble (16, 17) simulations]. Nonequilibrium proposals may drive a subset of degrees of freedom, the thermodynamic parameters characterizing the equilibrium distribution, or both, significantly expanding the repertoire of Monte Carlo moves that lead to high acceptance and efficient mixing in dense condensed-phase systems.

Equilibrium and Expanded Thermodynamic Ensembles

For physical systems in equilibrium, the probability of observing a microstate is given by the Boltzmann distribution,

$$\pi_\lambda(x) = Z_\lambda^{-1} e^{-u_\lambda(x)}; \qquad Z_\lambda = \int_\Gamma dx\, e^{-u_\lambda(x)}, \qquad [1]$$

where $x \in \Gamma$ denotes a microstate of the system (which may include coordinates, momenta, and other dynamical variables, such as simulation box dimensions), λ denotes a set of thermodynamic parameters whose values define a thermodynamic state, and Z_λ is a normalizing constant known as the partition function.

The reduced potential $u_\lambda(x)$ depends on the thermodynamic ensemble under consideration (18). For instance, in an isothermal-isobaric (NpT) ensemble, the reduced potential will assume the form,

$$u_\lambda(x) = \beta[H(x) + pV(x)], \qquad [2]$$

which depends on the Hamiltonian $H(x)$ (which may include an external biasing potential, and is presumed to be invariant under momentum inversion) and the system volume $V(x)$. In this ensemble, the vector of controllable thermodynamic parameters $\lambda \equiv \{\beta, H, p\}$ includes the inverse temperature β, the Hamiltonian $H(x)$, and external pressure p. Other thermodynamic parameters and their conjugate variables can be included or excluded to generate alternative physical (or unphysical) ensembles.

To allow sampling from multiple thermodynamic states within a single simulation, we also define an expanded ensemble (16, 17), which specifies a joint distribution for (x, λ) in a weighted mixture of thermodynamic states,

$$\pi(x, \lambda) = \frac{Z_\lambda \pi_\lambda(x) \omega_\lambda}{\sum_{\nu \in \mathcal{G}} \int_\Gamma dy\, Z_\nu \pi_\nu(y) \omega_\nu}, \qquad [3]$$

where $\omega_\lambda > 0$ specifies an externally imposed weight for state λ. Here, $\lambda \in \mathcal{G}$ may assume values in a discrete or continuous space $\mathcal{G}$. If the set $\mathcal{G}$ consists of a single value of λ, a single thermodynamic state is sampled, and $\pi(x, \lambda) = \pi_\lambda(x)$. These thermodynamic states may correspond to a variety of different states of interest, such as temperatures in a simulated tempering simulation (19), alchemical states in a simulated scaling simulation (20), or protonation states in a constant-pH simulation (21).

Nonequilibrium Candidate Monte Carlo

We first describe the general form of NCMC. At the start of an iteration, the current sample in the Markov chain, $(x^{(n)}, \lambda^{(n)})$, which is assumed to be drawn from $\pi(x, \lambda)$, is used to initialize a trajectory, $(x_0, \lambda_0) = (x^{(n)}, \lambda^{(n)})$. A candidate configuration (x_T, λ_T) is then proposed through a nonequilibrium process in which a set of degrees of freedom and/or thermodynamic parameters may be driven according to some protocol (22) selected with a probability dependent only on (x_0, λ_0). Even if we only wish to sample from a single thermodynamic state λ, we may use a protocol that transiently drives the thermodynamic parameters away from λ and back again (as in ref. 14). Finally, an acceptance probability is computed and used to decide whether the next sample in the Markov chain, $(x^{(n+1)}, \lambda^{(n+1)})$, is the candidate, (x_T, λ_T), or the momentum reversal of the initial sample, $(\tilde{x}^{(n)}, \lambda^{(n)})$.

An NCMC move begins by selecting a protocol Λ from a set of possible protocols with probability $P(\Lambda|x_0, \lambda_0)$, such that there exists a reverse protocol labeled as $\tilde{\Lambda}$ (to be defined momentarily) with $P(\tilde{\Lambda}|\tilde{x}_T, \lambda_T) > 0$. A protocol Λ specifies both a series of T perturbation kernels $\alpha_t(x, y)$ and propagation kernels $K_t(x, y)$, arranged in an alternating pattern such that $\Lambda \equiv \{\alpha_1, K_1, \alpha_2, K_2, \ldots, \alpha_T, K_T\}$. Both $\alpha_t(x, y)$ and $K_t(x, y)$ are conditional probabilities of $y \in \Gamma$ given any $x \in \Gamma$, and must satisfy the requirement that if $p(x, y) > 0$, then $p(y, x) > 0$, for p substituted by α_t and K_t.

Each perturbation kernel α_t drives some or all of the degrees of freedom x in a stochastic or deterministic way (e.g., by driving a torsion angle, a distance between two atoms, or the volume of the simulation cell). Similarly, each propagation kernel K_t propagates some or all of the coordinates of the system at fixed λ_t according to some form of MCMC or MD [e.g., Metropolis Monte Carlo (2, 3), velocity Verlet (23) deterministic dynamics, or overdamped Langevin stochastic dynamics (24, 25)] that may also depend on the time index t. Interleaving perturbation and propagation allows for energetically unfavorable interactions introduced by perturbation to be relaxed during propagation, potentially increasing acceptance rates relative to the instantaneous perturbations of standard Metropolis Monte Carlo.

The procedure by which a trajectory $X \equiv (x_0, x_1, \ldots, x_T)$ is generated from an initial microstate x_0 according to a protocol Λ can be illustrated by the scheme,

$$x_0 \xrightarrow{\alpha_1} x_1^* \xrightarrow{K_1} x_1 \longrightarrow \cdots \longrightarrow x_{T-1} \xrightarrow{\alpha_T} x_T^* \xrightarrow{K_T} x_T. \qquad [4]$$

Application of the perturbation α_t to x_{t-1} generates a perturbed configuration x_t^*, which is then propagated by the kernel K_t to obtain x_t.

The reverse protocol $\tilde{\Lambda} \equiv \{K_T, \alpha_T, \ldots, K_0, \alpha_0\}$ reverses the order in which the perturbation and propagation steps are applied,

generating the time-reversed trajectory $\tilde{X} \equiv \{\tilde{x}_T, \ldots, \tilde{x}_0\}$, where $\tilde{x}$ denotes x with inverted momenta,

$$\tilde{x}_T \xrightarrow{K_T} \tilde{x}_T^* \xrightarrow{\alpha_T} \tilde{x}_{T-1} \longrightarrow \cdots \longrightarrow \tilde{x}_1 \xrightarrow{K_1} \tilde{x}_1^* \xrightarrow{\alpha_1} \tilde{x}_0. \quad [5]$$

The next step in NCMC is to accept or reject (x_T, λ_T) as the next sample in the chain. To ensure that the stationary distribution $\pi(x, \lambda)$ is preserved, we enforce a strict pathwise form of detailed balance*,

$$\begin{aligned} &A(X|\Lambda)\Pi(X|x_0, \Lambda)P(\Lambda|x_0, \lambda_0)\pi(x_0, \lambda_0) \\ &\quad = A(\tilde{X}|\tilde{\Lambda})\Pi(\tilde{X}|\tilde{x}_T, \tilde{\Lambda})P(\tilde{\Lambda}|\tilde{x}_T, \lambda_T)\pi(\tilde{x}_T, \lambda_T). \end{aligned} \quad [6]$$

The quantity $A(X|\Lambda)$ is the NCMC acceptance probability, whereas $\Pi(X|x_0, \Lambda)$ and $\Pi(\tilde{X}|\tilde{x}_T, \tilde{\Lambda})$ denote the probability of generating trajectory X from initial configuration x_0 using protocol Λ, or $\tilde{X}$ from final configuration $\tilde{x}_T$ with protocol $\tilde{\Lambda}$, respectively,

$$\Pi(X|x_0, \Lambda) = \prod_{1 \le t \le T} \alpha_t(x_{t-1}, x_t^*)K_t(x_t^*, x_t) \quad [7]$$

$$\Pi(\tilde{X}|\tilde{x}_T, \tilde{\Lambda}) = \prod_{T \ge t \ge 1} \alpha_t(\tilde{x}_t^*, \tilde{x}_{t-1})K_t(\tilde{x}_t, \tilde{x}_t^*). \quad [8]$$

Summation of Eq. **6** over all trajectories starting with x_0 and ending with x_T recovers the standard detailed balance condition (see *Appendix* for proof).

We define the ratio of proposal kernels as

$$\frac{\alpha(\tilde{X}|\tilde{\Lambda})}{\alpha(X|\Lambda)} \equiv \prod_{t=1}^{T} \frac{\alpha_t(\tilde{x}_t^*, \tilde{x}_{t-1})}{\alpha_t(x_{t-1}, x_t^*)}, \quad [9]$$

and the ratio of propagation kernels as the exponentiated difference in forward and backward conditional path actions as

$$e^{-\Delta\mathcal{S}(X|\Lambda)} \equiv \prod_{t=1}^{T} \frac{K_t(\tilde{x}_t, \tilde{x}_t^*)}{K_t(x_t^*, x_t)}. \quad [10]$$

Using the above expressions and the momentum invariance property $\pi(x, \lambda) = \pi(\tilde{x}, \lambda)$, we may write the ratio of acceptance probabilities as

$$\begin{aligned} \frac{A(X|\Lambda)}{A(\tilde{X}|\tilde{\Lambda})} &= \frac{\pi(\tilde{x}_T, \lambda_T)}{\pi(x_0, \lambda_0)} \frac{P(\tilde{\Lambda}|\tilde{x}_T, \lambda_T)}{P(\Lambda|x_0, \lambda_0)} \frac{\Pi(X|\tilde{x}_T, \tilde{\Lambda})}{\Pi(X|x_0, \Lambda)} \\ &= \frac{\pi(x_T, \lambda_T)}{\pi(x_0, \lambda_0)} \frac{P(\tilde{\Lambda}|\tilde{x}_T, \lambda_T)}{P(\Lambda|x_0, \lambda_0)} \prod_{t=1}^{T} \frac{\alpha_t(\tilde{x}_t^*, \tilde{x}_{t-1})}{\alpha_t(x_{t-1}, x_t^*)} \frac{K_t(\tilde{x}_t, \tilde{x}_t^*)}{K_t(x_t^*, x_t)} \\ &\equiv \frac{\omega_T}{\omega_0} \frac{P(\tilde{\Lambda}|\tilde{x}_T, \lambda_T)}{P(\Lambda|x_0, \lambda_0)} \frac{\alpha(\tilde{X}|\tilde{\Lambda})}{\alpha(X|\Lambda)} e^{-\Delta\mathcal{S}(X|\Lambda) - \Delta u(X|\Lambda)}, \end{aligned} \quad [11]$$

where $\Delta u(X|\Lambda) \equiv u_T(x_T) - u_0(x_0)$ is the energy difference. Eq. **11** is the main result of this paper, and is highly general with regard to both the choice of protocol for perturbation and propagation. In subsequent sections, we discuss specific choices for these protocols that lead to particularly simple acceptance criteria.

Many choices of acceptance probabilities $A(X|\Lambda)$ that satisfy Eq. **11** are possible, including the well-known Metropolis–Hastings criterion (2, 3)

$$A(X|\Lambda) = \min\left\{1, \frac{\omega_T}{\omega_0} \frac{P(\tilde{\Lambda}|\tilde{x}_T, \lambda_T)}{P(\Lambda|x_0, \lambda_0)} \frac{\alpha(\tilde{X}|\tilde{\Lambda})}{\alpha(X|\Lambda)} e^{-\Delta\mathcal{S}(X|\Lambda) - \Delta u(X|\Lambda)}\right\}. \quad [12]$$

After generating (x_T, λ_T) and evaluating $A(X|\Lambda)$, we generate a uniform random variate U. If $A(X|\Lambda) > U$, then the candidate becomes the next value in the chain, $(x^{(n+1)}, \lambda^{(n+1)}) = (x_T, \lambda_T)$. Otherwise, it is rejected, we perform a momentum flip, and the next value becomes $(x^{(n+1)}, \lambda^{(n+1)}) = (\tilde{x}_0, \lambda_0)$. Alternately, we may perform a momentum flip upon acceptance, $(x^{(n+1)}, \lambda^{(n+1)}) = (\tilde{x}_T, \lambda_T)$ and preserve the momentum upon rejection, $(x^{(n+1)}, \lambda^{(n+1)}) = (x_0, \lambda_0)$. We cannot, however, ignore the momentum flip completely; as explained in *Appendix*, this flip is necessary to preserve the equilibrium distribution.

We note that NCMC need not be used exclusively to sample from $\pi(x, \lambda)$, but can be mixed with other MCMC moves or with MD (1). For example, one may reinitialize velocities from the Maxwell-Boltzmann distribution after each NCMC step; this is a Gibbs sampling MCMC move using the marginal distribution for velocities.

Perturbation Kernels

A large variety of choices are available for the perturbation kernels $\alpha_t(x, y)$. Through judicious selection of these kernels, a practitioner can design nonequilibrium proposals that carry some component of the system from one high-probability region to another with high acceptance rates. We briefly describe a few possibilities.

Stochastically Driven Degrees of Freedom. Suppose we wish to drive a torsion angle ϕ (an angle subtended by four bonded atoms) stochastically by rotating it to a new torsion angle ϕ' (holding angles and bonds fixed) according to some probability, such as the von Mises circular distribution centered on ϕ,

$$\eta(\phi'|\phi) = [2\pi I_0(\kappa)]^{-1} e^{\kappa\cos(\phi' - \phi)}, \quad [13]$$

with $I_0(\kappa)$ denoting the modified Bessel function of order zero and $\kappa > 0$ taking the role of a dimensionless force constant. Because the stochastic perturbation is made in a non-Cartesian coordinate, a Jacobian $J(\phi)$ must be included to compute $\alpha(x, y)$ in Cartesian coordinates, resulting in the ratio,

$$\frac{\alpha_t(\tilde{y}, \tilde{x})}{\alpha_t(x, y)} = \frac{\eta(\phi|\phi')J(\phi)}{\eta(\phi'|\phi)J(\phi')} = 1, \quad [14]$$

where $J(\phi') = J(\phi) = 1$ because the transformation (a rotation about a bond vector) preserves the Cartesian phase space volume.

Deterministically Driven Degrees of Freedom. Instead of perturbing the torsion angle stochastically, we can deterministically drive it in small, fixed increments $\Delta\phi$. In this case, we effectively define an invertible map $\mathcal{M}$ that takes $x \to y$, such that $y = \mathcal{M}x$ differs from x only in the rotation of the specified torsion ϕ by $\Delta\phi$. To implement this, we may choose a perturbation $\Delta\phi$ from a distribution where $\pm\Delta\phi$ have equal probability, and drive $\phi(x)$ from its current value ϕ_0 to a final value $\phi_T = \phi_0 + \Delta\phi$ over T steps in equal increments, such that $\phi(x_t)$ is constrained to $\phi_t \equiv (1 - t/T)\phi_0 + (t/T)\phi_T$. In this case, $\alpha_t(x, y) = \delta(y - \mathcal{M}x)$ $J(x)$, where the Jacobian $J(x)$ represents the factor by which Cartesian phase space is compressed on the application of the map $\mathcal{M}$, which is again unity for rotation about a torsion angle by $\Delta\phi$, and, due to the invertibility of the map, the ratio $\alpha_t(\tilde{y}, \tilde{x})/\alpha_t(x, y) = 1$.

Simulation Box Scaling. Another possible deterministic perturbation kernel is simulation box scaling. A barostat can be implemented by proposing propagation kernels that scale the molecular

*The described pathwise detailed balance condition is closely related to "super-detailed balance" (e.g., ref. 26), but also accounts for momentum reversal to extend its definition to include molecular dynamics integrators.

centers and box geometry by a factor $s = [(V(x) + \Delta V)/V(x)]^{1/3}$ with ΔV chosen uniformly from $[V - \Delta V_0, V + \Delta V_0]$ applied as a factor of $s^{1/T}$ over the course of T steps. In this case, the perturbation kernel $\alpha_t(x, y)$ is a delta function that compresses or expands the molecular centers and box geometry. Because the Jacobian is the ratio of infinitesimal volumes upon scaling, the ratio of perturbation kernels is $\alpha(\tilde{X}|\tilde{\Lambda})/\alpha(X|\Lambda) = s^{3N}$, where N denotes the number of molecular centers.

Thermodynamic Perturbation. In many driven nonequilibrium processes, there is no direct perturbation to the coordinates, such that $\alpha_t(x, y) = \delta(x - y)$ and the ratio $\alpha(\tilde{X}|\tilde{\Lambda})/\alpha(X|\Lambda) = 1$. Instead, only the thermodynamic parameters λ are varied in time, carrying the system out of equilibrium through action of the K_t propagation kernels. We recover Neal's method (7) if the reduced potential u_t is a simple linear interpolation such that $u_t(x) = (1 - t/T)u_0(x) + (t/T)u_T(x)$, the probability of choosing protocol Λ is symmetric with $\tilde{\Lambda}$, and MC (2, 3) is used for the propagation kernel K_t.

Propagation Kernels

The choice of propagation kernels available is also very broad. If strong driving is performed in selection of α, one may elect to choose a time-independent propagation kernel $K_t(x, y) \equiv K(x, y)$ that samples from a stationary distribution $\pi(x)$ of interest. Alternatively, a strongly time-dependent K_t could be selected to transiently drive the system out of equilibrium, or from the equilibrium distribution at one thermodynamic state to another. Some possible choices are described below.

Reversible Markov Chain Monte Carlo. One may propagate some or all of a system's degrees of freedom (e.g., those not affected by the perturbation kernel α_t) by a method that satisfies detailed balance in π_t,

$$\pi_t(x)K_t(x, y) = \pi_t(\tilde{y})K_t(\tilde{y}, \tilde{x}), \qquad [15]$$

where $\pi_t(x) \equiv Z_t^{-1}e^{-u_t(x)}$ for a specified $u_t(x)$. Many MCMC methods (1), including Metropolis (2, 3) and various hybrid Monte Carlo (HMC) algorithms that combine discrete-timestep integrators with Monte Carlo acceptance/rejection steps (27, 28), obey detailed balance and are commonly used for the simulation of physical systems.

By analogy with Crooks (29), we define a work w and heat q for the nonequilibrium driven process,

$$w(X|\Lambda) = \sum_{t=1}^{T}[u_t(x_t^*) - u_{t-1}(x_{t-1})] \qquad [16]$$

$$q(X|\Lambda) = \sum_{t=1}^{T}[u_t(x_t) - u_t(x_t^*)], \qquad [17]$$

such that $w(X|\Lambda) + q(X|\Lambda) = \Delta u(X|\Lambda)$, a restatement of the first law of thermodynamics.

The conditional path action difference can then be written in terms of the heat of the process, $q(X|\Lambda)$,

$$\Delta\mathcal{S}(X|\Lambda) = -\ln\prod_{t=1}^{T}\frac{\pi_t(x_t^*)}{\pi_t(x_t)} = -q(X|\Lambda), \qquad [18]$$

leading to an acceptance probability similar to standard MC, except that the work, $w(X|\Lambda)$, replaces the instantaneous potential energy difference,

$$\frac{A(X|\Lambda)}{A(\tilde{X}|\tilde{\Lambda})} = \frac{\omega_T}{\omega_0}\frac{P(\tilde{\Lambda}|\tilde{x}_T, \lambda_T)}{P(\Lambda|x_0, \lambda_0)}\frac{\alpha(\tilde{X}|\tilde{\Lambda})}{\alpha(X|\Lambda)}e^{-w(X|\Lambda)}. \qquad [19]$$

Deterministic Dynamics. When an isolated system is propagated by a symplectic integrator—a reversible, deterministic integrator that preserves phase space volume—the propagation kernels follow $K_t(x, y) = K_t(\tilde{y}, \tilde{x})$. Hence, $\Delta\mathcal{S}(X|\Lambda) = 0$ and the acceptance ratio is,

$$\frac{A(X|\Lambda)}{A(\tilde{X}|\tilde{\Lambda})} = \frac{\omega_T}{\omega_0}\frac{P(\tilde{\Lambda}|\tilde{x}_T, \lambda_T)}{P(\Lambda|x_0, \lambda_0)}\frac{\alpha(\tilde{X}|\tilde{\Lambda})}{\alpha(X|\Lambda)}e^{-\Delta u(X|\Lambda)}, \qquad [20]$$

where $\Delta u(X|\Lambda) \equiv u_T(x_T) - u_0(x_0)$ is the energy difference. The equivalence of the work and energy difference for volume-preserving integrators was previously recognized in the context of fluctuation theorem calculations (30, 31).

Symplectic integrators include velocity Verlet (23). These integrators are also symplectic when utilizing constraints through the application of algorithms such as RATTLE (32), provided that the constraints are iterated to convergence each timestep (33).

Stochastic Dynamics. Stochastic integrators such as velocity Verlet discretizations of Langevin dynamics (34, 35) sample a modified distribution that differs from the desired equilibrium distribution π_t in a timestep-dependent manner (36). Whereas this modified distribution may be difficult or impossible to compute to recover equilibrium properties by reweighting, computation of the relative action $\Delta\mathcal{S}(X|\Lambda)$ is relatively straightforward, and the NCMC acceptance criteria ensures that the NCMC-sampled configurations are distributed according to the desired equilibrium ensemble[†]. As examples, we compute $\Delta\mathcal{S}(X|\Lambda)$ for the overdamped Langevin (Brownian) dynamics integrator of Ermak and Yeh (24, 25) and the Brünger–Brooks–Karplus (BBK) Langevin integrator (38–40) in *Appendix*.

Illustrative Application: Bistable Dimer in a WCA Fluid

To demonstrate NCMC, we ran simulations of a bistable dimer (adapted from section 1.3.2.4 of ref. 36) in vacuum as well as a dense fluid. The dimer consists of a pair of "bonded" particles interacting via a double-well potential, with minima at $r = r_0$ (compact) and $r = 2r_0$ (extended), and a $5k_BT$ barrier (see Fig. 1). In the solvated simulations, the dimer was immersed in a dense bath (reduced density $\rho\sigma^3 = 0.96$) of particles that interact with the bonded particles and each other via the Weeks–Chandler–Andersen (WCA) soft repulsive potential (41). Each simulation "iteration" consisted of velocity reassignment from the Maxwell–Boltzmann distribution, 500 steps of generalized hybrid Monte Carlo (GHMC) dynamics (1, 28, 36, 37) (essentially, a Metropolis-corrected form of Langevin dynamics, henceforth referred to here as MD), optionally followed by either an instantaneous MC move or an NCMC move.

The rate at which effectively uncorrelated samples are generated can be quantified in terms of the correlation time τ for the dimer extension $r(t)$ (shown in Fig. 2). This time represents the asymptotic decay time constant for the correlation function $C(t) = \langle r(0)r(t)\rangle$, which will behave like

$$C(t) = C_\infty + (C_0 - C_\infty)e^{-t/\tau} \qquad [21]$$

for large t, where $C_0 = \langle r^2\rangle$ and $C_\infty = \langle r\rangle^2$. The correlation time is related to the statistical inefficiency, $g = 1 + 2\tau$, a factor that describes the number of iterations necessary to generate an effectively uncorrelated sample (42).

For the MD simulation in vacuum (Fig. 2, top trace), we observe slow hopping between compact and extended configura-

[†]Alternatives to using NCMC to correct stochastic integration include introducing a Metropolization correction after each step, as in the generalized hybrid Monte Carlo (GHMC) integrator we use in the example (1, 28, 36, 37).

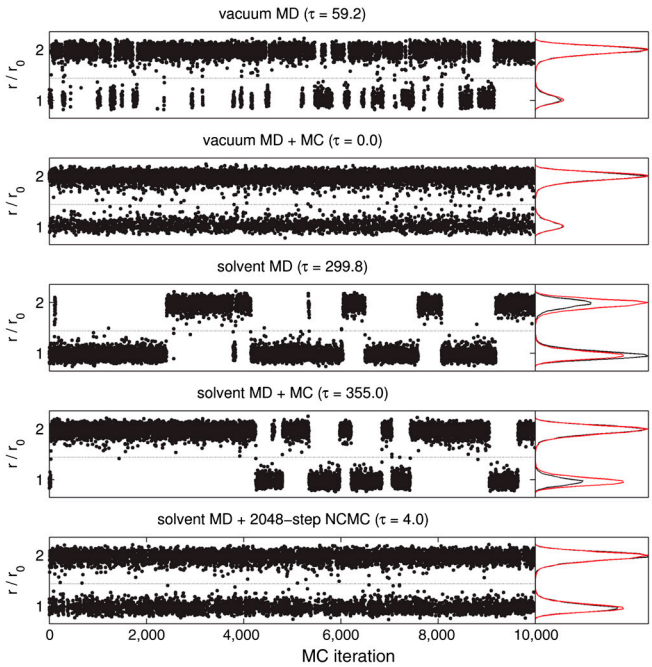

Fig. 2. Trajectories of WCA dimer system in vacuum and solvent. (*Left*) The dimer extension r as a function of simulation iteration. The dotted horizontal line denotes division between compact and extended configurations. The quantity τ printed above each plot indicates the estimated integrated autocorrelation time for the dimer extension r. (*Right*) Histogram accumulated over trajectory (black), with true equilibrium distribution (red). Plot titles denote whether simulation was run in vacuum (vacuum) or dense WCA fluid (solvent), and whether the simulation utilized only 500 GHMC steps per iteration (MD) or included instantaneous MC (MC) or 2,048-step NCMC moves (NCMC) following each iteration.

tions with a correlation time $\tau = 59.2$ iterations, resulting in slow convergence of the histogram. Introducing instantaneous MC dimer extension/contraction moves that modify the dimer extension by $\Delta r = \pm r_0$ reduces the correlation time to $\tau \approx 0.0$, such that an uncorrelated configuration is generated after each iteration of 500 MD steps and one instantaneous MC step (Fig. 2, second trace from top).

When the dimer is immersed in a dense fluid of WCA particles, however, iterations consisting of 500 MD steps alone result in extremely slow barrier crossings, requiring $g \approx 600$ iterations to produce an uncorrelated sample (Fig. 2, middle trace). Unlike in vacuum, the introduction of instantaneous MC moves does not significantly reduce the correlation time τ (Fig. 2, second trace from bottom). However, performing these same dimer expansion and contraction moves over 2,048-step NCMC moves (Fig. 2, bottom trace) allows the system to rapidly mix between both compact and extended states with a correlation time of $\tau = 4.0$ iterations. Whereas each iteration requires a fivefold increase in computational effort (500 MD steps + 2,048 NCMC switching steps = 2,548 force evaluations, versus 500 force evaluations for MD alone), a 67-fold reduction in correlation time is achieved, resulting in a remarkable order-of-magnitude gain in overall efficiency.

The length of the NCMC switching process is a free parameter that may be tuned to further improve efficiency. Toward this end, we estimated the acceptance probability of the extension/contraction moves in dense solvent as a function of switching length (Fig. 3). Whereas instantaneous MC proposals of $\pm r_0$ are only accepted with probability $\approx 10^{-27}$ (the error is this quantity is likely underestimated due to its extremity), dividing the move into smaller steps boosts the acceptance rate to a level useful in condensed-phase simulation. If the move is divided into a small number of steps (1–8), there is little to no increase in acceptance rate, but for an intermediate number of steps (16–1,024), there is a superlinear boost in the acceptance probability relative to the length of the switching process. The acceptance probability finally reaches useful levels around 2,000 steps, achieving an acceptance rate of 12% using nonequilibrium proposal trajectories of 2,048 steps, or 38% for 8,192 steps.

In general, there is no direct relationship between acceptance probability and efficiency. Under certain assumptions relevant to the bistable dimer, however, it is possible to link the NCMC acceptance probability to τ_{eff}, an indirect estimate of the correlation time,

$$\tau_{\text{eff}} = \tau_{\text{MD}} \left[\frac{\tau_{\text{NCMC}}}{\tau_{\text{MD}} + \tau_{\text{NCMC}}} \right], \qquad [22]$$

where τ_{MD} and τ_{NCMC} are correlation times for iterations consisting solely of MD or NCMC moves, respectively. The latter correlation time may be estimated from the average acceptance probability γ by $\tau_{\text{NCMC}} \approx -1/\ln(1-2\gamma)$ (see *Appendix* for derivation).

As shown in Fig. 4, the effective correlation time τ_{eff} is only diminished when the NCMC acceptance probability is large enough such that $\tau_{\text{NCMC}} \approx \tau_{\text{MD}}$, which occurs when $\gamma \geq 0.13\%$ (about 256 switching steps or more). For shorter switching times, even though the acceptance probability is high relative to instantaneous MC, it is still too small to significantly reduce τ_{eff}.

When comparing efficiency, we are most interested in the rate of generating uncorrelated configurations for a given amount of computational effort. Relative to MD alone, this rate is described by the efficiency gain,

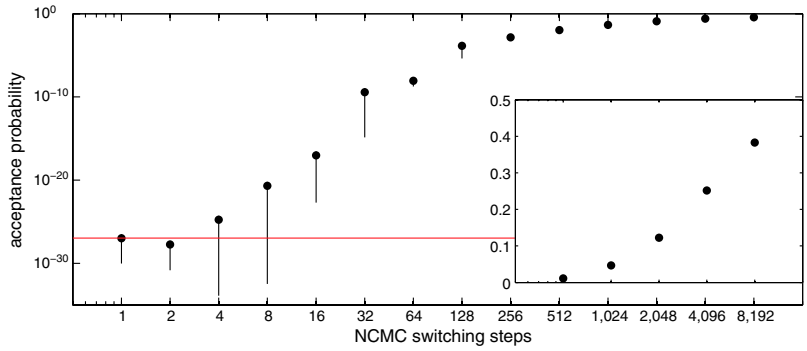

Fig. 3. Acceptance probabilities of NCMC proposals. (*Top*) Acceptance probability of NCMC proposals as a function of length of nonequilibrium proposal trajectory (black dots), compared with instantaneous MC proposal (red line). (*Inset*) Enlarged region with acceptance probability shown on linear scale. Estimated 95% confidence intervals are shown as vertical lines.

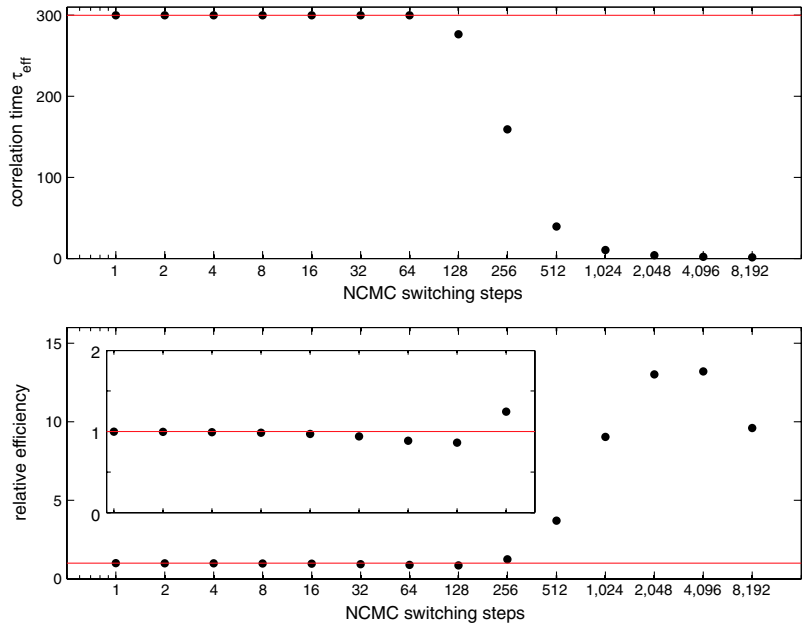

Fig. 4. Statistical efficiency gain of NCMC proposals relative to instantaneous MC proposals. (*Upper*) Effective correlation time τ_{eff}, in iterations, for MD+NCMC (black dots) compared to MD alone (red line). (*Lower*) Relative statistical efficiency of MD+NCMC, in terms of number of uncorrelated configurations generated for a fixed amount of computational effort, for MD+NCMC (black dots) relative to MD alone (red line).

$$E \equiv \frac{g_{\mathrm{MD}} T_{\mathrm{MD}}}{g_{\mathrm{NCMC}}(T_{\mathrm{MD}} + T_{\mathrm{NCMC}})}. \qquad [23]$$

Here, $T_{\mathrm{MD}} = 500$ steps per iteration, and T_{NCMC} is again varied from 1–8,192 steps. The results are shown in the bottom panel of Fig. 4. Surprisingly, there is actually a slight loss in efficiency at short switching times—dropping to a minimum of 86.9% the efficiency of MD alone at 128 steps—followed by a rapid gain in efficiency, plateauing at an efficiency gain of approximately 13× the efficiency of MD alone for 2,048–4,096-step NCMC proposals. (A similar plateau behavior is observed in the modified parallel tempering protocol of ref. 15.) After this point, longer switching times do not achieve as high of an efficiency gain; even though the acceptance rate continues to increase as the number of NCMC switching steps is doubled again to 8,192 steps, the reduction in correlation time is not sufficient to offset the additional cost of these moves.

Epilogue

We have described a procedure—nonequilibrium candidate Monte Carlo (NCMC)—by which nonequilibrium proposals can be used within MCMC simulations to enhance acceptance rates. In our illustration, we have demonstrated how its use can lead to large improvements in statistical efficiency—the rate at which uncorrelated samples are generated for a fixed amount of computational effort. In other applications, whether similarly large efficiency gains are achieved will depend on the precise nature of the system under study and the nonequilibrium proposals introduced. The most straightforward approach—borrowing Metropolis Monte Carlo proposals that are reasonable for one component of the system in isolation, and converting them to nonequilibrium proposals—is likely to be a fruitful avenue for generating efficient schemes, as was demonstrated here for a simple system.

More generally, the problem of selecting efficient nonequilibrium proposals is similar to the problem of choosing good reaction coordinates, in that it is desirable to drive the system along (possibly complex) slow collective coordinates where orthogonal degrees of freedom relax quickly. The search for such collective coordinates is a topic of active research (43–49). Given an initial nonequilibrium protocol, the issue of optimizing such a protocol to minimize dissipation (and maximize acceptance) is also a topic of active study, led by forays into the world of single-molecule measurement (50–52). Recent work has also suggested that estimating the thermodynamic metric tensor along the nonequilibrium parameter switching path (53–56), could prove useful in adaptively optimizing the switching protocol (57).

We note that switching trajectories contain potentially useful information. Indeed, several methods (56, 58, 59) have recently been developed to estimate equilibrium properties from nonequilibrium samples through the application of statistical estimator theory to nonequilibrium fluctuation theorems (30, 60, 61); these are particularly relevant to switching between multiple thermodynamic states. Though the development of efficient estimators that utilize both nonequilibrium switching trials and sampled equilibrium data generated in NCMC simulations remains an open challenge, it is at least straightforward to incorporate information from rejected NCMC proposals in the estimation of equilibrium averages (26, 62).

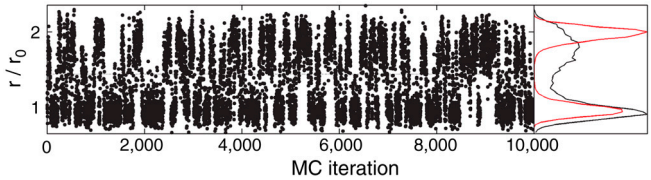

Fig. 5. Umbrella sampling simulation of the dimer in WCA solvent. (*Left*) The dimer extension r as a function of simulation iteration. (*Right*) The histogram accumulated over the trajectory, with the observed histogram in black and the reweighted histogram (corrected for the applied umbrella bias potential) in red.

Materials and Methods

The dimer system considered here consists of two particles that interact via a double-well bonded potential in the interatomic distance r,

$$U_{\mathrm{bond}}(r) = h\left[1 - \frac{(r - r_0 - s)^2}{s^2}\right]^2, \qquad [24]$$

with $h = 5k_BT$, $r_0 = r_{WCA}$, and $s = r_{WCA}/2$, where $r_{WCA} \equiv 2^{1/6}\sigma$. Simulations denoted as "vacuum" contain only these two particles, whereas simulations denoted as "solvated" also interact with a dense bath of particles via the WCA nonbonded potential (41),

$$U_{WCA}(r) = \begin{cases} 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] + \epsilon, & r < r_{WCA} \\ 0 & r \geq r_{WCA} \end{cases}, \quad [25]$$

with mass $m = 39.9$ amu, $\sigma = 3.4$ Å, and $\epsilon = 120\ k_BT$. The nonbonded WCA interaction is excluded between the two bonded particles. The solvated system contains a total of 216 WCA particles at a reduced density of $\rho\sigma^3 = 0.96$. For all simulations, the reduced temperature is $k_BT/\epsilon = 0.824$. A custom Python code making use of the graphics processing unit (GPU)-accelerated OpenMM package (63–65) and the PyOpenMM Python wrapper (66) was used to conduct the simulations. All scripts are available for download from http://simtk.org/home/ncmc.

To ensure that observed differences were not due to changes in the stationary distribution of the integrator, we used GHMC (1, 28, 36, 37) for all our simulations. GHMC is based on a velocity Verlet discretization (23) of Langevin dynamics—the two are equivalent in the limit of small timesteps—but includes an acceptance/rejection step to correct for errors introduced by finite timesteps so that the stationary distribution is the exact equilibrium distribution. We used a timestep of 0.002τ, where $\tau = \sqrt{\sigma^2 m/\epsilon}$, and the collision rate was set to τ^{-1}. With this timestep, the acceptance probability is $99.929 \pm 0.001\%$; the resulting dynamics closely approximates Langevin dynamics.

In simulations employing instantaneous Monte Carlo moves, a perturbation Δr to the interatomic distance r of the dimer was chosen according to

$$\Delta r = \begin{cases} +r_0 & \text{if } r < 1.5r_0 \\ -r_0 & \text{if } 1.5r_0 \leq r \leq 3r_0 \\ 0 & \text{otherwise} \end{cases}. \quad [26]$$

The dimer was contracted or expanded about the bond midpoint to generate a new configuration x_{new} with dimer extension r_{new} from the old configuration x_{old} with dimer extension r_{old}, and the move accepted or rejected with the Metropolis–Hastings criterion (3),

$$A(x_{new}|x_{old}) = \min\{1, e^{-\beta[U(x_{new})-U(x_{old})]}J_r(x_{old}, x_{new})\}, \quad [27]$$

where the Jacobian ratio term $J_r(x_{old}, x_{new}) = (r_{new}/r_{old})^2$ accounts for the expansion and contraction of phase space due to the Monte Carlo proposals.

For simulations employing T-step NCMC moves, proposals were made by selecting a new velocity vector from the Maxwell–Boltzmann distribution, integrating T steps of velocity Verlet dynamics (23) for all bath atoms as the dimer extension was driven from r_{old} to r_{new} in equal steps of size $\Delta r/T$, and accepting or rejecting based on the modified Metropolis criteria for symplectic integrators (Eqs. **12–20**),

$$A(X) = \min\{1, e^{-\beta[H(x_T)-H(x_0)]}J_r(x_0, x_T)\}. \quad [28]$$

The Jacobian ratio is also $(r_{new}/r_{old})^2$. This scheme corresponds to a choice for the perturbation kernel of

$$\alpha_t(x, y) = \left[\frac{r(y)}{r(x)}\right]^2 \delta([r(y) - r(x)] - [\Delta r/T]), \quad [29]$$

where $r(x)$ denotes the dimer separation of configuration x. The propagation kernel $K_\tau(x, y)$ corresponds to velocity Verlet dynamics where the dimer atoms are held fixed in space. The final configuration after the MC or NCMC rejection procedure was stored and plotted to generate Fig. 2.

The mean acceptance probabilities for each switching time τ can be estimated via the sample mean

$$\langle A\rangle_\tau \approx \frac{1}{N}\sum_{n=1}^{N} A(X_n). \quad [30]$$

For numerical stability, logarithms of $A(X_n)$ were stored, as $a_n \equiv \ln A(X_n)$. We then estimated $\ln\langle A\rangle_\tau$ (shown in Fig. 4) as

$$\ln\langle A\rangle_\tau \approx -\ln N + \ln b + \sum_{n=1}^{N} e^{a_n - b}, \quad [31]$$

where $b \equiv \max_n a_n$.

Integrated autocorrelation times were estimated using the rapid scheme described in section 5.2 of ref. 42.

The acceptance probabilities plotted in Fig. 4 were estimated from a trajectory consisting of 10,000 iterations of 2,048-step NCMC, with 500 steps of GHMC dynamics in between each NCMC trial, to ensure equilibrium sampling. Prior to each 2,048-step NCMC iteration, a velocity from the Maxwell–Boltzmann distribution was selected, and NCMC trial moves with varying switching times were conducted solely to accumulate statistics. The statistical error in the estimate of $\ln\langle A\rangle_\tau$ and the computed relative efficiency over instantaneous Monte Carlo was estimated by 1,000 bootstrap trials, in which the dataset of 10,000 work samples was resampled with replacement in each bootstrap trial and 95% confidence intervals computed from the distribution over bootstrap replicates.

The reference distribution for the interparticle distribution $\mathscr{P}(r)$ plotted in red on the right side of Fig. 2 was computed analytically for the vacuum system,

$$\mathscr{P}_{vac}(r) \propto 4\pi r^2 e^{-\beta U_{bond}(r)}. \quad [32]$$

For the solvated system, this distribution was estimated from an umbrella sampling simulation (Fig. 5) employing a modified bonded potential intended to remove the barrier in between compact and extended states,

$$U_{umbrella}(r) = k_BT\ln r^2 + \theta(r_{min} - r)(K/2)[r - r_{min}]^2 + \theta(r - r_{max})(K/2)[r - r_{max}]^2, \quad [33]$$

where $r_{min} = r_0$, $r_{max} = 2.05r_0$, and $K = k_BT/\eta^2$, with $\eta = 0.3$ Å, and $\theta(r)$ is the Heaviside function that assumes a value of unity for $r \geq 0$ and zero otherwise. The true solvated interparticle distribution $p(r)$ was estimated by reweighting the data produced from this simulation, using the relationship

$$\mathscr{P}_{sol}(r) \propto \frac{\sum_{n=1}^{N} \delta(r - r_n)e^{-\beta[U_{bond}(r_n)-U_{umbrella}(r_n)]}}{\sum_{n=1}^{N} e^{-\beta[U_{bond}(r_n)-U_{umbrella}(r_n)]}}, \quad [34]$$

where r_n denotes the bond separation for sample n, and a finite-width histogram bin was used instead of the delta function $\delta(r)$.

Appendix

Proof that NCMC Preserves the Equilibrium Distribution. Following the proof for GHMC in ref. 36, here we show that NCMC preserves the equilibrium distribution. The expected acceptance rate for NCMC moves initiated from (x, λ) is

$$\alpha(x, \lambda) \equiv \int d\Lambda \int dX \rho(X, \Lambda|x, \lambda)A(X|\Lambda). \quad [35]$$

Suppose that we have a variate $(x^{(n)}, \lambda^{(n)})$ drawn from the equilibrium distribution $\pi(x, \lambda)$. The probability density of the next value in the chain, $p(x^{(n+1)}, \lambda^{(n+1)})$, has contributions from two scenarios: when the candidate variate is rejected and when it is accepted. The contribution from rejecting the candidate and flipping the momentum such that $(x^{(n+1)}, \lambda^{(n+1)}) = (\tilde{x}^{(n)}, \lambda^{(n)})$ is

$$\int dx \sum_{\lambda} \pi(x, \lambda)[1 - \alpha(x, \lambda)]\delta(\tilde{x} - x^{(n+1)})\delta(\lambda - \lambda^{(n+1)}) = \pi(\tilde{x}^{(n+1)}, \lambda^{(n+1)})[1 - \alpha(\tilde{x}^{(n+1)}, \lambda^{(n+1)})]. \quad [36]$$

The latter contribution from accepting the candidate such that $(x^{(n+1)}, \lambda^{(n+1)}) = (x_T, \lambda_T)$ is,

$$\begin{aligned}
&\int dx \sum_{\lambda} \pi(x,\lambda) \int dX \int d\Lambda \rho(X,\Lambda|x,\lambda) A(X|\Lambda) \delta(x_T - x^{(n+1)}) \delta(\lambda_T - \lambda^{(n+1)}) \\
&= \int dx_0 \sum_{\lambda_0} \int dX \int d\Lambda [\pi(x_0,\lambda_0) \rho(X,\Lambda|x_0,\lambda_0) A(X|\Lambda)] \delta(x_T - x^{(n+1)}) \delta(\lambda_T - \lambda^{(n+1)}) \\
&= \int dx_T \sum_{\lambda_T} \int d\tilde{X} \int d\tilde{\Lambda} [\pi(\tilde{x}_T,\lambda_T) \rho(\tilde{X},\tilde{\Lambda}|\tilde{x}_T,\lambda_T) A(\tilde{X}|\tilde{\Lambda})] \delta(x_T - x^{(n+1)}) \delta(\lambda_T - \lambda^{(n+1)}) \\
&= \pi(\tilde{x}^{(n+1)}, \lambda^{(n+1)}) \alpha(\tilde{x}^{(n+1)}, \lambda^{(n+1)}),
\end{aligned} \tag{37}$$

where $\rho(X,\Lambda|x_0,\lambda_0) \equiv \Pi(X|x_0,\Lambda) P(\Lambda|x_0,\lambda_0)$ is the probability of generating the trajectory-protocol pair (X,Λ) from (x_0,λ_0), and the pathwise detailed balance condition (Eq. **6**) is used to produce the quantity in brackets.

Taking the sum of Eqs. **36** and **37**, we find that the equilibrium distribution is preserved

$$p(x^{(n+1)}, \lambda^{(n+1)}) = \pi(x^{(n+1)}, \lambda^{(n+1)}). \tag{38}$$

By analogous reasoning, maintaining the momentum upon rejection, $(x^{(n+1)}, \lambda^{(n+1)}) = (x^{(n)}, \lambda^{(n)})$, and flipping it upon acceptance, $(x^{(n+1)}, \lambda^{(n+1)}) = (\tilde{x}_T, \lambda_T)$ will also preserve the equilibrium distribution.

Acceptance Criteria for Overdamped Langevin (Brownian) Integrator of Ermak and Yeh. A common integrator for Brownian dynamics (the overdamped regime of Langevin dynamics), in which only coordinates x are explicitly integrated, is given by Ermak and Yeh (24, 25). In our notation, where the perturbed coordinates x_t^* are propagated by one step of the stochastic integrator to obtain x_t, application of the propagation kernel $K(x_t^*, x_t)$ can be written,

$$x_t = x_t^* - \frac{\Delta t}{\gamma m} F_t(x_t^*) + \sqrt{2}\left(\frac{\Delta t}{\gamma m}\right)^{1/2} \xi_t, \tag{39}$$

where m is the particle mass, $F_t(x) \equiv -(\partial/\partial x) H_t(x)$ is the (potentially time-dependent) systematic force, and γ is an effective collision frequency or friction coefficient with units of inverse time. The noise history ξ_t for each degree of freedom is a normal random variate with zero mean and variance β^{-1}, drawn from the distribution

$$\phi(\xi_t) = \frac{1}{\sqrt{2\pi\beta^{-1}}} \exp\left[-\frac{\beta}{2}\xi_t^2\right]. \tag{40}$$

In NCMC, every application of the propagation kernel K_t produces a transition $x_t^* \to x_t$ determined by the noise history variable ξ_t, there is a corresponding $\tilde{\xi}_t$ that generates the opposite step, $x_t \to x_t^*$. By noting

$$\begin{aligned}
x_t &= x_t^* - \frac{\Delta t}{\gamma m} F_t(x_t^*) + \sqrt{2}\left(\frac{\Delta t}{\gamma m}\right)^{1/2} \xi_t \\
x_t^* &= x_t - \frac{\Delta t}{\gamma m} F_t(x_t) + \sqrt{2}\left(\frac{\Delta t}{\gamma m}\right)^{1/2} \tilde{\xi}_t,
\end{aligned} \tag{41}$$

we can compute the relationship

$$\tilde{\xi}_t = \frac{1}{\sqrt{2}}\left(\frac{\Delta t}{\gamma m}\right)^{1/2} [F_t(x_t) + F_t(x_t^*)] - \xi_t. \tag{42}$$

Then, the ratio of transition kernels appearing in Eq. **10** can be written in terms of noise history ξ_t and the computed reverse noise history ξ_t^*

$$\begin{aligned}
\Delta\mathcal{S}(X) &= -\ln \prod_{t=1}^{T} \frac{K_t(x_t, x_t^*)}{K_t(x_t^*, x_t)} = -\ln \prod_{t=1}^{T} \frac{\phi(\tilde{\xi}_t)|\frac{\partial x_t^*}{\partial \tilde{\xi}_t}|}{\phi(\xi_t)|\frac{\partial x_t}{\partial \xi_t}|} = -\ln \prod_{t=1}^{T} \frac{\phi(\tilde{\xi}_t)}{\phi(\xi_t)} \\
&= -\ln \prod_{t=1}^{T} \exp\left[-\frac{\beta}{2}(\tilde{\xi}_t^2 - \xi_t^2)\right] = \frac{\beta}{2}\sum_{t=1}^{T}(\tilde{\xi}_t^2 - \xi_t^2),
\end{aligned} \tag{43}$$

where the tildes are dropped because the microstate x contains no momenta. The quantity $|\partial x_t/\partial \xi_t|$ represents the Jacobian for the change of variables from the ξ_t to x_t, and the Jacobians in the numerator and denominator cancel. The quantity in Eq. **43** can easily be computed during integration.

Acceptance Criteria for Langevin Integrator of Brooks, Brünger, and Karplus (BBK). The Brünger–Brooks–Karplus (BBK) stochastic integrator (38, 39) is a popular integrator for simulating Langevin dynamics. In our notation, where the perturbed coordinates x_t^* are propagated by one step of the stochastic integrator to obtain x_t, application of the propagation kernel $K(x_t^*, x_t)$ can be written

$$\begin{aligned}
v_t' &= v_t^* + \frac{\Delta t}{2m}\left(F_t(r_t^*) - \gamma m v_t^* + \sqrt{\frac{2\gamma m}{\Delta t}}\xi_t\right) \\
r_t &= r_t^* + \Delta t v_t' \\
v_t &= \frac{1}{1+\frac{\gamma\Delta t}{2}}\left[v_t' + \frac{\Delta t}{2m}\left(F_t(r_t) + \sqrt{\frac{2\gamma m}{\Delta t}}\xi_t'\right)\right],
\end{aligned} \tag{44}$$

where we have used a velocity Verlet discretization of the BBK integrator (36, 40). Here r_t and v_t denote the respective Cartesian position and velocity components of the microstate x_t, γ the effective collision frequency with units of inverse time, and m the particle mass. v_t' is an auxiliary variable used only in simplifying the mathematical representation of the integration scheme. Note that we require two random variates, ξ_t and ξ_t', per degree of freedom per timestep in order for this scheme to be able to generate both the forward trajectory X and its time-reverse $\tilde{X}$ (e.g., section 2.2.3.2 of ref. 36).

The noise history terms ξ_t and ξ_t' are normal random variates with zero mean and variance β^{-1}. Their joint distribution can therefore be written

$$\psi(\xi_t, \xi_t') = \frac{1}{2\pi\beta^{-1}} \exp\left[-\frac{\beta}{2}(\xi_t^2 + \xi_t'^2)\right]. \tag{45}$$

For every step $x_t^* \to x_t$, the positions and velocities undergo a transition $(r_t^*, v_t^*) \to (r_t, v_t)$ determined by the noise variables (ξ_t, ξ_t'). A corresponding choice of noise variables $(\tilde{\xi}_t, \tilde{\xi}_t')$ will generate the

reverse step, $\tilde{x}_t \to \tilde{x}_t^*$, carrying $(r_t, -v_t) \to (r_t^*, -v_t^*)$ With a little algebra, it is seen that

$$\tilde{\xi}_t = \xi'_t - \sqrt{2\gamma m \Delta t v_t}$$
$$\tilde{\xi}'_t = \xi_t - \sqrt{2\gamma m \Delta t v_t^*}. \quad [46]$$

To write the ratio of transition kernels appearing in Eq. **10** in terms of noise variables (ξ_t, ξ'_t) and the computed reverse noise variables $(\tilde{\xi}_t, \tilde{\xi}'_t)$, we must first compute the Jacobian $J(\xi_t, \xi'_t)$ because the random variates are not in Cartesian space

$$J(\xi_t, \xi'_t) \equiv \det \begin{bmatrix} \frac{\partial r_t}{\partial \xi_t} & \frac{\partial v_t}{\partial \xi_t} \\ \frac{\partial r_t}{\partial \xi'_t} & \frac{\partial v_t}{\partial \xi'_t} \end{bmatrix}, \quad [47]$$

which can be shown to be independent of ξ_t and ξ'_t. The conditional path action difference can now be computed

$$\Delta \mathcal{S}(X) = -\ln \prod_{t=1}^{T} \frac{K_t(\tilde{x}_t, \tilde{x}_t^*)}{K_t(x_t^*, x_t)} = -\ln \prod_{t=1}^{T} \frac{\psi(\tilde{\xi}_t, \tilde{\xi}'_t) J(\tilde{\xi}_t, \tilde{\xi}'_t)}{\psi(\xi_t, \xi'_t) J(\xi_t, \xi'_t)}$$
$$= \frac{\beta}{2} \sum_{t=1}^{T} [(\tilde{\xi}_t^2 + \tilde{\xi}_t'^2) - (\xi_t^2 + \xi_t'^2)], \quad [48]$$

where the ratio of Jacobians $J(\tilde{\xi}_t, \tilde{\xi}'_t)/J(\xi_t, \xi'_t)$ cancels because the Jacobians are independent of the noise variates.

Derivation of Effective Correlation Time for Mixed MD/NCMC Sampling. For simplicity, we make the assumption that the system of interest has two long-lived conformational states of equal population with dimer extensions r_c and r_e. This assumption holds to good approximation for the WCA dimer example considered here, and may apply to other systems of interest as well. We assume that the correlation time for a fixed number of MD simulation steps per iteration is given by τ_{MD}, and describe the probability of finding the system ends up in a given conformational state after one iteration by a 2×2 column-stochastic transition matrix $\mathbf{T}_{\mathrm{MD}}$

$$\mathbf{T}_{\mathrm{MD}} = \begin{bmatrix} 1-\alpha & \alpha \\ \alpha & 1-\alpha \end{bmatrix}. \quad [49]$$

For a 2×2 system whose time evolution is governed by the column stochastic transition matrix $\mathbf{T}$, we can write the autocorrelation function for the dimer extension r as

$$C(n\Delta t) = \langle r(0) r(n\Delta t) \rangle = [r_c \quad r_e] \mathbf{T}^n \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix} \begin{bmatrix} r_c \\ r_e \end{bmatrix}$$
$$= [r_c \quad r_e] \mathbf{U} \begin{bmatrix} 1 & 0 \\ 0 & \mu^n \end{bmatrix} \mathbf{U}^{-1} \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix} \begin{bmatrix} r_c \\ r_e \end{bmatrix}$$
$$= (C_0 - C_\infty)\mu^n + C_\infty, \quad [50]$$

where the transition matrix $\mathbf{T}$ has unitary eigenvalue decomposition $\mathbf{UMU}^{-1}$, and μ is the nonunit eigenvalue of $\mathbf{T}$. The constants are $C_0 = (1/2)(r_c^2 + r_e^2)$ and $C_\infty = (1/4)(r_c + r_e)^2$.

Relating this to the autocorrelation time τ estimated from a timeseries, intended to reflect the fit to

$$C(t) = (C_0 - C_\infty) e^{-t/\tau} + C_\infty, \quad [51]$$

we can see that $\tau = -1/\ln \mu$. We then determine that the correlation time $\tau_{\mathrm{MD}} = -1/\ln \mu_{\mathrm{MD}}$, with $\mu_{\mathrm{MD}} = 1 - 2\alpha$.

Similarly, we can write the probability that the NCMC step will carry the system from one conformational state to another in terms of the acceptance probability γ, which we assume to be symmetric,

$$\mathbf{T}_{\mathrm{NCMC}} = \begin{bmatrix} 1-\gamma & \gamma \\ \gamma & 1-\gamma \end{bmatrix}, \quad [52]$$

where we have correlation time $\tau_{\mathrm{NCMC}} = -1/\ln \mu_{\mathrm{NCMC}}$ and $\mu_{\mathrm{NCMC}} = 1 - 2\gamma$.

The effective transition matrix $\mathbf{T}_{\mathrm{eff}}$ for iterations consisting of MD simulation steps followed by an NCMC trial is given by

$$\mathbf{T}_{\mathrm{eff}} = \mathbf{T}_{\mathrm{MD}} \mathbf{T}_{\mathrm{NCMC}} = \begin{bmatrix} 1-\alpha & \alpha \\ \alpha & 1-\alpha \end{bmatrix} \begin{bmatrix} 1-\gamma & \gamma \\ \gamma & 1-\gamma \end{bmatrix}$$
$$= \begin{bmatrix} 1-(\alpha+\gamma) & (\alpha+\gamma) - 2\alpha\gamma \\ (\alpha+\gamma) - 2\alpha\gamma & 1-(\alpha+\gamma) \end{bmatrix}, \quad [53]$$

where the effective correlation time $\tau_{\mathrm{eff}} = -1/\ln \mu_{\mathrm{eff}}$, with $\mu_{\mathrm{eff}} = 1 - 2[(\alpha+\gamma) - 2\alpha\gamma]$. Substituting in $\alpha = (1 - e^{-1/\tau_{\mathrm{MD}}})/2$ and $\gamma = (1 - e^{-1/\tau_{\mathrm{NCMC}}})/2$, we obtain

$$\tau_{\mathrm{eff}} = -\frac{1}{\ln[1 - (1 - e^{-1/\tau_{\mathrm{MD}}}) - (1 - e^{-1/\tau_{\mathrm{NCMC}}}) + (1 - e^{-1/\tau_{\mathrm{MD}}})(1 - e^{-1/\tau_{\mathrm{NCMC}}})]}$$
$$= -\frac{1}{\ln[e^{-1/\tau_{\mathrm{MD}}} e^{-1/\tau_{\mathrm{NCMC}}}]} = \frac{1}{\tau_{\mathrm{MD}}^{-1} + \tau_{\mathrm{NCMC}}^{-1}} = \frac{\tau_{\mathrm{MD}} \tau_{\mathrm{NCMC}}}{\tau_{\mathrm{MD}} + \tau_{\mathrm{NCMC}}}. \quad [54]$$

As a check of the accuracy of Eq. **54**, we note that for MD with 2,048-step NCMC switching, we compute $\tau_{\mathrm{eff}} \approx 4.0$ iterations, using only $\tau_{\mathrm{MD}} = 299.8$ iterations and the NCMC acceptance probability $\gamma = 12.1\%$. The actual correlation time measured from a 10,000 iteration simulation is computed to be $\tau_{\mathrm{eff}} = 4.0$.

ACKNOWLEDGMENTS. We thank Gabriel Stoltz (CERMICS, École des Ponts ParisTech); Jed W. Pitera (IBM Almaden Research Center); Manuel Athenès (CEA Saclay); Firas Hamze (D-Wave Systems); Yael Elmatad, Anna Schneider, Paul Nerenberg, Todd Gingrich, David Chandler, David Sivak, Phillip Geissler, Michael Grünwald, and Ulf Rörbæck Pederson (University of California, Berkeley); Vijay S. Pande (Stanford University); and Suriyanarayanan Vaikuntanathan, Andrew J. Ballard, and Christopher Jarzynski (University of Maryland), and Huafeng Xu (D. E. Shaw Research) for enlightening discussions on this topic and constructive feedback on this manuscript, as well as the two anonymous referees for their helpful suggestions for improving clarity. J.P.N. was supported by the US Department of Energy (DOE) by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. G.E.C. was funded by the Helios Solar Energy Research Center, which is supported by the Director, Office of Science, Office of Basic Energy Sciences of the DOE under Contract DE-AC02-05CH11231. D.D.L.M. was funded by a Director's Postdoctoral Fellowship from Argonne National Laboratory. J.D.C. was supported through a distinguished postdoctoral fellowship from the California Institute for Quantitative Biosciences (QB3) at the University of California, Berkeley. Additionally, the authors are grateful to OpenMM developers Peter Eastman, Mark Friedrichs, Randy Radmer, and Christopher Bruns for their generous help with the OpenMM GPU-accelerated computing platform and associated PyOpenMM Python wrappers. This research was performed under the auspices of the DOE by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344 and by University of Chicago Argonne, LLC, Operator of Argonne National Laboratory (Argonne). Argonne, a USDE Office of Science laboratory, is operated under Contract DE-AC02-06CH11357.

1. Liu JS (2002) *Monte Carlo Strategies in Scientific Computing* (Springer, New York), 2nd Ed.
2. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E (1953) Equation of state calculations by fast computing machines. *J Chem Phys* 21:1087–1092.
3. Hastings WK (1970) Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* 57:97–109.
4. Rahman A (1964) Correlations in the motion of atoms in liquid argon. *Phys Rev* 136: A405–A411.
5. Swendsen RH, Wang J-S (1987) Nonuniversal critical dynamics in Monte Carlo simulations. *Phys Rev Lett* 58:86–88.
6. Wolff U (1989) Collective Monte Carlo updating for spin systems. *Phys Rev Lett* 62:361–364.
7. Neal RM (2004) Taking bigger Metropolis steps by dragging fast variables. (Department of Statistics, University of Toronto, Toronto), Technical Report 0411.
8. Andrieu C, Doucet A, Holenstein R (2010) Particle Markov chain Monte Carlo methods. *J R Stat Soc Series B Stat Methodol* 72:269–342.
9. Athènes M (2002) Computation of a chemical potential using a residence weight algorithm. *Phys Rev E* 66:046705.
10. Bürgi R, Kollman PA, van Gunsteren WF (2002) Simulating proteins at constant pH: An approach combining molecular dynamics and Monte Carlo simulation. *Proteins* 47:469–480.
11. Stern HA (2007) Molecular simulation with variable protonation states at constant pH. *J Chem Phys* 126:164112.
12. Nilmeier J, Jacobson MP (2009) Monte Carlo sampling with hierarchical move sets: POSH Monte Carlo. *J Chem Theor Comput* 1968–1984.
13. Opps SB, Schofield J (2001) Extended state-space Monte Carlo methods. *Phys Rev E* 63:56701.
14. Brown S, Head-Gordon T (2003) Cool walking: A new Markov chain Monte Carlo sampling method. *J Comput Chem* 24:68–76.
15. Ballard AJ, Jarzynski C (2009) Replica exchange with nonequilibrium switches. *Proc Natl Acad Sci USA* 106:12224–12229.
16. Lyubartsev AP, Martsinovski AA, Shevkunov SV, Vorontsov-Velyaminov PN (1992) New approach to Monte Carlo calculation of the free energy: Method of expanded ensembles. *J Chem Phys* 96:1776–1783.
17. Park S (2008) Comparison of the serial and parallel algorithms of generalized ensemble simulations: An analytical approach. *Phys Rev E* 77:016709.
18. Shirts MR, Chodera JD (2008) Statistically optimal analysis of samples from multiple equilibrium states. *J Chem Phys* 129:124105.
19. Marinari E, Parisi G (1992) Simulated tempering: A new Monte Carlo scheme. *Europhys Lett* 19:451–458.
20. Li H, Fajer M, Yang W (2007) Simulated scaling method for localized enhanced sampling and simultaneous "alchemical" free energy simulations: A general method for molecular mechanical, quantum mechanical, and quantum mechanical/molecular mechanical simulations. *J Chem Phys* 126:024106.
21. Mongan J, Case DA, McCammon JA (2004) Constant pH molecular dynamics in generalized Born implicit solvent. *J Comput Chem* 25:2038–2048.
22. Jarzynski C (2000) Hamiltonian derivation of a detailed fluctuation theorem. *J Stat Phys* 98:77–102.
23. Swope WC, Andersen HC, Berens PH, Wilson KR (1982) A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *J Chem Phys* 76:637–649.
24. Ermak DL, Yeh Y (1974) Equilibrium electrostatic effects on the behavior of polyions in solution: Polyion-mobile ion interaction. *Chem Phys Lett* 24:243–248.
25. Ermak DL (1975) A computer simulation of charged particles in solution. I. Technique and equilibrium properties. *J Chem Phys* 62:4189–4196.
26. Frenkel D (2004) Speed-up of Monte Carlo simulations by sampling of rejected states. *Proc Natl Acad Sci USA* 101:17571–17575.
27. Duane S, Kennedy AD, Pendleton BJ, Roweth D (1987) Hybrid Monte Carlo. *Phys Lett B* 195:216–222.
28. Lelièvre T, Stoltz G, Rousset M (2010) Langevin dynamics with constraints and computation of free energy differences. (Imperial College Press, London).
29. Crooks GE (1998) Nonequilibrium measurements of free energy differences for microscopically reversible markovian systems. *J Stat Phys* 90:1481–1487.
30. Jarzynski C (1997) Nonequilibrium equality for free energy differences. *Phys Rev Lett* 78:2690–2693.
31. Lechner W, Oberhofer H, Dellago C, Geissler PL (2006) Equilibrium free energies from fast-switching trajectories with large time steps. *J Chem Phys* 124:044113.
32. Andersen HC (1983) Rattle: A "velocity" version of the Shake algorithm for molecular dynamics calculations. *J Comput Phys* 52:24–34.
33. Leimkuhler B, Skeel RD (1994) Symplectic numerical integrators in constrained Hamiltonian systems. *J Comput Phys* 112:117–125.
34. Paterlini MG, Ferguson DM (1998) Constant temperature simulations using the Langevin equation with velocity Verlet integration. *Chem Phys* 236:243–252.
35. Izaguirre JA, Sweet CR, Pande VS (2010) Multiscale dynamics of macromolecules using normal mode Langevin. *Pac Symp Biocomput* 15:240–251.
36. Lelièvre T, Stoltz G, Rousset M (2010) *Free Energy Computations: A Mathematical Perspective* (Imperial College Press, London), 1st Ed.
37. Gustafson P (1998) A guided walk Metropolis algorithm. *Stat Comput* 8:357–364.
38. Brünger A, Brooks CL, III, Karplus M (1984) Stochastic boundary conditions for molecular dynamics simulations of ST2 water. *Chem Phys Lett* 105:495–500.
39. Pastor RW, Brooks Br, Szabo A (1988) An analysis of the accuracy of Langevin and molecular dynamics algorithms. *Mol Phys* 65:1409–1419.
40. Schlick T (2002) *Molecular Modeling and Simulation: An Interdisciplinary Guide* (Springer, New York), 1st Ed.
41. Weeks JD, Chandler D, Andersen HC (1971) Role of repulsive forces in determining the equilibrium structure of simple liquids. *J Chem Phys* 54:5237–5247.
42. Chodera JD, Swope WC, Pitera JW, Seok C, Dill KA (2007) Use of the weighted histogram analysis method for the analysis of simulated and parallel tempering simulations. *J Chem Theor Comput* 3:26–41.
43. Bolhuis PG, Chandler D, Dellago C, Geissler PL (2002) Transition path sampling: Throwing ropes over rough mountain passes, in the dark. *Annu Rev Phys Chem* 53:291–318.
44. Best RB, Hummer G (2005) Reaction coordinates and rates from transition paths. *Proc Natl Acad Sci USA* 102:6732–6737.
45. Ma A, Dinner AR (2005) Automatic method for identifying reaction coordinates in complex systems. *J Phys Chem B* 109:6769–6779.
46. E W, Ren W, Vanden-Eijnden E (2005) Finite temperature string method for the study of rare events. *J Phys Chem B* 109:6688–6693.
47. Berezhkovskii A, Szabo A (2005) One-dimensional reaction coordinates for diffusive activated rate processes in many dimensions. *J Chem Phys* 122:014503.
48. Peters B, Trout BL (2006) Obtaining reaction coordinates by likelihood maximization. *J Chem Phys* 125:054108.
49. Ensing B, de Vivo M, Liu Z, Moore P, Klein ML (2006) Metadynamics as a tool for exploring free energy landscapes of chemical reactions. *Acc Chem Res* 39:73–81.
50. Schmiedl T, Seifert U (2007) Optimal finite-time processes in stochastic thermodynamics. *Phys Rev Lett* 98:108301.
51. Then H, Engel A (2008) Computing the optimal protocol for finite-time processes in stochastic thermodynamics. *Phys Rev E* 77:041105.
52. Gomez-Marin A, Schmiedl T, Seifert U (2008) Optimal protocols for minimal work processes in underdamped statistical thermodynamics. *J Chem Phys* 129:024114.
53. Salamon P, Berry RS (1983) Thermodynamic length and dissipated availability. *Phys Rev Lett* 51:1127–1130.
54. Crooks GE (2007) Measuring thermodynamic length. *Phys Rev Lett* 99:100602.
55. Feng EH, Crooks GE (2009) Far-from-equilibrium measurements of thermodynamic length. *Phys Rev E* 79:012104.
56. Minh DDL, Chodera JD (2011) Multiple-timeslice nonequilibrium estimators: Estimating equilibrium ensemble averages from nonequilibrium experiments using multiple timeslices, with application to thermodynamic length. *J Chem Phys* 134:024111.
57. Shenfeld DK, Xu H, Eastwood MP, Dror RO, Shaw DE (2009) Minimizing thermodynamic length to select intermediate states for free-energy calculations and replica-exchange simulations. *Phys Rev E* 80:046705.
58. Minh DDL, Adib AB (2008) Optimized free energies from bidirectional single-molecule force spectroscopy. *Phys Rev Lett* 100:180602.
59. Minh DDL, Chodera JD (2009) Optimal estimators and asymptotic variances for nonequilibrium path-ensemble averages. *J Chem Phys* 131:134110.
60. Crooks GE (1999) Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Phys Rev E* 60:2721–2726.
61. Crooks GE (2000) Path-ensemble averages in systems driven far from equilibrium. *Phys Rev E* 61:2361–2366.
62. Athènes M, Marinica M-C (2010) Free energy reconstruction from steered dynamics without post-processing. *J Chem Phys* 229:7129–7146.
63. Friedrichs MS, et al. (2009) Accelerating molecular dynamic simulation on graphics processing units. *J Comput Chem* 30:864–872.
64. Eastman P, Pande VS (2010) OpenMM: A hardware-independent framework for molecular simulations. *Comput Sci Eng* 12:34–39.
65. Eastman P, Pande VS (2010) Efficient nonbonded interactions for molecular dynamics of a graphics processing unit. *J Comput Chem* 31:1268–1272.
66. Bruns CM, Radmer RA, Chodera JD, Pande VS (2010) PyOpenMM. http://simtk.org/home/pyopenmm.

Corrections

CHEMISTRY, BIOPHYSICS AND COMPUTATIONAL BIOLOGY

Correction for "Nonequilibrium candidate Monte Carlo is an efficient tool for equilibrium simulation," by Jerome P. Nilmeier, Gavin E. Crooks, David D. L. Minh, and John D. Chodera, which appeared in issue 45, November 8, 2011, of *Proc Natl Acad Sci USA* (108:E1009–E1018; first published October 24, 2011; 10.1073/pnas.1106094108).

The authors note that Reference 62 appeared incorrectly. The corrected reference appears below.

Additionally, the authors note that Equations 39, 41, and 42 appeared incorrectly. The corrected equations appear below.

$$x_t = x_t^* + \frac{\Delta t}{\gamma m} F_t(x_t^*) + \sqrt{2}\left(\frac{\Delta t}{\gamma m}\right)^{1/2} \xi_t, \qquad [39]$$

$$\begin{aligned} x_t &= x_t^* + \frac{\Delta t}{\gamma m} F_t(x_t^*) + \sqrt{2}\left(\frac{\Delta t}{\gamma m}\right)^{1/2} \xi_t \\ x_t^* &= x_t + \frac{\Delta t}{\gamma m} F_t(x_t) + \sqrt{2}\left(\frac{\Delta t}{\gamma m}\right)^{1/2} \tilde{\xi}_t, \end{aligned} \qquad [41]$$

$$\tilde{\xi}_t = -\frac{1}{\sqrt{2}}\left(\frac{\Delta t}{\gamma m}\right)^{1/2} [F_t(x_t) + F_t(x_t^*)] - \xi_t. \qquad [42]$$

62. Athènes M, Marinica M-C (2010) Free energy reconstruction from steered dynamics without post-processing. *J Comput Phys* 229:7129–7146.

www.pnas.org/cgi/doi/10.1073/pnas.1207617109

BIOPHYSICS AND COMPUTATIONAL BIOLOGY

Correction for "Successful prediction of the intra- and extracellular loops of four G-protein-coupled receptors," by Dahlia A. Goldfeld, Kai Zhu, Thijs Beuming, and Richard A. Friesner, which appeared in issue 20, May 17, 2011, of *Proc Natl Acad Sci USA* (108:8275–8280; first published May 2, 2011; 10.1073/pnas.1016951108).

The authors note that their conflict of interest statement was omitted during publication. The authors declare that R.A.F. is a founder of Schrodinger, Inc.

www.pnas.org/cgi/doi/10.1073/pnas.1207511109