A generalized linear response framework for expanded ensemble and replica exchange simulations

Brian K. Radak,¹ Donghyuk Suh,² and Benoît Roux^{3,4,a)}

¹Theoretical and Computational Biophysics Group, Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801-2325, USA

²Department of Chemistry, University of Chicago, Chicago, Illinois 60637-1454, USA

³Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, Illinois 60637-1454, USA

⁴Center for Nanoscale Materials, Argonne National Laboratory, Argonne, Illinois 60439-8643, USA

(Received 2 March 2018; accepted 15 May 2018; published online 1 June 2018)

Expanded ensemble simulation is a powerful technique for enhancing sampling over a range of thermodynamic parameters. However, although the premise is relatively simple, running successful simulations in practice still presents something of an *ad hoc* challenge. Three main difficulties exist: (1) the selection of the thermodynamic states, (2) the selection of the sampling weights, and (3) efficient sampling of the expanded parameter space. Here we consider these problems in the context of a pairwise linear response approach to the work fluctuation theorem. The approach offers comprehensive tactics for addressing the three difficulties and can be used as either an alternative or a complement to replica exchange simulations. Importantly, the results are trivially implemented for multi-dimensional parameter spaces and they recover results from the literature aimed at the special cases of simulated/parallel tempering and replica exchange umbrella sampling. Illustrative examples are shown using the NAMD simulation engine. *Published by AIP Publishing.* <https://doi.org/10.1063/1.5027494>

I. INTRODUCTION

Expanded ensembles are an enhanced sampling device applicable in both Monte Carlo (MC) and molecular dynamics (MD) simulations.¹ Whereas simulations in standard ensembles (e.g., the canonical ensemble) sample from a fixed phase space defined by the system coordinates and momenta, $\mathbf{x}$, expanded ensembles also permit variation of the ensemble parameters, $\boldsymbol{\theta}$. This dual dependence can be expressed by writing the reduced potential, $u(\mathbf{x}; \boldsymbol{\theta})$, as a function of both sets of variables. Here we will only consider a discrete set of ensemble parameters such that a specific parameter combination can be given an index k ; for brevity, $u_k(\mathbf{x}) \equiv u(\mathbf{x}; \boldsymbol{\theta}_k)$. Each index is then said to correspond to a “thermodynamic state.” By combining MD and MC in an appropriate manner, two sets of conditional probability distributions can then be sampled,²

$$\rho(\mathbf{x}; \boldsymbol{\theta}_k) = \frac{e^{-u_k(\mathbf{x})}}{\int d\mathbf{x} e^{-u_k(\mathbf{x})}}, \quad (1a)$$

$$\rho(\boldsymbol{\theta}_k; \mathbf{x}) = \frac{e^{g_k - u_k(\mathbf{x})}}{\sum_l e^{g_l - u_l(\mathbf{x})}}. \quad (1b)$$

Here $g_k \equiv g(\boldsymbol{\theta}_k)$ is an arbitrary dimensionless energy constant that only depends on the state and tunes the relative population. The integral in Eq. (1a) is over all possible configurations, while the sum in Eq. (1b) runs over the set of all possible states.

Chodera and Shirts² recently emphasized that sampling according to Eq. (1) can be cast as a Gibbs sampling scheme,

a long-established technique in the statistics community. A number of beloved techniques from the computational physics community can be recognized as Gibbs sampling algorithms.² This realization provides excellent new resources for developing sampling algorithms in the space of ensemble parameters. Unfortunately, it does little to address two problems of equal importance: how to select the set of states (i.e., the particular values of $\boldsymbol{\theta}$) and the “state weights,” g_k . Both of these issues have a dramatic effect on sampling efficiency.

Far and away, the most popular solution to the weighting problem is the use of replica exchange algorithms. That is, replica exchange eliminates the need for state weights by introducing multiple replicas that occupy all states concurrently and are simulated in *parallel*. A change in a state parameter thus requires at least two simulations to exchange their state; modifying the MC procedure to reflect this exchange leads to cancellation of the (still unknown) weights. Parallel tempering,³ in which the temperature changes between states, is probably the most widely recognized replica exchange algorithm. However, there are legitimate problems with the cost scaling of this approach. Specifically, the number of replicas scales rapidly with the size of the system, especially when explicit solvent models are employed. This problem is further exacerbated by the introduction of additional parameter dimensions (e.g., two-dimensional replica exchange). Furthermore, a significant and arguably unwarranted amount of effort is exerted simulating temperatures at which data are unreliable and/or unwanted.⁴ Nonetheless the approach continues to find use because no parameter tuning is necessary.

The alternative to the parallel exchange scheme is a *serial* approach, where only one copy of the system is simulated

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

and changes in the state are governed by a choice of the state weights. Here we note a recent work by Pan *et al.*⁵ which shows that ideas by Park and Pande⁶ are actually quite effective at selecting the sampling weights, at least when the ensemble parameter is temperature. The technique used in that study, simulated tempering,⁷ was also able to improve the convergence of protein folding simulations by as much as a few orders of magnitude. Park⁸ also later developed a rigorous linear response formulation to rationalize their approach and extend it to the case of alchemical coupling. The goal of both studies by Park was to equalize the probability of being in neighboring states so that the distribution across all states is uniform.

Here we extend the analysis by Park⁸ yet further and show that it can be used not only to estimate state weights in virtually any context but also to estimate the fluctuations that govern sampling efficiency between neighboring states. In conjunction with basic results from Markov modeling, the result is a clear approach to setting up and optimizing efficiency in expanded ensemble simulations in general. Throughout, it is shown that this linear response approach recovers seemingly unrelated analyses regarding the selection of temperature ranges and spacings^{9–12} for parallel tempering, as well as for replica exchange umbrella sampling.^{13–15}

II. THEORY

In what follows, we briefly outline existing ideas in the literature regarding expanded ensemble simulations before exploring the novel components proposed here. The background is necessarily brief in order to keep the scope of this work manageable and to retain emphasis. For a more exhaustive perspective on sampling, the interested reader should consider the recent work by Chodera and Shirts.² The notion of using linear response was originally articulated by Park and Pande⁶ (although it became known to us by a different route). However, the exposition in this work presents a rather different perspective by starting from general statistical mechanical relationships.

A. Expanded ensembles

The basic assumption in using the expanded ensemble is that either: (1) the relative behavior in each state is interesting (i.e., one desires to know a quantity as a function of θ) or (2) the efficiency in sampling $\mathbf{x}$ is better for some states than others (i.e., changing the state enhances sampling with respect to a few particular states of interest). This is relevant because the amount of time spent in each state depends critically on the choice of the sample weights, g_k . If no choice is made (i.e., $g_k = 0$ for all k), then the populations are dictated by the free energies of the system. This is often not useful in practice since the relative free energies between states may differ by one or more orders of magnitude. A *uniform* sampling of states would be achieved if the sample weights were exactly the free energies.^{6,7} While it is not obvious that this situation is necessarily ideal,^{5,16} any other distribution can be trivially chosen by simply adding shifts to the weights.

As discussed earlier, a noteworthy alternative to the weight selection process is the use of a “parallel” expanded ensemble

algorithm. In parallel algorithms, Eq. (1a) is independently sampled for *all* of the states (usually, but not necessarily, simultaneously¹⁷) while Eq. (1b) is expanded to sample *permutations* of the state indices for the current values of $\mathbf{x}$.² Since each parallel trajectory samples from the same space of $\mathbf{x}$ and θ , they are effectively replicates of each other. Furthermore, a simple way to sample permutations in θ is to perform pairwise swaps between replicas (hence this is frequently referred to as replica exchange, even though the name only indicates a specific variant). An important consequence of this pairing approach is that the g_k terms in Eq. (1b) cancel during sampling and therefore need not be selected. The states’ populations are, by construction, uniform, since each state is always occupied by a replica.

B. Linear response

Consider two thermodynamic states (with indices i and j , say) with reduced potentials $u_i(\mathbf{x})$ and $u_j(\mathbf{x})$ and relative dimensionless free energy Δf_{ij} in the direction $i \rightarrow j$. The following identity is true for any arbitrary function $\alpha(\mathbf{x})$:

$$e^{-\Delta f_{ij}} = \frac{\langle \alpha e^{-u_j} \rangle_i}{\langle \alpha e^{-u_i} \rangle_j}, \quad (2)$$

where the angle brackets indicate an average with respect to the distribution for the subscripted index and the $\mathbf{x}$ arguments have been suppressed for clarity. This relation seems to have been first reported by Bennett¹⁸ who used it to derive an estimator for the free energy. Interestingly, particular choices for $\alpha(\mathbf{x})$ can also be used to derive: (1) either of the relations now ascribed to Zwanzig,¹⁹ (2) the work equality due to Jarzynski,²⁰ and (3) the fluctuation theorem of Crooks²¹ (see also Shirts *et al.*²²). This connection has long been known in the literature^{23–25} and serves as a useful reminder that many fundamental equations are highly interrelated. A generalization of this relation for multiple states and finite sample sizes can also be used as a starting point for deriving multistate reweighting methods.²⁶

As a follow up to earlier work,⁶ Park⁸ showed that a cumulant expansion based on Eq. (2) leads to a simple, approximate estimator for the free energy based on the mean difference in energy between the two distributions,

$$\Delta f_{ij} = \frac{\langle w \rangle_i + \langle w \rangle_j}{2}, \quad (3)$$

where $w(\mathbf{x}) \equiv u_j(\mathbf{x}) - u_i(\mathbf{x})$. Although Eq. (3) is certainly well known in the literature, especially as a free energy estimator,²⁷ it first became known to us via a nonequilibrium linear response approach to Crooks’ theorem.^{28,29} The correspondence comes from realizing that w is simply the *instantaneous* work needed to move from state i to state j . If the work values were exactly Gaussian distributed, then the cumulant expansion formally vanishes beyond second order and Eq. (3) is exact.

In addition to Eq. (3), the Gaussian approximation also implies a *second* relation between the means of the two distributions. This relation, however, gives the instantaneous work variance,

$$\sigma_{ij}^2 = \langle w \rangle_i - \langle w \rangle_j. \quad (4)$$

While Eq. (3) has an obvious utility for quickly estimating Δf_{ij} (and, by extension, $g_j - g_i$), Eq. (4) will be shown to be useful in selecting the range and spacing of the states.

Equations (3) and (4) can also be obtained by considering the mean instantaneous work, $\langle w \rangle$, as a function of θ . The linear response approximation is recovered by considering a piecewise linear form for $\langle w \rangle$ on the set of points (Fig. 1). The free energy is obtained by trapezoidal integration of the mean work. This is just thermodynamic integration along the variable. Similarly, the variance is obtained by a central differencing,

$$\frac{\partial \langle w \rangle}{\partial \theta} = \frac{\langle w \rangle_j - \langle w \rangle_i}{\Delta \theta} = -\frac{\sigma_{ij}^2}{\Delta \theta}, \quad (5)$$

where $\Delta \theta \equiv \theta_j - \theta_i$. Both the integral and derivative are exact within piecewise linearity (or on a vanishingly small interval). These relations arise from the fact that $\partial u / \partial \theta = w / \Delta \theta$ within piecewise linearity, such that the free energy is similar to a generating function with

$$\frac{\partial \Delta f}{\partial \theta} = \frac{\langle w \rangle}{\Delta \theta} \quad \text{and} \quad \frac{\partial^2 \Delta f}{\partial \theta^2} = -\frac{\sigma^2}{\Delta \theta^2} \quad (6)$$

and so on for higher order cumulants.

The above arguments pertaining to thermodynamic integration speak to the conditions for accuracy of the linear response approach. First of all, this method for estimating the free energies and variances between neighboring states clearly becomes more accurate when: (1) the states are closer together in parameter space and (2) the mean work does not have a large curvature. The validity of condition two can be checked visually after an initial estimate. It is also worth reiterating that this approach is only being suggested here as a rapid method for estimating parameters at the start of a

FIG. 1. The linear response approximation can be viewed graphically by recognizing the relationship between thermodynamic derivatives and the work moments (i.e., mean and variance). In this view, linear response literally asserts that the mean instantaneous work needed to change some thermodynamic parameter, θ , is piecewise linear over small intervals, $\Delta \theta$ (middle panel). The integral of this function is then exactly a trapezoidal integration and corresponds to the free energy (top panel), while the derivative can be found by central differencing and corresponds to the work variance (bottom panel).

simulation. Once data have been obtained from an expanded ensemble simulation, rigorous reweighting using the weighted histogram analysis method³⁰ or any of its variants^{26,31} is likely superior.

C. Markov models and state selection

While the importance of the sampling weights in expanded ensemble simulations is rather obvious, the problem of selecting the parameter values that describe each state is also quite significant. If states are too far apart, they are unlikely to overlap and MC transitions between them will occur infrequently, regardless of the value of the weights. This can be understood by computing the mean acceptance probability between a pair of states based on the Metropolis criterion and the Gaussian work approximation.^{8,29} If the free energy between the states is zero (i.e., the sample weights are chosen to be exactly the free energies), the mean acceptance probability is simply

$$\overline{P}_{ij} = \text{erfc} \left(\frac{\sigma_{ij}}{2\sqrt{2}} \right). \quad (7)$$

This is only an upper bound for non-zero Δf (i.e., when the g_k are not estimated exactly).²⁹ During replica exchange, two variations on this equation occur. First, the zero free energy approximation becomes exact (for the same reason that sampling weights are not needed). Second, there are now *two* Gaussian distributions with identical variance, effectively yielding a single Gaussian with *double* that variance. The factor of $\sqrt{2}$ is therefore omitted in the case of replica exchange. As will be seen, Eq. (7) is essentially identical to the expression derived for parallel and simulated tempering.^{9–11} However, the assumptions in those derivations are quite different and are specific to the case that the temperature is changing.

Several authors have addressed the problem of how to select an optimal “ladder” of temperatures for both simulated⁷ (serial) and parallel³ tempering,^{9–12,32,33} as well as replica exchange umbrella sampling.^{13–15} These approaches have generally focused on the idea of maximizing the round trip rate or, equivalently, minimizing the mean passage time. This can be articulated formally by considering a linear Markov model with M states in which transitions occur between neighboring states. The first criterion for optimization is that the transition probability, P , between neighboring states is a global constant (Fig. 2).³³ Using standard methods, it can be deduced that the mean round trip rate is^{11,33}

$$k_s = \frac{P}{2M(M-1)}. \quad (8)$$

FIG. 2. The sequence of states sampled during an expanded ensemble simulation can be considered as a linear Markov model with a constant global transition probability, P , between neighboring states.

The model is completed by asserting that P can be well approximated by the mean transition probability, $\bar{P}_{ij}$, if the states are chosen such that σ_{ij}^2 is constant between neighboring states.

The key contribution of this work is that the definition of neighbor is now completely general. This is to say that the general linear response relations [Eqs. (4) and (7)] make no assumptions regarding the nature of the parameter(s) that is/are changing. This also means that they can be readily applied with multiple parameters in multiple dimensions.

Equation (8) is true for a *stochastic* process on the set of states. As noted by Lingenheil *et al.*,³³ this does not hold for deterministic algorithms that concertedly alternate between even/odd or up/down pairs of states, as is most commonly done for parallel algorithms. Instead, based on relative diffusion arguments, they proposed the relationship

$$k_d = \frac{k_s}{1 - P}, \quad (9)$$

which seems to be quite accurate for $M \geq 10$.³³

Yet another alternative is the “convective” scheme originally proposed by Spill, Bouvier, and Nilges³⁴ for parallel tempering. This approach is analogous to the deterministic scheme except that pairs of states only change when a selected replica completes an end-to-end trip. This scheme has not, to our knowledge, been used before in the serial context. In this case (but not in the parallel case), the mean round trip rate can be computed as

$$k_c = Mk_s. \quad (10)$$

Recently, Chodera and Shirts² re-examined an excellent approach originally proposed by Liu,³⁵ which they termed “Metropolized independence sampling.” This approach has much in common with the stochastic approach except that it expands the set of candidate states beyond nearest neighbors. Because multiple pairs are considered, an equation analogous to (8) cannot be written in terms of a single parameter. Empirically, however, the method corresponds to a roughly global enhancement in P by a factor of ~ 1.6 – 2.2 . This approximation clearly must fail for large M , since the mean transition probability must saturate at one. Nonetheless, it highlights that the standard stochastic approach and the Metropolized approach can essentially be treated equally except that the latter has a higher acceptance probability.

Equations (8)–(10) can be maximized under the above assumptions to obtain an expression for the optimal number of states, M_{opt} . The general expression is³⁶

$$-(M_{\text{opt}} - 1) \frac{d\sigma}{dM} \Big|_{M_{\text{opt}}} = f[\sigma(M_{\text{opt}})], \quad (11)$$

where

$$\begin{aligned} f_s[\sigma(M)] &\equiv 2\sqrt{2\pi} \operatorname{erfcx} \left[\frac{\sigma(M)}{2\sqrt{2}} \right], \\ f_d[\sigma(M)] &\equiv f_s[\sigma(M)] \operatorname{erf} \left[\frac{\sigma(M)}{2\sqrt{2}} \right], \\ f_c[\sigma(M)] &\equiv \frac{1}{2} f_s[\sigma(M)] \end{aligned}$$

in the stochastic (s) [Eq. (8)], deterministic (d) [Eq. (9)], and convective (c) [Eq. (10)] cases, respectively. Here

$\operatorname{erfcx}(x) \equiv \exp(x^2) \operatorname{erfc}(x)$ is the so-called scaled complementary error function.³⁷ For the parallel case, $\sigma(M)$ is replaced by $\sqrt{2}\sigma(M)$, since the effective variance is doubled.

The left-hand side of Eq. (11) can be expressed as a function of $\sigma(M)$ alone whenever $\sigma(M)$ is a monomial in $M - 1$. In that case, the equation has a single root with respect to σ and can be solved numerically to yield $\sigma_{\text{opt}} \equiv \sigma(M_{\text{opt}})$ and $\bar{P}_{\text{opt}} \equiv \bar{P}(M_{\text{opt}})$ as *constants*. In the specific and, as will be seen, surprisingly common case that $\sigma(M) = \sigma_2/(M - 1)$, for some constant σ_2 , Eq. (11) reduces to $\sigma_{\text{opt}} = f(\sigma_{\text{opt}})$. The values for this case are tabulated in Table I. Note that these values are universal and independent of σ_2 . However, M_{opt} and $k(M_{\text{opt}})$ do depend on σ_2 . Solving this particular form of Eq. (11) for M_{opt} yields

$$M_{\text{opt}} = 1 + \frac{\sigma_2}{\sigma_{\text{opt}}}, \quad (12)$$

where the appropriate value of σ_{opt} is used depending on the sampling scheme (stochastic, deterministic, or convective) and decreased by a factor of $\sqrt{2}/2$ for a parallel scheme.

In general, σ_2 will be the product of: (1) a “boundary” term involving the minimal and maximal values of θ and (2) a derivative of a thermodynamic average with respect to θ . The linear response relations here only dictate that such derivatives are constant on the interval between states. Even so, the derivative may be approximately constant on a wider interval such that a spacing rule can be derived between many states. If this is not the case, then the parameter range must be divided into different segments with slightly different spacing rules.

1. Simulated tempering/scaling

When the parameter is the temperature the following expression is obtained:

$$\sigma(M) = \sqrt{c} \frac{\alpha(M) - 1}{\sqrt{\alpha(M)}}, \quad (13)$$

where $\alpha(M) \equiv T_{i+1}/T_i = (T_M/T_1)^{1/(M-1)}$, $c \equiv d\langle U \rangle / d(k_B T)$ is the dimensionless excess heat capacity, and $\langle U \rangle$ is the mean potential energy. The exponential spacing rule implied by $\alpha(M)$ derives from the fact that this definition equalizes the mean acceptance probability for neighboring temperatures on the interval $[T_1, T_M]$ assuming that c is constant. Using a binomial expansion and inserting the resulting expression into Eq. (7) recovers the relation originally due to the work of Kofke.⁹ However, the rather unexpected (at least to us) approximation $(x - 1)/\sqrt{x} \approx \ln x$ is both more accurate and convenient

TABLE I. All three neighbor schemes have distinct constants arising from Eq. (12). The optimal mean transition rate, $\bar{P}_{\text{opt}}$, depends on the scheme being serial or parallel in a non-obvious way [see Eq. (7)], while σ_{opt} is simply smaller by a factor of $\sqrt{2}/2 \approx 0.707$ in the parallel case.

	σ_{opt}^a	$\bar{P}_{\text{opt}}^b$	
		Serial	Parallel
Stochastic	2.381	23.4	9.2
Deterministic	1.729	38.7	22.2
Convective	1.504	45.2	...

^aIn $k_B T$.

^bIn percent.

because it recovers Eq. (12) such that¹¹

$$\sigma_2 = \sqrt{c} \ln \frac{T_M}{T_1}. \quad (14)$$

Assuming a parallel scheme recovers the equation derived by Nadler and Hansmann¹¹ using rather different assumptions (i.e., assuming that c is large).

A similar derivation to the above can also be employed for simulated scaling³⁸ as well as for selective scaling methods, such as solute tempering/scaling.^{39,40} For example, if different interactions are scaled on different temperature ranges, the expression is

$$\sigma_2 = \sqrt{c_a} \ln \frac{T_M^a}{T_1} \sqrt{1 + \frac{c_b}{c_a} \left(\ln \frac{T_M^b}{T_M^a} \right)^2}, \quad (15)$$

where T_M^a and T_M^b represent different upper bounds and c_a and c_b are the heat capacities for two different sets of interactions (e.g., dihedral and non-bonded interactions). This just adds an extra factor to the standard case which, as expected, tends to unity when $T_M^a \approx T_M^b$ and/or $c_a \gg c_b$. Extension to more than two sets of interactions follows trivially.

A slightly different result is obtained when the thermostating or scaling is partitioned between different sets of particles. For example, solute scaling^{39,40} divides the system into solute and solvent particles. In this case one has

$$\sigma_2 = \sqrt{c_{uv}} \ln \frac{T_M}{T_1} \sqrt{\frac{1}{4} + \frac{c_{uu}}{c_{uv}}}, \quad (16)$$

where c_{uu} and c_{uv} are the heat capacities for the solute-solute and solute-solvent interactions, respectively. Note that the extra factor now tends to 1/2 as the system increases in size.

2. Linear coupling

A broad class of simulation algorithms use a linear perturbation yielding a reduced potential of the form

$$u(\mathbf{x}) = u_0(\mathbf{x}) + \lambda \Delta u(\mathbf{x}), \quad (17)$$

where $u_0(\mathbf{x})$ is some base potential, $\lambda \in [\lambda_1, \lambda_M]$ is a coupling constant, and $\Delta u(\mathbf{x})$ is a perturbing energy term. This is true for many alchemical free energy methods, particularly those involving charge perturbations. Once again, Eq. (12) is obtained but now with

$$\sigma_2 = \sqrt{\left| \frac{\partial^2 \Delta f}{\partial \lambda^2} \right|} (\lambda_M - \lambda_1). \quad (18)$$

Here Δf is the free energy between states 1 and M and its second derivative is assumed to be (approximately) constant on the interval $[\lambda_1, \lambda_M]$. The spacing between values is thus fixed at $(\lambda_M - \lambda_1)/(M - 1)$.

If the transformation is a change from a neutral to charged species inside an approximately spherical cavity, the expression for Δf can be replaced by the Born solvation expression leading to

$$\left| \frac{\partial^2 \Delta f}{\partial \lambda^2} \right| = \frac{k_C q^2}{R}, \quad (19)$$

where k_C is Coulomb's constant divided by $k_B T$, q is the charge, and R is the effective radius (including factors pertaining to the dielectric of the environment). This again rather

clearly articulates some obvious intuition; the optimal number of λ values increases with the magnitude of the charge and inversely as the square root of the extent over which that charge is distributed. Similar arguments hold when perturbing a point dipole or dispersive potential. The main difference is a change in the power to which R is raised.

Although the conceptual motivation is considerably different, the potential form for Gaussian accelerated molecular dynamics (GaMD) is also the same as Eq. (17).⁴¹ In this case, Δu is known as the “boost” potential and provides a global bias to enhance sampling of high energy states. If a series of boost potentials are considered up to some maximum boost, the free energy term becomes

$$\left| \frac{\partial^2 \Delta f}{\partial \lambda^2} \right| = \frac{1}{2} \sigma_{\Delta u}^2, \quad (20)$$

where $\sigma_{\Delta u}^2$ is the variance of the full boost potential at a given λ (the *scaled* boost is still used for dynamics). By construction, the variance will decrease monotonically as λ increases and so a reasonable approximation is to use the variance at the lowest λ value. This is typically an unbiased simulation corresponding to $\lambda = 0$, in which case simulations must already be performed under these conditions in order to estimate the boost parameters.⁴¹ A replica exchange variant of GaMD was recently reported; extension to a serial variant is largely trivial.

3. Harmonic biasing (“umbrella sampling”)

Many umbrella sampling protocols in common use employ a localizing bias, usually of a harmonic form, along a few selected collective variables. The potential takes the form

$$u(\mathbf{x}) = u_0(\mathbf{x}) + \frac{1}{2} k (\xi - \xi_0)^2, \quad (21)$$

where ξ is a collective variable to be localized about some bias center ξ_0 with force constant k . The overlap between states is governed by the strength of the bias, k (assumed to be constant for simplicity), as well as the underlying character of the potential of mean force (PMF) along ξ . If the PMF is locally linear, Eq. (12) is encountered yet again with

$$\sigma_2 = \sqrt{k} (\xi_M - \xi_1), \quad (22)$$

where $\xi_M - \xi_1$ is the interval on which the linear approximation holds. This is unfortunately unlikely to be accurate over very long distances. Extending the PMF approximation to a higher order (e.g., using a Taylor series) might be more robust but still likely to fail over large intervals. Nonetheless, some subtle conclusions can be drawn. Perhaps the most non-obvious result is that the optimal spacing mostly depends on the length of the interval and the bias strength, but not the gradient of the underlying PMF (although it would if quadratic terms were considered). This linear approximation is essentially the same as assuming that the curvature of the bias is much greater than that of the PMF on the interval.

One might also consider inverting the equation for the optimal number of biases to yield an equation for the optimal

force constant and spacing. Taking $\Delta\xi \equiv (\xi_M - \xi_1)/(M_{\text{opt}} - 1)$, the relationship becomes

$$k_{\text{opt}}\Delta\xi^2 = \sigma_{\text{opt}}^2. \quad (23)$$

In the case of replica exchange sampled with a deterministic scheme, the right-hand side is ≈ 1.494 , which is exactly the equation derived by Park, Kim, and Im.¹³ Expanding this and Eq. (22) to the case of variable k is also potentially instructive, but a detailed analysis is beyond the scope of this work. More comprehensive treatments have been presented elsewhere.^{13–15}

4. Constant pH

During a simulation at constant pH, the system potential is extended to include the number of protons which is coupled to a chemical potential specified by the pH value,

$$u(\mathbf{x}, \lambda) = u_0(\mathbf{x}; \lambda) + \ln 10 p H n_\lambda. \quad (24)$$

Here λ is a vector specifying if protonation sites are occupied (1) or unoccupied (0) and n_λ is the sum of its elements. Unfortunately, the variance cannot be cast as a product involving a constant derivative since this depends on the distribution of pK_a values in the system. Nonetheless, a constant upper bound can be easily obtained as

$$\sigma_2 \leq \frac{\ln 10}{2} \sqrt{N_s} (pH_M - pH_1), \quad (25)$$

where the N_s is the number of protonation sites that have pK_a values roughly between pH_1 and pH_M . This upper bound is only reached in the extremely rare case that all of the pK_a values are degenerate and so might be an overestimate in practice. Finally, it is noted that this once again indicates that evenly spaced pH values should be used, although this is less accurate at the extremes where only a few pK_a values are found.

III. PRACTICAL CONSIDERATIONS

The linear response relations given by Eqs. (3) and (4) combined with the state optimization criterion from Eq. (12) yield a straightforward algorithm for estimating optimal parameters for expanded ensemble simulations with essentially *any* set of thermodynamic parameters. In what follows, it will be assumed that the user has selected: (1) one or more parameter dimensions, (2) the ranges on which those parameters are to be sampled, and (3) a Gibbs sampling technique with which to sample them. The following procedure is then followed:

1. Estimate σ_2 , compute M_{opt} , and allocate the set of state parameters.
2. Run a short simulation with each parameter combination and accumulate averages for the instantaneous work, w , between neighboring states. Compute the pairwise free energies (i.e., estimate the state weights) and variances from linear response.
3. (optional) Make a new estimate for σ_2 ; go to (1).
4. Run an expanded ensemble simulation using the current weights.
5. (optional) Validate the weight and variance estimates using WHAM; go to (3).

Some of these points deserve further elaboration.

Using Eqs. (3) and (4) is relatively straightforward in most situations. In many cases, only a few energy terms need to be accumulated in order to compute several work values. For example, consider simulated scaling applied to the non-bonded terms of a molecular mechanics force field; the potential has the form

$$u_k(\mathbf{x}) = u_b(\mathbf{x}) + \gamma_k u_{\text{nb}}(\mathbf{x}), \quad (26)$$

where $u_b(\mathbf{x})$ and $u_{\text{nb}}(\mathbf{x})$ are the (reduced) bonded and non-bonded potentials, respectively, and γ_k is a scaling parameter on the interval (0, 1]. Setting $\gamma_k = T_1/T_k$ permits interpreting this as an inverted form of simulated tempering where attenuating forces have an effect equivalent to increasing velocities.³⁸ The mean instantaneous work at “temperature” i going to j is simply

$$\langle w \rangle_i = (\gamma_j - \gamma_i) \langle u_{\text{nb}} \rangle_i, \quad (27)$$

such that Eqs. (3) and (4) only require accumulating averages of $u_{\text{nb}}(\mathbf{x})$ during a simulation.

Once the necessary averages have been computed, it is trivial to compute estimates for Δf_{ij} and σ_{ij}^2 between neighboring states. Since the sample weights $g_i \approx \Delta f_{0i}$ are arbitrarily taken relative to state 0, these must be computed by sequential summation (e.g., $g_2 = \Delta f_{01} + \Delta f_{12}$). The state weights are thus recursively estimated from Eq. (3) such that

$$g_i = g_{i-1} + \frac{\langle w \rangle_i + \langle w \rangle_{i-1}}{2}, \quad (28)$$

where $w(\mathbf{x}) \equiv u_i(\mathbf{x}) - u_{i-1}(\mathbf{x})$. This is slightly more complicated in higher dimensions, since then the estimate may depend on more than one neighbor. While any sequential sum of free energies may yield a reasonable estimate, a self-consistent estimate is best obtained by averaging the pairwise free energies with respect to all neighbors. For example, in two-dimensions, the double-indexed state with weight g_{11} relative to the state $g_{00} \equiv 0$ has two neighbors with weights g_{10} and g_{01} . The doubly indexed weight estimation equation then becomes

$$g_{ij} = \frac{g_{(i-1)j} + g_{i(j-1)} + \Delta f_{ij,i(j-1)} + \Delta f_{ij,(i-1)j}}{2}, \quad (29)$$

where the free energies are estimated from the appropriate average work values along each parameter axis.

As discussed above, assuming that the variance is constant across a range of parameter values can be used to estimate the optimal number and placement of states on that interval. A reasonable way to estimate this is to simply take the average of pairwise variances on the interval. For example, for states $i = 1, 2, \dots, M$ with sequential estimates of $\sigma_{i,i+1}^2$, a single variance can be obtained as

$$\overline{\sigma^2} = \frac{1}{M-1} \sum_{i=1}^{M-1} \sigma_{i,i+1}^2. \quad (30)$$

The appropriate quantity in the expression for σ_2 (e.g., the heat capacity or second derivative of the free energy) can then be determined from $\overline{\sigma^2}$.

In order to check that $\overline{\sigma^2}$ is in fact a well-defined constant, one may also apply a linear fit to the averages such that the slope is a thermodynamic derivative. The exact quantities in question depend on the parameters. For example, the slope of a linear fit to the mean energy versus temperature is the heat

capacity. Most importantly, the correlation coefficient of the fit can be used to assess the fidelity of the linear approximation. It may also be possible to devise a linear expression for the state weights themselves as a function of the state parameter. For example, in the case of linear coupling (such as a charge mutation) it is reasonable to assume that the second derivative of the free energy is constant. This yields the expression

$$\Delta f(\theta) = \frac{1}{2} \left| \frac{\partial^2 \Delta f}{\partial \theta^2} \right| \theta^2, \quad (31)$$

which is linear in the quantity $\theta^2/2$. Since $\Delta f(0) \equiv 0$, the first state can be omitted from the fit and the intercept can be constrained to be zero. Furthermore, the squared standard error of the free energy estimates should increase linearly along the sequence so that a weighted regression can be performed. The three estimates, the mean variance, and the slopes of the linear and linearized fits should all be comparable and large deviations between the estimates may be taken as a sign that the linear approximation is not valid over the full interval. In such instances, multiple spacing rules may be necessary.

IV. COMPUTATIONAL METHODS

The serial expanded ensemble protocol outlined above was tested in NAMD 2.12⁴³ using its Tcl scripting interface. A series of one-dimensional expanded ensemble simulations were run using a linear alchemical coupling parameter as the state parameter. Acetic acid was modeled with the CGenFF force field and solvated in a ~ 20 Å truncated octahedron containing 216 rigid mTIP3P water molecules.⁴⁴ Coupling between the protonated and deprotonated forms was accomplished via a “dual-topology” paradigm with additional, zero-length bonds between equivalent but otherwise non-interacting atoms (force constant 100 kcal/mol-Å²).⁴⁵ Langevin dynamics was employed at 298 K with a friction coefficient of 1 ps⁻¹ and a 2 fs time step; all bonds with hydrogen were kept rigid. Periodic boundary conditions were applied using particle mesh Ewald (grid spacing of 1.5 Å) and force switching of the Lennard-Jones interactions between 8 and 10 Å, after which an isotropic long-range approximation was used. Calibration runs to determine the state weights and variances were run for 60 ps in each state with the first 10 ps discarded as equilibration and the instantaneous work was stored every ps. Production expanded ensemble simulations were run for 50 ns with MC moves every 5 ps and data saved every ps.

We previously reported results from two-dimensional expanded ensemble simulations of five terminally blocked amino acids in which the state parameters were the thermostat temperature and an alchemical coupling parameter for the protonated and deprotonated forms.⁴⁶ The specific MD settings of these simulations can be found in that work. Here we are primarily concerned with sampling behavior and the effectiveness of the linear response scheme in determining states and weights. Furthermore, the use of multiple temperatures also permits a detailed analysis of the enthalpic and entropic components of the free energy by using the van’t Hoff equation (see the [supplementary material](#)).

V. RESULTS AND DISCUSSION

A. Linear alchemical coupling: Solvated acetic acid

Acetic acid was chosen as a test system because transformation between its protonated and deprotonated states represents a net charging of a non-spherical molecule possessing a permanent dipole. This example is rather complex in comparison to charging, say, an alkali ion, for which linear response is already known to yield excellent results.⁴⁷ It is thus arguably non-obvious that a quadratic, Born-like relationship should hold for the relative free energy. Nonetheless, a linear least-squares fit of the state weights versus $(1 - (1 - \lambda)^2)/2$ yielded a correlation coefficient of 0.999 in all cases.

A single calibration of the state weights was run for each ladder of states, starting from the arbitrary, but reasonable, initial estimate of five coupling values. The estimated variance by all three methods (mean variance and both linear fits) indicated that the optimal number of states is 8, 11, and 12 (± 1) for the stochastic/Metropolized, deterministic, and convective methods, respectively. An expanded ensemble simulation for each method on each ladder of states was then run for comparison (Fig. 3). After the simulations were completed, a single variance [$230.6 (k_B T)^2$] was estimated using the final free energies from WHAM; this is used as the parameter for all comparisons with linear response. In addition, a constant scale factor of 1.6 was used for the results from Metropolized independent sampling.

In all cases, the observed and predicted performance metrics ($\bar{P}$ and k) are in good agreement, although there is some slight overestimation, especially at larger state counts. This lack of agreement is not surprising since (1) the linear response approximation is not exact and (2) the state weights are not

FIG. 3. Several serial expanded ensemble simulations of acetic acid are closely in line with the results from linear response. All four methods easily converge to the correct free energy (top panel), although appreciable decrease in the uncertainty occurs for larger numbers of states (*nb*: the x-coordinate has been shifted for clarity). The optimal mean acceptance probability, $\bar{P}$ (middle panel, black dashed lines), and round trip rate, k (bottom panel, solid colored lines), are predicted using the relative free energy from WHAM ($115.3 k_B T$). In order to aid visualization with the log-scale, the round trip rate curves have been scaled to minimize the root mean square deviation resulting in coefficients of 0.61, 0.52, 0.82, and 0.98, respectively.

perfectly estimated. The latter of these will clearly always diminish performance because the mean acceptance probability decreases with the free energy of the transformation, which is only zero when the state weight is exact.²⁹ While these discrepancies are not ideal, it is very encouraging to see that the disagreement is global and can be largely corrected by simply rescaling the estimates (Fig. 3, bottom panel). Furthermore, the observed maxima are indeed as predicted, although there is often a pronounced plateau as M increases. This plateau behavior is not surprising since the dependence of the round trip rate on M is generally quite small near the optimal value (see the [supplementary material](#)). It is therefore recommendable to start with a small number of states and increase after calibration, rather than try to remove states, which may yield false negatives. Finally, as predicted previously, the empirical correction for Metropolized independence sampling seems to fail at large M . This is because the algorithm permits large steps in λ , beyond nearest neighbors, and is thus less impeded by the introduction of extra states than the standard stochastic approach.

B. Multidimensional sampling: Titratable dipeptides

Multidimensional expanded ensembles are an attractive simulation approach because they permit more flexibility in the state parameters being explored. For example, one might combine sampling along a quantity of interest (e.g., a specific reaction coordinate or order parameter) with an enhanced sampling technique (e.g., simulated tempering). Since the number of states is roughly polynomial in the parameter dimension, a parallel scheme for this kind of simulation can be quite expensive, so the serial scheme is rather attractive, especially when faced with limited computational resources. As an illustration, we consider the alchemical deprotonation reaction of several titratable dipeptides. Besides an alchemical coupling parameter between the protonated and deprotonated states, the thermostat temperature was also varied between 290 and 325 K.

In addition to accelerating the relatively slow dihedral degrees of freedom, varying the temperature also permits computing the relative deprotonation free energy as a function of temperature and, by extension, the enthalpic and entropic components (Fig. 4). Just as with linear response, the enthalpic and entropic components are based on thermodynamic derivatives and/or fluctuations and are thus much more sensitive to statistical errors than the free energy. However, whereas the mean work depends on a derivative of the relative free energy, the enthalpy (or internal energy) depends on a derivative of a relative free energy difference. A finite differencing approximation for the enthalpy thus requires four free energies, all with non-zero statistical errors. In the present work, this leads to relative uncertainties $>100\%$, even when using minimum variance estimates from WHAM. As an alternative, because the relative free energy difference can be estimated over a temperature range, one can use a parametric approach to extract parameters at a fixed temperature (Fig. 4). Computing the relative internal energy at 298 K by this method leads to relative uncertainties of $\sim 15\%$, which is still substantial, but at least recovers one significant digit compared to finite differencing (see the [supplementary material](#)). Finally, it is worth noting that

FIG. 4. An intrinsic outcome of using WHAM on expanded ensemble data from a combined alchemical/simulated tempering simulation is the relative free energy between endpoints as a function of temperature. Here this is the deprotonation free energy of several titratable dipeptides. Data points represent the states that were directly sampled, but WHAM formally provides a continuous estimate that should be reliable across the whole temperature range. The solid lines represent parametric fits using the relative internal energy and relative heat capacity at 298 K. Error bars represent 99% confidence intervals.

this approach is not unique to the expanded ensemble approach used here; however, the necessary analysis from WHAM is itself an intrinsic part of the expanded ensemble workflow and, aside from the calibration of the state weights, requires no additional cost beyond running equivalent alchemical simulations at multiple fixed temperatures.

A key assumption in the above analysis is that the data are being *interpolated* rather than *extrapolated* to the endpoints at $\lambda = 0/1$. It is therefore critical that the simulation actually explores the full state space rather than localizing sampling in

FIG. 5. Representative state diffusion data from glutamate dipeptide indicate that the full state space in both alchemical coupling and temperature is rapidly explored. For clarity, the data have been taken at 50 MD/MC cycle intervals and colored in sequential blocks of 4 ns. In the top panel, the data have also been slightly shifted about the grid of discrete states in order to avoid overlap. This shows that the full range of states (42 total, bottom panel) is traversed multiple times and that the system is not “trapped” in a local region of state space, although sampling does seem to occur most heavily at median values (top panel).

a small region. This can be verified qualitatively by visualizing the state parameters as a function of time (Fig. 5). Indeed, the system does traverse the full range of states, even within a relatively small fraction of the full simulation time.

VI. CONCLUSION

Expanded ensemble simulations are a widely used and powerful approach for enhancing sampling and exploring properties along thermodynamic variables. However, obtaining good efficiency requires careful and often *ad hoc* calibration of the states and weights. Parallel approaches, such as replica exchange, mitigate part of this problem, but clearly have lower sampling efficiency and high cost scaling, especially in higher dimensions. Here we have presented generic linear response relationships that can be used to optimize both serial and parallel expanded ensemble simulations in any number of dimensions by choosing states that maximize the mean round trip rate for several sampling schemes. In the case of temperature and harmonic biases, the state selection rules reduce to derivations previously found in the literature for those specific cases. The equations also generalize past efforts to optimize the sampling weights when the variable is temperature or an alchemical coupling parameter. Taken together, a clear protocol is developed for optimizing a simulation procedure with any combination of state parameters in any number of dimensions.

SUPPLEMENTARY MATERIAL

See [supplementary material](#) for a detailed analysis of maximizing the round trip rate for multiple schemes and for the numerical procedure and fitting parameters for enthalpic/entropic analysis.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation through Grant No. MCB-1517221 (B.R.) and the National Institutes of Health Grant Nos. U54-GM087519 (B.R.) and P41-RR005969. Computational resources were provided by the University of Chicago Research Computing Center and the Argonne Leadership Computing Facility.

- ¹A. P. Lyubartsev, A. A. Martsinovski, S. V. Shevkunov, and P. N. Vorontsov-Velyaminov, *J. Chem. Phys.* **96**, 1776 (1992).
- ²J. D. Chodera and M. R. Shirts, *J. Chem. Phys.* **135**, 194110 (2011).
- ³Y. Sugita and Y. Okamoto, *Chem. Phys. Lett.* **314**, 141 (1999).
- ⁴D. M. Zuckerman and E. Lyman, *J. Chem. Theory Comput.* **2**, 1200 (2006).
- ⁵A. C. Pan, T. M. Weinreich, S. Piana, and D. E. Shaw, *J. Chem. Theory Comput.* **12**, 1360 (2016).
- ⁶S. Park and V. S. Pande, *Phys. Rev. E* **76**, 016703 (2007).
- ⁷E. Marinari and G. Parisi, *Europhys. Lett.* **19**, 451 (1992).
- ⁸S. Park, *Phys. Rev. E* **77**, 016708 (2008).
- ⁹D. A. Kofke, *J. Chem. Phys.* **117**, 6911 (2002); **121**, 1167 (2004).
- ¹⁰C. Predescu, M. Predescu, and C. V. Ciobanu, *J. Chem. Phys.* **120**, 4119 (2004).
- ¹¹W. Nadler and U. H. E. Hansmann, *J. Phys. Chem. B* **112**, 10386 (2008).
- ¹²R. Denschlag, M. Lingenheil, and P. Tavan, *Chem. Phys. Lett.* **473**, 193 (2009).
- ¹³S. Park, T. Kim, and W. Im, *Phys. Rev. Lett.* **108**, 108102 (2012).

- ¹⁴S. Park and W. Im, *J. Chem. Theory Comput.* **10**, 2719 (2014).
- ¹⁵D. S. Dashti and A. E. Roitberg, *J. Chem. Theory Comput.* **9**, 4692 (2013).
- ¹⁶E. Rosta and G. Hummer, *J. Chem. Phys.* **131**, 165102 (2009).
- ¹⁷B. K. Radak, M. Romanus, E. Gallicchio, T.-S. Lee, O. Weidner, N.-J. Deng, P. He, W. Dai, D. M. York, R. M. Levy, and S. Jha, in *Proceedings of the Conference on Extreme Science and Engineering Discovery Environment: Gateway to Discovery, XSEDE '13* (ACM, San Diego, CA, USA, 2013), pp. 1–8; B. K. Radak, M. Romanus, T.-S. Lee, H. Chen, M. Huang, A. Treikalis, V. Balasubramanian, S. Jha, and D. M. York, *J. Chem. Theory Comput.* **11**, 373 (2015).
- ¹⁸C. H. Bennett, *J. Comput. Phys.* **22**, 245 (1976).
- ¹⁹R. W. Zwanzig, *J. Chem. Phys.* **22**, 1420 (1954).
- ²⁰C. Jarzynski, *Phys. Rev. Lett.* **78**, 2690 (1997).
- ²¹G. E. Crooks, *Phys. Rev. E* **60**, 2721 (1999).
- ²²M. R. Shirts, E. Bair, G. Hooker, and V. S. Pande, *Phys. Rev. Lett.* **91**, 140601 (2003).
- ²³K. S. Shing and K. E. Gubbins, *Mol. Phys.* **46**, 1109 (1982).
- ²⁴N. Lu, J. K. Singh, and D. A. Kofke, *J. Chem. Phys.* **118**, 2977 (2003).
- ²⁵H. Oberhofer, C. Dellago, and P. L. Geissler, *J. Phys. Chem. B* **109**, 6902 (2005).
- ²⁶M. R. Shirts and J. D. Chodera, *J. Chem. Phys.* **129**, 124105 (2008).
- ²⁷J. Gore, F. Ritort, and C. Bustamante, *Proc. Natl. Acad. Sci. U. S. A.* **100**, 12564 (2003).
- ²⁸D. A. Sivak and G. E. Crooks, *Phys. Rev. Lett.* **108**, 190602 (2012).
- ²⁹B. K. Radak and B. Roux, *J. Chem. Phys.* **145**, 134109 (2016).
- ³⁰S. Kumar, D. Bouzida, R. H. Swendsen, P. A. Kollman, and J. M. Rosenberg, *J. Comput. Chem.* **13**, 1011 (1992); M. Souaille and B. Roux, *Comput. Phys. Commun.* **135**, 40 (2001); E. Gallicchio, M. Andrec, A. K. Felts, and R. M. Levy, *J. Phys. Chem. B* **109**, 6722 (2005); J. D. Chodera, W. C. Swope, J. W. Pitera, C. Seok, and K. A. Dill, *J. Chem. Theory Comput.* **3**, 26 (2007).
- ³¹Z. Tan, E. Gallicchio, M. Lapelosa, and R. M. Levy, *J. Chem. Phys.* **136**, 144102 (2012).
- ³²W. Nadler and U. H. E. Hansmann, *Phys. Rev. E* **76**, 065701R (2007).
- ³³M. Lingenheil, R. Denschlag, G. Mathias, and P. Tavan, *Chem. Phys. Lett.* **478**, 80 (2009).
- ³⁴Y. G. Spill, G. Bouvier, and M. Nilges, *J. Comput. Chem.* **34**, 132 (2013).
- ³⁵J. S. Liu, *Biometrika* **83**, 681 (1996).
- ³⁶This requires the minor approximation $2M - 1 \approx 2M$. Since M is an integer greater than one, this is frequently exact.
- ³⁷*Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*, Applied Mathematics, edited by M. Abramowitz and I. A. Stegun (U.S. Government Printing Office, Washington, D.C., 1964).
- ³⁸A. E. Mark, W. F. van Gunsteren, and H. J. C. Berendsen, *J. Chem. Phys.* **94**, 3808 (1991); H. Tsujishita, I. Moriguchi, and S. Hirono, *J. Phys. Chem.* **97**, 4416 (1993).
- ³⁹R. Denschlag, M. Lingenheil, P. Tavan, and G. Mathias, *J. Chem. Theory Comput.* **5**, 2847 (2009).
- ⁴⁰P. Liu, B. Kim, R. A. Friesner, and B. J. Berne, *Proc. Natl. Acad. Sci. U. S. A.* **102**, 13749 (2005); L. Wang, R. A. Friesner, and B. J. Berne, *J. Phys. Chem. B* **115**, 9431 (2011).
- ⁴¹Y. Miao, V. A. Feher, and J. A. McCammon, *J. Chem. Theory Comput.* **11**, 3584 (2015).
- ⁴²Y. M. Huang, J. A. McCammon, and Y. Miao, *J. Chem. Theory Comput.* **14**, 1853 (2018).
- ⁴³J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, and K. Schulten, *J. Comput. Chem.* **26**, 1781 (2005).
- ⁴⁴K. Vanommeslaegh and A. D. MacKerell, Jr., *J. Chem. Inf. Model.* **52**, 3144 (2012); K. Vanommeslaegh, E. P. Raman, and A. D. MacKerell, Jr., *ibid.* **52**, 3155 (2012); A. D. MacKerell, Jr., D. Bashford, M. Bellott, R. L. Dunbrack, Jr., J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher III, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus, *J. Phys. Chem. B* **102**, 3586 (1998).
- ⁴⁵P. H. Axelsen and D. Li, *J. Comput. Chem.* **19**, 1278 (1998).
- ⁴⁶B. K. Radak, C. Chipot, D. Suh, S. Jo, W. Jiang, J. C. Phillips, K. Schulten, and B. Roux, *J. Chem. Theory Comput.* **13**, 5933 (2017).
- ⁴⁷*Free Energy Calculations: Theory and Applications in Chemistry and Biology*, Springer Series in Chemical Physics Vol. 86, edited by C. Chipot and A. Pohorille (Springer, New York, 2007).