

# Variationally derived intermediates for correlated free-energy estimates between intermediate states

 Martin Reinhardt and Helmut Grubmüller \*

Max Planck Institute for Biophysical Chemistry, 37077 Göttingen, Germany



(Received 29 July 2020; accepted 28 September 2020; published 13 October 2020)

Free-energy difference calculations based on atomistic simulations generally improve in accuracy when sampling from a sequence of intermediate equilibrium thermodynamic states that bridge the configuration space between two states of interest. For reasons of efficiency, usually the same samples are used to calculate the stepwise difference of such an intermediate to both adjacent intermediates. However, this procedure violates the assumption of uncorrelated estimates that is necessary to derive both the optimal sequence of intermediate states and the widely used Bennett acceptance ratio estimator. In this work, via a variational approach, we derive the sequence of intermediate states and the corresponding estimator with minimal mean-squared error that account for these correlations and assess its accuracy.

 DOI: [10.1103/PhysRevE.102.043312](https://doi.org/10.1103/PhysRevE.102.043312)

## I. INTRODUCTION

Free-energy calculations are widely used to investigate physical and chemical processes [1–7]. Their accuracy is essential to biomedical applications such as computational drug development [8–11] or material design [12–15]. Among the most widely used methods based on simulations with atomistic Hamiltonians are alchemical equilibrium techniques, including the free-energy perturbation (FEP) [16] and thermodynamic integration (TI) [17] methods. These techniques determine the free-energy difference between two states, representing, for example, two different ligands bound to a target, by sampling from intermediate states whose Hamiltonians are constructed from those of the end states. The free-energy difference between the end states is then determined via a stepwise summation of the differences between the intermediate states.

The choice of these intermediates critically affects the accuracy of the free-energy estimates [18–20] by determining which parts of the configuration space are sampled to which extent [21], thereby performing a function similar to importance sampling [22]. In addition, different estimators that determine the free-energy differences between these intermediates and the end states have been developed, most prominently the Zwanzig formula [16] for FEP, the Bennett acceptance ratio method (BAR) [23], and multistate BAR (MBAR) [24].

We have recently derived [25] the sequence of discrete intermediate states—the variationally derived intermediates (VI)—that yield, for finite sampling, the lowest mean-squared

error (MSE) of the free-energy estimates with respect to the exact value. Their form differs from the most common scheme, which, for  $N$  states, linearly interpolates between the end states Hamiltonians  $H_1(\mathbf{x})$  and  $H_N(\mathbf{x})$ , respectively, along a path variable  $\lambda_s$ ,

$$H_s(\mathbf{x}) = (1 - \lambda_s)H_1(\mathbf{x}, \lambda_s) + \lambda_s H_N(\mathbf{x}, \lambda_s), \quad \lambda_s \in [0, 1], \quad (1)$$

where  $\mathbf{x} \in \mathbb{R}^{3M}$  denotes the coordinate vector of all  $M$  particles in the system. All states are labeled by an integer  $s$  with  $1 \leq s \leq N$  and  $\lambda_s$  corresponding to state  $s$ . The additional  $\lambda_s$  argument of the end states Hamiltonians indicates the common use of soft-core potentials [26–28] to avoid divergences for vanishing particles. Other approaches involve the interpolation of exponentially weighted Hamiltonians of the end states, such as enveloping distribution sampling [29] (and variants thereof [30,31]) or the minimum variance path [32,33] for TI.

In contrast, the VI are not directly defined via the end states; instead, the optimal form of each intermediate  $s$  is determined by the form of the adjacent ones  $s - 1$  and  $s + 1$ . For the setup shown in Fig. 1(a), which consists of two types of intermediates, sampling is conducted in the first type labeled with even numbered  $s$  and indicated by the solid lines with yellow points. These are governed by the optimal Hamiltonian,

$$H_s(\mathbf{x}) = -\frac{1}{2} \ln \left[ e^{-2H_{s-1}(\mathbf{x})} r_{s-1,s}^{-2} + e^{-2H_{s+1}(\mathbf{x})} r_{s+1,s}^{-2} \right], \quad (2)$$

where  $r_{s,t} = Z_s/Z_t$  denotes the ratio of the configurational partition sums of states  $s$  and  $t$ . Virtual intermediates are the second type and labeled with odd numbers  $s$  with  $2 < s < N - 1$  and indicated by the dashed lines in Fig. 1(a). For these,

$$H_s(\mathbf{x}) = \ln \left[ e^{H_{s-1}(\mathbf{x})} r_{s-1,s} + e^{H_{s+1}(\mathbf{x})} r_{s+1,s} \right]. \quad (3)$$

Virtual intermediates are used as target states to evaluate the difference in free energy, too, and no sampling is conducted in those. Due to the coupling of the VI, the optimal MSE sequence of Hamiltonians  $H_2(\mathbf{x}) \cdots H_{N-1}(\mathbf{x})$  is determined by solving the system of  $N - 2$  equations of Eqs. (2) and (3).

\*hgrubmu@gwdg.de

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Open access publication funded by the Max Planck Society.

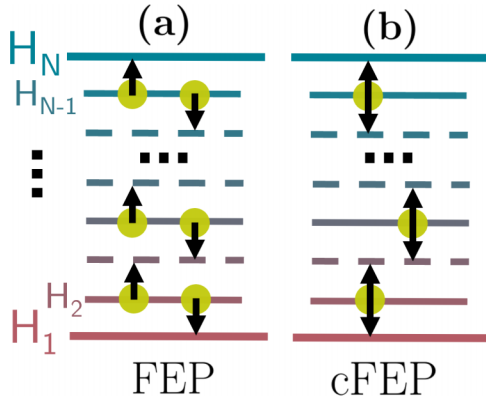


FIG. 1. Two schemes of free-energy calculation. The arrows indicate the Zwanzig formula is used to evaluate the free-energy difference to the adjacent state based on sample sets represented through yellow dots. The dashed lines represent virtual intermediate states that no sampling is conducted in. (a) Separate and uncorrelated sample set are used to calculate the free-energy difference of the respective intermediate to the state above and below. (b). The same sample set is used for this purpose.

The variational MSE minimization has been conducted based on the Zwanzig formula [16],

$$\Delta G_{s,s+1} = -\ln\langle e^{-[H_{s+1}(\mathbf{x})-H_s(\mathbf{x})]} \rangle_s, \quad (4)$$

being used to calculate the difference between two adjacent states, as indicated by the arrows in Fig. 1. However, using the virtual target states described by Eq. (3) is equivalent to using BAR directly between two sampling states [25,34], and, therefore, Eq. (3) also describes the optimal intermediates for BAR.

Note that the Hamiltonians of the optimal intermediates, Eqs. (2) and (3), depend on the ratios of the partition sums, i.e., the desired quantity. Therefore, the system of equations has to be solved iteratively. The principle is the same as for BAR, where the estimator depends on an estimate of the free-energy difference, and the optimal estimator is, therefore, determined iteratively. However, BAR is, in practice, mostly used with sampling states that are governed by Eq. (1) and a user chosen  $\lambda$  value. The VI method generalizes the BAR principle and determines not only the estimator but also the form of all intermediates through such an iterative optimization by using the information of the free-energy estimates between these states. The resulting form differs from Eq. (1) and does not require any additional user choice of a  $\lambda$  variable.

However, for both BAR and VI to be optimal for multiple states, the free-energy estimates to the states above and below an intermediate in the sequence have to be based on separate, uncorrelated sample points. This is illustrated by the separate yellow points in Fig. 1(a) that we refer to as the regular FEP setup, which was the topic of our previous work [25]. Yet it would be twice as efficient to use the same sample points in both directions, as illustrated by Fig. 1(b), and as generally done in practice. However, this introduces correlations between the estimates to both adjacent intermediates, thereby violating the assumptions underlying the derivation of Eqs. (2) and (3). Therefore, in this case, BAR and the above variational intermediates are not optimal anymore. Due to these correla-

tions, we refer to the Fig. 1(b) as the correlated FEP (cFEP) setup, which is the topic of this work.

Here we derive the minimal MSE sequence of intermediate states and the corresponding estimators for cFEP, as used in practice, that take these correlations properly into account. This is in contrast to the derivation of VI [25] for FEP, where these correlations do not occur. As will be shown below, what might seem as a minor technical twist, markedly changes the shape of the optimal intermediates and considerably improves the accuracy of the obtained free-energy estimates.

## II. THEORY

For the cFEP scheme shown in Fig. 1(b), we aim to derive the sequence of intermediate Hamiltonians  $H_2(\mathbf{x}) \cdots H_{N-1}(\mathbf{x})$  that optimizes the MSE,

$$\text{MSE}(\Delta G^{(n)}) = \mathbb{E}[(\Delta G - \Delta G^{(n)})^2], \quad (5)$$

along similar lines as for the regular FEP scheme [25], shown in Fig. 1(a). Here  $\Delta G_{1,N}^{(n)}$  denotes the free-energy estimate based on a finite number of sample points  $n$  and  $\Delta G_{1,N}$  the exact difference between the end states 1 and  $N$ .

For the optimization metrics, different choices are possible, such as the Kullback-Leibler divergence [35] or the Fisher information metric [18,36] that measure the (dis)similarity between configuration space densities. Instead, here we chose to directly optimize the MSE, as it quantifies the average accuracy with respect to the exact free-energy difference, which is the relevant measure for most practical applications. Furthermore, as the MSE can be decomposed into variance plus bias squared, we account for both of these contributions that are oftentimes optimized separately in the literature [37,38].

The cFEP variant in Fig. 1(b) only uses sampling in the intermediate states. Setups that, in addition, involve sampling in the end states, can also be treated with the formalism below. However, first, as we have tested, the accuracy for a given computational effort does not increase in this case. Second, mixing two different types of sample points (the ones used to evaluate  $\Delta H$  to only one adjacent state vs. to both adjacent states) further complicates the analysis.

For cFEP, the estimated difference is

$$\Delta G^{(n)} = \sum_{s=2}^{N-2} \sum_{s \text{ even}} [\Delta G_{s \rightarrow s+1}^{(n)} - \Delta G_{s \rightarrow s-1}^{(n)}]. \quad (6)$$

As in Fig. 1(b), the arrows point from sampling to target states, i.e., either the end states or the virtual intermediates. Assuming for each sample state  $s$  a set of  $n$  independent sample points  $\{\mathbf{x}_i\}$ , drawn from  $p_s(\mathbf{x}) = e^{-H_s(\mathbf{x})}/Z_s$ , with partition function  $Z_s$ , expanding Eq. (5) with the use of Eq. (6) reads

$$\begin{aligned} \text{MSE}(\Delta G_{1,N}^{(n)}) &= (\Delta G_{1,N})^2 + \sum_{s=2}^{N-2} \sum_{s \text{ even}} \mathbb{E}[(\Delta G_{s \rightarrow s+1}^{(n)})^2 + (\Delta G_{s \rightarrow s-1}^{(n)})^2] \end{aligned}$$

$$\begin{aligned}
 & -2\Delta G_{1,N} \left( \sum_{\substack{s=2 \\ s \text{ even}}}^{N-2} (\mathbb{E}[\Delta G_{s \rightarrow s+1}^{(n)}] - \mathbb{E}[\Delta G_{s \rightarrow s-1}^{(n)}]) \right) \\
 & - \sum_{\substack{s=2 \\ s \text{ even}}}^{N-2} \sum_{\substack{t=2 \\ t \text{ even}}}^{N-2} \mathbb{E}[2\Delta G_{s \rightarrow s+1}^{(n)} \Delta G_{t \rightarrow t-1}^{(n)}]. \quad (7)
 \end{aligned}$$

The first two lines of Eq. (7) have already been processed in Ref. [25], but the last term differs. Previously, as in the regular FEP scheme in Fig. 1(a), these last expectation values were originally derived from independent sample sets and were, therefore, uncorrelated. In the present context of cFEP, however, these estimates are correlated. Therefore, the term needs to be split in two sums, distinguishing between the pairs with samples from the same state and the ones from different states,

$$\begin{aligned}
 & \sum_{\substack{s=2 \\ s \text{ even}}}^{N-2} \sum_{\substack{t=2 \\ t \text{ even}}}^{N-2} \mathbb{E}[2\Delta G_{s \rightarrow s+1}^{(n)} \Delta G_{t \rightarrow t-1}^{(n)}] \\
 & = 2 \sum_{\substack{s=2 \\ s \text{ even}}}^{N-2} \mathbb{E}[\Delta G_{s \rightarrow s+1}^{(n)} \Delta G_{s \rightarrow s-1}^{(n)}] \\
 & + 2 \sum_{\substack{s=2 \\ s \text{ even}}}^{N-2} \sum_{\substack{t=2 \\ t \text{ even} \\ t \neq s}}^{N-2} \mathbb{E}[\Delta G_{s \rightarrow s+1}^{(n)}] \mathbb{E}[\Delta G_{t \rightarrow t-1}^{(n)}], \quad (8)
 \end{aligned}$$

where the expectation value of the product between the two estimates based on different sample sets has been separated, as these are uncorrelated.

As we are only interested in the intermediates that optimize the MSE, and not in the absolute value of the MSE, we focus on the terms that will not drop out in the optimization below.

Continuing with the expression inside the sum of the first term on the right-hand side of Eq. (8),

$$\begin{aligned}
 & \mathbb{E}[\Delta G_{s \rightarrow s+1}^{(n)} \Delta G_{s \rightarrow s-1}^{(n)}] \\
 & = - \int p_s(\mathbf{x}_1) d\mathbf{x}_1 \dots \int p_s(\mathbf{x}_n) d\mathbf{x}_n \quad (9) \\
 & \ln \left\{ \frac{1}{n} \sum_{i=1}^n e^{-[H_{s+1}(\mathbf{x}_i) - H_s(\mathbf{x}_i)]} \right\} \\
 & \ln \left\{ \frac{1}{n} \sum_{i=1}^n e^{-[H_{s-1}(\mathbf{x}_i) - H_s(\mathbf{x}_i)]} \right\}. \quad (10)
 \end{aligned}$$

As in the derivation of Ref. [25], the Hamiltonians are now shifted by a constant offset  $C_s$ , i.e.,  $H'_s(\mathbf{x}) = H_s(\mathbf{x}) - C_s$ . This offset will cancel out for a given shape of an intermediate when calculating the accumulated free-energy difference in Eq. (6). However, as the intermediate states will turn out

to be coupled, these offsets do influence the shape of these intermediates. The offsets can now be chosen such that the terms inside the logarithms of Eq. (10) are close to one. In this case,  $\mathbb{E}[\Delta G_{s' \rightarrow (s+1)'}^{(n)}] = \Delta G_{s',(s+1)'}$  [25], and, therefore, the two linear terms arising from Eq. (10) can be expressed in terms of the exact free-energy differences.

Next, the product of the two sums in Eq. (10) is split into terms based on the same and different sample points, respectively,

$$\begin{aligned}
 & \mathbb{E}[\Delta G_{s' \rightarrow (s+1)'}^{(n)} \Delta G_{s' \rightarrow (s-1)'}^{(n)}] \\
 & = -\frac{1}{n^2} \int p_s(\mathbf{x}_1) d\mathbf{x}_1 \dots \int p_s(\mathbf{x}_n) d\mathbf{x}_n \quad (11) \\
 & \left( \left\{ \sum_{i=1}^n e^{-[H'_{s+1}(\mathbf{x}_i) - H'_s(\mathbf{x}_i)]} \right\} \left\{ \sum_{\substack{j=1 \\ j \neq i}}^n e^{-[H'_{s-1}(\mathbf{x}_j) - H'_s(\mathbf{x}_j)]} \right\} \right) \\
 & + \sum_{i=1}^n e^{-H'_{s+1}(\mathbf{x}_i) - H'_{s-1}(\mathbf{x}_i) + 2H'_s(\mathbf{x}_i)} \\
 & + f_{s'}(\Delta G_{s' \rightarrow (s-1)'}, \Delta G_{s' \rightarrow (s+1)'}), \quad (12)
 \end{aligned}$$

where the terms that can be expressed solely based on (constant) free-energy differences are summarized by the term  $f_s$ . Again, the first two terms of Eq. (12) can be expressed in terms of the free-energy differences between  $s$  and  $s+1$  as well as between  $s$  and  $s-1$ , respectively.

Collecting all terms arising from Eq. (7),

$$\begin{aligned}
 & \text{MSE}(\Delta G_{1,N}^{(n)}) \\
 & = \sum_{\substack{s=2 \\ s \text{ odd}}}^{N-2} \frac{1}{n} \left\{ \int p_s(\mathbf{x}) d\mathbf{x} e^{-2[H'_{s+1}(\mathbf{x}) - H'_s(\mathbf{x})]} \right. \\
 & + \int p_{s+2}(\mathbf{x}) d\mathbf{x} e^{-2[H'_{s+1}(\mathbf{x}) - H'_{s+2}(\mathbf{x})]} \\
 & + \int p_{s+1}(\mathbf{x}) d\mathbf{x} e^{-H'_{s+2}(\mathbf{x}) - H'_s(\mathbf{x}) + 2H_{s+1}(\mathbf{x})} \\
 & \left. + g'_s(\Delta G_{s',(s+1)'}, \Delta G_{(s+2)',(s+1)'}, \Delta G_{1',N'}) \right\}, \quad (13)
 \end{aligned}$$

where the function  $g'_s$  serves the same purpose as  $f'_s$  and can be dropped in the optimization below.

The condition of small  $\Delta G_{s' \rightarrow (s+1)'}$  is fulfilled by setting  $C_s = -\ln Z_s$ . By variation of the MSE from Eq. (13),

$$\frac{\partial}{\partial H_s(\mathbf{x})} \left\{ \text{MSE}(\Delta G_{1,N}^{(n)}) + \nu \int [e^{-H_s(\mathbf{x})} - Z_s] d\mathbf{x} \right\} \stackrel{!}{=} 0, \quad (14)$$

where  $\nu$  is a Lagrange multiplier, the optimal sequence of Hamiltonians is obtained. For  $s$  even, we obtain

$$\begin{aligned}
 H_s(\mathbf{x}) & = -\frac{1}{2} \ln \left[ e^{-2H_{s-1}(\mathbf{x})} r_{s-1,s}^{-2} + e^{-2H_{s+1}(\mathbf{x})} r_{s+1,s}^{-2} \right. \\
 & \left. - 2e^{-H_{s-1}(\mathbf{x}) - H_{s+1}(\mathbf{x})} r_{s-1,s}^{-1} r_{s+1,s}^{-1} \right] \quad (15)
 \end{aligned}$$

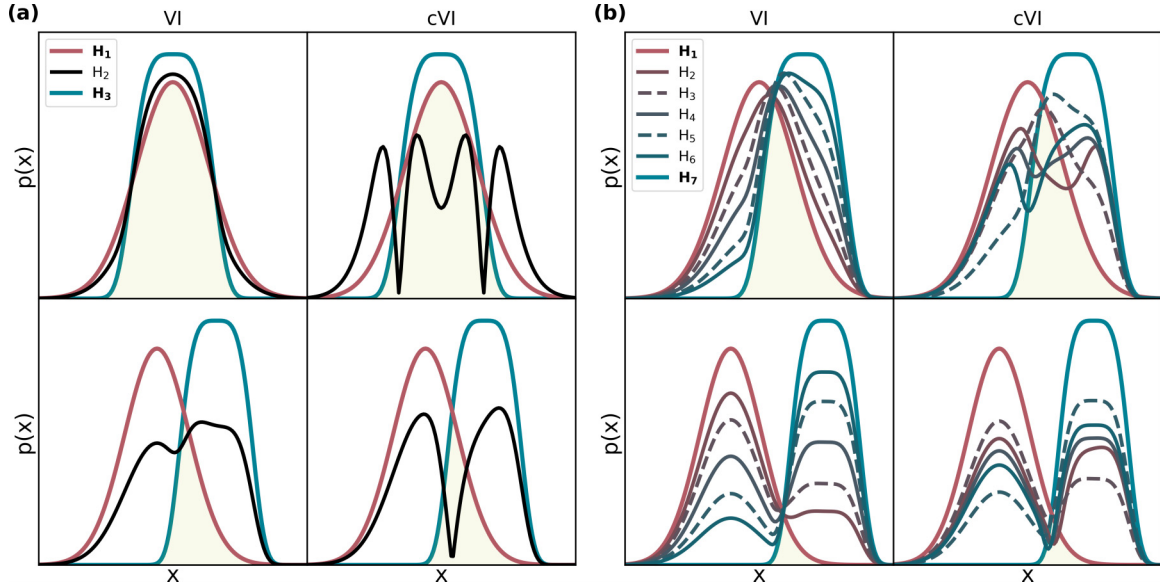


FIG. 2. Configuration space densities of VI (left column) and cVI (right column) for (a)  $N = 3$  and (b)  $N = 7$  states. The individual rows show different shifts in  $x$  direction between the minima of the harmonic,  $H_1(\mathbf{x})$  (red toward the left in each panel), and the quartic,  $H_N(\mathbf{x})$  (blue toward the right), potentials of the end states, thereby showing setups with different configuration space density overlap  $K$  between the end states, indicated by the yellow (light) area. Sampling is conducted in the even numbered intermediates. The dashed lines in (b) indicate the (odd numbered) virtual intermediate target states that no sampling is conducted in.

For  $s$  odd and  $2 < s < N - 1$ :

$$H_s(\mathbf{x}) = \ln[e^{H_{s-1}(\mathbf{x})}r_{s-1,s} + e^{H_{s+1}(\mathbf{x})}r_{s+1,s}] - \ln[e^{-H_{s-2}(\mathbf{x})+H_{s-1}(\mathbf{x})}r_{s-1,s-2} + e^{-H_{s+2}(\mathbf{x})+H_{s+1}(\mathbf{x})}r_{s+1,s+2}], \quad (16)$$

where, as in Eqs. (2) and (3), the ratios  $r_{s,t}$  of the partition sums between states  $s$  and  $t$  have to be determined iteratively. The above sequence, Eqs. (15) and (16), that we refer to as the correlated variational intermediates (cVI), yield the minimal MSE estimates for cFEP.

Figure 2 shows the resulting configuration space densities of the above intermediates for the example of a start state with a harmonic Hamiltonian,  $H_1(\mathbf{x}) = \frac{1}{2}x^2$ , and an end state with a quartic one,  $H_N(\mathbf{x}) = (x - x_0)^4$ . Figure 2(a) shows the VI that are optimal for the regular FEP scheme in Fig. 1(a). Figure 2(b) shows the cVI, optimal for cFEP.

The yellow (light) areas in Fig. 2, Eq. (17), provide a simple measure of the configuration space density overlap  $K$  between the end states 1 and  $N$ ,

$$K = \int_{-\infty}^{+\infty} d\mathbf{x} \min(p_A(\mathbf{x}), p_B(\mathbf{x})), \quad (17)$$

Here  $K = 0$  indicates two separate distributions without any overlap, and  $K = 1$  full overlap, i.e., identical configuration space densities.

The two rows in Figs. 2(a) and 2(b) depict the result for two different values of  $x_0$ , and correspondingly, varying  $K$ .

As can be inferred from Eq. (15), for  $N = 3$ ,  $H_2(\mathbf{x})$  diverges at the points where  $p_1(\mathbf{x}) = p_3(\mathbf{x})$ , and therefore  $p_2(\mathbf{x}) = 0$  at these points, as can also be seen for the intermediate sampling state shown in Fig. 2(a). More generally,  $H_2(\mathbf{x})$  of cVI “directs” sampling away from the overlap regions and toward the ones that are only relevant for one but not both end

states. For instance, the tails of the start state in the upper row of Fig. 2(a) are sampled more for cVI than for VI. For larger horizontal shifts of  $x_0$ , i.e., low values of  $K$ , the two variants become increasingly similar, as the additional term in Eq. (15) with respect to Eq. (2) becomes smaller compared to the first term.

For  $N = 7$  states, Fig. 2(b) shows the converged resulting configuration space densities. The case of  $x_0 = 0$ , as shown in Fig. 2(a), was omitted in Fig. 2(b) as the visualization is more difficult in this case due to the higher number of states. In Fig. 2(b), the additional changes from VI to cVI become more complex. As in Fig. 2(a), the sampling states have smaller densities  $p(\mathbf{x})$  in the overlap regions of the end states, but, in contrast to Fig. 2(a), still differ between VI and cVI for smaller values of overlap  $K$ . The reason is that while the overlap between the end states vanishes with decreasing  $K$ , an overlap between adjacent intermediate states remains that affects the shape of the intermediates. Note that the divergences mentioned above introduce instabilities in solving the system of Eqs. (15) and (16). Hence, for  $N > 3$  the factor 2 of the additional term in the logarithm Eq. (15) has been replaced by a factor  $\kappa$  that was set to slightly below 2 ( $\kappa = 1.95$ ) in case of Fig. 2(b). See Appendix for details.

### cBAR Estimator

As mentioned above, using the Zwanzig formula [16] to evaluate the free-energy difference between two sampling states with respect to the virtual intermediate, Eq. (3), of VI is equivalent to BAR [25,34]. Correspondingly, the virtual intermediate defined by Eq. (16) of cVI also corresponds to an estimator that is optimal for the sampling states of cFEP and that we will refer to as correlated BAR (cBAR).

To derive cBAR, we use the relation between the two approaches. Determining the free-energy difference between two sampling states labeled  $s - 1$  and  $s + 1$  by using the virtual intermediate  $s$  to evaluate the difference between the adjacent states yields

$$\Delta G_{s-1,s+1}^{(n)} = -\ln \frac{\langle e^{-[H_s(\mathbf{x})-H_{s+1}(\mathbf{x})]} \rangle_{s+1}}{\langle e^{-[H_s(\mathbf{x})-H_{s-1}(\mathbf{x})]} \rangle_{s-1}}. \quad (18)$$

Using the approach of Bennett [23] instead,

$$\Delta G_{s-1,s+1}^{(n)} = \ln \frac{\langle w(H_{s-1}(\mathbf{x}), H_{s+1}(\mathbf{x})) e^{-H_{s-1}(\mathbf{x})} \rangle_{s+1}}{\langle w(H_{s-1}(\mathbf{x}), H_{s+1}(\mathbf{x})) e^{-H_{s+1}(\mathbf{x})} \rangle_{s-1}}. \quad (19)$$

where  $w(H_{s-1}(\mathbf{x}), H_{s+1}(\mathbf{x}))$  is a weighting function. From Eqs. (21) and (19) follows that the two approaches are equivalent if the weighting function relates to the Hamiltonian of the virtual intermediate state through

$$w(H_{s-1}(\mathbf{x}), H_{s+1}(\mathbf{x})) = e^{-H_s(\mathbf{x})+H_{s-1}(\mathbf{x})+H_{s+1}(\mathbf{x})}. \quad (20)$$

Therefore, any Hamiltonian of a virtual intermediate state corresponds to a weighting function. Bennett optimized the weighting function with respect to the variance yielding the famous BAR result

$$\Delta G_{s-1,s+1}^{(n)} - C = \ln \frac{\langle f[H_{s-1}(\mathbf{x}) - H_{s+1}(\mathbf{x}) - C] \rangle_{s+1}}{\langle f[H_{s+1}(\mathbf{x}) - H_{s-1}(\mathbf{x}) + C] \rangle_{s-1}}, \quad (21)$$

where  $C \approx \Delta G_{s-1,s+1}$  has to be determined iteratively and  $f(x)$  is the Fermi function. This result is equivalent to using the virtual intermediate of Eq. (3) with Eq. (18). Note that the relation of a virtual intermediate to BAR result had already been obtained by Lu *et al.* [34], albeit through a different formalism, and that using the hyperbolic secant function [Eq. (10), p. 2980] in their overlap sampling approach [34,39] is equivalent to Eq. (20).

Next, for cFEP, using the Hamiltonian of the virtual intermediate from Eq. (16) in Eq. (20) yields the weighting function of cBAR,

$$\begin{aligned} & w(H_{s-2}(\mathbf{x}), H_{s-1}(\mathbf{x}), H_{s+1}(\mathbf{x}), H_{s+2}(\mathbf{x}), \\ & C_{s-2,s-1}, C_{s-1,s+1}, C_{s+1,s+2}) \\ & = [e^{-H_{s-2}(\mathbf{x})+H_{s-1}(\mathbf{x})+C_{s-2,s-1}} \\ & e^{-H_{s+2}(\mathbf{x})+H_{s+1}(\mathbf{x})+C_{s+1,s+2}} / \\ & [e^{H_{s-1}(\mathbf{x})-H_{s+1}(\mathbf{x})-C_{s-1,s+1}} + 1]], \end{aligned} \quad (22)$$

where the MSE of the resulting estimates is minimal if all  $C_{s,t} \approx \Delta G_{s,t}$ . A numerator of 1 in Eq. (22) would yield the original BAR result.

Note that  $H_{s-2}(\mathbf{x})$ , and  $H_{s+2}(\mathbf{x})$ , are also virtual intermediates determined by Eq. (16). As such, the result is a system of weighting functions, i.e., one for every pair of adjacent sampling states. The optimal estimate can, therefore, only be found by iteratively solving for the free-energy estimates between all sampling states at once. In this regard, the procedure is similar to MBAR [24].

### III. TEST SIMULATIONS

To assess to what extent our variational scheme improves accuracy, we consider the one-dimensional system with a harmonic and a quartic end state shown in Fig. 2. Rejection

sampling is used to obtain uncorrelated sample points. The free-energy estimate, obtained from these finite sample sets, is compared to the exact free-energy difference. The MSE, Eq. (5), is then calculated by averaging over one million of such realizations. With this procedure, different combinations of overlap  $K$ , numbers of states  $N$ , and sample points  $n$  are considered.

We compare three variants. First, using VI, Eqs. (2) and (3), with FEP, i.e., the scheme in Fig. 1(a). Here the estimates to both adjacent states are based on separate sample sets and, therefore, not correlated. Second, also using VI but now with cFEP, shown in Fig. 1(b). In contrast to variant 1, these estimates are based on the same sample sets and, therefore, correlated. In order to keep the total computational effort constant, the number of sample points per set (i.e., per yellow point in Fig. 1) is two times larger for cFEP than for FEP. Third, using cVI, Eqs. (15) and (16), that accounts for these correlations, also with cFEP.

### IV. RESULTS

For  $N = 3$  states, Fig. 3(a) shows the MSEs of the three variants for different numbers of sample points. Here, for the quartic end state,  $x_0 = 0$ , corresponding to  $K = 0.85$ , was used. The corresponding configuration space densities of VI and cVI are shown in the upper row of Fig. 2(a).

As can be seen, cVI with cFEP, shown by the dark blue (lower) line, yields the best MSE for all numbers of sample points except very few ones. The other two variants, i.e., VI with FEP (dashed green line) and cFEP (red upper line) yield very similar MSEs. As such, the gain in information from evaluating the Hamiltonians to both adjacent states for all sample points yields only a very small improvement compared to using separate sample sets for this purpose.

In order to quantify the improvement of cVI compared to VI for cFEP, Fig. 3(b) shows the ratio of the MSEs of the two variants, again in relation to the number of sample points per set. The dark orange (upper) curve ( $K = 0.85$ ), corresponds to the MSEs shown in (a) (i.e., the values of the red curve divided by the blue curve). The improvement in the MSE plateaus slightly above two for more than two hundred sample points per state. In addition, the improvements for setups with different overlap  $K$  between the end states are shown (orange to yellow). This improvement becomes smaller for smaller values of  $K$ , but the qualitative dependence on the number of sample points remains the same.

For a constant number of sample points  $n = 200$  (and  $n = 100$  per set for VI with FEP, shown by the dashed green line), Fig. 3(c) shows how the MSEs of the three variants improve with increasing  $K$ . The MSEs converge at low  $K$ , which is in agreement with the observation from Fig. 2(a) that the phase space densities of the intermediate state become more similar in this case.

Figure 3(d) shows the MSEs for  $N = 7$  states. The corresponding configuration space densities for two different values of  $K$  are shown in Fig. 2(b). Here VI with FEP and cFEP still yield similar MSEs, whereas cVI with cFEP, in contrast to  $N = 3$ , now yields the best MSE for all  $K$ . The improvement to VI ranges from around 20% for low  $K$ , to around 50% for large  $K$ . This is in line with the observation

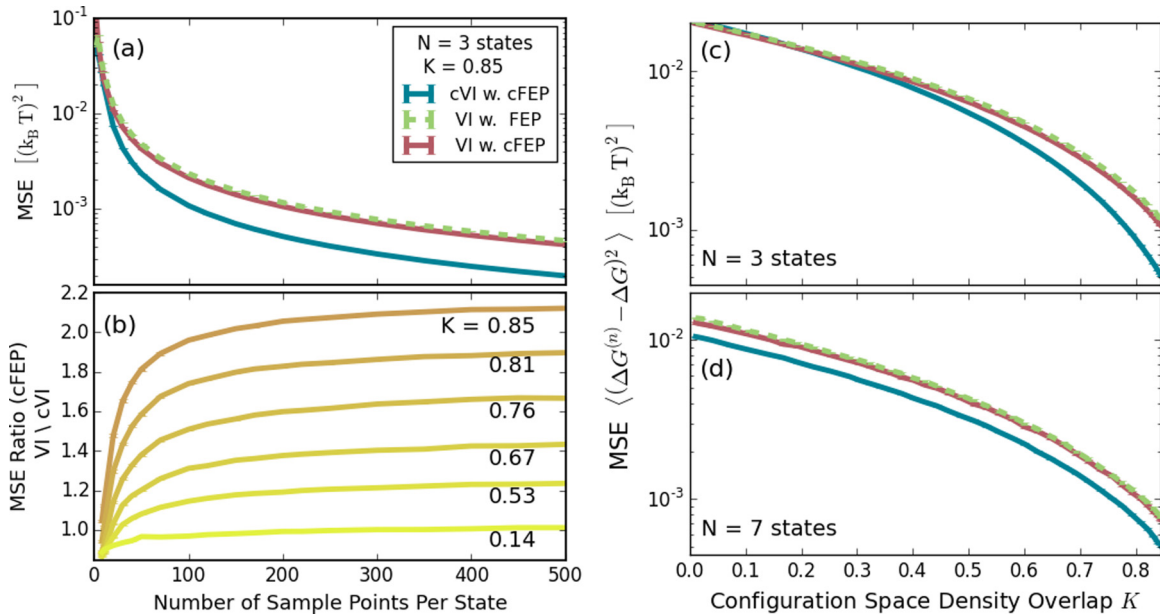


FIG. 3. Comparison of the accuracy of VI and cVI using the schemes of Fig. 1. The accuracies were obtained from test simulations based on the setups shown in Fig. 2. (a) Using  $N = 3$  states and comparing three variants of free-energy calculations: Using cVI with cFEP (blue lower solid line), VI with cFEP (red upper solid line) and VI with FEP (green dashed line). The MSEs of free-energy calculations are shown for different number of sample points. (b) The ratio of the MSEs, and, therefore, the improvement, of using cVI compared to VI for cFEP. The dark orange (upper) line ( $K = 0.85$ ) corresponds to the ratio between the red and the blue (solid) lines in (a). In addition, the results for different configuration space density overlaps  $K$  between the end states are shown (orange to yellow). (c) Using  $n = 200$  sample points, the MSEs of the three variants from (a) are shown over the full range of  $K$ . (d) As in (c), but with  $N = 7$  states. The computational effort was kept constant by reducing the number of sample points per state.

from Fig. 2(b) that the configuration space densities between VI and cVI become more similar but do not fully converge for a larger number of states in the limit of small  $K$ .

Last, the cBAR estimator can be used with any choice of intermediate states for cFEP. To assess how much the cBAR estimator improves the accuracy of free-energy estimates compared to BAR for cFEP, we conducted test simulations where the sampling states were chosen as in Eq. (1), i.e., by linear interpolation between the Hamiltonians of the end states. Test simulations were conducted at varying values of  $K$  and at  $N = 5$  and  $N = 7$ . Evaluating the MSE, we found a statistically significant improvement, however, only in the range of 1–2% (data therefore not shown here). The improvement was independent of  $K$  and similar for both numbers of  $N$ .

Considering that the MSEs of cVI and VI can improve up to an order of magnitude compared to the linear intermediates defined in Eq. (1) (for a detailed comparison between VI and linear intermediates, see Ref. [25]), the large majority of improvements is not due to an improved estimator but due to the way samples are generated.

## V. DISCUSSION AND CONCLUSION

In summary, we have derived a variant of variational intermediates (cVI) that yield the optimal free-energy estimate with minimal MSE when using the same sample points to evaluate the differences between the adjacent states above and below in the sequence (cFEP). This procedure is commonly used in free-energy simulations, as it is computationally much

cheaper to evaluate sample points at different Hamiltonians than to generate these. However, the resulting correlations between these estimates have not been considered yet.

Our test simulations for a one-dimensional Hamiltonian show that cVI with cFEP yields an improved MSE compared to the optimal sequence (VI) with FEP, i.e., using different sample points for estimates to states above and below in the sequence. For  $N = 3$  states, the first variant improved the MSE by more than a factor of two for end states with high configuration space density overlap  $K$ , whereas at low  $K$  the MSEs were similar. For  $N = 7$  states, the MSE improved between 20% (low  $K$ ) and 50% (large  $K$ ).

Interestingly, due to the correlations mentioned above, using VI with FEP yields only slightly worse MSEs for all  $K$  as using VI with cFEP, even though the latter involves twice as many evaluations of Hamiltonians from adjacent states. Only for cVI, thereby accounting for these correlations, the additional gain in information translates into a marked improvement of the MSE.

Similarly to most other theoretical analyses and derivations of free-energy calculation methods, we also needed to assume that all sample points within each intermediate state are uncorrelated. If atomistic simulations are used for sampling, then the resulting time-correlations reduce the number of essentially independent sample points. Unfortunately, for our one-dimensional systems, cVI increases barrier heights, thereby increasing correlation times. We have so far not tested our method on any complex biomolecular systems, so it is unclear if these barriers can be circumvented or what the expected increase in correlation times is. However, to avoid such

correlations between sample points in atomistic simulations, usually only a small subset of all sample points is used to calculate free-energy differences. Based on our findings and in contrast to common practice, we therefore recommend to use different subsets to evaluate the free-energy differences to different adjacent states.

The above derivation provides an example on how optimal intermediates and estimators with minimal MSE can be derived for different types of setups based on finite sampling that may help to incorporate a variety of assumptions and models into future theoretical approaches.

#### APPENDIX: AVOIDING NUMERICAL INSTABILITIES

The divergence in Eq. (15) at all  $\mathbf{x}$  for which

$$\begin{aligned} & e^{-2H_{s-1}(\mathbf{x})}r_{s-1,s}^{-2} + e^{-2H_{s+1}(\mathbf{x})}r_{s+1,s}^{-2} \\ &= 2e^{-H_{s-1}(\mathbf{x})-H_{s+1}(\mathbf{x})}r_{s-1,s}^{-1}r_{s+1,s}^{-1} \end{aligned} \quad (\text{A1})$$

causes numerical instabilities in solving the system of Eqs. (15) and (16). Replacing the factor 2 in Eq. (15) in the logarithm with a factor  $\kappa$ , i.e., for  $s$  even,

$$\begin{aligned} H_s(\mathbf{x}) = & -\frac{1}{2} \ln \left[ e^{-2H_{s-1}(\mathbf{x})}r_{s-1,s}^{-2} + e^{-2H_{s+1}(\mathbf{x})}r_{s+1,s}^{-2} \right. \\ & \left. - \kappa e^{-H_{s-1}(\mathbf{x})-H_{s+1}(\mathbf{x})}r_{s-1,s}^{-1}r_{s+1,s}^{-1} \right], \end{aligned} \quad (\text{A2})$$

and setting, e.g.,  $\kappa = 1.95$ , avoids these complications. As can be easily validated, the inside of the logarithm in Eq. (A2) is larger than zero for  $0 < \kappa < 2$  for all  $H_{s-1}(\mathbf{x})$  and  $H_{s+1}(\mathbf{x})$ . As shown for cVI in Fig. 2(b),  $\kappa < 2$  prevents  $p_s(\mathbf{x})$  to go to zero at the crossing points of  $p_{s-1}(\mathbf{x})$  and  $p_{s+1}(\mathbf{x})$  of the neighboring states but is still lowered at these points.

- 
- [1] C. Chipot and A. Pohorille (eds.), *Free Energy Calculations: Theory and Applications in Chemistry and Biology*, Springer Series in Chemical Physics, Vol. 86 (Springer, Berlin, 2007).
- [2] D. M. Zuckerman, Equilibrium sampling in biomolecular simulations, *Annu. Rev. Biophys.* **40**, 41 (2011).
- [3] R. Jinnouchi, F. Karsai, and G. Kresse, Making free-energy calculations routine: Combining first principles with machine learning, *Phys. Rev. B* **101**, 060201(R) (2020).
- [4] T. Sun, J. P. Brodholt, Y. Li, and L. Vočadlo, Melting properties from *ab initio* free energy calculations: Iron at the Earth's inner-core boundary, *Phys. Rev. B* **98**, 224301 (2018).
- [5] H. Ge and H. Qian, Mesoscopic kinetic basis of macroscopic chemical thermodynamics: A mathematical theory, *Phys. Rev. E* **94**, 052150 (2016).
- [6] P. V. Klimovich, M. R. Shirts, and D. L. Mobley, Guidelines for the analysis of free energy calculations, *J. Comput.-Aided Mol. Des.* **29**, 397 (2015).
- [7] C. D. Christ, A. E. Mark, and W. F. van Gunsteren, Basic ingredients of free energy calculations: A review, *J. Comput. Chem.* **31**, 1569 (2009).
- [8] C. D. Christ and T. Fox, Accuracy assessment and automation of free energy calculations for drug design, *J. Chem. Inf. Model.* **54**, 108 (2014).
- [9] M. De Vivo, M. Masetti, G. Bottegoni, and A. Cavalli, Role of molecular dynamics and related methods in drug discovery, *J. Med. Chem.* **59**, 4035 (2016).
- [10] Z. Cournia, B. Allen, and W. Sherman, Relative binding free energy calculations in drug discovery: Recent advances and practical considerations, *J. Chem. Inf. Model.* **57**, 2911 (2017).
- [11] B. J. Williams-Noonan, E. Yuriev, and D. K. Chalmers, Free energy methods in drug design: Prospects of “alchemical perturbation” in medicinal chemistry, *J. Med. Chem.* **61**, 638 (2018).
- [12] J. M. Rickman and R. LeSar, Free-energy calculations in materials research, *Annu. Rev. Mater. Res.* **32**, 195 (2002).
- [13] T. D. Swinburne and M.-C. Marinica, Unsupervised Calculation of Free Energy Barriers in Large Crystalline Systems, *Phys. Rev. Lett.* **120**, 135503 (2018).
- [14] R. Freitas, R. E. Rudd, M. Asta, and T. Frolov, Free energy of grain boundary phases: Atomistic calculations for  $\Sigma 5(310)[001]$  grain boundary in Cu, *Phys. Rev. Mater.* **2**, 093603 (2018).
- [15] M. de Koning, A. Antonelli, and S. Yip, Optimized Free-Energy Evaluation Using a Single Reversible-Scaling Simulation, *Phys. Rev. Lett.* **83**, 3973 (1999).
- [16] R. W. Zwanzig, High temperature equation of state by a perturbation method. I. Nonpolar gases, *J. Chem. Phys.* **22**, 1420 (1954).
- [17] J. G. Kirkwood, Statistical mechanics of fluid mixtures, *J. Chem. Phys.* **3**, 300 (1935).
- [18] D. K. Shenfeld, H. Xu, M. P. Eastwood, R. O. Dror, and D. E. Shaw, Minimizing thermodynamic length to select intermediate states for free-energy calculations and replica-exchange simulations, *Phys. Rev. E* **80**, 046705 (2009).
- [19] D. M. Zuckerman and T. B. Woolf, Theory of a Systematic Computational Error in Free Energy Differences, *Phys. Rev. Lett.* **89**, 180602 (2002).
- [20] D. M. Zuckerman and T. B. Woolf, Systematic finite-sampling inaccuracy in free energy differences and other nonlinear quantities, *J. Stat. Phys.* **114**, 1303 (2004).
- [21] T. T. Pham and M. R. Shirts, Identifying low variance pathways for free energy calculations of molecular transformations in solution phase, *J. Chem. Phys.* **135**, 034114 (2011).
- [22] A. Gelman and X.-I. Meng, Simulating normalizing constants: From importance sampling to bridge sampling to path sampling, *Stat. Sci.* **13**, 163 (1998).
- [23] C. H. Bennett, Efficient estimation of free energy differences from Monte Carlo data, *J. Comput. Phys.* **22**, 245 (1976).
- [24] M. R. Shirts and J. D. Chodera, Statistically optimal analysis of samples from multiple equilibrium states, *J. Chem. Phys.* **129**, 124105 (2008).
- [25] M. Reinhardt and H. Grubmüller, Determining free-energy differences through variationally derived intermediates, *J. Chem. Theory Comput.* **16**, 3504 (2020).
- [26] T. C. Beutler, A. E. Mark, R. C. van Schaik, P. R. Gerber, and W. F. van Gunsteren, Avoiding singularities and

- numerical instabilities in free energy calculations based on molecular simulations, *Chem. Phys. Lett.* **222**, 529 (1994).
- [27] M. Zacharias, T. P. Straatsma, and J. A. McCammon, Separation-shifted scaling, a new scaling method for Lennard-Jones interactions in thermodynamic integration, *J. Chem. Phys.* **100**, 9025 (1994).
- [28] T. Steinbrecher, D. L. Mobley, and D. A. Case, Nonlinear scaling schemes for Lennard-Jones interactions in free energy calculations, *J. Chem. Phys.* **127**, 214108 (2007).
- [29] C. D. Christ and W. F. van Gunsteren, Enveloping distribution sampling: A method to calculate free energy differences from a single simulation, *J. Chem. Phys.* **126**, 184110 (2007).
- [30] J. W. Perthold and C. Oostenbrink, Accelerated enveloping distribution sampling: Enabling sampling of multiple end states while preserving local energy minima, *J. Phys. Chem. B* **122**, 5030 (2018).
- [31] G. König, N. Glaser, B. Schroeder, A. Kubincová, P. H. Hünenberger, and S. Riniker, An alternative to conventional  $\lambda$ -intermediate states in alchemical free energy calculations:  $\lambda$ -enveloping distribution sampling, *J. Chem. Inf. Model.* (2020), doi: [10.1021/acs.jcim.0c00520](https://doi.org/10.1021/acs.jcim.0c00520).
- [32] A. Blondel, Ensemble variance in free energy calculations by thermodynamic integration: Theory, optimal “Alchemical” path, and practical solutions, *J. Comput. Chem.* **25**, 985 (2004).
- [33] T. T. Pham and M. R. Shirts, Optimal pairwise and non-pairwise alchemical pathways for free energy calculations of molecular transformation in solution phase, *J. Chem. Phys.* **136**, 124120 (2012).
- [34] N. Lu, J. K. Singh, and D. A. Kofke, Appropriate methods to combine forward and reverse free-energy perturbation averages, *J. Chem. Phys.* **118**, 2977 (2003).
- [35] I. Bilonis and P. Koutsourelakis, Free energy computations by minimization of Kullback-Leibler divergence: An efficient adaptive biasing potential method for sparse representations, *J. Comput. Phys.* **231**, 3849 (2012).
- [36] F. Weinhold, Metric geometry of equilibrium thermodynamics, *J. Chem. Phys.* **63**, 2479 (1975).
- [37] J. Gore, F. Ritort, and C. Bustamante, Bias and error in estimates of equilibrium free-energy differences from nonequilibrium measurements, *Proc. Natl. Acad. Sci. USA* **100**, 12564 (2003).
- [38] M. R. Shirts and V. S. Pande, Comparison of efficiency and bias of free energies computed by exponential averaging, the Bennett acceptance ratio, and thermodynamic integration, *J. Chem. Phys.* **122**, 144107 (2005).
- [39] N. Lu, D. Wu, T. B. Woolf, and D. A. Kofke, Using overlap and funnel sampling to obtain accurate free energies from nonequilibrium work measurements, *Phys. Rev. E* **69**, 057702 (2004).