

# Dynamical reweighting: Improved estimates of dynamical properties from simulations at multiple temperatures

John D. Chodera,<sup>1,\*</sup> William C. Swope,<sup>2,†</sup> Frank Noé,<sup>3,‡</sup>  
Jan-Hendrik Prinz,<sup>4,3,§</sup> Michael R. Shirts,<sup>5,¶</sup> and Vijay S. Pande<sup>6,\*\*</sup>

<sup>1</sup>California Institute of Quantitative Biosciences (QB3), University of California,  
Berkeley, 260J Stanley Hall, Berkeley, California 94720, USA

<sup>2</sup>IBM Almaden Research Center, San Jose, CA 95120

<sup>3</sup>DFG Research Center Matheon, FU Berlin, Arnimallee 6, 14195 Berlin, Germany

<sup>4</sup>IWR, University of Heidelberg, INF 368, 69120 Heidelberg, Germany

<sup>5</sup>Department of Chemical Engineering, University of Virginia, VA 22904

<sup>6</sup>Department of Chemistry, Stanford University, Stanford, CA 94305

(Dated: January 18, 2011)

Dynamical averages based on *functionals* of dynamical trajectories, such as time-correlation functions, play an important role in determining kinetic or transport properties of matter. At temperatures of interest, the expectations of these quantities are often dominated by contributions from rare events, making the precise calculation of these quantities by molecular dynamics simulation difficult. Here, we present a reweighting method for combining simulations from multiple temperatures (or from simulated or parallel tempering simulations) to compute an optimal estimate of the dynamical properties at the temperature of interest without the need to invoke an approximate kinetic model (such as the Arrhenius law). Continuous and differentiable estimates of these expectations at any temperature in the sampled range can also be computed, along with an assessment of the associated statistical uncertainty. For rare events, aggregating data from multiple temperatures can produce an estimate of the desired precision at greatly reduced computational cost compared with simulations conducted at a single temperature. Here, we describe use of the method for the canonical (NVT) ensemble using four common models of dynamics (canonical distribution of Hamiltonian trajectories, Andersen thermostating, Langevin, and overdamped Langevin or Brownian dynamics), but it can be applied to *any* thermodynamic ensemble provided the ratio of path probabilities at different temperatures can be computed. To illustrate the method, we compute a time-correlation function for solvated terminally-blocked alanine peptide across a range of temperatures using trajectories harvested using a modified parallel tempering protocol.

*Keywords:* Weighted histogram analysis method (WHAM); multiple histogram reweighting; multistate Bennett acceptance ratio (MBAR); extended bridge sampling (EBS); temperature reweighting; temperature-dependent dynamics; parallel tempering; molecular dynamics.

## I. INTRODUCTION

Dynamical properties, such as diffusion constants, position and velocity autocorrelation functions, rotational correlation times, frequency-dependent dielectric constants, and reaction or isomerization rates play a critical role in our understanding of various phenomena in chemistry and biology. Besides providing physical insight, calculation of these properties from simulations is often necessary for making comparison with spectroscopic experiments (such as FTIR, 2DIR, NMR, dynamic light scattering, or neutron correlation spectroscopy) or to make predictions about material or structural properties under conditions difficult to access experimentally or of substances yet to be created in physical form.

When these phenomena involve enthalpic barriers or entropic bottlenecks in systems obeying classical statistical mechanics, averages of these properties can be extremely slow to converge in standard molecular dynamics simulations, requiring tediously long simulations to produce estimates with the desired statistical precision. Additionally, the temperature-dependence of these properties is often of interest, requiring either the use of simulations at numerous temperatures, where each simulation must be long enough to ensure the difference in estimates between different temperatures is statistically significant, or the assumption of a kinetic model, such as Arrhenius behavior for rate processes.

In recent years, a number of algorithmic advances have helped ameliorate difficulties for the computation of equilibrium expectations or thermodynamic properties caused by the presence of significant enthalpic and entropic barriers. Chief among these have been *simulated* [1–3] and *parallel* [4–7] *tempering*, in which the time required to cross enthalpic (entropic) barriers is reduced by allowing the system to access higher (lower) temperatures during the simulation. The use of random temperature-switching proposals and a Metropolis-like criterion for their acceptance embeds the simulations in

---

\* Corresponding author; [jchodera@berkeley.edu](mailto:jchodera@berkeley.edu)

† [swope@us.ibm.com](mailto:swope@us.ibm.com)

‡ [frank.noe@fu-berlin.de](mailto:frank.noe@fu-berlin.de)

§ [jan.prinz@fu-berlin.de](mailto:jan.prinz@fu-berlin.de)

¶ [michael.shirts@virginia.edu](mailto:michael.shirts@virginia.edu)

\*\* [pande@stanford.edu](mailto:pande@stanford.edu)

a Markov chain that ensures, in the long run, the stationary distribution at thermal equilibrium is sampled at each temperature [8]. The advantages of this procedure are twofold: convergence times for averages at a *single* temperature can often be reduced by appropriate choice of temperatures to enhance mixing [8–11], and data from all temperatures can be combined to produce superior estimates of equilibrium expectations over a range of temperatures using histogram reweighting [12–15] or histogram-free [16, 17] statistical analysis methods.

While helpful in the calculation of equilibrium expectations of mechanical observables that are functions only of the atomic coordinates and momenta, the data produced by simulated and parallel tempering simulations are generally not directly useful for the computation of dynamical quantities. This is due to the fact that replica trajectories include unphysical changes in temperature, and permuting the data to separate trajectories by temperature results in discontinuous trajectories for each temperature, with a complicated correlation structure entangling the temperatures [15, 18]. The short trajectory segments generated between exchanges, however, are valid dynamical trajectories that can be used to estimate dynamical properties, as in [19].

Here, we show how trajectories from multiple temperatures (including those harvested from tempering simulations) can be reweighted to produce optimal estimates of dynamical quantities at the temperature(s) of interest, and how these estimates might be superior to those from single-temperature simulation if the crossing of enthalpic barriers or entropic bottlenecks is accelerated at some of the replica temperatures. Application of this reweighting scheme to simulated and parallel tempering simulations requires little, if any, modification of the simulation protocol. We require only that the time between exchanges is long enough to compute dynamical expectations of interest, the model of dynamics used to propagate the replica samplers in between exchanges is of a form amenable to reweighting, and the temperature spacing is close enough to permit estimation of appropriate normalization constants. We illustrate this approach by estimating a slowly-convergent property—the normalized fluctuation autocorrelation function for a conformational state of terminally-blocked solvated alanine peptide that is only sparsely populated at 300 K—and show that dynamical reweighting provides significant advantages over standard estimates from single temperatures alone.

## II. THEORY

We now lay out the main theoretical tools necessary for estimating dynamical expectations by making use of simulations at multiple temperatures. In Section II A, we provide a precise definition of the dynamical expectations we can estimate through reweighting schemes.

Next, in Section II B, we review the generalized path estimator for making optimal use of trajectories sampled from multiple ensembles. As this estimator requires we compute path action differences at different temperatures for every trajectory sampled, Section II C presents simplified, convenient forms of these quantities for several common models of dynamics within the canonical ensemble. Finally, Section II D describes a modified parallel tempering protocol that can be used to easily sample trajectories from multiple temperatures in a way that they can easily be used with this estimation-by-reweighting procedure.

### A. Dynamical expectations

The equilibrium thermodynamic expectation of some static (non-kinetic) property  $A$  for a system obeying classical statistical mechanics can be written as

$$\langle A \rangle \equiv \int dx p(x) A(x) \quad (1)$$

where  $x$  denotes the instantaneous configuration of the system (such as the coordinates and momenta of all particles),  $A(x)$  is often referred to as a *phase function* or *mechanical observable*, and  $p(x)$  is the equilibrium probability density, given by

$$p(x) = Z^{-1} q(x) ; \quad Z \equiv \int dx q(x) \quad (2)$$

where  $Z$  is a normalizing constant or *partition function*, and  $q(x) > 0$  is an unnormalized probability density. (Here, we will choose  $q(x)$  to contain terms that depend explicitly on  $x$ ;  $x$ -independent multiplicative terms will generally be absorbed into  $Z$  throughout.) In the canonical ensemble, for example, the system is in contact with a heat bath and held at fixed volume, and we have  $q(x) = e^{-\beta H(x)}$ , where  $H(x)$  is the Hamiltonian and  $\beta = (k_B T)^{-1}$  is the inverse temperature.

We can write an analogous expression for equilibrium *dynamical expectations* of a kinetic property  $\mathcal{A}$  as

$$\langle \mathcal{A} \rangle = \int dX p[X] \mathcal{A}[X] \quad (3)$$

where  $\mathcal{A}[X]$  now denotes a *functional* of a trajectory  $X \equiv x(t)$ ,  $p[X]$  is the probability density of trajectories, and the integral is taken over the space of all such trajectories with respect to an appropriate measure  $dX$ . Analogous to the case of phase space probability densities,  $p[X]$  can also be written in terms of an unnormalized density  $q[X] > 0$ ,

$$p[X] = Z^{-1} q[X] ; \quad Z \equiv \int dX q[X] \quad (4)$$

As we will see in Section II C, the precise definition of  $p[X]$  will depend on the dynamical model under consideration.

While the expectation of *any* trajectory functional can, at least formally, be computed this way, the most common dynamical quantities of interest are *time-correlation functions* of the form

$$C_{AB}(t) \equiv \langle A(0) B(t) \rangle \quad (5)$$

for some pair of phase functions  $A$  and  $B$ . This corresponds to the choice

$$\mathcal{A}[X] \equiv A(x(0)) B(x(t)) \quad (6)$$

For practical purposes, we will henceforth presume that the functional  $\mathcal{A}[X]$  is *temporally local*, in that it can be expressed in a way that operates on a finite time interval that can be bounded by some fixed duration  $\tau$ . We can then restrict ourselves to considering trajectory objects  $X \equiv x(t)$  of fixed length  $\tau$ , so that  $x(t)$  is defined only on the interval  $t \in [0, \tau]$ . In the case of time-correlation functions  $C_{AB}(t)$  above, for example,  $t \leq \tau$ .

Expectations with respect to some altered trajectory probability density  $p_*[X]$  can also be computed within this equilibrium expectation framework by simply incorporating the weighting factor  $p_*[X]/p[X]$  into the trajectory functional  $\mathcal{A}$ . For example, to compute time-correlation functions such as Eq. 6 with respect to *nonequilibrium* initial conditions,  $\mathcal{A}[X]$  could be redefined to include a factor of  $\rho[X]/p[X]$  to account for non-equilibrium (or even non-canonical) initial phase space density  $\rho[X]$  [20].

## B. Dynamical reweighting

The groundwork for reweighting *trajectories* sampled from multiple thermodynamic states in order to produce an optimal estimate of some dynamical expectation was laid in work by Minh and Chodera [21]. Though presented in the context of nonequilibrium experiments (in which the consequences of the Crooks fluctuation theorem [22, 23] were explored), the estimator is sufficiently general that it applies to equilibrium trajectories sampled from different equilibrium thermodynamic states within the same thermodynamic ensemble, producing an asymptotically optimal estimate of properties within the thermodynamic state of interest. The generalized path ensemble estimator [21] is in turn based on the statistical inference framework of extended bridge sampling [24–26], which provides a solid statistical foundation for earlier estimation and reweighting schemes found in statistical physics and chemistry [12–14, 16, 27].

Here, we briefly review the general estimator formalism and examine its application to common schemes used to model dynamics at constant temperature. While we restrict our consideration to the canonical (NVT) ensemble, extension to other thermodynamic ensembles (such as the isobaric and semigrand-canonical ensembles) and other schemes for generating trajectories

within the canonical and other thermodynamic ensembles is operationally straightforward.

Suppose we have  $K$  path ensembles at different thermodynamic conditions, indexed by  $i \in \{1, \dots, K\}$ , characterized by trajectory probability densities

$$p_i[X] = Z_i^{-1} q_i[X] ; Z_i \equiv \int dX q_i[X] \quad (7)$$

where  $q_i[X] > 0$  is an unnormalized density and  $Z_i$  an unknown normalization constant, from which we have collected  $N_i$  trajectories of duration  $\tau$ . We denote these trajectories  $X_n$ , where the trajectory index  $n$  runs from 1 to  $N \equiv \sum_{k=1}^K N_k$ , with the trajectories from different path ensembles indexed in *arbitrary* order. The association of a trajectory with the path ensemble from which it was sampled will not be relevant in the estimating equations.

The optimal estimator for a dynamical expectation

$$\langle \mathcal{A} \rangle_i = \int dX p_i[X] \mathcal{A}[X] \quad (8)$$

was shown by Minh and Chodera [21] to be given by

$$\hat{\mathcal{A}}_i = \sum_{n=1}^N w_{ni} \mathcal{A}[X_n] \quad (9)$$

where the  $N \times K$  weight matrix  $W \equiv (w_{ni})$  containing the appropriate trajectory weights for all trajectories  $n$  in all path ensembles  $i$  is given by

$$w_{ni} = \hat{Z}_i^{-1} \left[ \sum_{k=1}^K N_k \hat{Z}_k^{-1} q_k[X_n]/q_i[X_n] \right]^{-1}. \quad (10)$$

Here,  $w_{ni}$  denotes the weight contribution from trajectory  $n$  in the aggregated pool of  $N$  trajectories for estimating expectations for state  $i$ . The presence of the leading normalizing factor  $\hat{Z}_i^{-1}$  ensures that the weights are normalized such that  $\sum_{n=1}^N w_{ni} = 1$  for any state  $i$ . The unnormalized path densities  $q_k[X_n]/q_i[X_n]$  represent the ratio of how likely a particular trajectory  $X_n$  is to appear in path ensemble  $k$  over path ensemble  $i$ , up to some ratio of normalizing constants  $Z_i/Z_k$  that is trajectory-independent. Importantly, the estimator expression in Eqs. 9–10 holds even if no trajectories are sampled from some of the path ensembles, such that  $N_i = 0$  for these unsampled ensembles, but there are still samples from other path ensembles such that  $N > 0$ .

The normalizing constants  $\{\hat{Z}_i\}$ , determined only up to an arbitrary multiplicative constant, are determined by solving the coupled set of  $K$  nonlinear equations for  $i = 1, \dots, K$ , under the constraint that  $\sum_{n=1}^N w_{ni} = 1$ :

$$\hat{Z}_i = \sum_{n=1}^N \left[ \sum_{k=1}^K N_k \hat{Z}_k^{-1} q_k[X_n]/q_i[X_n] \right]^{-1} \quad (11)$$

This can be done efficiently through a number of methods [17, 21]. For example, the simplest such approach is

to iterate a form of Eq. 11 to self-consistency. Suppose we choose an initial guess  $\hat{Z}_i^{(0)} = 0$ ,  $i = 1, \dots, K$ . We can employ an iterative update procedure,

$$\hat{Z}_i^{(n+1)} = \sum_{n=1}^N \left[ \sum_{k=1}^K N_k [\hat{Z}_k^{(n)}]^{-1} q_k[X_n] / q_i[X_n] \right]^{-1} \quad (12)$$

to generate a new set of estimates  $\{\hat{Z}_i^{(n+1)}\}$  from a previous set of estimates  $\{\hat{Z}_i^{(n)}\}$ . Each iteration of the update procedure requires only that we be able to compute the ratio  $q_k[X_n]/q_i[X_n]$  for all trajectories in the pool of  $N$  trajectories; these ratios can be precomputed and stored. This iterative procedure is continued until these estimates converge to within some specified tolerance [17, 21]. For numerical stability, it is convenient to work with  $\ln Z_i^{(n)}$  instead of  $Z_i^{(n)}$  directly [17].

In the canonical ensemble, the probability distributions are parameterized by a temperature  $T$ , or equivalently, the inverse temperature  $\beta \equiv (k_B T)^{-1}$ . The expectation of observable  $\mathcal{A}$  at some *arbitrary* inverse temperature  $\beta$ ,  $\langle \mathcal{A} \rangle_\beta$ , estimated from simulations at fixed inverse temperatures  $\beta_1, \dots, \beta_K$ , can then be estimated as

$$\hat{\mathcal{A}}(\beta) = \sum_{n=1}^N w_n(\beta) \mathcal{A}[X_n] \quad (13)$$

where the temperature-dependent trajectory weights  $w_n(\beta)$  are

$$w_n(\beta) = \hat{Z}(\beta)^{-1} \left[ \sum_{k=1}^K N_k \hat{Z}_k^{-1} q[X_n|\beta_k] / q[X_n|\beta] \right]^{-1} \quad (14)$$

with normalization constants  $Z_k \equiv Z(\beta_k)$ , and

$$\hat{Z}(\beta) = \sum_{n=1}^N \left[ \sum_{k=1}^K N_k \hat{Z}_k^{-1} q[X_n|\beta_k] / q[X_n|\beta] \right]^{-1} \quad (15)$$

Once the normalizing constants  $\{\hat{Z}_k\}$  have been determined by solving the coupled nonlinear equations in Eq. 11, no further nonlinear equation solution iterations are necessary to estimate  $\hat{\mathcal{A}}(\beta)$  at other temperatures of interest.

Perhaps surprisingly, Eqs. 10 and 14 do not contain any information linking a trajectory  $X_n$  with the path ensemble  $p_j[X]$  from which it was sampled; this fact is a direct, if unobvious, consequence of the extended bridge sampling formalism [25, 26].

The unnormalized trajectory probability density  $q[X|\beta]$  will depend on the model of dynamics within the canonical ensemble used in the simulations; we give expressions for several popular models in Section II C. In all the cases treated here, however,  $q[X|\beta]$  is a continuous and differentiable function of  $\beta$ , meaning that  $\mathcal{A}(\beta)$  will also be continuous and differentiable. The utility of this estimate will depend on a number of contributing circumstances; the estimate is only expected to be

reliable within the range of temperatures sampled, but estimates can, in principle, be obtained for any temperature.

The statistical uncertainty in  $\hat{\mathcal{A}}(\beta)$  can be computed in a straightforward manner from an estimate of the asymptotic variance of  $\hat{\mathcal{A}}(\beta)$  [17, 21]. Briefly, the estimating equations are linearized about the optimal estimator and the variance in the estimating equations is propagated into the corresponding variance in the estimator (see, for example, the Appendix in [26].) Operationally, the  $N \times K$  weight matrix  $W$  is augmented by two additional columns, indexed (for convenience) by  $a$  and  $A$ , consisting of

$$w_{na} = w_n(\beta) ; w_{nA} = \frac{\mathcal{A}[X_n]}{\hat{\mathcal{A}}(\beta)} w_{na} \quad (16)$$

where  $\beta$  is the temperature of interest.

The variance in  $\hat{\mathcal{A}}(\beta)$  can be computed from the asymptotic covariance matrix  $\hat{\Theta}$ , which is estimated by

$$\hat{\Theta} \equiv \mathbf{W}^T \cdot [\mathbf{I}_N - \mathbf{W}\mathbf{N}\mathbf{W}^T]^+ \cdot \mathbf{W} \quad (17)$$

where  $\mathbf{I}_N$  is the  $N \times N$  identity matrix,  $\mathbf{N} = \text{diag}(N_1, N_2, \dots, N_K, 0, 0)$ , and  $\mathbf{A}^+$  denotes the pseudoinverse of a matrix  $\mathbf{A}$ . For cases where the columns of  $\mathbf{W}$  are linearly independent (which will be the case when  $\beta \neq \beta_i$  for all  $i \in \{1, \dots, K\}$  and  $\beta_i \neq \beta_j$  for  $i \neq j$ ), a simpler expression that requires pseudoinversion of only  $\mathcal{O}(K) \times \mathcal{O}(K)$  can be used:

$$\hat{\Theta} = [(\mathbf{W}^T \mathbf{W})^+ - \mathbf{N}]^+ \quad (18)$$

Efficient computation of  $\hat{\Theta}$  by other means is discussed in the Appendix of [17].

The uncertainty in  $\hat{\mathcal{A}}(\beta)$  is then estimated as

$$\begin{aligned} \text{var } \hat{\mathcal{A}}(\beta) &\equiv \left\langle \left( \hat{\mathcal{A}}(\beta) - \langle \hat{\mathcal{A}}(\beta) \rangle \right)^2 \right\rangle \\ &\approx [\hat{\mathcal{A}}(\beta)]^2 \left[ \hat{\Theta}_{aa} - 2\hat{\Theta}_{aA} + \hat{\Theta}_{AA} \right] \end{aligned} \quad (19)$$

Near  $\hat{\mathcal{A}}(\beta) \equiv 0$ , the uncertainty estimation in Eq. 13 may run into numerical issues; in this case, it is recommended that the relative uncertainty  $\text{var } \hat{\mathcal{A}}(\beta) / [\hat{\mathcal{A}}(\beta)]^2$  be computed instead.

Similarly, the covariance between two estimates  $\hat{\mathcal{A}}(\beta)$  and  $\hat{\mathcal{B}}(\beta')$  at respective temperatures  $\beta$  and  $\beta'$  (which may be the same or different) can be computed by augmenting  $W$  with additional columns [17, 21],

$$w_{nb} = w_n(\beta') ; w_{nB} = \frac{\mathcal{B}[X_n]}{\hat{\mathcal{B}}(\beta')} w_{nb} \quad (20)$$

yielding

$$\begin{aligned} \text{cov} \left( \hat{\mathcal{A}}(\beta), \hat{\mathcal{B}}(\beta') \right) &\equiv \left\langle \left( \hat{\mathcal{A}}(\beta) - \langle \hat{\mathcal{A}}(\beta) \rangle \right) \left( \hat{\mathcal{B}}(\beta') - \langle \hat{\mathcal{B}}(\beta') \rangle \right) \right\rangle \\ &\approx \hat{\mathcal{A}}(\beta) \hat{\mathcal{B}}(\beta') \left[ \hat{\Theta}_{ab} - \hat{\Theta}_{aB} - \hat{\Theta}_{Ab} + \hat{\Theta}_{AB} \right] \end{aligned} \quad (21)$$

This can be extended to more than two expectations if the covariance of many such pairs is desired simultaneously [17, 21].

For purposes of numerical stability, it is convenient to work with the logarithms of quantities like  $q[X|\beta_i]/q[X|\beta_j]$  and  $\hat{Z}_i/\hat{Z}_j$ . We therefore define a trajectory *action* functional  $S[X|\beta]$  as

$$S[X|\beta] \equiv -\ln q[X|\beta] \quad (22)$$

Note that this definition differs from standard definitions (e.g. [28]) in that only terms that depend on *both* trajectory  $X$  and temperature  $\beta$  are included in  $q[X|\beta]$  and  $S[X|\beta]$ —any remaining multiplicative terms in the path probability density  $p[X|\beta]$  are subsumed by the normalization constant  $Z(\beta)$  to simplify the subsequent development. Inclusion of explicit multiplicative terms that depend on  $\beta$  but not  $X$  could reduce the variance of dynamical reweighting estimates, but this is not done here for simplicity.

For all of the models of dynamics considered in Section II C, the trajectory- and temperature-dependent action  $S[X|\beta]$  is *linear* in the inverse temperature  $\beta$ , having the form

$$S[X|\beta] = \beta\mathcal{H}[X] \quad (23)$$

where the quantity  $\mathcal{H}[X]$  plays the role of a generalized path Hamiltonian that operates on trajectory space, analogous to the role of the standard Hamiltonian  $H(x)$  that operates on microstates in phase space. We note that the path Hamiltonian can be trivially computed from the action  $S[X|\beta]$  or unnormalized path density  $q[X|\beta]$  as

$$\mathcal{H}[X] = \beta^{-1}S[X|\beta] = -\beta^{-1}\ln q[X|\beta] \quad (24)$$

We also define the dimensionless path free energy  $f_i$  of path ensemble  $i$  as

$$f_i \equiv -\ln Z_i \quad (25)$$

The logarithms of the trajectory weights  $w_n(\beta)$  can then be written in terms of the path Hamiltonians  $\mathcal{H}[X]$  and dimensionless free energies  $f_i$  as

$$\ln w_n(\beta) = \hat{f}(\beta) - \ln \sum_{k=1}^K N_k e^{\hat{f}_k - (\beta_k - \beta)\mathcal{H}[X_n]} \quad (26)$$

where the temperature-dependent dimensionless free energy  $\hat{f}(\beta)$  is given by

$$\hat{f}(\beta) = -\ln \sum_{n=1}^N \left[ \sum_{k=1}^K N_k e^{\hat{f}_k - (\beta_k - \beta)\mathcal{H}[X_n]} \right]^{-1} \quad (27)$$

and the  $\hat{f}_k \equiv \hat{f}(\beta_k)$ . Practical notes about working with the estimation equations in a numerically stable and efficient manner are given in [17, 21]. Due to the analogous nature of estimating static equilibrium expectations using the Bennett acceptance ratio (MBAR)

method [17], `pymbar`, the Python package for performing MBAR analysis (available at <http://simtk.org/home/pymbar>), can be used for the computation of dimensionless path free energies  $\hat{f}_i$ , expectations  $\hat{A}$ , and covariances  $\text{cov}(\hat{A}, \hat{B})$  by using path actions  $S[X]$  in place of the reduced potential energy  $u(x)$ . This scheme was used for the calculations presented here.

### C. Models of dynamics in the canonical ensemble

While the path ensemble estimator (described for temperature-dependent dynamics in Section II B, and more generally in [21]) can be applied to dynamics in *any* thermodynamic ensemble in which the unnormalized trajectory probability density ratios  $q_i[X]/q_j[X]$  are finite and nonzero over the same domain, we restrict our consideration to the canonical (NVT) ensemble in classical statistical mechanics. As the concept of a thermodynamic ensemble is a purely *equilibrium* construct, it specifies only the relative probability  $p(x|\beta)$  with which a configuration or phase space point  $x$  is observed at equilibrium. To write a probability density over equilibrium trajectories at inverse temperature  $\beta$ ,  $p[X|\beta]$ , we must further specify a particular scheme for modeling dynamics within this ensemble. The choice of model used to simulate dynamics within this ensemble is not unique—many schemes can be employed that will generate the same equilibrium density in phase space  $p(x|\beta)$ , but will have different trajectory probability densities  $p[X|\beta]$ . This choice must therefore be guided by both a desire to mimic the relevant physics for the system of interest (such as whether the system must be allowed to exchange energy with an external heat reservoir during the course of dynamical evolution), balanced with computational convenience (e.g. the use of a stochastic thermostat in place of explicitly simulating a large external reservoir).

We consider four common models for simulating dynamics within the NVT ensemble for which the ratio of unnormalized trajectory probability densities  $q[X|\beta_i]/q[X|\beta_j]$  can be computed: (i) Hamiltonian dynamics with canonically-distributed initial conditions, (ii) an Andersen thermostat [29], (iii) Langevin dynamics, and (iv) overdamped Langevin (sometimes called *Brownian*) dynamics. Notably, the deterministic nature of the Nosé-Hoover thermostat [30, 31] and the way that the inverse temperature  $\beta$  appears in its equations of motion means a trajectory  $X \equiv x(t)$  generated at inverse temperature  $\beta_i$  will have identically zero probability  $p[X|\beta_j]$  at inverse temperature  $\beta_j \neq \beta_i$ , preventing its use in the reweighting scheme described here. The same is true of the Berendsen weak-coupling algorithm [32], though any usage of this form of thermal control is highly discouraged because of its failure to generate a canonical ensemble [33–35] or be ergodic [34].

### 1. Canonical distribution of Hamiltonian trajectories

Consider a collection of trajectories of length  $\tau$  in which initial phase space points  $x_0 \equiv x(0)$  are sampled from the canonical density  $p(x_0|\beta)$ , and whose dynamical evolution is governed by Hamiltonian mechanics. Physically, this corresponds to the statistics of a situation where the system is initially in contact with a thermal reservoir at inverse temperature  $\beta$ , but the evolution time  $\tau$  is short enough that effectively no heat is exchanged with the reservoir during this time. Practically, an ensemble of these trajectories can be generated in a number of ways: For example, the initial configurations  $x_0$  could be generated from the desired canonical ensemble  $p(x_0|\beta) \propto e^{-\beta H(x_0)}$  by a form of Metropolis-Hastings Monte Carlo [36] or hybrid Monte Carlo [37].

The trajectory probability density  $p[X|\beta]$  is given by

$$p[X|\beta] = Z(\beta)^{-1} e^{-\beta H(x_0)} \quad (28)$$

where  $H(x)$  denotes the Hamiltonian for the system and  $x_0 \equiv x(0)$  is the initial phase space point. While the normalization constant can be written

$$Z(\beta) = \int dx e^{-\beta H(x)} \quad (29)$$

we do not need to compute it for use in dynamical reweighting.

The path Hamiltonian  $\mathcal{H}[X]$  can then be identified (via Eq. 24) as being identical to the Hamiltonian

$$\mathcal{H}[X] = H(x_0). \quad (30)$$

We note that, while the Hamiltonian is invariant along the trajectory for true Hamiltonian dynamics, numerical integrators will not exactly preserve this property.

### 2. Andersen thermostat

The Andersen thermostat [29, 38] allows the system to exchange heat with an external thermal reservoir through stochastic collisions with virtual bath particles. These ‘‘collisions’’ simply cause the real particle’s velocity to be reassigned from a Maxwell-Boltzmann distribution corresponding to the thermostat temperature, ensuring that the canonical ensemble is sampled [29].

Two schemes are commonly used. In the first scheme, the system undergoes so-called *massive collisions*, in which all momenta are reassigned from the Maxwell-Boltzmann distribution at fixed intervals; evolution between collisions occurs by normal Hamiltonian dynamics. This collision interval can have a large effect on the kinetic properties of the system, so this scheme is usually used only for studying static thermodynamic properties.

In the second scheme, each *particle* has a fixed probability per unit time  $\nu$  of having its momenta reassigned. This corresponds to a scenario in which each particle

undergoes independent stochastic collisions with fictitious bath particles, with the system evolving according to Hamiltonian dynamics between collisions. Andersen suggested a physically-motivated collision frequency  $\nu$  could be selected based on the thermal conductivity  $\kappa$  of the sample [29]

$$\nu = \frac{2a\kappa}{3k_B\rho^{1/3}N_p^{2/3}} \quad (31)$$

where  $a$  is a dimensionless constant that depends on the shape of the system,  $\rho$  is the density, and  $N_p$  is the number of particles. Note that, as the system size increases with  $N_p$ , the collision frequency  $\nu$  decreases as  $N_p^{-2/3}$ , such that the dynamics will approach that of a Hamiltonian system for large systems. Further elaborations of this thermostating scheme attempt to model physically realistic dynamics of large solutes, such as solvated biological macromolecules, by coupling the thermostat only to solvent degrees of freedom (see, e.g., Ref. [39]). Systems with constraints will require the velocity components along the constraints to be removed prior to integration by an algorithm such as RATTLE [40].

For either thermostating scheme, the probability of sampling a trajectory  $X \equiv x(t)$  at inverse temperature  $\beta$  is therefore determined by the probability of selecting the initial configuration  $x_0 \equiv x(0)$  from equilibrium and the probability of generating the resulting sequence of random collisions. Up to irrelevant temperature-independent multiplicative constants that simply determine the probability that a given sequence of  $N_c$  particle collisions occurred, this probability is

$$p[X|\beta] \propto e^{-\beta H(x_0)} \prod_{n=1}^{N_c} \pi_v(\mathbf{v}_n|m_n, \beta) \quad (32)$$

where the index  $n$  in the product runs over all particle collisions that occurred during integration (either at regular intervals or stochastically determined), indexed in arbitrary order.  $\mathbf{v}_n$  is the velocity of the particle after the  $n$ th massive collision,  $m_n$  its mass, and  $\pi_v(\mathbf{v}|m, \beta)$  is the Maxwell-Boltzmann distribution for the velocity vector of a particle of mass  $m$  at inverse temperature  $\beta$ :

$$\pi_v(\mathbf{v}|\beta, m) \propto e^{-\beta m v^2/2} \quad (33)$$

The path Hamiltonian  $\mathcal{H}[X]$  can therefore be identified as the Hamiltonian for the initial phase space point  $x_0$  plus the new kinetic energies of all particles that have undergone collisions during the course of integration:

$$\mathcal{H}[X] = H(x_0) + \frac{1}{2} \sum_{n=1}^{N_c} m_n v_n^2 \quad (34)$$

where again,  $\mathbf{v}_n$  is the new velocity of the particle that had undergone collision indexed by  $n$ , and  $m_n$  is its mass.

### 3. Langevin dynamics

In Langevin dynamics, stochastic forces are used to mimic the influence of degrees of freedom that are not explicitly represented by particles in the system, such as solvent molecules whose influence is incorporated into the potential by mean-field interactions. The memoryless Langevin equations of motion for phase space point  $x \equiv (\mathbf{r}, \mathbf{v})$ , consisting of Cartesian atomic positions  $\mathbf{r}$  and velocities  $\mathbf{v}$ , are

$$\begin{aligned} \dot{r}_i(t) &= v_i(t) \\ m_i \dot{v}_i(t) &= F_i(\mathbf{v}(t)) - \gamma_i v_i(t) + R_i(t) \end{aligned} \quad (35)$$

where  $m_i$  is the mass associated with degree of freedom  $i$ ,  $F_i(\mathbf{q}) \equiv -\partial H/\partial q_i$  is the systematic force component, and  $\gamma_i$  is an effective collision frequency or friction coefficient for that degree of freedom, with units of inverse time (e.g. 91 ps<sup>-1</sup> for water-like viscosity).  $\gamma_i$  is related to the temperature-dependent diffusion constant  $D_i(\beta)$  by

$$D_i(\beta) = (\beta \gamma_i m_i)^{-1} \quad (36)$$

The stochastic force  $R_i(t)$  has zero mean and satisfies the fluctuation-dissipation theorem,

$$\langle R_i(t) R_j(t') \rangle = 2\beta^{-1} m_i \gamma_i \delta_{ij} \delta(t - t') \quad (37)$$

The Langevin Leapfrog integrator [41] is an accurate and stable algorithm for integrating the Langevin equations of motion. Updating of positions  $\mathbf{r}_t$  and velocities  $\mathbf{v}_t$  by a discrete timestep  $\Delta t$  is handled by the following scheme:

$$\begin{aligned} v_{i,t+\frac{1}{2}} &= a_i v_{i,t} + b_i \Delta t m_i^{-1} F_i(X_t) + c_i m_i^{-1/2} \xi_{it} \\ r_{i,t+1} &= r_{i,t} + \Delta t v_{i,t+\frac{1}{2}} \\ v_{i,t+1} &= a_i v_{i,t+\frac{1}{2}} + b_i \Delta t m_i^{-1} F_i(X_{t+1}) + c_i m_i^{-1/2} \xi_{i(t+1)} \end{aligned} \quad (38)$$

where we have defined the dimensionless constants  $a_i$ ,  $b_i$ , and  $c_i$  associated with each degree of freedom  $i$  as

$$\begin{aligned} a_i &\equiv e^{-\gamma_i \Delta t/2} \\ b_i &\equiv (1 - e^{-\gamma_i \Delta t/2})/(\gamma_i \Delta t) \\ c_i &\equiv (1 - e^{-\gamma_i \Delta t})^{1/2} \end{aligned}$$

As the collision rate  $\gamma_i \rightarrow 0$ , we have  $a_i \rightarrow 1$ ,  $b_i \rightarrow 1/2$ , and  $c_i \rightarrow 0$ , resulting in the standard leapfrog integrator scheme.

$\xi_{it}$  is a unit-bearing random normal variate with zero mean and variance  $\beta^{-1} = k_B T$ , simply denoted

$$\xi_{it} \sim \mathcal{N}(0, \beta^{-1}). \quad (39)$$

Note that  $\xi_{it}$  has units of the square-root of energy.

Propagation for  $N_t$  steps of Langevin leapfrog integration from initial configuration  $x_0 \equiv (\mathbf{r}_0, \mathbf{v}_0)$  for a system with  $N_d$  degrees of freedom requires generating the sequence of random variates  $\xi_{it}$ . Ignoring multiplicative factors not involving both trajectory  $X$  and temperature  $\beta$ , the unnormalized equilibrium probability of

sampling a trajectory  $X$  which originates at  $x_0$  and has noise history  $\xi_{it}$  at inverse temperature  $\beta$  is given by

$$q[X|\beta] = e^{-H(x_0)} \prod_{t=0}^{N_t} \prod_{i=1}^{N_d} \phi(\xi_{it}; 0, \beta^{-1}) \quad (40)$$

where  $\phi(x; \mu, \sigma^2)$  is the normal probability density

$$\phi(y; \mu, \sigma^2) \propto \exp\left[-\frac{1}{2\sigma^2}(y - \mu)^2\right] \quad (41)$$

The path Hamiltonian can therefore be computed from the initial phase space point  $x_0$  and unit-bearing noise history  $\xi_{it}$ , as

$$\mathcal{H}[X] = H(x_0) + \frac{1}{2} \sum_{t=0}^{N_t} \sum_{i=1}^{N_d} \xi_{ti}^2 \quad (42)$$

which can easily be accumulated during integration.

Note that, if  $\gamma_i = 0$  for a degree of freedom  $i$ , the corresponding noise term  $\xi_{ti}^2$  does not contribute to the integration in Eq. 38 and can be omitted from the path Hamiltonian above.

### 4. Overdamped Langevin (Brownian) dynamics

At high friction, momentum relaxation becomes fast compared to particle motion, such that inertial motion may be neglected. A common integrator for this overdamped regime, in which only coordinates  $\mathbf{q}$  are explicitly integrated, is given by Ermak and Yeh [42, 43]

$$r_{t+1,i} = r_{ti} + \frac{\Delta t}{\gamma_i m_i} F_i(x_t) + \sqrt{2} \left( \frac{\Delta t}{\gamma_i m_i} \right)^{1/2} \xi_{ti} \quad (43)$$

where the stochastic component  $\xi_{ti}$  is again a unit-bearing normal random variate with zero mean and variance  $\beta^{-1} = k_B T$ ,

$$\xi_{ti} \sim \mathcal{N}(0, \beta^{-1}) \quad (44)$$

with units of the square-root of energy, and  $\gamma_i$  is a collision rate of friction constant for degree of freedom  $i$ , with units of inverse time, where the temperature-dependent diffusion constant  $D_i$  is again related by Eq. 36.

Ignoring irrelevant multiplicative factors that do not depend on both trajectory  $X$  and temperature  $\beta$ , the unnormalized probability functional for a realization of  $N_t$  steps of this process is then

$$q[X|\beta] = e^{-\beta U(x_0)} \prod_{t=1}^{N_t} \prod_{i=1}^{N_d} \phi(\xi_{ti}; 0, \beta^{-1}) \quad (45)$$

where  $U(x) = U(\mathbf{r})$  denotes the potential energy and  $\phi(y; \mu, \sigma^2)$  again denotes the normal probability density. This allows us to identify the path Hamiltonian as

$$\mathcal{H}[X] = U(x_0) + \frac{1}{2} \sum_{t=1}^{N_t} \sum_{i=1}^{N_d} \xi_{ti}^2 \quad (46)$$

#### D. Modified parallel tempering protocol

An especially convenient way to harvest trajectories from multiple temperatures is through the use of parallel tempering [4–7]. As illustrated by Buchete and Hummer [44], the trajectory segments generated in between exchange attempts are valid dynamical trajectories sampled from the corresponding replica temperatures. Simulated tempering can likewise be employed, though we will not discuss this here.

Below, we enumerate several considerations that must be taken into account for the use of a parallel tempering protocol to collect trajectories for use in dynamical reweighting.

First, the dynamics in between exchange attempts must use a scheme amenable to reweighting, such as one of the models of dynamics within the canonical ensemble described in Section II C. While this means some popular choices, such as Nosé-Hoover dynamics [30, 31] and the Berendsen weak-coupling algorithm [45] cannot be used, considerable flexibility remains in which dynamical schemes are permitted, including Hamiltonian, Andersen, Langevin, and overdamped Langevin (Brownian) dynamics, as discussed in the previous section.

Second, the time between exchange attempts should be at least as long as  $\tau$ , the time over which the computed dynamical observable  $\mathcal{A}[X]$  is temporally local. For example, if a correlation function  $C_{AB}(t) \equiv \langle A(0)B(t) \rangle$  is desired for  $t \in [0, \tau]$ , the time between exchanges should be at least  $\tau$ . This may, in some circumstances, impact the efficiency of parallel tempering in producing uncorrelated samples. For functions  $\mathcal{A}$  that require a large  $\tau$ , much of the benefit of the exchanges in parallel tempering may be lost if few exchange attempts occur during the simulation. Often, the best choice is to make this time *exactly*  $\tau$ , since the enhanced mixing properties of parallel tempering simulations diminish as the time between exchanges grows (for fixed total simulation time) [46].

Thirdly, it is necessary that there is sufficient overlap in the distribution of path Hamiltonians sampled from neighboring temperatures for dynamical reweighting to be effective. While the standard replica-exchange acceptance criteria [7] will lead to the correct distribution of trajectories at each temperature, the fraction of accepted exchange attempts reports on the quality of overlap of potential energies, rather than the overlap of path Hamiltonians; as a result, poor reweighting performance may be obtained despite a high exchange acceptance fraction. A simple way to ensure both good overlap in path Hamiltonian between neighboring temperatures and correct sampling of the equilibrium distribution is to use a modified exchange criterion based

on the trajectory action:

$$P_{\text{exch}}(\beta_i, \beta_j) = \min \left\{ 1, \frac{p[X_i|\beta_j]p[X_j|\beta_i]}{p[X_i|\beta_i]p[X_j|\beta_j]} \right\} \\ = \min \left\{ 1, e^{-(\beta_j - \beta_i)(\mathcal{H}[X_i] - \mathcal{H}[X_j])} \right\} \quad (47)$$

where, again,  $\mathcal{H}[X] = \beta^{-1}S[X|\beta] = -\beta^{-1} \ln q[X|\beta]$  denotes the path Hamiltonian for trajectory  $X$  defined in Eq. 24. The resulting procedure can be considered a form of parallel tempering transition path sampling [47], with the difference that new trajectories are generated from old ones in the more conventional fashion of integrating equations of motion from the final configuration  $x_T$  of the previous trajectory, rather than some form of transition path sampling [28, 48]. Note that, even for the case of a canonical distribution of Hamiltonian trajectories, the modified acceptance criteria of Eq. 47 differs from the standard acceptance criteria [7], as the total Hamiltonian in the modified criteria replaces the potential energy in the standard criteria.

Attempting to swap temperatures  $\beta_i$  and  $\beta_j$  for an arbitrarily chosen pair of replicas  $(i, j)$  will ensure that the trajectories are correctly distributed from equilibrium at their new (or old) temperatures after the exchange attempt guarantees each replica visiting temperature  $\beta_i$  samples from  $p[X|\beta_i]$  in the long run [8, 49, 50]; as a result, their individual timeslices  $x_t$  will also be distributed from the corresponding equilibrium distribution  $p(x|\beta_i)$ . If there is poor overlap between the trajectory action distributions, then the exchange attempts between these temperatures will rarely succeed; on the other hand, if the overlap is good, then the exchange rate will be accepted with significant probability. Because the cost of a single exchange attempt is insignificant compared to the computational cost of propagating all  $K$  replicas by another time  $\tau$ , it is recommended that many swaps between pairs  $(i, j)$  chosen with uniform probability be attempted in order to ensure the replicas are well-mixed. Monitoring of attempted and accepted exchanges early in the simulation can help diagnose whether there is sufficient coupling between all temperatures for successful reweighting.

While it is not strictly necessary to employ this modified acceptance probability, some care must be taken to ensure that a canonical distribution is actually obtained. For example, swapping using an exchange criteria based only on potential energies (as suggested by Sugita et al. [7]) and failing to rescale velocities may lead to an incorrect distribution. On the other hand, rescaling the velocities by  $(\beta_{\text{old}}/\beta_{\text{new}})^{1/2}$  as recommended [7], or subsequently redrawing the initial velocities from a Maxwell-Boltzmann distribution at the new temperature [6, 51]—a process equivalent to “massive collisions” for the Andersen thermostat, is not problematic. However, extra care must be taken to ensure the temperatures are spaced such that there is sufficient overlap in trajectory action between neighboring temperatures.

Finally, we note that the temporally sequential trajec-



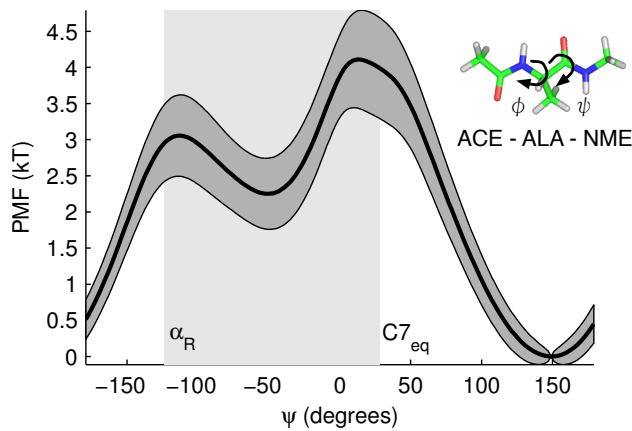


FIG. 1. **Terminally-blocked alanine in explicit solvent and potential of mean force of  $\psi$  torsion at 300 K.** The dark shaded region denotes a 95% confidence interval (two standard errors) in the free energy difference between the plotted  $\psi$  torsion and the lowest point on the potential of mean force (near  $\psi = 150$  degrees). The labeled light shaded region denotes the  $\alpha_R$  conformation. The PMF and error bars were estimated from the parallel tempering dataset using MBAR as described in the text.

tories produced by each replica will have some degree of correlation in the observable  $\mathcal{A}[X]$ . If this correlation is significant, the uncertainty estimate produced by Eqs. 19 will be an underestimate of the true statistical error, as the samples used in the estimation are assumed to be *uncorrelated*. A simple way to remove this correlation is to construct the timeseries  $A_t \equiv \mathcal{A}[X_t]$  over the sequentially-sampled trajectories  $X_t$  from a single replica (*without* having permuted the replicas to collect trajectory data by temperature) and estimate the statistical inefficiency  $g_A$  from the integrated autocorrelation time of the timeseries  $A_t$ ,  $t = 0, 1, \dots$  [17, 52–55]. If an equally-spaced subset of  $N/g_A$  trajectories from the replicas is generated, these samples will be effectively uncorrelated. Alternatively, the time between exchanges  $\tau$  can be increased to ensure that sequential  $A_t$  are decorrelated.

### III. CORRELATION FUNCTIONS OF A SOLVATED TERMINALLY-BLOCKED ALANINE PEPTIDE

To illustrate the application of dynamical reweighting to a condensed-phase system, we consider the estimation of the normalized fluctuation autocorrelation function for a long-lived conformational state of terminally-blocked alanine peptide in water that is sparsely populated at 300 K. Additional simulation details are given in Appendix A. All code, datasets, analysis scripts, and plotting scripts used for this application are made available for download at <http://simtk.org/home/dynamical-reweighting>.

The modified parallel tempering protocol described

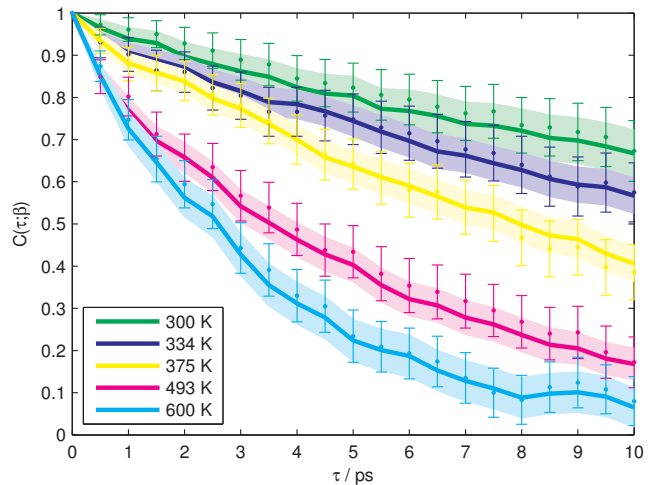


FIG. 2. **Normalized fluctuation autocorrelation functions for indicator function on  $\alpha_R$  conformation.** In both plots, single-temperature estimates are shown as points with error bars, and the dynamical reweighting estimate is shown as a solid line with shaded error interval. Error bars denote 95% confidence intervals (two standard errors).

in Section IID was used to sample many 10 ps trajectories from each of 40 exponentially-spaced temperatures spanning 300–600 K, using the “canonical distribution of Hamiltonian trajectories” model of dynamics described in Section IIC1. Replica temperatures were chosen to be exponentially spaced so that  $T_k = (T_{\max} - T_{\min}) \exp[(k-1)/(K-1)] + T_{\min}$  for  $k = 1, \dots, K$ ; this is equivalent to geometrically-spaced replicas, which would achieve equal acceptance rates among replicas if the system being simulated was an ideal gas. **These temperatures may not be optimal in terms of reducing replica correlation times or equalizing acceptance probabilities for the system studied here. While numerous schemes for intelligently choosing or adapting temperatures exist [56? ? –60], no attempt was made to do so here.** The modified protocol makes use of the exchange criteria of Eq. 47, where path Hamiltonians instead of potential energies are used in computing the probability of accepting a proposed temperature swap between replicas; this amounts to simply using the *total* energies  $H(x)$  in the exchange criteria, the sum of potential and kinetic energies, instead of only the potential energies in standard parallel tempering protocols. A total of 6720 iterations were conducted (67 ns/replica, or 2.7  $\mu$ s aggregate), where each iteration consisted of a temperature exchange phase, assignment of new velocities from the Maxwell-Boltzmann distribution at the appropriate replica temperature, and generation of a 10 ps trajectory segment.

To ensure thorough mixing of the replica associations with temperatures,  $40^3 = 64\,000$  exchange attempts among randomly-selected pairs of replicas were attempted each parallel tempering iteration. Because

exchanges are attempted among *all* temperatures, and not just neighboring ones, the probability any two temperatures are exchanged during the simulation is diagnostic of whether there is sufficient overlap among temperatures for successful reweighting; in particular, if no swaps occur between specified subsets of temperature, reweighting will produce highly uncertain estimates. Over the course of the simulation, 31.5% of proposed neighbor swaps were accepted, 4.5% of next-neighbor swaps, and 1.1% of more remote proposed exchanges accepted. The second-largest eigenvalue of the empirical temperature exchange matrix was less than unity—estimated to be 0.99801—indicating that the set of temperatures did not decompose into two or more subsets with extremely poor exchange properties between them; were there two subsets of temperatures with no transitions in between, this eigenvalue would have been unity. This indicates that a sufficient number of replicas were chosen to span the range of temperatures with reasonably good exchange among temperatures.

A statistically uncorrelated subset of trajectories generated from the modified parallel tempering simulation were extracted according to the procedure described in Section II D. The statistical inefficiency  $g$  of the time-series  $u_n$  was computed, where  $u_n$  is defined as the *effective reduced potential* for the replica ensemble, defined in terms of the instantaneous joint configuration of all replica conformations  $\mathcal{X} \equiv \{X_1, \dots, X_K\}$ :

$$u_n \equiv u(\mathcal{X}_n) = \sum_{k=1}^K \beta_k \mathcal{H}[X_{kn}] \quad (48)$$

Here,  $X_{kn}$  denotes the trajectory sampled from the replica at temperature  $\beta_k$  at iteration  $n$  of the parallel tempering simulation, and  $\mathcal{H}[X]$  the path Hamiltonian. As the quantity  $u_n$  denotes the log-probability of the overall replica-exchange joint probability distribution function  $P(X_1, \dots, X_K | \beta_1, \dots, \beta_K)$ , and the statistical inefficiency of the timeseries  $u_n$ ,  $n = 1, \dots, N$  roughly corresponds to the number of parallel tempering iterations required to generate an independent set of samples. After discarding the first 50 iterations to equilibrium, the statistical inefficiency was estimated to be 3.1 iterations.

We define the  $\alpha_R$  conformation as  $\psi \in [-124, 28)$  degrees, based on examination of the one-dimensional potential of mean force (PMF) in the  $\psi$  torsion at 300 K (Figure 1), which was estimated using the multistate Bennett acceptance ratio (MBAR) [17] and a Gaussian kernel density estimator with a bandwidth of 20 degrees. This state has a low population at 300 K ( $9.6 \pm 0.5\%$  as estimated by MBAR), but a relatively long lifetime compared to other conformational transitions in this peptide model. As an observable, we chose the indicator function for the  $\alpha_R$  state:

$$h(x) = \begin{cases} 1 & \psi(x) \in [-124, 28) \text{ degrees} \\ 0 & \text{else} \end{cases} \quad (49)$$

We compute the normalized fluctuation correlation function  $C(\tau; \beta)$ ,

$$C(\tau; \beta) \equiv \frac{\langle \delta h(0) \delta h(\tau) \rangle_\beta}{\langle \delta h^2 \rangle_\beta} \quad (50)$$

where we use the shorthand  $\delta h(t) \equiv h(x(t)) - \langle h \rangle_\beta$ . We note that, while the correlation function  $C(\tau; \beta)$  describes the relaxation dynamics out of the defined region  $\psi \in [-124, 28)$  degrees, this correlation function may contain multiple exponential timescales due to the simplistic definition of the  $\alpha_R$  stable state. Estimation of accurate rate estimates would require calculation of a quantity robust to imperfect definition of dividing surface, such as examination of the reactive flux correlation function [61, 62] or the Markov model described in a companion paper [63].

Because  $C(\tau; \beta)$  is not easily expressed as a single expectation, we write it as a combination of two elementary path expectations. For a fixed choice of  $\tau$ , we define the path functionals  $\mathcal{A}$  and  $\mathcal{B}$  as

$$\begin{aligned} \mathcal{A}[X] &\equiv h(x(0))h(x(\tau)) \\ \mathcal{B}[X] &\equiv h(x(0)) \end{aligned} \quad (51)$$

The correlation function  $C(\tau; \beta)$  can then be written as

$$C(\tau; \beta) = \frac{\langle \mathcal{A} \rangle_\beta - \langle \mathcal{B} \rangle_\beta^2}{\langle \mathcal{B} \rangle_\beta - \langle \mathcal{B} \rangle_\beta^2} \quad (52)$$

and hence its estimator, in terms of  $\hat{A}(\beta)$  and  $\hat{B}(\beta)$ , is

$$\hat{C}(\tau; \beta) = \frac{\hat{A}(\beta) - \hat{B}(\beta)^2}{\hat{B}(\beta) - \hat{B}(\beta)^2} \quad (53)$$

where we have used the fact that  $(\mathcal{B}[X])^2 = \mathcal{B}[X]$  because  $[h(x)]^2 = h(x)$ . The statistical uncertainty in  $\hat{C}(\tau; \beta)$  is determined by simple propagation of error. Suppressing the functional dependence on  $\beta$ , we have

$$\delta^2 \hat{C} = \left[ \frac{\partial \hat{C}}{\partial \hat{A}} \right]^2 \delta^2 \hat{A} + \left[ \frac{\partial \hat{C}}{\partial \hat{B}} \right]^2 \delta^2 \hat{B} + 2 \left[ \frac{\partial \hat{C}}{\partial \hat{A}} \right] \left[ \frac{\partial \hat{C}}{\partial \hat{B}} \right] \delta \hat{A} \delta \hat{B} \quad (54)$$

where the partial derivatives are given by

$$\frac{\partial \hat{C}}{\partial \hat{A}} = \frac{1}{\mathcal{B}(1 - \mathcal{B})} ; \quad \frac{\partial \hat{C}}{\partial \hat{B}} = \frac{\mathcal{A} - 1}{(\mathcal{B} - 1)^2} - \frac{\mathcal{A}}{\mathcal{B}^2} \quad (55)$$

Where the quantities  $\delta^2 \hat{A} \equiv \text{var } \hat{A}$ ,  $\delta^2 \hat{B} \equiv \text{var } \hat{B}$ , and  $\delta \hat{A} \delta \hat{B} \equiv \text{cov}(\hat{A}, \hat{B})$  represent the statistical errors (variance or covariance) of the estimators  $\hat{A}$  and  $\hat{B}$ .

Estimates for  $C(\tau; \beta)$  were computed using either the pool of uncorrelated trajectories at each replica temperature, or for a superset of temperatures (including the simulation temperatures) by dynamical reweighting as described in Section II B. For the former case, statistical error estimates were computed using sample

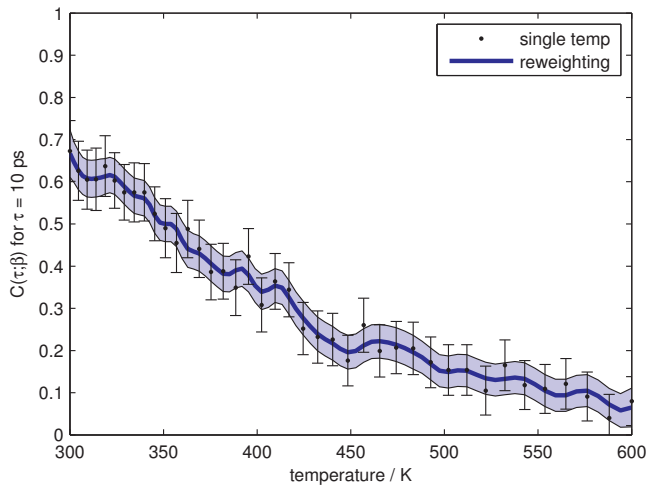


FIG. 3. **Temperature dependence of the normalized fluctuation autocorrelation function**  $C(\tau; \beta)$  for  $\tau = 10$  ps. Single-temperature estimates are shown as points with error bars denoting a 95% confidence interval; the dynamical reweighting estimate is shown as a solid line with shaded confidence interval.

(co)variances; for the latter case, statistical errors were computed using covariances obtained using Eqs. 19 and 21. Figure 2 compares correlation functions estimated for a few temperatures using single-temperature data and multiple-temperature data using dynamical reweighting.

To demonstrate how dynamical reweighting produces a smooth estimate throughout the temperature range, even at temperatures not sampled, Figure 3 shows the estimate of  $C(\tau; \beta)$  for  $\tau = 10$  ps. Clearly, the reweighted estimate produces a smoother, continuous estimate across all temperatures, with a uniformly reduced statistical uncertainty compared to the estimates computed from individual temperatures. While some features in the reweighted estimate are likely spurious, the 95% confidence interval envelope (shaded region) still permits a smooth, well-behaved curve within its boundaries.

Finally, we compare the bias and variance of single-temperature estimates with the dynamical reweighting estimate in a manner independent of the statistical error estimator. The replica-exchange trajectory was divided into 20 contiguous blocks of equal length, and  $C(\tau; \beta)$  estimated independently on each block. Using the dynamical reweighting estimate over the *entire* dataset as a reference, the bias and variance of the single-temperature and dynamical reweighting estimators were assessed by computing the estimators for each of 20 consecutive subsets of the simulation, and computing the mean deviation or standard deviation of the estimates over the blocks. The bias, which is defined as the expectation over many independent realizations of the experiment of the deviation of the estimate from the true

value, was estimated as the sample mean of the block estimates minus the estimate computed using all data; the variance, which represents the expected squared-deviation of the estimator from its expectation over many independent realizations of the experiment, was estimated by the sample variance over the blocks. Figure 4 depicts the estimated bias and variance in  $\hat{C}(\tau; \beta)$  for all times  $\tau \in [0, 10]$  and all simulation temperatures  $\beta_k$ . The advantage of dynamical reweighting is clear. Bias is minimized by virtue of the fact that contributions to the dynamical property can occur at any temperature and are incorporated with their appropriate weight, rather than relying upon small-number statistics where the events of interest may occur few or no times within the simulation at the temperature of interest. Similarly, inclusion of data from multiple temperatures reduces the overall variance component of the statistical error, resulting in improved estimates at all temperatures.

Figure 5 shows the total trajectory weight contributing from each sampled temperature as a function of target temperature  $\beta$ , where the per-trajectory weight  $w_n(\beta)$  from Eq. 14 is summed over all trajectories from the contributing sampled temperature and plotted on a base-10 logarithmic scale, normalized so that the target temperature has a total contributing weight of unity (zero on log scale). After the target temperature, neighboring temperatures contribute the most trajectory weight to averages at the target temperature, with the contribution rapidly falling off for more distant temperatures. **Note that this does not mean that distant temperatures are unhelpful, as they may still serve to reduce the effective replica correlation time if correlation times at high temperatures are especially short [8].**

Normalizing the total trajectory weight contributed from each temperature to the weight contributed from the target temperature allows us to roughly estimate how much useful data in total we are extracting by making use of reweighting in this system by summing this total relative weight over all temperatures. This quantity, averaged over all target temperatures, is  $\sim 2.4$ , suggesting that approximately 2.4 times as much data is provided in this system by making use of reweighting instead of just the data collected from the replicas that visit the target temperature. Increasing the quantity of data collected by a factor of 2.4 should decrease the standard error by  $\sim \sqrt{2.4} \approx 1.6$ , an amount consistent with the difference in confidence interval widths depicted in Figure 2.

#### IV. DISCUSSION

The dynamical reweighting scheme outlined here provides a convenient way to estimate equilibrium dynamical properties from simulations at multiple temperatures, complementing reweighting schemes for estimating static expectations (such as the related multi-state Bennett acceptance ratio, or MBAR, method [17],

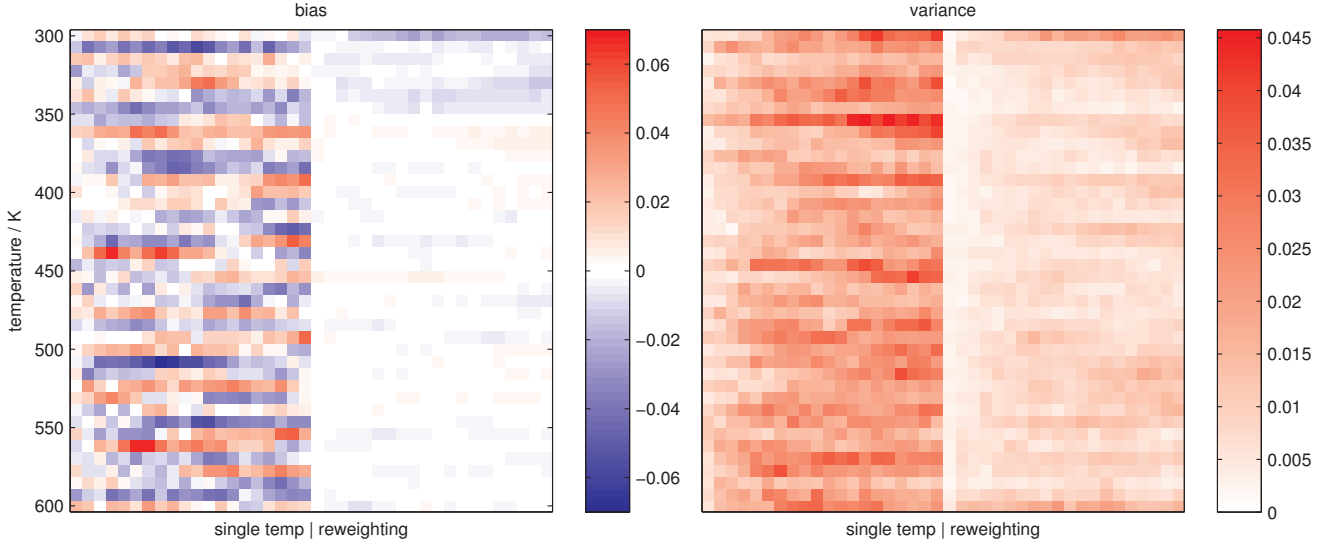


FIG. 4. **Bias and variance of single-temperature estimate versus dynamical reweighting estimates.** The estimated bias (left) and variance (right) of the estimated normalized fluctuation correlation function estimated by dividing the simulation into 20 blocks of equal length. The left half of each color block shows the statistic for single-temperature estimates, while the right half shows the statistic for dynamical reweighting estimates employing data from all temperatures. Each horizontal row represents the estimated bias or variance for the correlation function  $C(\tau; \beta)$  over  $\tau \in [0, 10]$  ps for one temperature; temperatures are ordered from lowest (top) to highest (bottom).

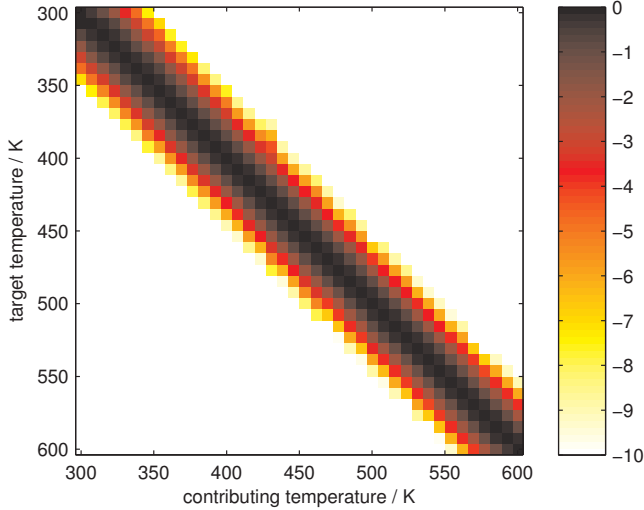


FIG. 5. **Logarithmic total weight contributed from all trajectories at each sampled temperature to various temperatures.** The base-10 logarithm of the sum of weights  $w_n(\beta)$  from all trajectories sampled from each contributing temperature is shown for each target (reweighted) temperature corresponding to inverse temperature  $\beta$ . The contribution from the target temperature is normalized in this plot to have a total weight of unity (zero logarithm).

or the histogram-based WHAM [13, 15, 18]). Dynamical reweighting provides a way to make use of *all* the data from a parallel or simulated tempering simulation, provided the dynamical model employed is

amenable to reweighting by the unnormalized density ratios  $q[X|\beta]/q[X|\beta']$  being both finite and nonzero. This condition is fulfilled for the common models of dynamics in the canonical ensemble—canonical distribution over Hamiltonian trajectories; Andersen, Langevin, and overdamped Langevin / Brownian stochastic dynamics, reviewed in Section IIC—but not Nosé-Hoover or Berendsen.

While not all properties will benefit greatly from the use of reweighting, some dynamical expectations—particularly those that involve large contributions from trajectories that are rare at the temperature of interest, but more plentiful at elevated (or reduced) temperatures—will especially benefit in terms of reduced variance. For example, though not shown here, transitions along the  $\phi$  torsion for terminally-blocked alanine involve conformational states with high free energy which may not even be *sampled* at some temperatures in simulated or parallel tempering simulations of typical length, necessitating the use of reweighting to provide an estimate of rates involving these states (illustrated in detail in a companion paper [63]). Even in the case where little or no temperature-dependent enhancement of the phenomena is expected, neighboring temperatures from a simulated or parallel tempering simulation often carry a reasonably large amount of information about the temperature of interest, and their inclusion will further reduce the expected statistical error in the estimated expectation if the data is available anyway.

The variety of models of dynamics within the canon-

ical ensemble presented in Section II C, all of which lead to the same static equilibrium distribution of configurations, naturally leads one to ask which model is most appropriate. In some regimes—such as the low-friction regime of Langevin dynamics and the canonical distribution of Hamiltonian trajectories containing large baths of explicit solvent—these models may give nearly identical expectations, while in other regimes—those where the Andersen collision rate  $\nu$  or Langevin collision rate  $\gamma$  is large—the expectations for dynamical properties may differ considerably. Which situation is more physically reasonable will undoubtedly depend on both the system under study and the properties of interest, but a detailed investigation of this is beyond the scope of this work.

The efficiency of reweighting—that is, the degree to which temperatures  $\beta' \neq \beta$  contribute to expectations at temperature  $\beta$ —will depend both on the size of the system and, for stochastic dynamics, the trajectory length. To understand this behavior, consider a standard parallel tempering simulation using the conventional exchange criteria based on potential energies [7]. In order for the ensembles at two temperatures to have good potential energy overlap, and hence good exchange acceptance rates, their temperatures should be spaced roughly the half-width of the potential energy distribution,  $\delta U \sim \langle (U - \langle U \rangle)^2 \rangle^{1/2} = T\sqrt{k_B C_v}$ . The heat capacity for the potential energy,  $C_v \equiv (\partial/\partial T) \langle U \rangle_T$ , tells us that the temperature shift  $\delta T$  corresponding to this energy shift  $\delta U$  is roughly  $T\sqrt{k_B/C_v}$ .

Analogously, good path Hamiltonian overlap (and good exchange acceptance rates using the modified acceptance criterion in Eq. 47) can be assured with a temperature spacing of roughly  $T\sqrt{k_B/C_v}$ , where  $C_v$  is a generalized form of the heat capacity for path Hamiltonians  $\mathcal{H}[X]$ :

$$C_v \equiv \frac{\partial \langle \mathcal{H} \rangle_T}{\partial T} = k_B \beta^2 \langle (\mathcal{H} - \langle \mathcal{H} \rangle)^2 \rangle_\beta \quad (56)$$

For the canonical distribution of Hamiltonian trajectories, this generalized path heat capacity  $C_v$  is identical to the heat capacity for the total energy,  $(\partial/\partial T) \langle H \rangle_T = C_v + N_d k_B/2$ , which grows linearly with the number of degrees of freedom  $N_d$ .

For Andersen, Langevin, and Brownian dynamics, the generalized path heat capacity  $C_v$  differs from the thermodynamic heat capacity  $C_v$ . Because the variance of a sum of uncorrelated random variables is simply the sum of the variances, we can compute this difference analytically. For Andersen dynamics,  $C_v = C_v + N_c k_B/2$ , where  $N_c$  is the number of collisions that occur during realization of the trajectory, where roughly  $N_c \sim N_d N_t \nu$ , with  $\nu$  is the per-step collision probability and  $N_t$  is the number of integration timesteps in the trajectory. For Langevin dynamics,  $C_v = C_v + N_d(N_t + 1)k_B/2$ , and for Brownian dynamics,  $C_v = C_v + N_d N_t k_B/2$  (where  $C_v$  here denotes the heat capacity for the *potential energy* contribution only for Brownian dynamics). For con-

stant, nonzero collision rates, the generalized heat capacity is therefore extensive in  $N_d N_t$ ; as the systems grow larger or the trajectories grow longer, the efficiency of reweighting is therefore expected to diminish.

The form of the path Hamiltonians suggests a further analogy with the equilibrium statistical mechanics governing single configurations. For Hamiltonian dynamics, a path is uniquely identified by its initial phase space point. For the stochastic forms of dynamics (Andersen, Langevin, Brownian), a trajectory given an initial phase space point can be thought of as an ideal polymer consisting of  $N_t$  “monomers” that are *replicates* of the  $N_p$ -particle system, one replicate per trajectory timeslice, where the only interactions are “bonds” between corresponding atoms in sequential timeslices. The interaction energies for these “bonds” are harmonic in the noise variables  $\xi_{ti}$  corresponding to the atomic displacement between sequential timeslices.

Surprisingly, the path Hamiltonians and generalized path heat capacities are independent of the choice of collision rate  $\gamma_i$  for Langevin and Brownian dynamics, except when  $\gamma_i$  is identically zero, at which point the particles with zero associated collision rates no longer contribute to the path Hamiltonian or heat capacity since they evolve by deterministic dynamics. This is analogous to the fact that the heat capacity of an ideal polymer is independent of the spring constant and particle masses. Only in the case of Andersen dynamics will the collision rate  $\nu$  modulate the heat capacity through determining the number of particle collisions observed in a trajectory of fixed length  $N_t$ .

It should be noted that there is often additional information available that could be incorporated into the reweighting scheme through the use of control variates [26]. For example, for the model of Hamiltonian trajectories with randomized momenta, the normalization constants  $\hat{Z}_i$  that are inferred during the procedure actually contain the temperature-dependent momentum partition function, which can be computed analytically. Incorporation of this as a constraint, or other integrals of the dynamics that are known exactly, can further reduce the variance in the estimated properties of interest, though the degree to which this may occur will be problem-dependent.

## V. ACKNOWLEDGMENTS

The authors would like to thank Hans C. Andersen (Stanford University), Nicolae-Viorel Buchete (University College Dublin), Ken Dill (University of California, San Francisco), Peter Eastman (Stanford), Phillip L. Geissler (University of California, Berkeley), Gerhard Hummer (National Institutes of Health), Jesús Izaguirre (University of Notre Dame), Antonia S. J. S. Mey (University of Nottingham), David D. L. Minh (Argonne National Laboratory), and Jed W. Pitera (IBM Almaden Research Center) for stimulating discussions on

this topic. JDC gratefully acknowledges support from the HHMI and IBM predoctoral fellowship programs, NIH grant GM34993 through Ken A. Dill (UCSF), and NSF grant for Cyberinfrastructure (NSF CHE-0535616) through Vijay S. Pande (Stanford), and a QB3-Berkeley Distinguished Postdoctoral Fellowship at various points throughout this work. JHP and FN acknowledge funding from the German Science Foundation (DFG) through the Research Center Matheon, the International Graduate College 710, and grant NO 825/2. VSP ac-

knowledges NIH grant U54 GM072970. JDC thanks OPENMM developers Peter Eastman, Mark Friedrichs, Randy Radmer, and Christopher Bruns for their generous help with the OPENMM GPU-accelerated computing platform and associate PYOPENMM Python wrappers. This research was supported in part by the National Science Foundation through TeraGrid resources provided by NCSA Lincoln under grant number TGMCB100015. We specifically acknowledge the assistance of John Estabrook of NCSA Consulting Services in aiding in the use of these resources.

- 
- [1] E. Marinari and G. Parisi, *Europhys. Lett.*, **19**, 451 (1992).
- [2] A. P. Lyubartsev, A. A. Martynov, S. V. Shevkunov, and P. N. Vorontsov-Velyaminov, *J. Chem. Phys.*, **96**, 1776 (1992).
- [3] A. Mitsutake and Y. Okamoto, *Chem. Phys. Lett.*, **332**, 131 (2000).
- [4] C. J. Geyer and E. A. Thompson, *J. Am. Stat. Assoc.*, **90**, 909 (1995).
- [5] K. Hukushima and K. Nemoto, *J. Phys. Soc. Jpn.*, **65**, 1604 (1996).
- [6] U. H. E. Hansmann, *Chem. Phys. Lett.*, **281**, 140 (1997).
- [7] Y. Sugita and Y. Okamoto, *Chem. Phys. Lett.*, **314**, 141 (1999).
- [8] J. S. Liu, *Monte Carlo strategies in scientific computing*, 2nd ed. (Springer-Verlag, New York, 2002).
- [9] S. Trebst, M. Troyer, and U. H. E. Hansmann, *J. Chem. Phys.*, **124**, 174903 (2006).
- [10] J. Machta, *Phys. Rev. E*, **80**, 056706 (2009).
- [11] E. Rosta and G. Hummer, *J. Chem. Phys.*, **2009**, 165102 (2009).
- [12] A. M. Ferrenberg and R. H. Swendsen, *Phys. Rev. Lett.*, **63**, 1195 (1989).
- [13] S. Kumar, D. Bouzida, R. H. Swendsen, P. A. Kollman, and J. M. Rosenberg, *J. Comput. Chem.*, **13**, 1011 (1992).
- [14] C. Bartels and M. Karplus, *J. Comput. Chem.*, **18**, 1450 (1997).
- [15] E. Gallicchio, M. Andrec, A. K. Felts, and R. M. Levy, *J. Phys. Chem. B*, **109**, 6722 (2005).
- [16] C. Bartels, *Chem. Phys. Lett.*, **331**, 446 (2000).
- [17] M. R. Shirts and J. D. Chodera, *J. Chem. Phys.*, **129**, 124105 (2008).
- [18] J. D. Chodera, W. C. Swope, J. W. Pitera, C. Seok, and K. A. Dill, *J. Chem. Theor. Comput.*, **3**, 26 (2007).
- [19] N.-V. Buchete and G. Hummer, *Phys. Rev. E*, **7**, 030903(R) (2008).
- [20] R. D. Skeel, *SIAM J. Sci. Comput.*, **31**, 1363 (2009).
- [21] D. D. L. Minh and J. D. Chodera, *J. Chem. Phys.*, **131**, 134110 (2009).
- [22] G. E. Crooks, *Phys. Rev. E*, **60**, 2721 (1999).
- [23] G. E. Crooks, *Phys. Rev. E*, **61**, 2361 (2000).
- [24] X.-L. Meng and W. H. Wong, *Statistica Sinica*, **6**, 831 (1996).
- [25] A. Kong, P. McCullagh, X.-L. Meng, D. Nicolae, and Z. Tan, *J. Royal Stat. Soc. B.*, **65**, 585 (2003).
- [26] Z. Tan, *J. Am. Stat. Assoc.*, **99**, 1027 (2004).
- [27] C. H. Bennett, "Algorithms for chemical computations," (*Am. Chem. Soc.*, Washington, D.C., 1977) Chap. 4, pp. 63–97.
- [28] D. Chandler, "Classical and quantum dynamics in condensed phase simulations," (World Scientific, 1998) Chap. Finding transition pathways: throwing ropes over rough mountain passes, in the dark, pp. 51–66.
- [29] H. C. Andersen, *J. Chem. Phys.*, **72**, 2384 (1980).
- [30] W. G. Hoover, *Phys. Rev. A*, **31**, 1695 (1985).
- [31] D. J. Evans and B. L. Holian, *J. Chem. Phys.*, **83**, 4069 (1985).
- [32] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, *J. Chem. Phys.*, **81**, 3684 (1984).
- [33] T. Morishita, *J. Chem. Phys.*, **113**, 2976 (2000).
- [34] B. Cooke and S. C. Schmidler, *J. Chem. Phys.*, **129**, 164112 (2008).
- [35] E. Rosta, N.-V. Buchete, and G. Hummer, *J. Chem. Theor. Comput.*, **5**, 1393 (2009).
- [36] W. K. Hastings, *Biometrika*, **57**, 97 (1970).
- [37] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, *Phys. Lett. B*, **195**, 216 (1987).
- [38] W. E and D. Li, *Commun. Pure Appl. Math*, **61**, 0096 (2008).
- [39] J. Juraszek and P. G. Bolhuis, *Proc. Nat. Acad. Sci. USA*, **103**, 15859 (2006).
- [40] H. C. Andersen, *J. Comput. Phys.*, **52**, 24 (1983).
- [41] J. A. Izaguirre, C. R. Sweet, and V. S. Pande, *Pac. Symp. Biocomp.*, **15**, 240 (2010).
- [42] D. L. Ermak and Y. Yeh, *Chem. Phys. Lett.*, **24**, 243 (1974).
- [43] D. L. Ermak, *J. Chem. Phys.*, **62**, 4189 (1975).
- [44] N.-V. Buchete and G. Hummer, *J. Phys. Chem. B*, **112**, 6057 (2008).
- [45] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, *J. Chem. Phys.*, **81**, 3684 (1984).
- [46] D. Sindhikara, Y. Meng, and A. E. Roitberg, *J. Chem. Phys.*, **128**, 024103 (2008).
- [47] T. J. H. Vlugt and B. Smit, *Phys. Chem. Comm.*, **2**, 1 (2001).
- [48] P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler, *Annu. Rev. Phys. Chem.*, **53**, 291 (2002).
- [49] C. J. Geyer, in *Computing Science and Statistics: The 23rd symposium on the interface*, edited by E. Keramigas (Interface Foundation, 191) pp. 156–163.
- [50] K. Hukushima and K. Nemoto, *J. Phys. Soc. Japan*, **65**, 1604 (1996).
- [51] J. W. Pitera, I. Haque, and W. C. Swope, *J. Chem. Phys.*, **124**, 141102 (2006).
- [52] W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson, *J. Chem. Phys.*, **76**, 637 (1982).
- [53] H. Flyvbjerg and H. G. Petersen, *J. Chem. Phys.*, **91**, 461 (1989).

- [54] W. Janke, in *Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms*, Vol. 10, edited by J. Grotendorst, D. Marx, and A. Murmatsu (John von Neumann Institute for Computing, 2002) pp. 423–445.
- [55] J. D. Chodera, W. C. Swope, J. W. Pitera, C. Seok, and K. A. Dill, *J. Chem. Theor. Comput.*, **3**, 26 (2007).
- [56] D. A. Kofke, *J. Chem. Phys.*, **117**, 6911 (2002).
- [57] H. G. Katzgraber, S. Trebst, D. A. Huse, and M. Troyer, *J. Stat. Mech.*, **2006**, 03018 (2006).
- [58] S. Trebst, M. Troyer, and U. H. E. Hansmann, *J. Chem. Phys.*, **124**, 174903 (2006).
- [59] D. Gront and A. Kolinski, *J. Phys.: Condens. Matter*, **19**, 036225 (2007).
- [60] S. Park and V. S. Pande, *Phys. Rev. E*, **76**, 016703 (2007).
- [61] D. Chandler, *J. Chem. Phys.*, **68**, 2959 (1978).
- [62] J. A. M. Jr., D. Chandler, and B. J. Berne, *J. Chem. Phys.*, **70**, 4056 (1979).
- [63] J.-H. Prinz, J. D. Chodera, V. S. Pande, W. C. Swope, J. C. Smith, and F. Noé, Submitted to JCP. (2010).
- [64] W. Zhang, T. Hou, C. Schafmeister, W. S. Ross, and D. A. Case, “Leap from ambertools 1.2,” (2009).
- [65] P. A. Kollman, *Acc. Chem. Res.*, **29**, 461 (1996).
- [66] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, *J. Chem. Phys.*, **79**, 926 (1983).
- [67] D. C. Liu and J. Nocedal, *Mathematical Programming*, **45**, 503 (1989).
- [68] J. Nocedal, *Mathematics of Computation*, **35**, 773 (1980).
- [69] K. Beauchamp, C. Bruns, P. Eastman, M. Friedrichs, J. P. Ku, V. Pande, R. Radmer, and M. Sherman, “Openmm user manual and theory guide, release 2.0,” (2010).
- [70] P. Eastman and V. S. Pande, *J. Chem. Theor. Comput.*, **6**, 434 (2010).
- [71] S. Miyamoto and P. A. Kollman, *J. Comput. Chem.*, **13**, 952 (1992).
- [72] T. A. Darden, D. M. York, and L. G. Pedersen, *J. Chem. Phys.*, **98**, 10089 (1993).
- [73] M. P. Allen and D. J. Tildesley, *Computer simulation of liquids* (Clarendon Press, Oxford, 1991).
- [74] M. S. Friedrichs, P. Eastman, V. Vaidyanathan, M. Houston, S. LeGrand, A. L. Beberg, D. L. Ensign, C. M. Bruns, and V. S. Pande, *J. Comput. Chem.*, **30**, 864 (2009).
- [75] P. Eastman and V. S. Pande, *Computing in Science and Engineering*, **12**, 34 (2010).
- [76] P. Eastman and V. S. Pande, *J. Comput. Chem.*, **31**, 1268 (2010).
- [77] C. M. Bruns, R. A. Radmer, J. D. Chodera, and V. S. Pande, “Pyopenmm,” (2010).
- [78] W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson, *J. Chem. Phys.*, **76**, 637 (1982).
- [79] L. Verlet, *Phys. Rev.*, **159**, 98 (1967).
- [80] L. Verlet, *Phys. Rev.*, **165**, 201 (1968).

### Appendix A: Additional simulation details for alanine dipeptide

Using the LEaP program from the AmberTools 1.2 molecular mechanics package [64], a terminally-blocked alanine peptide (sequence ACE-ALA-NME, see Figure 1) was generated in the extended conformation, with

peptide force field parameters taken from the AMBER parm96 parameter set [65]. The system was subsequently solvated with 749 TIP3P water molecules [66] in a cubic simulation box with dimensions chosen to ensure all box boundaries were at least 9 Å from any atom of the extended peptide. The system was subjected to energy minimization using L-BFGS [67, 68] to reduce the root-mean-square force to less than 1 kJ/mol/nm, and then equilibrated for 1 ns with a leapfrog Langevin integrator [41] at a control temperature of 300 K, using a 2 fs timestep and collision rate of 5/ps. During equilibration, an isotropic Monte Carlo barostat with a control pressure of 1 atm was applied, with volume moves attempted every 25 timesteps, and the range of volume change proposals automatically adjusted to give an approximate acceptance probability of approximately 50% [69]. All bonds to hydrogen, and hydrogen-hydrogen distances in waters, were constrained by the CCMA [70] and SETTLE [71] algorithms, as appropriate. The particle-mesh Ewald (PME) method [72] was used to treat electrostatics, using a real-space cutoff of 9 Å. PME parameters were automatically selected by an algorithm that attempts to find the most efficient set of parameters for which an error bound is less than a specified error tolerance [69], which was set to the default of  $5 \cdot 10^{-4}$ . Lennard-Jones interactions were truncated at 9 Å without switching, and a homogeneous analytical long-range dispersion correction [73] was employed to compensate for dispersion interactions outside the cutoff. The resulting system had a box volume of 23 073.7 Å<sup>3</sup>, and the box volume was fixed in the subsequent parallel tempering simulation.

A custom Python code making use of the GPU-accelerated OPENMM package [74–76] and the PY-OPENMM Python wrapper [77] was used to conduct the simulations. Because OPENMM lacks a velocity Verlet integrator, a hybrid velocity Verlet [78] / leapfrog [79, 80] integration scheme was used for integration of the equations of motion, implemented as follows. At the beginning of the dynamical propagation phase of parallel tempering iteration, on-step velocities were first generated from the Maxwell-Boltzmann distribution at the current replica temperature. Components of the velocity along constrained bonds (all bonds to hydrogen and the hydrogen-hydrogen bonds of TIP3P water) were removed using the RATTLE algorithm [40] with  $10^{-16}$  relative tolerance, and a 1 fs “half-kick backwards” applied to the modified velocities using the force evaluated at the current positions. The OPENMM Verlet integrator was then used to evolve the positions by 0.5 ps (250 steps with a 2 fs timestep) on the GPU, and a 1 fs “half-kick forwards” was applied to the velocities to synchronize them with the positions after again applying the RATTLE algorithm. During leapfrog integration, the CCMA and SETTLE algorithms were used to constrain bonds, as appropriate.