JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

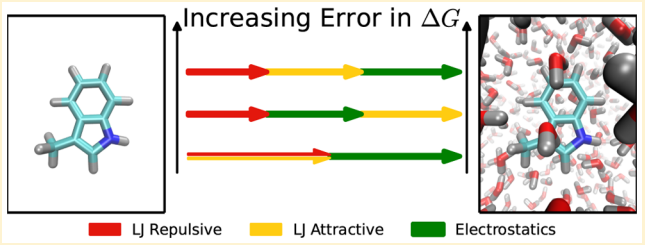
Linear Basis Function Approach to Efficient Alchemical Free Energy Calculations. 2. Inserting and Deleting Particles with Coulombic Interactions

Levi N. Naden and Michael R. Shirts*

Department of Chemical Engineering, University of Virginia, Charlottesville, Virginia 22904, United States

Supporting Information

ABSTRACT: We extend our previous linear basis function approach for alchemical free energy calculations to the insertion and deletion of charged particles in dense fluids. We compute a near optimal statistical path to introduce Coulombic interactions into various molecules in solution and find that this near optimal path is only marginally more efficient than simple linear coupling of electrostatics in all cases where a repulsive core is already present. We also explore the order in which nonbonded forces are coupled to the environment in alchemical transformations. We test two sets of Lennard-Jones basis functions, a Weeks−Chandler−Andersen (WCA) and a 12−6 decomposition of the repulsive and attractive forces turned on in sequence along with changes in charge, to determine a statistically optimized order in which forces should be coupled. The WCA decomposition has lower statistical uncertainty as coupling the attractive r^{-6} basis function contributes non-negligible statistical error. In all cases, the charge should be coupled only after the repulsive core is fully coupled, and the WCA attractive portion can be coupled at any stage without significantly changing the efficiency. The statistical uncertainty of two of the basis function approaches with charged particles is nearly identical to the soft core approach for decoupling electrostatics, though the correlation times for sampling are often longer for a soft core electrostatics approach than the basis function approach. The basis function approach for introducing or removing molecules or functional groups thus represents a useful alternative to the soft core approach with a number of clear computational advantages.



1. INTRODUCTION

Calculations of free energy differences by computer simulation can provide valuable insight into molecular thermodynamics without experimental measurements. Rigorous statistical methods allow computer simulations to estimate free energy differences through clearly defined thermodynamic states and sampling schemes.[1] Alchemical transformations are a common computational technique which provide the difference of free energies at two end states along a thermodynamic path connecting them. Sampling this path provides the information needed for statistical analysis and estimation of the free energy difference. A popular application of alchemical methods is the computation of drug binding in order to understand the molecular details of drug action or to reduce the number of molecules required to be synthesized.[2−7] However, there are a wide range of other applications for alchemical methods where knowledge of the chemical potential of a small molecule in some environment may be useful, such as estimating membrane permeability or solubility.

There can be significant computational costs associated with alchemical free energy calculations due to the number of intermediate states along the thermodynamic pathway that must be sampled to achieve sufficiently high precision. Intermediate states are chosen to provide larger phase space overlap between adjacent states which reduces the uncertainty in evaluating free energies and thermodynamic expectation values. A range of methods can be used to estimate free energies including exponential averaging (EXP),[8] thermodynamic integration (TI), the weighted histogram analysis method (WHAM),[9,10] the Bennett acceptance ratio (BAR),[11] and its multistate version (MBAR).[12] Each of these methods has its own advantages and disadvantages.[13,14] Estimating free energies of solvation and absolute binding affinities typically require numerous intermediate states because the phase space overlap is low between the state where the entire, large solute is fully interacting with its environment and the state where the solute is fully noninteracting.

A way to reduce this computational cost is to design thermodynamic paths which improve the phase space overlap between intermediate states.[15−19] The "soft core" approach is one way to define a pathway of potential energy functions from a fully interacting state to a fully noninteracting state.[20,21] The pathways in potential energy space that can be described by this approach provide decent phase space overlap between neighboring states, which reduces the statistical uncertainty of free energy calculations performed along the path and thus the number of samples needed to obtain a given level of

Received: November 23, 2014

DOI: 10.1021/ct501047e
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

precision.[11] Minimizing the statistical uncertainty of calculations performed along an alchemical path is equivalent to minimizing the thermodynamic length of the path.[17−19,22−24] The path which minimizes the total uncertainty for the transformation is one which has an equal contribution to the uncertainty across every point along the path, though the minimum is not necessarily achievable via pairwise potentials.[19]

To clarify the objectives of the article, we must introduce some statistical definitions. Using thermodynamic integration (TI) to calculate free energy differences involves calculating the average of $\partial u/\partial\lambda$ at each value of λ and then integrating this average numerically from $\lambda = 0$ to $\lambda = 1$ to obtain the total free energy of the transformation. In the standard case when simulations at each state are performed independently, the variances in the mean of $\partial u/\partial\lambda$, which are statistically independent, weighted by the scalar factors used in the numerical integration method, will add in the square to produce the variance of the total calculation of the free energy difference. The square root of this variance is the standard deviation of the overall calculation. By adjusting the pathway of potential energy functions that connect the end states, we can lower the variance of the overall calculation by increasing the overlap between individual Hamiltonians simulated.

When comparing pathways, the most statistically efficient path is roughly independent of the method performing the analysis such as TI, BAR, or MBAR[19,24] Since the variance of the mean in $\partial u/\partial\lambda$ at any simulated state will be proportional to the number of samples at that state, the overall variance of the entire process will also be proportional to the number of samples used for the entire process at all states, assuming a fixed proportion of samples per state and samples that are statistically decorrelated. We can thus normalize this variance in the calculation by the total number of decorrelated samples used in the calculation to obtain a measure of the statistical efficiency of the pathway that is independent of the length of the simulations. We will call this measure the variance of the pathway (or variance of the path).

Choosing the most statistically efficient pathway (i.e., the path with lowest variance of the pathway) through alchemical space is nontrivial. A number of efforts have been made to minimize the variance of the soft core path by changing the parameters in the potential energy function,[15,16,18,22,25−28] and an efficient formalism has been developed for calculating minimal variance paths within a given function space.[18,19,24]

The functional form of the soft core potentials adds significant complexity when implementing free energy methods on new, highly parallelized architectures such as GPUs or many integrated core (MIC) platforms. As described in our previous study,[24] soft core methods require writing special code to handle the alchemical modifications to atoms, and implementing these modifications on new platforms have lagged behind the implementation of other molecular mechanics methods. If the alchemical-specific code can be removed from the inner force loop of molecular simulations, the time needed to implement free energy methods on these new platforms could be drastically reduced.

As shown in our previous article, representing the thermodynamic path with a linear combination of basis functions makes it possible to remove the alchemical specific code from inner force loops.[24] The basis function approach splits the potential energy into basis functions involving only the spatial coordinates of a system and alchemical switches depending only on the transformation coordinate. We also showed how this basis function approach made it possible to find a near statistically optimal thermodynamic path computed from a single trial simulation entirely in postprocessing. This optimized linear basis function method was just as statistically efficient as soft core methods for uncharged Lennard-Jones particles.

Our previous study[24] looked only at the introduction and removal of Lennard-Jones interactions. This study here examines the most statistically efficient way to couple electrostatics. It also identifies the lowest uncertainty sequence in which all nonbonded forces can be coupled to the rest of the system, as this order becomes an additional variable when there are multiple basis functions. To complete the analysis of the transformation of alchemical nonbonded components, the following design questions must be addressed: (1) What combination of basis functions and alchemical switches produces the minimum variance pathway to modify the charge of a molecule, including reducing the charge to zero? (2) To which sequence should the entire set of nonbonded interactions be alchemically coupled to give minimum variance pathways for insertion or deletion of a charged molecule?

The answers to these questions extend the low variance linear basis function approach to nearly all nonbonded alchemical changes. This study looks at the statistically efficient way to alchemically couple electrostatics to the environment by analyzing the charging of ions with a fully repulsive core, inversion of charge on a dipole, and turning on the charges in ethanol and 1,4-butanediol molecules in water. This study also answers the question as to which is the most statistically efficient order in which the nonbonded forces are coupled. This order, or "schedule" as it is referred to here, becomes variable in the process when forces can be alchemically changed separately and can be successfully implemented in many different ways.[3,16,29−32] This study identifies the most statistically efficient schedules across different sets of basis functions and compares them to a soft core scheme in which all forces are coupled together using the tryptophan side chain analogue, 3-methylindole. This study focuses on inserting and deleting entire atomic sites as this class of alchemical transformation is the most challenging.[3,15,16,20,21]

Results for optimal charging pathways are valid for both absolute and relative free energy calculations. Although we only benchmark the approaches for the introduction or deletion of sites with absolute free energy calculations in water, the same pathways will also be efficient for the insertion or deletion into other dense fluids. The same principles will apply for relative free energy calculations in which multiple heavy atoms are inserted, as the fluid density around these sites is mostly independent of the fluid density around the unchanging parts of the molecule. The proposed transformations may not be maximally efficient for relative free energy calculations with small changes, but such calculations require many fewer intermediate states and thus require much less optimization.

2. THEORY

The notations in this study are identical to those in our previous study,[24] with only the addition of the notation for electrostatic terms. σ_{ij} and ϵ_{ij} are the Lennard-Jones parameters between sites i and j, and q_i and q_j are the charges of particle i and j, respectively. ϵ_0 is the permittivity of vacuum. Subscripts E, R, A, and C denote electrostatic, Lennard-Jones repulsive, Lennard-Jones attractive, and capping potential energy terms, respectively, e.g., $u_E(r)$ is the electrostatic basis function. The

DOI: 10.1021/ct501047e
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

capping potential is the basis function which transitions from uncapped repulsive potential to a capped repulsive near $r = 0$.

The theory is presented in the following manner. A review of the linear basis function approach and relevant equations are presented in section 2.1. Determination of optimal thermodynamic paths is discussed in section 2.2, followed by selecting the equations to find the electrostatic switches leading to minimum variance pathways in section 2.3.

2.1. Basis Function Methods and Equations. Before we consider the electrostatic force with the linear basis function approach, we first review the relevant basis function methods and equations to establish a common framework.

The linear combination of basis functions separate the potential energy into alchemical functions, which depend on the coupling parameter λ, and configurational functions, which depend on r. This separation allows for the removal of alchemical functions from the simulation's inner force loops.[24] With this approach, the nonbonded potential takes the form

$$\begin{aligned} u(r, \lambda) &= \sum_{i}^{n} h_i(\lambda) u_i(r) + u_{\text{unaffected}}(r) \\ &= \mathbf{h}(\lambda)\mathbf{u}^T(r) + u_{\text{unaffected}}(r) \end{aligned} \tag{1}$$

where $h_i(\lambda)$ is the alchemical switch controlled by the alchemical transformation parameter λ, $u_i(r)$ is the basis function dependent on the configuration of the system but not on λ, n is the total number of basis function and alchemical switch combinations, and $u_{\text{unaffected}}$ is the system's potential energy not dependent on the alchemical change. The matrix form of the potential is written with $\mathbf{h}(\lambda) = [h_1(\lambda),h_2(\lambda),...,h_n(\lambda)]$, and $\mathbf{u}(r) = [u_1(r),u_2(r),...,u_n(r)]$ with $\mathbf{u}^T$ the transpose. Because the inner force loops only depend on the basis functions, $u_i(r)$, the force and potential of these terms can be evaluated with the default simulation code, then simply scaled by the alchemical switches, $h_i(\lambda)$, through vector operations. The basis functions can also be computed through linear algebra by re-evaluating the sampled configurations' energies in states where different sets of $h_i(\lambda) = 0$ and $h_{j\neq i}(\lambda) = 1$, even if they are not explicitly the thermodynamic path. For example, we could compute the repulsive Lennard-Jones basis function by the energy difference between a state where only the repulsive and unaffected basis functions are present and a state where only the unaffected basis function is present. Neither of these states would be along a thermodynamic path, which only removes charge from a Lennard-Jones particle, but energies can be evaluated at the states regardless to determine the basis functions.

A staged coupling approach with the linear basis functions avoids the numerical problems with traditional linear alchemical potentials, which are numerically unstable near an uncoupled state.[3,16] The basis function approach modifies a particle which does not interact with surrounding fluid by scaling in a Lennard-Jones interaction site with a finite maximum, or cap, until the sites have similar phase space occupancy at the temperatures of interest to unmodified Lennard-Jones sites.[28,33] The cap is then linearly replaced with the full Lennard-Jones interaction. The relative speed (in units of inverse coupling parameter) at which these interactions are scaled can be adjusted to minimize the variance of the path.[24] A pathway which couples the particles will be assumed here because the reverse process of decoupling a particle is statistically equivalent.

We note that different alchemical switches, $h(\lambda)$, are effectively an alternate way adjust the $\Delta\lambda$ between sampled states. Both the switches and λ are defined on a $[0,1]$ domain, and $h(\lambda)$ is simply a remapping of λ along this interval. For example, consider the simple case such that the energy equations are terms of $\lambda u(r)$. Now, suppose sampling is done with uneven $\Delta\lambda$ such that $\lambda = \{0.0, 0.04, 0.16, 0.36, 0.64, 1.00\}$. This uneven $\Delta\lambda$ is equivalent to sampling the energy function $h(\lambda)\ u(r) = \lambda^2\ u(r)$ with uniform $\Delta\lambda = 0.2$.

The Lennard-Jones basis functions in the previous study were a modified Weeks−Chandler−Andersen (WCA)[33] decomposition of the potential. The total potential for this decomposition with a cap is written as

$$u_{\text{WCA}}(r, \lambda) = h_R(\lambda)u_R(r) + h_A(\lambda)u_A(r) + h_C(\lambda)u_C(r) \tag{2}$$

The attractive basis is the WCA attractive term

$$\begin{aligned} u_A(r) &= -\epsilon_{ij} \quad \text{for } r < 2^{1/6}\sigma_{ij} \\ &= u_{\text{LJ}}(r) \quad \text{for } r \geq 2^{1/6}\sigma_{ij} \end{aligned} \tag{3}$$

where $u_{\text{LJ}}(r)$ is the Lennard-Jones potential. In order to provide numerical stability, $u_R(r)$ is set to a finite maximum and has a different functional form from standard WCA decomposition to allow for connecting the shifted Lennard-Jones potential to the finite maximum with a finite and continuous first derivative as follows:

$$\begin{aligned} u_R &= u_{\text{cap}} && \text{for } r < f_{\text{cap}}\sigma_{ij} \\ &= Ar^4 + Br^3 + Cr^2 + Dr + E && \text{for } f_{\text{cap}}\sigma_{ij} \leq r < f_{\text{switch}}\sigma_{ij} \\ &= u_{\text{LJ}}(r) + \epsilon_{ij} && \text{for } f_{\text{switch}}\sigma_{ij} \leq r < 2^{1/6}\sigma_{ij} \\ &= 0 && \text{for } r \geq 2^{1/6}\sigma_{ij} \end{aligned} \tag{4}$$

Here, f_{cap} and f_{switch} control where the smooth curve connects the Lennard-Jones and the maximum. $u_{\text{cap}} = u_{LJ}(f_{\text{cap}}\sigma_{ij}) + \epsilon_{ij}$ is the capped, constant potential; the constants A−E are the polynomial constants set by the value and first derivatives at $f_{\text{cap}}\sigma_{ij}$ and $f_{\text{switch}}\sigma_{ij}$, and the second derivative at $f_{\text{switch}}\sigma_{ij}$ to ensure a monotonic function. The numerical stability and statistical uncertainty in coupling the repulsive interaction are very sensitive values of f_{cap} and f_{switch}, so these values must be chosen with care as described in a previous study.[24] $u_C(r)$ is a step to remove the capped repulsive potential and recover the unmodified Lennard-Jones potential at $\lambda = 1$. This is written as follows:

$$u_C(r) = u_{R,\text{WCA}}(r) - u_R(r) \tag{5}$$

where $u_{R,\text{WCA}}(r)$ is the unmodified repulsive WCA term of

$$\begin{aligned} u_{R,\text{WCA}}(r) &= u_{\text{LJ}}(r) + \epsilon_{ij} \quad \text{for } r < 2^{1/6}\sigma_{ij} \\ &= 0 \quad \text{for } r \geq 2^{1/6}\sigma_{ij} \end{aligned} \tag{6}$$

2.1.1. Minimizing the Basis Function Method's Variance from a Single Simulation. The basis function approach provides an efficient method to identify low variance pathways of potential energy functions. Any $h_i(\lambda)$ with a sufficient number of intermediates shares the same phase space as any other $h_i(\lambda)$ since $h_i(\lambda)$ is just a remapping of λ on the $[0,1]$

DOI: 10.1021/ct501047e
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

interval. Therefore, all thermodynamic information about any $h_i(\lambda)$ is contained in any one $h_i(\lambda)$. A variance-minimization technique was designed around this fact, which operates from a single simulation, entirely in postprocessing.[24] This section reviews the variance minimization method for the basis functions which will be used to find optimized alchemical switches for electrostatics. This section starts by assuming that phase space overlap is shared for a set of switches and addresses the case where the phase space is not shared afterward.

Alchemical switches can be mapped onto each other if they share the same phase space. The variance mapping equation estimates the variance along an unsampled pathway, λ_g, from data collected along a sampled pathway, λ_h. The subscripts h and g denote two particular sampling pathways and are not summation indices. An example of this approach can be seen in our previous study's Figure 8 where the data from a sampled pathway and set of switches, $h_i(\lambda)$, can predict the variance in $\partial u/\partial\lambda$ of an unsampled set of switches, $g_i(\lambda)$, with high accuracy,[24] using the same basis functions.

The variance in the mean of $\partial u/\partial\lambda$ at a specific point along λ_h is

$$\mathrm{Var}\left(\frac{\partial u}{\partial \lambda_h}\right) = \sum_{i,j} \mathrm{Cov}\left(\frac{\partial u}{\partial h_i}, \frac{\partial u}{\partial h_j}\right)\frac{\partial h_i}{\partial \lambda_h}\frac{\partial h_j}{\partial \lambda_h} \tag{7}$$

which is a double sum over all the basis functions and alchemical switch combinations, $\mathrm{Var}(\cdot)$ is the variance in the mean, and $\mathrm{Cov}(\partial u/\partial h_i, \partial u/\partial h_j)$ is the covariance between the derivative of the potential energy with respect to a given alchemical switch. Because the basis functions, $u_i(r)$, do not depend on the alchemical switches, $h_i(\lambda_h)$, the derivatives cancel out the covariance term, and eq 7 reduces to

$$\mathrm{Var}\left(\frac{\partial u}{\partial \lambda_h}\right) = \sum_{i,j} \frac{\partial h_i}{\partial \lambda_h}\frac{\partial h_j}{\partial \lambda_h}\mathrm{Cov}(u_i, u_j) \tag{8}$$

where the covariance is now only between basis functions. This can be written in condensed matrix form as

$$\mathrm{Var}\left(\frac{\partial u}{\partial \lambda_h}\right) = \mathbf{h}'(\lambda_h)\cdot\mathrm{Cov}(\mathbf{u}, \mathbf{u})(\lambda_h)\cdot\mathbf{h}'^{T}(\lambda_h) \tag{9}$$

where $\mathbf{h}'(\lambda) = [\partial h_1/\partial\lambda, \partial h_2/\partial\lambda,...]$, and $\mathrm{Cov}(\mathbf{u},\mathbf{u})$ is a covariance matrix.

Samples from one alchemical switch can be used to compute expectations and covariances that would be obtained with another alchemical switch. We choose monotonic alchemical switches on their shared $[0,1]$ domain. Monotonic switches mean each point along a sampled $h_i(\lambda_h)$ has a unique map to any unsampled switch $g_i(\lambda_g)$ and vice versa. As such, the covariances from a sampled path can be mapped to an unsampled path.[24] Mapping eq 8 to a set of unsampled alchemical switches and covariance, we found that the variance along an unsampled path λ_g is then

$$\begin{aligned}\mathrm{Var}\left(\frac{\partial u}{\partial \lambda_g}\right) &= \mathbf{g}'(\lambda_g)\cdot\mathrm{Cov}(\mathbf{u}, \mathbf{u})(\lambda_g)\cdot\mathbf{g}'^{T}(\lambda_g)\\ &= \mathbf{g}'(\lambda_g)\cdot\mathrm{Cov}(\mathbf{u}, \mathbf{u})(\mathbf{h}^{-1}(\mathbf{g}(\lambda_g)))\cdot\mathbf{g}'^{T}(\lambda_g)\end{aligned} \tag{10}$$

where the $\mathbf{g}'(\lambda_g) = [\partial g_1/\partial\lambda_g, \partial g_2/\partial\lambda_g,...]$ and $\mathbf{h}^{-1}(\lambda) = [h_1^{-1}(\lambda), h_2^{-1}(\lambda),...]$ terms are element-wise functional inversions which return points along λ_h, derived as eq 21 of the previous study.[24] Since $h_i^{-1}(g_i(\lambda_g))$ may be an unsampled point on λ_h, the potential energy at unsampled λ_h values can be calculated using the basis function. The only requirements for using eq 10 is that both pathways λ_h and λ_g use the same basis functions and share phase space overlap.

Equation 10 is only valid when phase space is shared between alchemical switches. Consider trying to map an electrostatic switch $h_E(\lambda_h)$ from a sampled λ_h where the repulsive interaction is present to a λ_g where the repulsive force is removed. There will be little to no phase space overlap between these two systems. $h_E(\lambda_h)$ will therefore not contain all of the thermodynamic information about $h_E(\lambda_g)$, making the mapping process impossible to use. This scenario is an example of two separate alchemical schedules. The choice of alchemical schedule and how to compare them is discussed in sections 2.2 and 4.2, respectively. When considering only the case of shared phase space, eq 10 can be used to find a variance minimized pathway.

Because we can often accurately estimate the variance of one path using sampling along another path, an optimized path can be found in postprocessing using a trial simulation. The total variance along any pathway is the integral of eq 8 along λ. An estimate for the minimum variance path is found by treating this integral as an objective function for a functional minimization. For this study, an 11-knot monotonic cubic spline is employed to represent the pathway where the value of $g_i(\lambda_g)$ is adjusted at fixed λ_g values to capture subtle changes in the shape. Because every alchemical switch provides information on every other alchemical switch, a simulation along a thermodynamic path with any arbitrary trial switch can be run first, and then variance of the path resulting from any proposed switch can be estimated.

Switches leading to low variance pathways can be found with this procedure. The simplest initial switch is a linear one where $h_i(\lambda) = \lambda$. In this study, we run simulations with these linear switches, then use functional minimization on the space of switches to find the switches leading to low pathway variances using eq 8 and the minimization procedure. The new switches may not lead to the absolute minimum variance path as (1) they are only minimized given the samples collected, and (2) we are approximating the functional space of the switches with a spline. The variance along the true minimum variance path is constant and may require potentials involving many-body terms.[18,19,23]

Applications, limitations, and full derivations of these equations and methods are provided in the previous study.[24]

2.2. Selecting a Statistically Efficient Alchemical Schedule. An alchemical "schedule" is defined by what order the nonbonded forces are coupled to the environment. Previous work[24] with the basis function method assumed a schedule where the Lennard-Jones repulsive and attractive interactions changed simultaneously, so $h_R(\lambda)$ varied with $h_A(\lambda)$ along the entire $0 \le \lambda \le 1$ interval in eq 2, followed by a small step where $h_C(\lambda)$ was fully coupled to complete the process. Changing interactions independent of each other causes each $h_i(\lambda)$ to change at different rates with λ, such as when decoupling electrostatics before decoupling Lennard-Jones. Examples of two different schedules are illustrated in Figure 1. The upper schedule shows attractive (h_A) and repulsive (h_R) switches being fully coupled before electrostatics (h_E), and the lower schedule shows repulsive force coupling continuously as electrostatic and attractive forces are alternately coupled. The height of each line shows the value of the respective $h_i(\lambda)$.

DOI: 10.1021/ct501047e
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

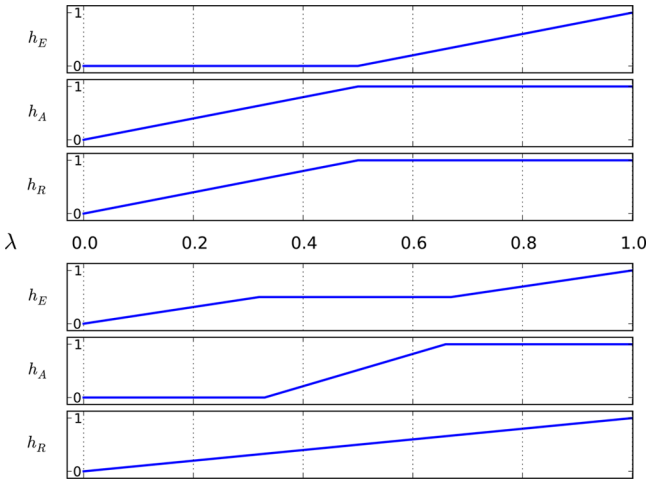


Figure 1. Two examples of alchemical switches with different schedules. Switch labels are shown to the left of each plot. Each switch is allowed to change over different ranges of the λ parameter. The height of each line is a relative value of $h_i(\lambda)$. The upper schedule shows full coupling of repulsive and attractive forces before electrostatics. The lower schedule has the repulsive potential continuously increasing but electrostatics and attractive switches alternating.

Some interactions may become fully coupled before $\lambda = 1$ as is the case for h_R in the upper schedule and h_A in both schedules.

2.2.1. Choosing Reasonable Alchemical Schedules. To simplify the analysis, we first identify schedules which provide converged numerical answers and are reasonable to implement. As a first simplification, the alchemical schedules assume that each type of interaction is fully coupled or decoupled over a single, continuous space of λ monotonically. This assumption ignores some less likely possibilities, e.g., removing half of the charge, removing the attractive term, then removing the other half of the charge; or removing the attractive term and then decreasing the charge while increasing the attractive term. Section 2.2.2 revisits these nonmonotonic pathways and argues that they cannot substantially improve the statistical error associated with each pathway.

Examining the possible orders to couple the three types of interactions (repulsive, attractive, and electrostatic), either sequentially or simultaneously, leads to 24 possible permutations of alchemical schedules. As an example of these permutations, for the upper schedule of Figure 1, the first set of switches is $\{h_R,h_A\}$ followed by $\{h_E\}$. Other alternatives include changing the switches simultaneously, $\{h_R,h_A,h_E\}$, or each switch could be changed one at a time, with six possible orderings.

This study distinguishes the separate schedules with a shorthand notation. A sequence of letters denotes the schedule along which force is coupled to the environment, and stages in the schedule are separated by a slash (/). The electrostatic, repulsive, and attractive switches are denoted by capital E, R, and A respectively. Additionally, each schedule must have a step to remove the cap on the repulsive potentials imposed by the linear basis approach, denoted by C. As the soft core methods do not have a cap stage, they will not have a C in their schedule. Although the schedules, derivations, and discussions shown here are for coupling, the process of decoupling is equivalent and would simply be the reverse ordering of the schedules. For example, the upper schedule in Figure 1 would be a soft core schedule and denoted AR/E since there is no explicit removal of the cap. All three forces in the soft core changing together would be denoted by AER, which is equivalent to RAE, EAR, and so forth.

We can quickly eliminate many of the 24 possible permutations. Basis function schedules must have a separate stage for coupling C, so AER would be invalid for the basis function approach. As an example linear basis function approach with a cap: coupling electrostatics, then attractive, then the capped repulsive, and last removing the cap would be denoted E/A/R/C. The repulsive force is always capped when it is coupled, then uncapped in the next step, i.e., C will always be in the stage after R. Such an ordering of R and C is required to have well-converged results. Adding these requirements reduce the 24 permutations to 13 unique ones. These unique permutations are summarized in Table 1 and grouped into classes discussed in the Appendix.

Table 1. Only Four Alchemical Schedules Are Reasonable to Use[a]

reasonable	attractive core	large coupling variance
AR/C/E[b]	A/E/R/C	ARE/C
R/C/AE	A/ER/C	E/AR/C
R/C/A/E	A/R/C/E	ER/C/A
R/C/E/A	AE/R/C	E/R/C/A
	E/A/R/C	

[a]Nonbonded forces of electrostatics (E), attractive (A), and repulsive (R) with a cap (C) can be coupled in 13 unique schedules, but logical arguments reduce the number of reasonable schedules to 4. The schedules can be "reasonable," "attractive core" where a negative infinite potential is generated at $r = 0$, or "large coupling variance" where the schedule is statistically inefficient due to a large capped repulsive interaction.[24] Schedules are shown by order of coupling, e.g. R/C/AE fully couples the capped repulsive force, removes the cap, then fully couples the attractive and electrostatic interactions together. [b]AR/C/E is only reasonable using the WCA decomposition.

Further analysis shows that there are only 4 of the 13 unique schedules that need to be examined in detail to identify the lowest variance pathways, which are R/C/A/E, R/C/E/A, R/C/AE, and AR/C/E. These are all paths in which the capped repulsive basis is first turned on, then the cap is removed, then the electrostatics can be turned on. The only difference is the timing of the addition of the Lennard-Jones attractive term. It could be turned on between the capping term and electrostatics term, after the electrostatics term, or during either the capped repulsive basis or the electrostatics term. The details of the logic restricting our search to these four sequences is given in the Appendix.

2.2.2. Nonmonotonic or Nonsequential Pathways Will Have Larger Variances than Monotonic and Sequential Pathways. We argue that we need only consider monotonic changes of λ. The large curvature of nonmonotonic switches increases the variance of any pathway, independent of the alchemical schedule. The variance of each schedule is directly proportional to the curvature in $\partial u/\partial\lambda$,[17,22−24] which itself is a function of $\partial^2 h_i/\partial\lambda^2$. Any force controlled by a nonmonotonic switch would naturally have more curvature and thereby pathway variance than any monotonic switches. The 13 unique schedules in Table 1 assume that each force is fully coupled to the system in one stage, that it remains fully coupled, and that the alchemical switch controlling that force is monotonic.

DOI: 10.1021/ct501047e
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

We also must consider whether it might make sense to couple a single force over multiple stages. The lower schedule in Figure 1 shows a process where the electrostatics are coupled only halfway in the first stage of the transformation, held constant in the second stage, and then coupled the remainder of the way in the last stage. This is in contrast to all of the suggested schedules where each force is coupled in a single stage. We can remove the need to test schedules where individual forces are coupled in multiple stages with two logical arguments.

(1) The repulsive force must be fully coupled in the first stage. The previous section and the Appendix argued that electrostatics should be coupled only after the repulsive force is fully coupled. The attractive force can only be coupled with the repulsive force if the potential has a WCA decomposition, but the repulsive force must still be the first stage to prevent an attractive core from forming.

(2) Identifying nonintuitive covariance terms would be the only way to lower the variance below fully coupling a force in a single stage. Since electrostatics and repulsive force must be coupled separately from each other, the only force which could be coupled in multiple stages is the attractive force. Any reduction in variance from coupling the attractive force in multiple stages would come from the covariance terms in eq 8. The attractive component can only be coupled with the repulsive force in the WCA decomposition, and we will show that it contributes a small amount to the variance in this case. As such, any $\mathrm{Cov}(u_R,u_A)$ terms will only provide small reductions in the total variance of the pathway. Since electrostatics can have both attractive and repulsive interactions, $\mathrm{Cov}(u_E,u_A)$ is system dependent, so no general trend can be established for changing $h_A(\lambda)$ in nonmonotonic ways.

Because coupling the attractive force across multiple stages will only provide marginal reduction in the variance at best and because repulsive and electrostatic forces must be coupled separately, we believe there is not enough value in designing schedules where any force is coupled over more than one stage.

2.3. Basis Functions for Electrostatics and Alternate Lennard-Jones Basis Functions. The standard form of a pairwise electrostatic interaction for point charges is the basis function studied here. Since this study is not examining schedules which require capped electrostatics, there is no need to create a cap as is done for the WCA decomposed Lennard-Jones basis functions.[24] This makes the electrostatic basis

$$u_E(r) = \frac{q_i q_j}{4\pi\epsilon_0 r} \tag{11}$$

The corresponding, minimal pathway variance alchemical switch, $h_E(\lambda)$, still needs to be designed, but this can be done with a trial switch, eq 10, and the method summarized in section 2.1.

Decomposing the electrostatics requires implementation based considerations. If the electrostatics are decomposed into short- and long-range contributions such as reaction field or particle mesh Ewald (PME),[34] then additional basis functions may be required. Specifically for PME, there is a term in the potential energy which scales as $h_E^2(\lambda)$ which cannot be neglected due to long-range interactions of each atom with its periodic copies. Because the alchemical implementation may be different based on software, a general solution cannot be provided here, but this study's specific implementation is discussed in Section S.2 in the Supporting Information.[35]

Exploring an alternate set of basis functions to the WCA decomposition of Lennard-Jones interactions can help generalize the basis function method. In the previous study and Section 2.2, a capped WCA decomposition was assumed which takes the form of eqs 2−6. An alternate set of Lennard-Jones functions can be defined which splits the repulsive and attractive interactions based on their respective exponent of r^{-n}. These 12−6 basis functions can be written as follows:

$$u_{12-6}(r) = u_{12}(r) + u_6(r) \tag{12}$$

where the subscripts 12 and 6 denote the repulsive and attractive terms, respectively. The 12−6 basis functions are simply

$$u_{12}(r) = \frac{4\epsilon_{ij}\sigma_{ij}^{12}}{r^{12}} \tag{13}$$

$$u_6(r) = \frac{-4\epsilon_{ij}\sigma_{ij}^{6}}{r^{6}} \tag{14}$$

The u_{12} basis will need to be capped in the same way as the WCA decomposed basis functions and is defined as follows:

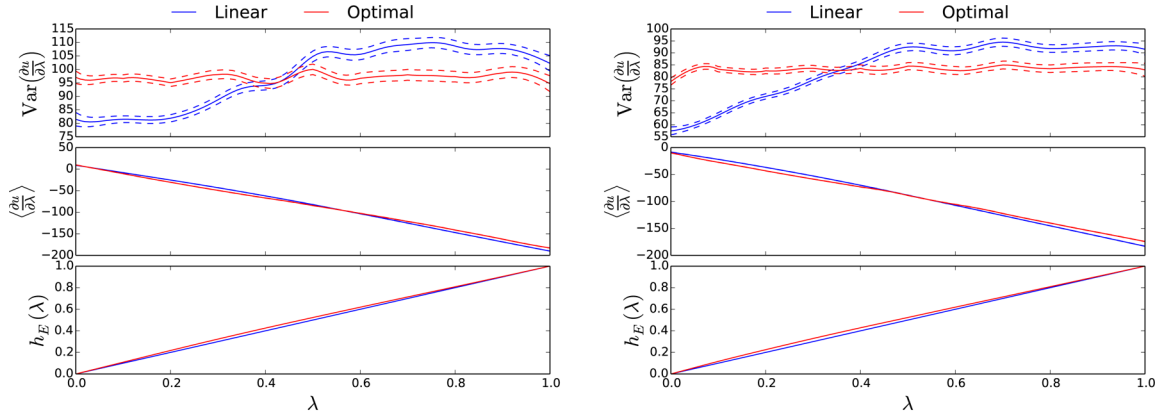
$$\begin{aligned} u_{12,\mathrm{cap}} &= u_{\mathrm{cap}} && \text{for } r < \mathrm{f}_{\mathrm{cap}}\sigma_{ij} \\ &= Ar^4 + Br^3 + Cr^2 + Dr + E && \text{for } \mathrm{f}_{\mathrm{cap}}\sigma_{ij} \le r < \mathrm{f}_{\mathrm{switch}}\sigma_{ij} \\ &= u_{12}(r) && \text{for } r \ge \mathrm{f}_{\mathrm{switch}}\sigma_{ij} \end{aligned} \tag{15}$$

where $u_{\mathrm{cap}} = u_{12}(f_{\mathrm{cap}}\sigma_{ij})$, $f_{\mathrm{cap}} = 0.82$, and $f_{\mathrm{switch}} = 0.92$ to keep the cap height between 3.5 k_BT and 8.8 k_BT and minimize the variance in the pathway from switching between an infinite potential and the capped potential.[24] The constants A−E are determined by ϵ_{ij}, σ_{ij}, f_{cap}, and f_{switch} in the same way as the constants of eq 4. This set of basis functions is numerically stable for all schedules in Table 1 except AR/C/E as denoted since the attractive interaction is not capped. Each schedule will be affixed with either "-WCA" or "-12−6" in its name to distinguish between the WCA and the 12−6 and basis functions, respectively.

3. EXPERIMENTAL DESIGN

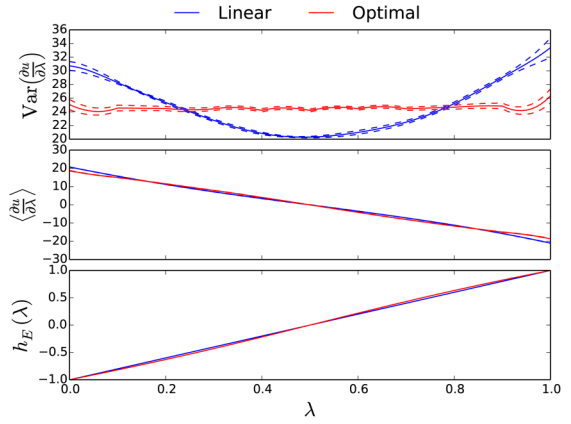
Molecular dynamics for the solvation of molecules with electrostatic interactions were carried out with YANK,[36,37] which was built on GPU accelerated OpenMM v4.1.1[38−42] in explicit TIP3P water. This includes the charging of positive and negative ions, dipole inversion, charging of all-atom (AA) ethanol, charging of AA 1,4-butanediol, and insertion of the AA side-chain analogue to tryptophan, 3-methylindole. OpenMM provides a platform to rapidly implement arbitrary basis functions and alchemical switches. However, it does not directly compute derivatives of the energy with respect to the coupling parameter, which are required to calculate variance in the mean at each point. In the case of a basis function pathway, these derivatives can easily be recomputed from the potential energy differences. This is not possible for the soft core approach, and so, soft core alchemical simulations with all forces changing at once were carried out with GROMACS v4.6.4.[43,44] The validation that the thermodynamic properties being computed are directly comparable between the two simulation packages is discussed below.

DOI: 10.1021/ct501047e
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX



(a) Charging a sodium ion, Na^+

(b) Charging a chlorine ion, Cl^-



(c) Dipole inversion

Figure 2. Linear alchemical switch is nearly optimal for coupling charged particles. The variance, $\langle \partial u/\partial \lambda \rangle$, and electrostatic alchemical switch for charging ions and inverting a dipole are shown. Simulations were run with a linear switch (blue curves) and the minimized variance curve computed in postprocessing (red curves). In all three cases, the minimal variance was nearly flat, indicating an approach to minimum variance over all possible (even nonpairwise) potentials. The optimized switches did not change significantly from a linear switch. Total variance reduction using an optimized path is less than 0.5%. This implies a linear switch is nearly sufficient for simple particles. Units are $(\mathrm{kcal/mol})^2$ for variance and kcal/mol for $\partial u/\partial \lambda$; error is shown as dashed lines around curves.

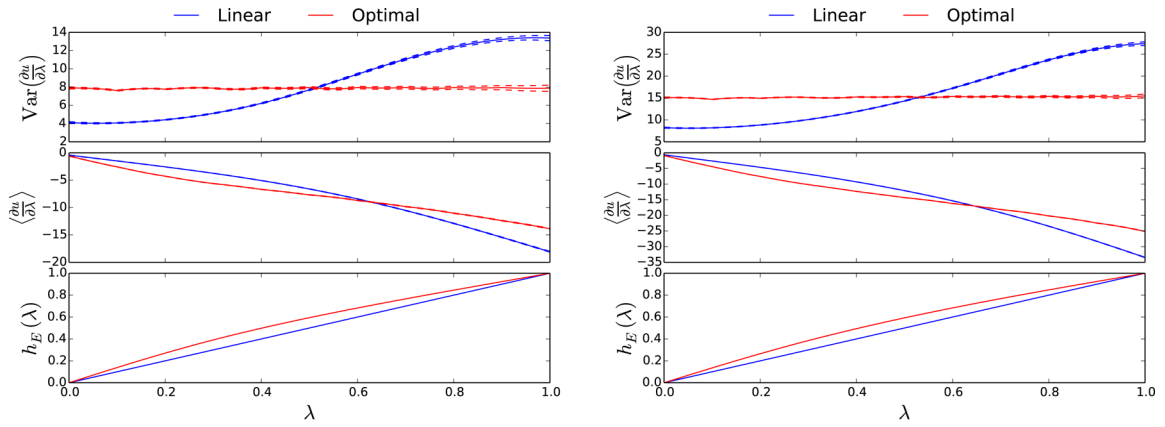
Solvation free energy simulations and charging-only transformations were tested with a variety of different molecules. Charging of a molecule was carried out on Na^+, Cl^-, a dipole made from united atom (UA) ethane with +0.5/−0.5 charge, AA ethanol, and AA 1,4-butanediol. Full insertion of a molecule with partial charges was tested with AA 3-methylindole. The charged species were chosen to test if opposite signs have different optimized paths in a polar solvent such as water. Ethanol was chosen to see if asymmetric charge density changes the optimized switch. 1,4-Butanediol was selected as a molecule with strong propensity to internally hydrogen bond in vacuum, which would be disrupted by the introduction of polar solvent. 3-Methylindole was selected as it is the small molecule analogue of the largest protein side chain.

All molecules were initially constructed using AMBERTOOLS's LEaP with OPLS-AA force field parameters for all molecules except the dipole. The dipole's starting geometries were from taken from Paliwal and Shirts.[13] 1,4-Butanediol structure was imported from PubChem (CID 8064). 3-Methylindole's starting molecular geometry was acquired from the Supporting Information from Mobley et al.[45] All molecules were inserted in a periodic cubic box of TIP3P water with boundaries 1.2 nm from the solute. The number of waters in ion systems were 633 and 707 for Na^+ and Cl^-, respectively. The UA ethane dipole had 1405 waters, the ethanol system had 769 waters, the 1,4-butanediol system had 872 waters, and the 3-methylindole system had 961 waters.

YANK simulations were carried out under isothermal−isobaric (NPT) conditions at 298 K and 1 atm. Each state was simulated for 2 ns with a 2 fs time step, samples collected every 1 ps, and Hamiltonian replica exchange[46] between all states attempted every 1 ps with Gibbs sampling to improve replica mixing efficiency.[47] 3-Methylindole's bonded hydrogens were constrained by the SHAKE algorithm,[48] and water was constrained by the SETTLE algorithm.[49] Pressure control was handled by a Monte Carlo barostat[50,51] and temperature control through Langevin dynamics. The nonbonded cutoff was 0.9 nm, and interactions outside this cutoff were handled by PME with a relative error tolerance of $5{\cdot}10^{-4}$. Although Hamiltonian replica exchange was not required for these systems, there was negligible computational effort to running with it in YANK.

GROMACS simulations were run with binaries compiled in double precision. NVT equilibration was carried out for 100 ps

DOI: 10.1021/ct501047e
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX



(a) Charging ethanol (b) Charging 1,4-butanediol

Figure 3. Linear switch is near minimal variance, even for a molecule which strongly interacts with the solute or itself. The variance in $\partial u/\partial\lambda$, $\langle \partial u/\partial\lambda \rangle$, and electrostatic alchemical switch for charging ethanol and 1,4-butanediol are shown. The linear switch (blue curves) is not as close to the minimal variance switches (red curves) as is the case in the charging ions and dipole inversion. The theoretical minimal variance is approached as seen from the flat variance; however, the improvement in variance is less than 5% for the ethanol and less than 4.4% for the 1,4-butanediol. Since this is not a significant reduction in variance, the best general purpose alchemical switch is a linear one when considering that linear switches are also nearly optimized for particles carrying full charge. Units are $(\mathrm{kcal/mol})^2$ for variance of $\partial u/\partial\lambda$ and kcal/mol for $\langle \partial u/\partial\lambda \rangle$; error is shown as dashed lines around curves.

followed by NPT equilibration for 500 ps before NPT production simulations of 6 ns per simulated state. Temperature was held at 290 K and maintained through Langevin dynamics. Pressure (for NPT) was maintained at 1 atm with a Parrinello-Rahman barostat,[52,53] a time constant of 5 ps, and a compressibility of 4.5×10^{-5} bar^{-1}. Alchemical sampling for the soft core path was done at 21 evenly spaced λ values from $\lambda = 0$ to $\lambda = 1$. Replica exchange was not done with GROMACS; however, only decorrelated samples were examined in both simulations which should provide the same thermodynamic results.

The two simulation packages were validated to ensure that the thermodynamic properties being computed are directly comparable. The potential energy from an NVE trajectory in OpenMM was computed by both GROMACS and OpenMM and was found to be identical to machine precision with the simulation settings used in this study. Although thermostats and barostat were different, only algorithms preserving the correct energy distributions were used[54] so that ensembles were directly comparable. Different thermostats do affect the correlation time of samples, so all trajectories were decorrelated before comparison.[55]

The analysis code and our implementation of the linear basis function method for OpenMM and YANK can be found on GitHub at https://github.com/shirtsgroup/basisanalyze.[56]

4. RESULTS AND DISCUSSION

4.1. Coupling Electrostatics. Linear coupling of electrostatics is almost exactly as statistically efficient as the minimum variance optimal path. Figure 2a and b shows the sample variance of $\partial u/\partial\lambda$, $\langle \partial u/\partial\lambda \rangle$, and optimized alchemical switches for coupling the electrostatics of the Na^+ and Cl^- ions. The optimized switch is recalculated in postprocessing with the technique outlined in section 2.1 and detailed in the original basis function study.[24] There is no significant difference in the shape between the optimized switches of the two ions and virtually no difference between either optimized alchemical switch and the simplest linear switch $h(\lambda) = \lambda$. The variances of the pathway using these optimized switches are nearly flat and approach the theoretical minimum of a perfectly flat variance of the path over the transformation.[18,19] Total variance for the optimized path is less than 0.5% lower than linearly coupling the electrostatics. Thus, there is no reason to use anything other than linear coupling of electrostatics for ions given how little reduction there is in the variance of the pathway, especially given the fact that the optimized switch would require implementing some more complicated functional form.

The dipole inversion behaves similarly to the ion solvation in that the linear coupling path has nearly the same pathway variance in the calculation as the minimal variance path. The sample variance in $\partial u/\partial\lambda$, $\langle \partial u/\partial\lambda \rangle$, and the electrostatic switch for the dipole inversion are shown in Figure 2c. This transformation is symmetric around $\lambda = 0.5$ as required by the symmetry of the system. The total variance along the optimized path is only marginally better than the linear transformation, reducing the total variance by only 0.5% with respect to the linear path. We again conclude that alchemically changing the charge with a linear switch will be suitable for most applications involving changes in charge of simple particles.

The optimized charging path for molecules with a more complicated charge distribution again results in only a small reduction in variance of the pathway compared to the linear path. Unlike the ions or the dipole, the optimized path for charging ethanol and 1,4-butanediol is not as close to linear as shown in Figure 3a and b. However, the minimum variance switch still reduces the variance less than 5% from the linear switch in both cases. Implementing this optimized switch could probably be done with some sort of low dimensional polynomial, but it is also not clear how this polynomial might change on a system-by-system basis. We note that in the cases of ethanol and 1,4-butanediol, the total variance in $\partial u/\partial\lambda$ associated with charging calculation is rather small compared to the variance of charging a large dipole, and even smaller than that with respect to the removal of a repulsive site.[24] This is likely because any nonlinear terms relate to the higher moments of the charge distribution, which will contribute less energy than the monopole and dipole components. As such, the

DOI: 10.1021/ct501047e
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

rest of the analysis will be carried out assuming a linear switch for electrostatic coupling, as the additional marginal efficiency gained will never be worth the complication.

4.2. Identifying the Optimal Alchemical Schedule. The variance of the pathway to calculate the free energy difference is not simply the sum of the sample variances in $\partial u/\partial\lambda$ at each point. Because the sample variance in $\partial u/\partial\lambda$ and free energy are estimated along a TI path, the variance of the calculation of free energy is computed by adding variances in the mean of $\partial u/\partial\lambda$, weighted by the scalar factors used in the numerical integration. Up to this point, the variances of $\partial u/\partial\lambda$ discussed have simply been the sample variances. For example, the sample variances in $\partial u/\partial\lambda$ computed for the AR/C/E-WCA path are shown in Figure 4.

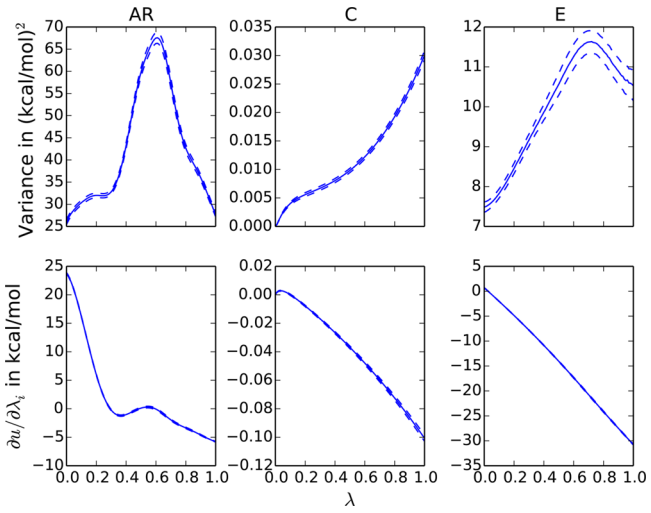


Figure 4. Sample variance of $\partial u/\partial\lambda$ in each stage of the transformation is not additive on its own. The variance in $\partial u/\partial\lambda$ (upper) and $\langle\partial u/\partial\lambda\rangle$ (lower) for the AR/C/E-WCA basis is shown. Only by dividing by the number of samples drawn from each state can the variances shown be added. Each stage in the transformation is given by its shorthand letter above the column described in the text. The capping stage (C) contributes very little to the total variance, even without dividing by the number of samples drawn. Error is shown as dashed lines around the figure and is often smaller than the thickness of the line.

The integral of such a variance curve is related to the thermodynamic length, or Riemannian length ($\mathcal{L}$),[17−19,22−24] between the states. The thermodynamic length can be computed by

$$\mathcal{L} = \int_0^1 \sqrt{\mathrm{Var}\left(\frac{\mathrm{d}u}{\mathrm{d}\lambda}\right)}\,\mathrm{d}\lambda \tag{16}$$

The variance minimization efforts up to this point have focused on minimizing this length, which happens when the sample variance in $\partial u/\partial\lambda$ is flat. However, as the number of decorrelated samples from each schedule is different, and the distribution of states is different between the basis functions and the soft core path, we cannot compare the sample variances in $\partial u/\partial\lambda$ alone to determine which path is the most statistically efficient.

A fair comparison of thermodynamic pathways requires distributing discrete samples in a statistically optimal manner for that path. If sampling is done at discrete states, the variance of the total calculation of the free energy, $\mathrm{Var}(\Delta F)$ is related to sample variance in $\partial u/\partial\lambda$, s^2, as

$$\mathrm{Var}(\Delta F) = \sum_{i=1}^{N_{\mathrm{stage}}} \sum_{j=1}^{K_i} \frac{w_{i,j}^2 s_{i,j}^2}{n_{s,i}} \tag{17}$$

where i loops over each of the N_{stage} stages in the transformation, j loops over each of the discrete K_i states in stage i, and $s_{i,j}^2$ is the sample variance from state j in stage i. $n_{i,j}$ is the number of samples drawn from state j and stage i. $w_{i,j}$ are the weights of the quadrature method used in the TI integration to calculate the free energy.[57] By choosing a consistent distribution of states and number of samples for each state, a fair comparison between each schedule can be made by comparing the variance of the overall free energy calculation from eq 17.

There are two limiting cases to consider when distributing samples between states for a fair comparison of the alchemical methods. In the maximally efficient case, sampling is done proportional to the anticipated sample standard deviation of $\partial u/\partial\lambda$.[17] This choice will maximize statistical efficiency of the entire calculation. However, this requires a priori knowledge of the sample variance, which is generally only partially known by analogy to other similar systems before running the simulation. This method is nearly mathematically equivalent to adjusting the spacing between states along the path. As $\Delta\lambda$ between states is reduced, additional samples are placed in nearby state space, ideally in the regions of high variance. The switch optimization procedure adjusting $h_i(\lambda)$ is also an equivalent process to adjusting the spacing between states since flattening $h_i(\lambda)$ results in additional sampling at the same intermediate states, also effectively increasing sampling at states with high sample variance. Given that the switch optimization procedure effectively adjusts the spacing between states, we will only look at uniformly spaced states along λ and instead distribute discrete samples to each state.

The other extreme case is simply performing uniform sampling at all states, which is a fairly common approach as it requires no additional knowledge beforehand. The realistic case is between these two situations, as some sense of the λ with large sample variance is often qualitatively known.

Finally, a fair comparison of statistical efficiency requires methods eliminating bias from free energy differences and analyzing only uncorrelated samples. Since each alchemical method and schedule are estimating the same transformation, the free energy differences should be within error of each other. If the free energy differences are not within error of each other, it indicates insufficient sampling and poor estimates at best, or fundamental implementation issues at worst. The value of the free energy differences in this study are correctly within statistical error of each other and tabulated in Table S.1 of the Supporting Information.[35] This statistical efficiency comparison relies only on uncorrelated samples, so each trajectory is subsampled to extract only uncorrelated samples using the time series functions of the pymbar Python module.[12] The computational cost to draw uncorrelated samples from the different states will be addressed later in this study.

4.2.1. Comparison of Statistical Efficiency from Different Sampling Schemes. The maximum number of samples to distribute among alchemical states will be fixed for each schedule and alchemical method. The different sampling schemes will assume 21000 uncorrelated samples are distributed to each schedule for both limiting cases. This number was chosen so that 1000 samples can be uniformly distributed to the 21 sampled states in the soft core

DOI: 10.1021/ct501047e
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

electrostatics case. Samples are distributed to a fixed number of states in each schedule. The basis function path has 101 evenly spaced states at each stage since we can estimate the variance at arbitrary λ values.[24] There is an overlapping state between each stage of a schedule, so care must be taken not to double count these states, e.g., in the AR/C/E schedule, the thermodynamic state defined by $h_R(\lambda) = h_A(\lambda) = 1$ is identical to the state $h_C(\lambda) = 0$. The soft core electrostatics pathway will only be estimated at the 21 sampled states as sample variance data are only available through direct simulation at each state. Equation 17 is valid even if the spacing is not uniform, though it will be held uniform in this study.

Both the AR/C/E-WCA and the R/C/AE-WCA paths have nearly the same statistical efficiency as the soft core electrostatics path within error when sampled proportionally with the sample standard deviation. Figure 5 shows the variance

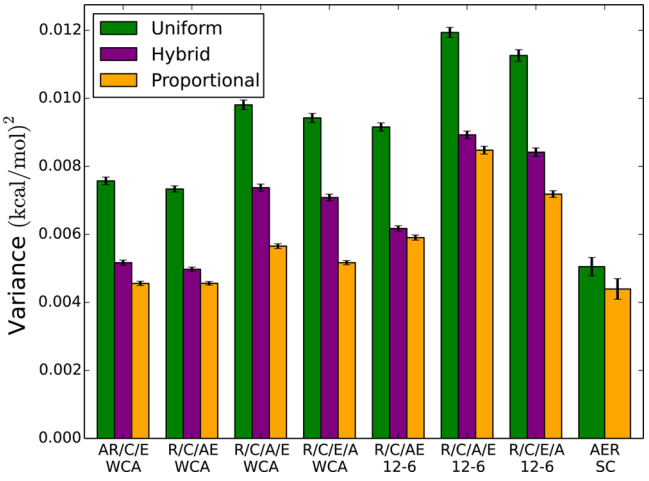


Figure 5. Hybrid sampling scheme with 3-step, WCA basis functions is as statistically efficient as soft core methods. The variance of the calculation for free energy is shown for three sampling schemes applied to all tested basis function schedules and the soft core path. Samples were distributed either uniformly, proportional to the sample standard deviation, or in a hybrid approach which is uniform except for the capping stage, where only end points are sampled. The hybrid scheme provides comparable variance of the calculation to uniformly sampled soft core methods. "WCA" and "12−6" labels distinguish the Lennard-Jones basis functions and "SC" soft core method applied to all forces with 1−1−6 parametrization.

of the calculation of free energy for each pathway for both uniform sampling (green, left most bars) and proportional sampling with the sample standard deviation (orange, right most bars) along with total error in the estimate of the variance of the free energy shown with the error bar at the top. The middle bars (purple) are a hybrid sampling scheme and will be discussed below. As expected, proportional sampling is more statistically efficient than uniform sampling, seeming to imply the sample variance must be known ahead of time for the basis function approach to be as efficient as the soft core.

However, most of this reduction in uncertainty can be gained by simply moving samples out of the capping stage to other stages. Because sampling at intermediates states of the cap contributes near negligible uncertainty as seen in Figure 4, sampling only the end points allows distributing more samples into the repulsive and electrostatics stages. These end points also overlap with the AR, R, E, and/or AE stages and must be sampled anyway. In this "hybrid" sampling, the C step is only sampled at the end points, but other stages are uniformly sampled. This hybrid pathway lowers the variance of the calculation in free energy for the AR/C/E-WCA and R/C/AE-WCA pathways to nearly the same statistical value as that of the soft core electrostatics (shown in Figure 5) but without the need for a priori knowledge of all the sample variance curves. Sampling proportional to the standard deviation still provides lower variance of the calculation, but at most by 13% for the AR/C/E-WCA and R/C/AE-WCA pathways. Neglecting sampling, the C stage does introduce a small amount of numerical integration error. However, the numerical bias introduced is less than 3% of the total variance of the free energy calculation, so makes no change statistically in the results.

The 12−6 and the four-step basis function schedules have significantly larger variances in the free energy than the three-step WCA basis function schedules. Because of this, there appears to be little benefit to splitting up the Lennard-Jones forces in the 12−6 basis unless the contribution from the individual dispersion and repulsion components is of interest to the study.

The larger variance of free energy calculation on the 12−6 schedules is due primarily to the uncapped attractive component, which contributes significantly more to the variance than the capped WCA attractive component. The attractive component in the 12−6 pathways will contribute more to the numerator in eq 17 for the uniform case, and reduce the number of samples available to the repulsive step, decreasing the denominator of eq 17 in the proportional case. Any of the four-step processes also suffer from having fewer samples available for the denominator in the repulsive step, which causes an increase in variance of the calculation as well.

The contribution of each stage to the variance of the calculation in Figure 5 can be visualized in the Supporting Information[35] in Figure S.8. The tabulated variance at each stage, free energy, sampled uncertainty, and numeric values for variance of the calculation from each sampling scheme are also included in the Supporting Information[35] in Tables S.1 and S.2.

In summary, there are four heuristics we can make about sampling with the basis function approach. From these heuristics, a schedule and sampling scheme can be chosen without explicit knowledge of the basis function sample variances. (1) Couple attractive Lennard-Jones forces simultaneously with either the repulsive or electrostatic term. (2) A WCA decomposition of Lennard-Jones terms is somewhat more efficient than a 12−6 decomposition. (3) Only the end states of the capping stage need to be sampled. (4) Uniform distribution of states to every stage and uniform sampling of each state for all stages except for the capping stage will provide a low statistical error, approximately as low as a soft core pathway where all forces (including electrostatics) are coupled simultaneously.

4.3. Decorrelation Times along the Soft Core Pathway Are Somewhat Longer than the Basis Function Pathway. The cost of generating uncorrelated samples is an additional constraint that we must also consider when looking at the efficiency of a schedule. Some states might have intrinsically slower configurational sampling. An extreme case was noted by Pham et al.,[18] where the provably lowest variance path actually resulted in extremely slow kinetics for transitioning between two dominant sets of allowed configurations, requiring tens or hundreds of nanoseconds for transitions. This resulted in a recommendation to sample a

DOI: 10.1021/ct501047e
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

somewhat higher variance path with much faster conformational sampling between the resulting configurations, as the computational efficiency of this approach outweighed the statistical inefficiency.

The two limiting cases in the comparison of schedules in the previous section assumed that a fixed number of uncorrelated samples were available to distribute across states but that the simulation time required to generate those uncorrelated samples will be different depending on the conformational dynamics created by the potential function. To explore how the correlation time of the sampling affects the efficiency of these schedules, we will look at the recommended AR/C/E-WCA and the soft core electrostatics paths, using ideas from the difficulty index method introduced by Schultz and Kofke for comparing methods accounting for statistical uncertainty and computational time.[58]

The difficulty index method proposes a difficulty, D, to compute a given property as

$$D = t^{1/2}\sigma \tag{18}$$

where t is a specified required time to compute a property, and σ is the uncertainty in that property, after time t. The difficulty index is the logarithm of the difficulty, scaled by a normalizing factor to compare the net computational effort required to compute different properties, accounting for both statistical efficiency and time to calculate of the property. As the same property is being compared here, only the difficulty, not the difficulty index, will be examined. σ in this case will be the standard deviation of the free energy, computed as the square root of the values in Figure 5, tabulated in the Supporting Information[35] as Table S.2. Assuming that identical hardware and software are used to generate these samples, the only difference between time to collect an uncorrelated sample is the correlation time. This removes any differences in CPU-clock time from different software, enhanced sampling method such as replica exchange, and hardware such as different processors or GPU vs CPU. For this study, the units of difficulty will be $ps^{1/2}$ kcal/mol.

Correlation times were calculated by running a NVE simulation of both the AR/C/E-WCA basis function pathway, and the soft core electrostatics pathway. Using the NVE simulation removes any effect temperature or pressure coupling has on the correlation time,[59] which is particularly important since GROMACS and OpenMM are being compared and have somewhat different implementations of the thermostats used in this study. Since energies computed by both simulation packages are identical to machine precision in NVE, the differences between correlation times will only be due to the different pathways in an NVE simulation.

The AR/C/E-WCA simulation, selected as one of the two best basis function-based paths and simulated in OpenMM, was run without replica exchange to match the soft core electrostatic simulations run in GROMACS. The autocorrelation time[55] of the $\partial u/\partial\lambda$ time series was computed for each method. The autocorrelation time of the $u(r,\lambda)$ time series in the fully coupled and fully decoupled states should be the same between both OpenMM and GROMACS simulations and is path independent, so it was used as a validation check. Because sampling was only done at a limited number of states, the correlation for the basis function time series is shown as a total progress of the entire coupling process and not separated by individual stages. Because there will be two values of $\partial u/\partial\lambda$ at a state where basis function stages overlap, the maximum autocorrelation time of the two was chosen to be conservative.

More sampling is required near the decoupled state of the soft core path relative to the basis function path for a practical implementation as shown by Figure 6. The correlation times at

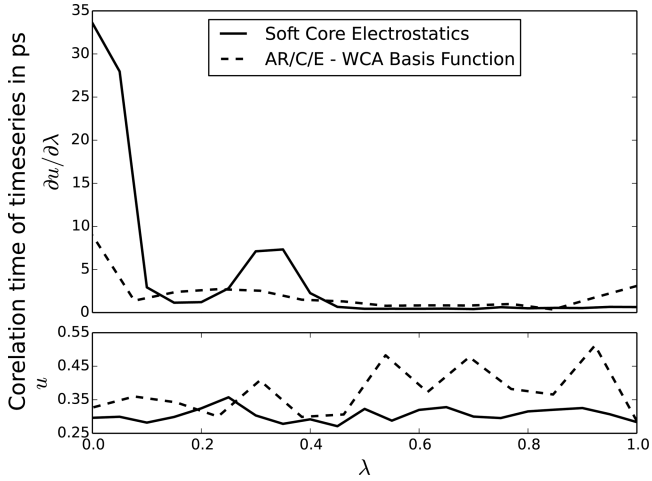


Figure 6. Basis function method has lower correlation times than soft core electrostatics. Correlation times of soft core electrostatics (solid lines) and a AR/C/E-WCA schedule (dashed lines) sampled from NVE simulations are shown. The basis function curve shows all three stages along a single λ parameter. The top image shows the correlation time for $\partial u/\partial\lambda$, and the bottom one shows the correlation time for $u(r,\lambda)$, which allow us to validate the simulation packages are generating the same dynamics at the coupled and decoupled states. The decoupled state for the soft core electrostatics would require nearly four times as many samples relative to the basis functions to generate the same number of uncorrelated samples. Correlation times for $u(r,\lambda)$ and $\partial u/\partial\lambda$ are roughly equal for $\lambda > 0.5$ and are shown in as separate plots for clarity. Lines are drawn between discretely sampled states and serve only to guide the eye.

the end states for the $u(r,\lambda)$ (bottom) time series agree within 10% showing the two simulation packages are generating comparable dynamics, as they should. At most intermediates, the correlation times in the two pathways are similar. However, the correlation time of $\partial u/\partial\lambda$ (top) for the soft core electrostatic path is almost four times that of the basis function path at the decoupled state. Since the remainder of the correlation times are comparable, this state at $\lambda = 0$ will be used to compute difficulty. From eq 18, the difficulty of the soft core path near the decoupled state is 0.38, whereas the basis function difficulty is 0.12. Although the decoupled state may take 3.5 times additional CPU effort for the soft core method, the remainder of the sampled states will take roughly equal CPU effort between the soft core and linear basis function approaches. This increased computational effort at the end state will result in a moderate overall increase in simulation time for simulating with the soft core approach instead of the basis function approach.

Finally, and most importantly, we note that the correlation time for calculations of complex systems, such as protein−ligand binding, is dominated by the large scale motions of the system, which are independent of the much shorter time scale of these alchemical changes. Reducing the variance of the path reduces the number of uncorrelated data points that must be collected for a given accuracy of calculation, where the correlation times are those of the longest time scale motions

DOI: 10.1021/ct501047e
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

of interest in the system, which are usually not the ones involving the alchemical transformation.

5. CONCLUSIONS

We have incorporated and tested electrostatics into our linear basis function formalism. Alchemical electrostatic methods in their current form are easily included in the basis function formalism when the charge is coupled after Lennard-Jones type interactions. We have shown in this study that the most statistically efficient switch for charging is nearly linear and provides too marginal (5% maximum) a reduction in variance for the pathway (and therefore statistical uncertainty) that is too small to justify the complexity of implementing improved switches. Simply turning on charges linearly appears to be the most effective option in essentially all cases.

We recommend either the AR/C/E-WCA or R/C/AE-WCA basis functions with linear charging be considered for future implementations of alchemical transformations. We have carefully examined multiple basis function coupling methods and compared them to soft core electrostatics in terms of statistical and computational efficiency. Several heuristics that do not rely on explicit knowledge of the variance were provided which should help those wishing to implement the basis function method. The overall statistical efficiency for the basis function with the heuristics is comparable to traditional soft core methods; however, the basis functions are slightly more computationally efficient when the time required to decorrelate sampling is included. The basis functions have the added benefit of being able to remove the alchemical forces from the inner force loop of the simulation and quickly compute energies at unsampled states, as well as avoiding code to explicitly compute the analytical derivative with respect to the coupling parameter.

APPENDIX: NARROWING THE CHOICES FOR LOW VARIANCE SCHEDULES

A.1. Attractive Force Must Be Coupled after the Repulsive Force to Prevent an Attractive Core

We note that the core of a particle must have a repulsive interaction to ensure that an infinite attractive core does not form, which would cause integration instabilities.[3,16,18−21,32] This applies not only at $r = 0$ but also for any off-atom sites (like lone pairs or the oxygen charge in 4 point water models) as they are protected by the repulsive cores of nearby atoms. Any schedule which has the attractive force coupled before the repulsive force need not be considered further, eliminating schedules such as AE/R/C or A/R/C/E. However, schedules where repulsive interactions are coupled simultaneously with attractive interactions (i.e., E/RA/C or RA/C/E) are retained since they can be decomposed by a Weeks−Chandler−Andersen (WCA)[33] decomposition, which eliminates the infinite attraction of the r^{-6} term. Schedules which have attractive cores are categorized in Table 1 as "attractive core" schedules. Removing these reduces the total unique schedules to eight.

A.2. Repulsive Core Should Be Fully Coupled before the Electrostatics

The basis functions and schedules must ensure that the magnitude of an attractive electrostatic potential never exceeds the magnitude of the Lennard-Jones repulsion at $r = 0$ to avoid an attractive core causing simulation instability.[32] There are two main ways to ensure this: cap the electrostatic interaction at finite values, similar to eq 4, or only couple electrostatics after the repulsive interactions have been uncapped. Although capped electrostatics can provide numerically stable simulations,[60−62] they do not provide low variance coupling pathways. The total capped potential must be between 3.5 k_BT and 8.8 k_BT at $r = 0$, in order to provide a low variance pathway.[24] As both attractive and repulsive electrostatic interactions are possible, we must design basis functions to handle both, while also keeping the total capped potential inside the target range.

In order to design numerically stable electrostatic basis functions, we must first examine how electrostatics are handled in the soft core, where repulsive and attractive interactions are treated identically. Soft core potentials require constrained constants to maintain numerical stability. Soft core methods must be designed to prevent a double minimum from forming in the potential energy separated by large energy barriers. The primary cause of this double minimum is the electrostatic cap dominating the repulsive cap near $r = 0$ for cases of a soft core alchemical atom interacting with an atom holding charge, but no Lennard-Jones core, e.g., the hydrogen in TIP3P water.[32] The soft core Lennard-Jones potential takes the general form

$$u_{\mathrm{SC,LJ}}(r, \lambda) = 4\epsilon_{ij}\lambda^{a}\left[\left(\frac{1}{\alpha(1-\lambda)^{b} + (r/\sigma_{ij})^{c}}\right)^{12/c} - \left(\frac{1}{\alpha(1-\lambda)^{b} + (r/\sigma_{ij})^{c}}\right)^{6/c}\right] \tag{A.1}$$

The electrostatic potential has a soft core form of

$$u_{\mathrm{SC,E}}(r, \lambda) = \lambda\frac{q_i q_j}{4\pi\epsilon_0[(1-\lambda)\beta + r^{m}]^{1/m}} \tag{A.2}$$

In eqs A.1 and A.2, a, b, c, and m are positive, usually integer, constants, and α and β are positive free parameters. Choosing $a = b = 1$ is a statistically efficient choice,[18] leaving the other constants free for variance minimization and numerical stability analysis. In order to prevent a double minimum from forming using both soft core equations together, the following constraints must be true:[32]

$$c = m \tag{A.3}$$

$$\alpha^{1/c}\sigma_{ij} \leq \beta^{1/m} \tag{A.4}$$

The basis functions must also prevent double minimum potentials but have arbitrary functional form. Contrary to the soft core potentials, the basis functions do not require any specific functional form. However, a single set of basis functions is better than multiple, system dependent sets, for simplicity in both use and implementation. Since a capped electrostatic potential is required to use with the capped repulsive basis function, the functional form for the electrostatics should be similar to eq 4 and is controlled by the same f_{cap} and f_{switch}, removing the option to define different parameters for different systems. This constraint could be removed, but then separate potentials would need to be defined for attractive and repulsive electrostatic interactions as will be shown.

To determine the necessary parameter values, and therefore the magnitude of the cap, we examine the limiting case of Li^+ and Fl^- ions interacting. These ions have small σ_{ij} terms but also carry full ±1 charges. The small σ_{ij} will allow the particles

DOI: 10.1021/ct501047e
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

to approach closer and therefore feel a stronger electrostatic force than typical solute−solvent systems with larger atom cores and partial charges. The required cap height in this extreme case will be more than sufficient for almost all other physically relevant systems.

The parameter f_{cap} must be smaller with an electrostatic cap than with a repulsive cap alone for the transformation to have low statistical error. The OPLS-AA parameters for Li^+ are σ_{ii} = 2.13 Å, ϵ_{ii} = 0.0183 kcal/mol, and for Fl^- are σ_{ii} = 2.73 Å and ϵ_{ii} = 0.720 kcal/mol.[63] Choosing the basis functions for attractive and repulsive Lennard-Jones interactions such as eq 3 and eq 4 respectively, the potential of the cap for either the repulsive Lennard-Jones or electrostatic potential is

$$u_{i,\mathrm{cap}} = u_i(f_{\mathrm{cap}}\sigma_{ij}) \tag{A.5}$$

where i is either R or E, u_i is the uncapped basis function for the respective force. The standard point charge interaction formula is chosen for the electrostatic basis such that

$$u_E(r) = \frac{q_i q_j}{4\pi\epsilon_0 r} \tag{A.6}$$

With these choices, a cap at $f_{cap} \approx 0.6$ is required to preserve a repulsive core at the capped values. This value ensures that the capped repulsive Lennard-Jones interaction dominates the capped attractive electrostatic potential and keeps the total potential in the target range at $r = 0$.

Repulsive interactions should be fully coupled with their caps removed before electrostatic interactions are coupled into the system. Although the total cap for an attractive electrostatic interaction is within our low variance target range, a repulsive electrostatic interaction is not. The Li^+/Li^+ interaction has a total cap of 397 k_BT at $f_{cap} \approx 0.6$, which can significantly increase the variance of the free energy calculation along such a path.[24] To correct this issue, a separate f_{cap} and f_{switch} would be required to bring the total cap back into the range of 3.5 k_BT and 8.8 k_BT. As an arbitrary system could have both repulsive and attractive electrostatic interactions, the basis functions would need to be designed and implemented with rules checking each interaction and using the appropriate parameters. This is a less desirable solution since it adds complexity to the implementation and usability of the basis function method. It is much simpler if the electrostatics are coupled after the repulsive interactions.

■ ASSOCIATED CONTENT

Supporting Information

Variance and $\partial u/\partial\lambda$ for the R/C/AE-WCA, R/C/AE-12-6, R/C/A/E-WCA, R/C/A/E-12-6, R/C/E/A-WCA, and R/C/E/A-12-6 basis function schedules; variance and $\partial u/\partial\lambda$ for the AER-SC (soft core) alchemical method; decomposition of each stage's contribution to the variance of the calculation of free energy for all sampling schemes; sample variance and free energy of solvation for 3-methylindole between seven different basis function pathways and the soft core electrostatics path; variance in the calculation for free energy for coupling 3-methylindole between seven different basis function pathways and the soft core electrostatics path; and implementing long-range electrostatic interactions with PME. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/ct501047e.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: michael.shirts@virginia.edu.

Funding

We acknowledge the support of NSF grant CHE-1152786.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank David Mobley, Lingle Wang, and Mark Abraham for helpful discussions and feedback.

■ REFERENCES

(1) Shirts, M. R.; Mobley, D. L. Methods in Molecular Biology. In *Biomolecular Simulations*. Monticelli, L., Salonen, E., Eds.; Humana Press: Totowa, NJ, 2013; Vol. *924*, pp 271−311.
(2) Kollman, P. *Chem. Rev.* **1993**, *93*, 2395−2417.
(3) Deng, Y.; Roux, B. J. *Phys. Chem. B* **2004**, *17*, 16567−16576.
(4) Shirts, M. R.; Mobley, D. L.; Chodera, J. D. *Annu. Rep. Comput. Chem.* **2007**, *3*, 41−59.
(5) Chodera, J. D.; Mobley, D. L.; Shirts, M. R.; Dixon, R. W.; Branson, K.; Pande, V. S. *Curr. Opin. Struct. Biol.* **2011**, *21*, 150−60.
(6) Shirts, M. Methods in Molecular Biology. In *Computational Drug Discovery and Design*; Baron, R., Ed.; Springer: New York, 2012; Vol. *819*, pp 425−467.
(7) Baron, R., Ed. *Computational Drug Discovery and Design*; Humana Press: New York, 2012; p 628.
(8) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420.
(9) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 1011−1021.
(10) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. *J. Comput. Chem.* **1995**, *16*, 1339−1350.
(11) Bennett, C. H. *J. Comput. Phys.* **1976**, *268*, 245−268.
(12) Shirts, M. R.; Chodera, J. D. *J. Chem. Phys.* **2008**, *129*, 124105.
(13) Paliwal, H.; Shirts, M. R. *J. Chem. Theory Comput.* **2011**, 4115−4134.
(14) de Ruiter, A.; Oostenbrink, C. *J. Chem. Theory Comput.* **2012**, *8*, 3686−3695.
(15) Blondel, A. *J. Comput. Chem.* **2004**, *25*, 985−93.
(16) Steinbrecher, T.; Mobley, D. L.; Case, D. A. *J. Chem. Phys.* **2007**, *127*, 214108.
(17) Shenfeld, D.; Xu, H.; Eastwood, M.; Dror, R.; Shaw, D. *Phys. Rev. E* **2009**, *80*, 1−4.
(18) Pham, T. T.; Shirts, M. R. *J. Chem. Phys.* **2011**, *135*, 034114.
(19) Pham, T. T.; Shirts, M. R. *J. Chem. Phys.* **2012**, *136*, 124120.
(20) Beutler, T. C.; Mark, A. E.; van Schaik, R. C. *Chem. Phys. Lett.* **1994**, *222*, 529−539.
(21) Zacharias, M.; Straatsma, T. P.; McCammon, J. A. *J. Chem. Phys.* **1994**, *100*, 9025.
(22) Crooks, G. E. *Phys. Rev. Lett.* **2007**, *99*, 10−13.
(23) Gelman, A.; Meng, X.-L. *Stat. Sci.* **1998**, *13*, 163−185.
(24) Naden, L. N.; Pham, T. T.; Shirts, M. R. *J. Chem. Theory Comput.* **2014**, *10*, 1128−1149.
(25) Shirts, M. R.; Pande, V. S. *J. Chem. Phys.* **2005**, *122*, 134508.
(26) Hritz, J.; Oostenbrink, C. *J. Chem. Phys.* **2008**, *128*, 144121.
(27) Riniker, S.; Christ, C. D.; Hansen, H. S.; Hünenberger, P. H.; Oostenbrink, C.; Steiner, D.; van Gunsteren, W. F. *J. Phys. Chem. B* **2011**, *115*, 13570−7.
(28) Buelens, F. P.; Grubmüller, H. *J. Comput. Chem.* **2011**, 25−33.
(29) Mobley, D. L.; Dumont, E.; Chodera, J. D.; Dill, K. A. *J. Phys. Chem. B* **2007**, *111*, 2242−54.
(30) Liu, H.; Mark, A.; van Gunsteren, W. F. *J. Phys. Chem.* **1996**, *3654*, 9485−9494.
(31) Oostenbrink, C. *J. Comput. Chem.* **2009**, *30*, 212−221.
(32) Steinbrecher, T.; Joung, I.; Case, D. A. *J. Comput. Chem.* **2011**, *32*, 3253−63.

DOI: 10.1021/ct501047e
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

(33) Weeks, J. D.; Chandler, D.; Andersen, H. C. *J. Chem. Phys.* **1971**, *54*, 5237.
(34) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577.
(35) For explicit numbers, please see the tabulated data in the Supporting Information.
(36) Wang, K.; Chodera, J.; Yang, Y.; Shirts, M. *J. Comput.-Aided Mol. Des.* **2013**, *27*, 989−1007.
(37) YANK. http://simtk.org/home/yank/ (accessed Jul 8, 2012), available through SimTK.
(38) Friedrichs, M.; Eastman, P.; Vaidyanathan, V.; Houston, M.; LeGrand, L.; Beberg, A.; Ensign, D.; Bruns, C.; Pande, V. S. *J. Comput. Chem.* **2009**, *30*, 864−872.
(39) Eastman, P.; Pande, V. S. *J. Comput. Chem.* **2009**, *31*, 1268−1272.
(40) Eastman, P.; Pande, V. *Comput. Sci. Eng.* **2010**, *12*, 34−39.
(41) Eastman, P.; Pande, V. *J. Chem. Theory Comput.* **2010**, 434−437.
(42) OpenMM. http://simtk.org/home/openmm (accessed Jul 8, 2012), available through SimTK.
(43) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E. *Bioinformatics* **2013**, *29*, 845−54.
(44) GROMACS. http://www.gromacs.org/ (accessed Jan 6, 2012).
(45) Mobley, D. L.; Bayly, C. I.; Cooper, M. D.; Shirts, M. R.; Dill, K. A. *J. Chem. Theory Comput.* **2009**, *5*, 350−358.
(46) Fukunishi, H.; Watanabe, O.; Takada, S. *J. Chem. Phys.* **2002**, *116*, 9058.
(47) Chodera, J. D.; Shirts, M. R. *J. Chem. Phys.* **2011**, *135*, 194110.
(48) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. *J. Comput. Phys.* **1977**, *23*, 327−341.
(49) Miyamoto, S.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 952−962.
(50) Chow, K.-H.; Ferguson, D. M. *Comput. Phys. Commun.* **1995**, *91*, 283−289.
(51) Åqvist, J.; Wennerström, P.; Nervall, M.; Bjelic, S.; Brandsdal, B. R. O. *Chem. Phys. Lett.* **2004**, *384*, 288−294.
(52) Parrinello, M.; Rahman, A. *J. Appl. Phys.* **1981**, *52*, 7182.
(53) Nosé, S.; Klein, M. L. *Mol. Phys.* **1983**, *50*, 1055−1076.
(54) Shirts, M. R. *J. Chem. Theory Comput.* **2013**, *9*, 909−926.
(55) Chodera, J. D.; Swope, W. C.; Pitera, J. W.; Seok, C.; Dill, K. A. *J. Chem. Theory Comput.* **2007**, *3*, 26−41.
(56) Analysis code and example implementation of basis functions for OpenMM and YANK can be found in the basisanalyze repository on GitHub at https://github.com/shirtsgroup/basisanalyze. The results in this paper used commit hash d2d18e440c.
(57) Lartillot, N.; Philippe, H. *Syst. Biol.* **2006**, *55*, 195−207.
(58) Schultz, A. J.; Kofke, D. A. *J. Chem. Theory Comput.* **2014**, *10*, 5229−5234.
(59) Basconi, J. E.; Shirts, M. R. *J. Chem. Theory Comput.* **2013**, *9*, 2887−2899.
(60) Chandler, D.; Andersen, H. C. *J. Chem. Phys.* **1971**, *54*, 26.
(61) Ciach, A.; Góźdź, W.; Stell, G. *Phys. Rev. E* **2007**, *75*, 051505.
(62) Patsahan, O.; Ciach, A. *Phys. Rev. E* **2007**, *75*, 051505.
(63) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *7863*, 11225−11236.

DOI: 10.1021/ct501047e
J. Chem. Theory Comput. XXXX, XXX, XXX−XXX