Supporting information for:

ACES: Optimized Alchemically Enhanced Sampling

Tai-Sung Lee, Hsu-Chun Tsai, Abir Ganguly, and Darrin M. York*

Laboratory for Biomolecular Simulation Research, Institute for Quantitative Biomedicine and Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, NJ 08854, USA

E-mail: Darrin.York@rutgers.edu

*To whom correspondence should be addressed

Contents

1 Notations

Table S1: Definition of potential energy terms and their abbreviations used as subscripts.[a]

Energy Term	Index	Description	Collective Term
U_{bond}	7	Bond stretch	Bonded (b) $U_b = U_{bond} + U_{ang} + U_{tor}$
U_{ang}	6	Angle bend	
U_{tor}	5	Torsion rotate (proper/improper)	
U_{LJ}	4	Lennard-Jones	Non-bonded (nb) $U_{nb} = U_{dir} + U_{1-4Ele} + U_{LJ} + U_{1-4LJ}$
U_{1-4LJ}	3	1-4 Lennard-Jones	
U_{dir}	2	PME direct/real space	
U_{1-4Ele}	1	1-4 Electrostatic	
U_{rec}	0	PME reciprocal space	

[a]Absence of an energy term subscript indicates all energy terms (i.e., summation over all energy terms, index $t = 0, \cdots 7$). The bonded and non-bonded terms are "short" and "intermediate" ranged, respectively, and kept track of with array lists, the latter for which is set by a distance cut-off and updated dynamically when needed. Note under these definitions, the 1-4 electrostatic and LJ terms are considered part of the non-bonded terms, and the total electrostatic energy (U_{Ele}) is not purely a "non-bonded" term as it contains also the non-local reciprocal space term. i.e., $U_{Ele} = U_{dir} + U_{1-4Ele} + U_{rec}$.

Each of the energy terms in Table S1, with the exception of U_{recip}, involves a straight-forward summation over the relevant sets of atoms to compute 2-body, 3-body or 4-body interactions. In the case of free energy simulations, we need to further distinguish between contributions to the energy that are made from different non-overlapping sets of atoms. Specifically, we need to subdivide the system into two main subdivisions: one region is alchemically transforming, whereas the remainder of the system is immutable (I), i.e., not transforming. Within the hybrid single-dual-topology, the immutable region is represented by a single "topology" and set of coordinates. The transforming region of the system is represented by a formal dual topology and separate sets of coordinates for each state, and is further subdivided into constrained coordinate/common core (CC) and the separable-coordinate/softcore (SC) regions. The CC region has corresponding atoms in each topology constrained to have the

same positions in order to facilitate phase space overlap between states during the alchemical transformation. The SC region, on the other hand, has separable independent coordinates for each topology that can adopt different conformations and do not directly interact with one another.

For example, if two drug molecules involved in an alchemical transformation share a common chemical core of atoms such as an aromatic ring and differ only by certain attached substituent for which atoms between the topologies cannot easily be mapped, then the CC region would contain the atoms of the common aromatic ring and the SC region would contain the atoms of the different substituent. In the alchemical transformation of $\lambda : 0 \rightarrow 1$, the SC atoms of state "0" are "turned off" by mutating the real atoms of state "0" into so-called "dummy atoms",[S1] while at the same time the SC atoms of state "1" are being "turned on" in a synchronous counter-transformation. The dummy atoms do not interact with their environment, with the exception of certain bonded interactions that must obey the constraint conditions that they introduce no net potential of mean force onto any non-dummy atom. Often this separable dual-coordinate approach requires the introduction of "softcore potentials",[S2,S3] i.e., explicit non-linear λ-dependent terms to "soften" the interaction of these atoms with their surroundings. These are most often employed for non-bonded interactions such as LJ and Ele (or in the case of PME electrostatics, often just the U_{dir} term), but other forms have also been developed for bonds and other energy terms.[S2,S4,S5]

Thus the system can be divided into non-overlapping regions described in Table S2: I (immutable), CC (transforming constrained/common core) and SC (transforming separable/softcore). The I region has the same atomic coordinates, parameters and internal potential energy for both states 0 and 1. The CC region can have different parameters between states 0 and 1, but the coordinates of mapped atoms are constrained to be the same. The SC region also can have different parameters between states 0 and 1, but unlike the CC region each state has its own separable set of atomic coordinates.

Table S2: Energy decomposition based on non-overlapping sets of atoms comprising the SC, CC and I regions[a].

Energy	Region/Interactions	Description
U^{SC}	Internal energy of the Transforming: Separable coordinate/softcore (SC) region	Each of the contributing bonded or non-bonded internal energy terms arises from a set of atoms that are contained within SC region; i.e., all atoms of the term belong to the SC region.
U^{CC}	Internal energy of the Transforming: Constrained coordinate/common core (CC) region	Each of the contributing bonded or non-bonded internal energy terms arises from a set of atoms that are contained within CC region; i.e., all atoms of the term belong to the CC region.
U^{I}	Internal energy of the Immutable (I) region	Each of the contributing bonded or non-bonded internal energy terms arises from a set of atoms that are contained within I region; i.e., all atoms of the term belong to the I region.
$U^{(CC+I)}$	Internal energy of the combined CC and I regions; $U^{CC+I} = U^{CC} + U^{CC/I} + U^{I}$	Each of the contributing bonded or non-bonded internal energy terms arises from a set of atoms that are contained within (CC+I) region; i.e., all atoms of the term belong to the (CC+I) region.
$U^{SC/(CC+I)}$	Interaction energy between SC and (CC+I) regions; $U^{SC/(CC+I)} = U^{SC/CC} + U^{SC/I}$	Each of the contributing bonded or non-bonded interaction energy terms arises from a set of atoms that span the SC and combined (CC+I) regions; i.e., some belong to the SC region, while others in the same term belong to the (CC+I) region.
$U^{(SC+CC+I)}$	Total internal of the system; $U = U^{(SC+CC+I)}$.	Total potential energy of the system. This can be decomposed into contributions within the SC and (CC+I) regions, and interaction energy between regions as: $U = U^{(SC+CC+I)} = U^{(CC+I)} + U^{SC} + U^{SC/(CC+I)}$.

[a]The three regions are as follows. 1) SC, Transforming: Separable coordinate/softcore; 2) CC, Transforming: Constrained coordinate/common core; and 3) I, Immutable (not transforming). Regions can be combined as a union of atom sets with the "+" operator; e.g., (SC+CC) combines the SC and CC regions, and (SC+CC+I) would imply all regions (and hence atoms) of the system. The energy decomposition involves using the superscripts U^{X} to indicate either an internal energy within the region "X", or an interaction energy $U^{X/Y}$ between regions "X" and "Y". Thus in its most expanded form, the total potential energy can be written as $U = U^{SC} + U^{CC} + U^{I} + U^{SC/CC} + U^{SC/I} + U^{CC/I}$. Note the absence of subscripts indicates a summation over all energy terms in Table S1; however, the superscript notation can also be applied to energy terms individually. It is assumed that the SC, CC and I regions are defined in such a way that no individual 3-body or 4-body term spans all three regions (e.g., there is no angle bending term that has one atom in each of the SC, CC and I regions).

2 Smoothstep functions

Consider the family of smoothstep functions of orders P ($P = 0, 1, 2, \cdots$) defined as the polynomial functions (up to $P = 4$ shown):

$$\text{for } 0 \le x \le 1:$$

$$S_0(x) = x,$$

$$S_1(x) = -2x^3 + 3x^2,$$

$$S_2(x) = 6x^5 - 15x^4 + 10x^3,$$

$$S_3(x) = -20x^7 + 70x^6 - 84x^5 + 35x^4,$$

$$S_4(x) = 70x^9 - 315x^8 + 540x^7 - 420x^6 + 126x^5,$$

$$\text{and}$$

$$S_P(x \le 0) = 0; S_P(x \ge 1) = 1, \forall \ P \in \mathbb{N} \tag{1}$$

The smoothstep functions are monotonically increasing functions that have desirable 0 and 1 endpoint values and vanishing endpoint derivative properties:

$$\left[\frac{d^k S_P(x)}{dx^k}\right]_{x=0} = \left[\frac{d^k S_P(x)}{dx^k}\right]_{x=1} = 0 \ \forall \ k \in \mathbb{N}, \ 0 < k \le \mathrm{P} \tag{2}$$

In addition, the smoothstep functions obey the symmetry condition

$$S_P(1 - x) = 1 - S_p(x) \tag{3}$$

A smoothstep function with a higher order will have a smoother function curve and smaller derivatives near 0 and 1 but a larger derivative in between. The zero-order ($P = 0$) smoothstep function is in fact simply linear with constant slope, including at the endpoints, which can lead to endpoint catastrophe problems. As illustrated in previous work,[S3] the second order smoothstep function ($P = 2$) overall offers a good balance between smooth vanishing

derivatives at the endpoints, and modest derivatives for intermediate values of λ. AMBER20 offers the flexibility to choose different smoothstep functions through the λ-scheduling mechanism described below.

3 Form of the λ-dependent weight functions

We now describe a general form for the weight functions $W(\lambda)$, where we only retain the 0 and 1 subscript to indicate the state. The weight functions are defined in term of the smoothstep functions as

$$W_0(\lambda) = 1 - S_P(\lambda) = S_P(1-\lambda) \tag{4}$$

$$W_1(\lambda) = S_P(\lambda) \tag{5}$$

Previous work has illustrated that use of smoothstep functions of order greater than 0 (i.e., a weight function that goes beyond the simple linear λ-dependence and has vanishing derivatives at the endpoints), affords improvement of the the transformation pathway, particularly at the endpoints where large variation in $< \partial U/\partial\lambda >_\lambda$ can occur.[S3] These weight functions both operate within the range $0 \leq \lambda \leq 1$ (they have constant endpoint values outside of this range), and satisfy the normalization condition:

$$W_0(\lambda) + W_1(\lambda) = 1 \tag{6}$$

and the symmetry condition:

$$W_0(1-\lambda) = W_1(\lambda) \tag{7}$$

4 λ-scheduling of weight functions

In some cases, it is desirable to have the flexibility to apply more complicated λ schedules that operate over a subinterval of λ values between 0 and 1. The generalized λ scheduling weight for W_0 can be defined so that it is changing only within the interval $\lambda_{min} \leq \lambda \leq \lambda_{max}$ as

$$\begin{aligned} W_0(\lambda) &= 1 - S_P(z(\lambda)) \\ z(\lambda) &\begin{cases} = 0, & \text{if } \lambda \leq \lambda_{min} \\ = \frac{\lambda - \lambda_{min}}{\lambda_{max} - \lambda_{min}}, & \text{if } \lambda_{min} \leq \lambda \leq \lambda_{max} \\ = 1, & \text{if } \lambda_{max} \leq \lambda \end{cases} \end{aligned} \tag{8}$$

where $0 \leq \lambda_{min} \leq \lambda_{max} \leq 1$. In the current framework, the complementary weight function $W_1(\lambda)$ can be selected to either satisfy the normalization condition (eq 6), or the symmetry condition (eq 7) above. Only if the interval $z(\lambda)$ is centered at $\lambda = 0.5$ are both the normalization and symmetry conditions simultaneously satisfied. AMBER22 allows flexible λ scheduling of this form for different energy components. The lambda-scheduling can be enabled by setting the input control `gti_lam_sch`=1 and the scheduling is defined in the lambda-scheduling control file (the file name has default value of "lambda.sch" and can be specified by the command line argument `-lambda_sch`). The detailed usage can be found in the updated AMBER22 manual.[S6]

5 New Softcore Potentials

In the current framework, softcore potentials are developed for both the LJ and non-bonded electrostatic interactions (i.e., the direct/real space component of the PME method). Hence, the main terms that are affected by the softcore potentials are those contained in the non-bonded interactions between the SC and (CC+I) regions, i.e., those terms present in $U_{nb}^{SC/(CC+I)}$. Formally, these terms can also be present in the internal energy of the SC

region, if these terms are being "turned off" to form the "dummy state". In principle, the internal potential energy interactions in the dummy state are arbitrary so long as they are treated consistently in different legs of the thermodynamic cycle that are subtracted. However, in practice, choice of the interactions in the dummy state are important, and should be selected to minimize the volume of phase space needed to sample the dummy state while at the same time avoiding sampling traps (multiple distinct free energy basins separated by high barriers) that could lead to inconsistent results. In fact, proper choice of potential energy interactions in the dummy state, together with generalized ensemble methods such as Hamiltonian replica exchange, can lead to powerful new alchemical enhanced sampling methods. For the present paper, the dummy state was created by scaling (i.e., "turning off") electrostatic interactions, and in some cases also torsion angle and 1-4 LJ terms, but keeping other bonded and normal LJ terms in place. Hence, the major influence of the new form of the softcore potential will affect the $U_{nb}^{SC/(CC+I)}$ term, which as results presented later in the paper will show has a profound affect on the free energy estimates.

The LJ, Coulomb and PME direct-space interactions for a set of interacting point particles i and j separated by a distance r_{ij} are given by

$$U_{\mathrm{LJ}}(r_{ij}) = 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \tag{9}$$

$$U_{\mathrm{Coul}}(r_{ij}) = \left(\frac{q_i q_j}{4\pi\epsilon_0}\right)\frac{1}{r_{ij}} \tag{10}$$

and

$$U_{\mathrm{dir}}(r_{ij}) = \left(\frac{q_i q_j}{4\pi\epsilon_0}\right)\frac{erfc(\kappa r_{ij})}{r_{ij}} \tag{11}$$

where σ_{ij} and ϵ_{ij} are the pairwise LJ contact distance and well depth, respectively, and q_i and q_j are the partial charges of particles i and j, respectively, $erfc()$ is the complementary error function and κ is the Ewald coefficient.

To soften these pairwise interactions particles, the introduction of a parametric form for

scaling with an adjustable parameter is utilized to modify the effective interaction distance. A commonly used form of these modifications[S2,S4] is shown as

$$r_{ij}^{\mathrm{LJ}}(\lambda;\alpha) = \left[r_{ij}^{n} + \lambda\alpha\sigma_{ij}^{n}\right]^{1/n} \tag{12}$$

and

$$r_{ij}^{\mathrm{Coul}}(\lambda;\beta) = \left[r_{ij}^{m} + \lambda\beta\right]^{1/m} \tag{13}$$

where n and m are positive integers and α and β are adjustable positive semi-definite parameters (note that α is unitless whereas β has units of distance raised to the power of m). The value of $n = 6$ and $m = 2$ are often used, and have been the default values in AMBER, until recently a modified form of the separation-shifted scaling leads to considerable improvement.[S7]

We introduce a new form of the interaction distance with separation-shifted scaling, given as

$$r_{ij}^{\mathrm{LJ*}}(\lambda;\alpha^{\mathrm{LJ*}}) = \left[r_{ij}^{n} + \alpha^{\mathrm{LJ}} \cdot f_{SW}(r_{ij})S_2(\lambda)\sigma_{ij}^{n}\right]^{1/n} \tag{14}$$

and

$$r_{ij}^{\mathrm{Coul*}}(\lambda;\alpha^{\mathrm{Coul}}) = \left[r_{ij}^{m} + \alpha^{\mathrm{Coul}} \cdot f_{SW}(r_{ij})S_2(\lambda)\sigma_{ij}^{m}\right]^{1/m} \tag{15}$$

where α^{LJ} and α^{Coul} are the corresponding unitless parameters, S_2 is the 2nd-order smooth-step function in eq 1, and $f_{SW}(r_{ij})$ is a switching function designed to smoothly return to the normal r_{ij}, and thus long-ranged behavior, by the end of the cutoff

$$f_{SW}(r_{ij}) \equiv 1 - S_2\left(\frac{r_{ij} - R_{cut,i}}{R_{cut,f} - R_{cut,i}}\right) \tag{16}$$

where $R_{cut,i}$ is the distance that the switching function begins switching and $R_{cut,f}$ is the final distance where the switching ends (returning the effective interaction distance to be r_{ij}). Henceforth, we will set $R_{cut,f} = R_{cut}$, and $R_{cut,i} = R_{cut} - 2$, respectively.

The form of the new softcore potential is thus

$$U_{LJ}(r_{ij};\lambda) = U_{LJ}[r_{ij}^{\mathrm{LJ}*}(\lambda;\alpha^{\mathrm{LJ}*})] \tag{17}$$

$$U_{C}(r_{ij};\lambda) = U_{C}[r_{ij}^{\mathrm{Coul}*}(\lambda;\alpha^{\mathrm{Coul}*})] \tag{18}$$

and

$$U_{dir}(r_{ij};\lambda) = U_{dir}[r_{ij}^{\mathrm{Coul}*}(\lambda;\alpha^{\mathrm{Coul}*})] \tag{19}$$

Most free energy simulations in the condensed phase are performed under periodic boundary conditions and use the PME method[S8,S9] to treat long-ranged electrostatic interactions, in which case the electrostatic softcore potentials described here apply to the U_{dir} term (eq 19). If, on the other hand, PME electrostatics is not used, then the electrostatic softcore potential apply to the U_C term (eq 18).

6 REST2-like enhanced sampling implementation

A REST2 like enhanced sampling scheme has been implemented in the AMBER22 package. As already discussed in the main text. The REST2 formula in fact is very similar to an alchemical transformation. Here we describe our REST2-like implementation, which is similar to REST2 in spirit but it provides more controls over how the weight functions behavior.

A region to be enhanced sampled is designed as the RE region. A parameter, τ , similar to λ in alchemical transformation, is used to define the reduction of interactions involved in the RE region. A set of τ-dependent weighting functions can be utilized to form the

τ-dependent total potential energy $U(\mathbf{r}^N;\tau)$ can be written as

$$U(\mathbf{r}^N;\tau) = U_{rec}(\mathbf{r}^N; W_{rec}(\tau)\cdot\mathbf{q}^{\mathbf{RE}}) + \sum_{t\neq rec} W_t^{RE}(\tau)\cdot U_t^{RE}(\mathbf{r}^N;\tau) + \sum_{t\neq rec} W_t^{RE/Env}(\tau)\cdot U_t^{RE/Env}(\mathbf{r}^N;\tau) \quad (20)$$

where $U(\mathbf{r}^N;\tau)$ is the system total potential energy. The subscription "rec" represents the PME reciprocal term and "t" represents all general terms (Table S1). The superscript "RE" represents the interactions within the RE region while "RE/Env" the interactions between the RE region and its surrounding environment. "$\mathbf{q}^{\mathbf{RE}}$" is the charge set of all atoms in the RE region. In eq 20, the τ-dependency of the PME reciprocal term is through the scaling of the charges of the RE region atoms by the weight function $W_{rec}(\tau)$, and other terms through alchemical transformation-like weight functions. In the default linear dependency scheme, the weight functions are

$$\begin{aligned} W_{rec}(\tau) = W_t^{RE/Env}(\tau) &= (1-\tau) \\ W_t^{RE}(\tau) &= (1-\tau)^2 \\ 0 \leq \tau, W_{rec}(\tau), W_t^{RE/Env}(\tau), W_t^{RE}(\tau) &\leq 1. \end{aligned} \quad (21)$$

It is clear that $\tau = 0$ corresponds to unscaled interactions while $\tau = 1$ the totally vanished interactions (similar to setting temperature as infinity). Table S3 shows the corresponding REST2 temperatures for different τ and weight functions. In our implementation, these weight functions can be modified just as the alchemical transformation weight functions, including non-linear dependency using smoothstep functions and onset/offset at different τ values (τ scheduling).

Table S3: The corresponding REST2 temperatures for different τ and weight functions, assuming that the simulation is perform at 298.0K

τ	$W_t^{RE/Env}(\tau) = (1-\tau)$	$W_t^{RE}(\tau) = (1-\tau)^2$	REST2 temperature (K)
0	1	1	298.00
0.1	0.9	0.81	367.90
0.2	0.8	0.64	465.63
0.3	0.7	0.49	608.16
0.4	0.6	0.36	827.78
0.5	0.5	0.25	1192.00
0.6	0.4	0.16	1862.50
0.7	0.3	0.09	3311.11

Furthermore, these weight functions can be turned on or off to accommodate different situations. Similar to the `gti_add_sc` input control for the alchemical transformation, the `gti_add_re` input control is shown in Table S4.

Table S4: The scaling behavior τ-dependence of the weight functions in eq 21, controlled by the `gti_add_re` flag, for different energy terms and regions/interactions in AMBER22.

Weight Symbol	Energy Term Abbreviation	Region / Interaction	**gti_add_re** flag 1	2	3	4	5	6	7
W_{bond}^{RE}	bond	RE	P	P	P	P	P	P	P
W_{ang}^{RE}	ang	RE	P	P	P	P	P	P	S
W_{tor}^{RE}	tor	RE	P	P	P	S	S	S	S
W_{dir}^{RE}	dir	RE	P	S	S	P	S	S	S
W_{1-4Ele}^{RE}	1-4 Ele	RE	P	S	S	S	S	S	S
W_{LJ}^{RE}	LJ	RE	P	P	S	P	P	S	S
W_{1-4LJ}^{RE}	1-4 LJ	RE	P	P	S	S	S	S	S
$W_{dir}^{RE/Env}$	dir	RE/Env	S	S	S	S	S	S	S
$W_{1-4Ele}^{RE/Env}$	1-4 Ele	RE/Env	S	S	S	S	S	S	S
$W_{LJ}^{RE/Env}$	LJ	RE/Env	S	S	S	S	S	S	S
$W_{1-4LJ}^{RE/Env}$	1-4 LJ	RE/Env	S	S	S	S	S	S	S
W_{rec}	rec	all	S	S	S	S	S	S	S

Energy terms are defined in text (and also Table S1). RE: the RE region internal interactions, RE/Env the interactions between the RE region and the environment.
Flags:
S: Scaled with τ. The corresponding λ-dependent weight function will be used.
P: Not scaled with τ. The corresponding weight function is simply 1.

The implemented module will be enabled by setting the input control "`ifreaf`=1" and the parameter τ is defined by the input "`reaf_tau`", e,g.,`reaf_tau`=0.001, The RE region is defined by the standard AMBER mask selection, e.g., "`reaf_mask1`='@1-8'". The "`reaf_mask1`" is used to define the RE region which are in a regular MD or at the $\lambda = 0$ in an alchemical transformation. This implementation can be used in conjunction with alchemical transformation, similar to FEP/REST. When used together with alchemical transformation, "`reaf_mask2`" is used to define the RE region at the state of $\lambda = 1$. Details are described in the AMBER22 manual.[S6]

7 Preliminary tests of different lambda spacing schemes

The Hamiltonian replica exchange efficiency will critically impact the ACES results, which leads to a related often-asked question: the best choice of lambda points and spacing. As a preliminary effort towards the optimal choice of lambda points and spacing, We repeated the XX simulation with different lambda-spacing schemes. Several schemes are defined:

21W: 21 evenly-spaced lambda windows from 0 to 1, spacing = 0.05

11W: 11 evenly-spaced lambda windows from 0 to 1, spacing = 0.10

13W(S1): 13 lambda windows from 0 to 1: 0.000, 0.177, 0.259, 0.326, 0.387, 0.444, 0.500, 0.556, 0.613, 0.674, 0.741, 0.823, 1.000

13W(S2): 13 lambda windows from 0 to 1: 0.000, 0.230, 0.303, 0.359, 0.409, 0.455, 0.500, 0.545, 0.591, 0.641, 0.697, 0.770, 1.000

13W(S1+S2): 13 lambda windows from 0 to 1: 0.000, 0.203, 0.282, 0.344, 0.399, 0.450, 0.500, 0.550, 0.601, 0.656, 0.718, 0.797, 1.000

We calculated several quantities for neighbor lambda window pairs, including, the exchanged probability $p^{ex}_{i,i-1}$, the exchange probability between the i^{th} replica and the $i-1^{th}$ replica, as well as $(\frac{1}{p^{ex}_{i,i-1}}-1)$, the standard errors and the variance of MBAR, the overlap between neigh windows (in percentage). Comparisons between these calculated quantities are shown in the following sections.

Figure S1 shows the exchange probabilities from different lambda schemes. The 21W scheme has all exchange probabilities > 0.3, although with smaller values in the middle range, apparently due to the small spacing between windows. The 11W scheme, on the other hand, has exchange probabilities less <0.1 between 0.3 and 0,7. All schemes with 13 windows, designed to smooth out the exchange probability distribution over the lambda axis, have narrow distributions between 0.2 and 0.5.

Figure S2 shows the relation between the exchange probabilities and the MBAR errors

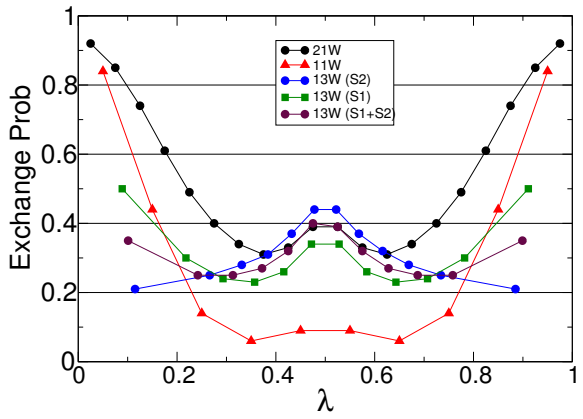

Figure S1: The exchange probabilities from different lambda schemes.

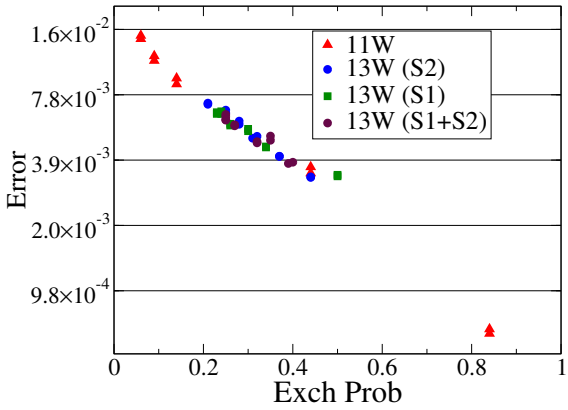

Figure S2: Exchange probabilities vs the MBAR error with different lambda schemes.

(in the logarithm scale) and demonstrates that the error is highly (anti-)correlated to the exchange probability. Note that in the 11W scheme, which has high exchange probabilities near the end states and low exchange probabilities in between, the MBAR errors are very small near the end stats but much larger otherwise. Similar situations are also seen in Figures S3 and S4.

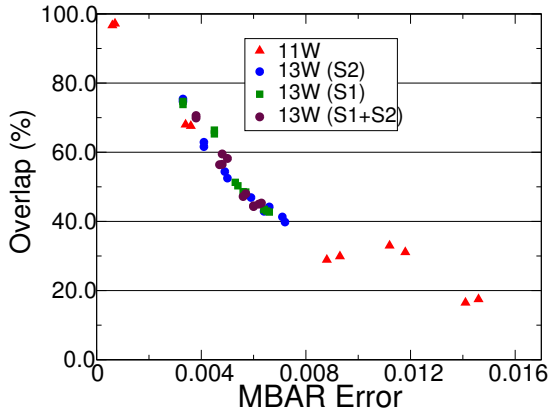

Figure S3: Phase space overlap (in percentage) vs the MBAR error with different lambda schemes.

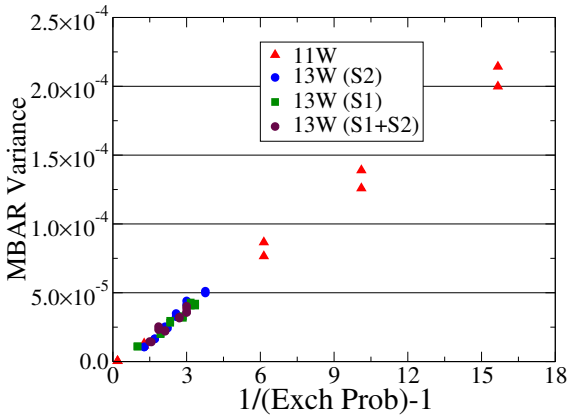

Figure S4: $(\frac{1}{p^{ex}_{i,i-1}} - 1)$ vs the MBAR variance with different lambda schemes.

Similarly, the Phase space overlap (in percentage) vs the MBAR error and $(\frac{1}{p^{ex}_{i,i-1}} - 1)$

vs the MBAR variance (the square of the MBAR errors) are shown in Figures S3 and S4, respectively. While all figures show strong correlations, Figure S4 shows a particularly strong linear relation between $(\frac{1}{p^{ex}_{i,i-1}} - 1)$ and the MBAR variance. The term $(\frac{1}{p^{ex}_{i,i-1}} - 1)$ is the core part of the analytic solution of the round trip time derived from an ideal Markov Chain Monte Carlo model, $E[T]$:[S10]

$$E[T] = 2N \left\{ 1 + \sum_{i=2}^{N} \left(\frac{1}{p^{ex}_{i,i-1}} - 1 \right) \right\} \tag{22}$$

where N is the total number of replicas, i is the replica index running from 1 to N. The strong linear correlation between $(\frac{1}{p^{ex}_{i,i-1}} - 1)$ and the MBAR variance suggests that $(\frac{1}{p^{ex}_{i,i-1}} - 1)$ could be a very useful measure of the efficiency of replica exchange or ACES simulations.

References

(S1) Fleck, M.; Wieder, M.; Boresch, S. *J. Chem. Theory Comput.* **2021**, *17*, 4403–4419.

(S2) Steinbrecher, T.; Joung, I.; Case, D. A. *J. Comput. Chem.* **2011**, *32*, 3253–3263.

(S3) Lee, T.-S.; Lin, Z.; Allen, B. K.; Lin, C.; Radak, B. K.; Tao, Y.; Tsai, H.-C.; Sherman, W.; York, D. M. *J. Chem. Theory Comput.* **2020**, *16*, 5512–5525.

(S4) Beutler, T. C.; Mark, A. E.; René C. van Schaik and Paul R. Gerber and Wilfred F. van Gunsteren, *Chem. Phys. Lett.* **1994**, *222*, 529–539.

(S5) Steinbrecher, T.; Mobley, D. L.; Case, D. A. *J. Chem. Phys.* **2007**, *127*, 214108.

(S6) Case, D. A.; Aktulga, H. M.; Belfon, K.; Ben-Shalom, I. Y.; Berryman, J.; Brozell, S. R.; Cerutti, D. S.; Cheatham III, T. E.; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.; Giambasu, G.; Gilson, M. K.; Gohlke, H.; Goetz, A. W.; Harris, R.; Izadi, S.; Izmailov, S. A.; Kasavajhala, K.; Kaymak, M. C.; King, E.; Kovalenko, A.; Kurtzman, T.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Machado, M.; Man, V.; Manathunga, M.; Merz, K. M.; Miao, Y.; Mikhailovskii, O.; Monard, G.; Nguyen, H.; O'Hearn, K. A.; Onufriev, A.; Pan, F.; Pantano, S.; Qi, R.; Rahnamoun, A.; Roe, D.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shajan, A.; Shen, J.; Simmerling, C. L.; Skrynnikov, N. R.; Smith, J.; Swails, J.; Walker, R. C.; Wang, J.; Wang, J.; Wei, H.; Wolf, R. M.; Wu, X.; ; Xiong, Y.; Xue, Y.; York, D. M.; Zhao, S.; Kollman, P. A. AMBER22. University of California, San Francisco: San Francisco, CA, 2022.

(S7) Tsai, H.-C.; Lee, T.-S.; Ganguly, A.; Giese, T. J.; York, D. M. *in press* **2022**,

(S8) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089–10092.

(S9) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577–8593.

(S10) Syed, S.; Bouchard-Côté, A.; Deligiannidis, G.; Doucet, A. *J. R. Stat. Soc. Series B Stat. Methodol.* **2021**, *0*, 1–30.