

On the Calculation of Free Energies over Hamiltonian and Order Parameters via Perturbation and Thermodynamic Integration

Fernando A. Escobedo*

Robert Frederick Smith School of Chem. and Biomolecular Engineering, Cornell University, Ithaca, New York 14853, USA.

*Email: fe13@cornell.edu

ABSTRACT

In this work, complementary formulas are presented to compute free-energy differences via perturbation (FEP) methods and thermodynamic integration (TI). These formulas are derived by selecting only the most statistically significant data from the information extractable from the simulated points involved. On the one hand, commonly used FEP techniques based on overlap sampling leverage the full information contained in the overlapping macrostate probability distributions. On the other hand, conventional TI methods only use information on the first moments of those distributions, as embodied by the first derivatives of the free-energy. Since the accuracy of simulation data degrades considerably for high-order moments (for FEP) or free-energy derivatives (for TI), it is proposed to consider, consistently for both methods, data up to second order moments/derivatives. This provides a compromise between the limiting strategies embodied by common FEP and TI and leads to simple, optimized expressions to evaluate free-energy differences. The proposed formulas are validated with an analytically solvable harmonic Hamiltonian (for assessing systematic errors), an atomistic system (for computing a potential of mean force with coordinate dependent order parameter), and a binary-component coarse-grained model (for tracing a solid-liquid phase diagram in an ensemble sampled through alchemical transformations). It is shown that the proposed FEP and TI formulas are straightforward to implement, perform similarly well, and allow robust estimation of free-energy differences even when the spacing of successive points does not guarantee them to have proper overlapping in phase space.

I. INTRODUCTION

Numerous studies, book chapters, and monographs have been reported on the methodology for free-energy calculations via molecular simulations,¹⁻⁷ a research area that can be considered mature, with well-established formulas extensively and successfully deployed. Nonetheless, improvement of such methods is still an active area of research, as the effectiveness and ease of use of a given method tends to strongly depend on the characteristics of the system at hand.²⁻¹⁴ Two widely used methods to obtain free-energy differences between two states, and by extension over an arbitrary multi-state thermodynamic path, are the free-energy perturbation (FEP) and thermodynamic integration (TI)²⁻⁴ methods. FEP encompasses such methods as simple perturbations,^{1,2} overlap sampling like Bennett’s acceptance ratio method (BAR),¹⁵ weighed histogram analysis (or multihistogram reweighting) WHAM¹⁶⁻¹⁹ and multistate BAR (MBAR) methods,²⁰ and the closely related transition matrix methods.²¹⁻²³ The conceptual and practical connections between FEP and TI are well known and while they can successfully be deployed interchangeably for many systems, one of them is often the method of choice for specific classes of problems.^{2,5,6} For example, in simulating the free-energy of ordered assemblies, overlap sampling and TI were found to have roughly equivalent efficiency, while the MBAR method offered no advantage,¹⁴ while in the simulation of benchmark systems involving insertion/deletion of atomic sites and changing partial charges, the BAR and MBAR methods were found to be consistently among the top performers.²⁴ For concreteness, we will not consider non-equilibrium work based methods²⁵⁻³⁰ despite their connection to FEP and TI methods.

Free-energy calculations can also be classified based on the type of independent variable(s) “ λ ” used to define the different states. These variables can be either (i) Hamiltonian and state parameters (HPs) if they appear as parameters in the ensemble Boltzmann’s factor and hence are trivial to set, or (ii) order parameters (OPs) or “reaction coordinates” if they depend on (a subset of) the coordinates (or phase space) of the system and are generally more difficult to fix.² In

general, free-energy calculations based on HPs are more straightforward to implement with both TI and FEP, while those using OPs often require specialized methods or combination of methods, such as umbrella sampling³¹ with WHAM when obtaining free-energy profiles or surfaces.^{18,19}

The most effective FEP methods rely on overlap sampling between neighboring states; i.e., if states A and B are simulated, then important regions of the phase space of state A are also sampled by state B and vice versa.^{9,10,32} It is well-known that FEP overlap methods can fail if the target and reference states do not share a sufficiently large set of relevant configurations; i.e., for large systems.¹³ When simulating along multiple state points, it is customary to employ statistics of all such points leveraging the collected distribution of microstates or of “work” (virtual transitions to neighboring states) henceforth to be denoted as the “ Π distribution”. The best-known and widely used method to implement this calculation is BAR for the two-state case,¹⁵ and MBAR²⁰ or equivalently, WHAM,¹⁶⁻¹⁹ for the many-states case. However, if states A and B are nearest neighbors in phase space, then in computing the free-energy difference between them, using statistics from more distant states may not be worth the extra effort, especially if those other states only overlap on the less relevant and noisier “outer” tails of the A and B Π -distributions.³³⁻

36

Likewise, when using TI it has been customary to use data from multiple points, but from first-order derivatives of the free-energy only, to evaluate the integrand. Indeed, as per typical formulas like the Simpson’s rule or Gaussian quadratures,³⁷ polynomial functions are fitted to such first-derivative data. More than 2 points are needed to capture the curvature of the integrand function which helps not only with the integration (and interpolation process involved) but also with any needed extrapolation,⁷ which is crucial when advancing λ in a particular direction as in TI intended to stepwise trace a phase coexistence line,³⁸⁻⁴² referred to as Gibbs-Duhem integrations “GDI”. However, fitting numerous points to a function is not beneficial for stiff integrands, and it may place too much weight on points distant from the interval of interest, or unduly smooth-out

local features of the free-energy function, thus reducing the nimbleness of the extrapolating function to adapt to quick changes of slope. In this work, we advocate the use of simulation data from only those two points, say A and B , constituting the bounds of a single TI step, when evaluating the integral between them, with the tenet that unless one is using very narrowly spaced λ values, then other distant points are unlikely to provide essential information that is not already contained in points A and B . To compensate for the lack of integrand curvature information when using only such two points, it is proposed to use simulation data on second derivatives of the free energy,⁴² which are typically as readily accessible as first derivatives with no or minimal computation overhead.⁴¹⁻⁴³ We show that using such 2nd order TI brings the two-point integration formula closer to FEP overlap methods. We disregard higher than 2nd order derivatives since in practice the accuracy of the simulated data degrades for such high order effects. While this strategy has been already embodied by the FENEX method⁴² – a variant of GDI - here we present additional optimized formulas for polynomial fitting and in a general context for broader applicability.

We also show that FEP formulas that only employ up to 2nd Π -moments data (related to 2nd order free-energy derivatives) are easier to use than, e.g., BAR, which requires significantly more bookkeeping and a numerical procedure to solve an integral equation (with numerical issues compounding for WHAM⁴⁴). Further, such 2nd-order FEP methods do not rely on overlapping neighboring Π 's to work as intended nor in properly sampling the Π distribution tails. Indeed, considering that even first moment Π data can exhibit significant statistical and systematic errors (i.e., due to ergodic issues), it would seem unwarranted to assume that using the full Π distribution will invariably lead to a significant gain in robustness or accuracy in the calculated free-energies. Cumulant expansions of the free-energy, which when truncated to second order correspond to an underlying Gaussian Π , have been successfully used in the past to derive perturbation approximations,² e.g., the Born's formula for the free-energy of an ion in a liquid.⁴⁵ Most recently, 2nd order expansions of Jarzynski's equality have been deployed for non-equilibrium simulation

methods;²⁷⁻²⁹ e.g., for steering simulations where it was shown that for stiff springs, the work is Gaussian-distributed regardless of the speed of the process simulated.²⁷

It is important to keep in mind that in optimizing the statistical efficiency of free-energy estimations (e.g., see Sec. A of Appendix), one should consider the interplay between a “breadth” strategy that uses more points and finer staging, and a “depth” strategy that uses fewer points but leverages as much information from each point as possible. Indeed, a fixed computational budget could be equally allocated by either shorter but more numerous runs using the most accurate first-order data only, or by longer but fewer runs using higher order data. Regardless of number of points, however, the most elementary calculation involves two “successive” points and hence optimizing such calculations should always pay off, help strike a balance between breadth and depth, and provide a cornerstone to complement multi-state or higher-order methods.

Machine Learning (ML) methods⁴⁶ have been used to aid countless problems seeking to unveil complex correlations between variables, including the construction and prediction of free-energy surfaces from simulation data.⁴⁷⁻⁵² Indeed, ML models can be trained on datasets generated by molecular simulations such as those obtained from TI, FEP, WHAM and other similar approaches, as a means to create “black-box” correlations that can be used to efficiently interpolate data or extrapolate results to unexplored conditions. Such data-driven methods, however, are arguably most indicated in data-rich situations,⁵³ i.e., once extensive free-energy data has already been generated. In contrast, the strategies discussed here focus on making the most out of limited data; they are relevant in cases where little is known about the free-energy landscape of a system and where only a limited cross-section is of interest. Hence, while ML could still be used in such data-scarce situations, it would be a most fruitful complement at later stages in the analysis.

The rest of paper is organized as follows. In Section II we summarize the main definitions and overarching strategy to be followed. In Sections III and IV (plus the Appendix) we describe the proposed formulas developed for FEP and TI, respectively. In Section V we present three sample

applications of the proposed FEP and TI formulas and in Section VI we provide concluding remarks.

II. BASIC DEFINITIONS

The independent variable defining the path over which free-energy calculations are performed can be classified as being either: (i) An explicit parameter in the ensembles Boltzmann’s factor, which may be a thermodynamic field or a force-field parameter and (ii) a collective property that depends on (a subset of) the atomic coordinates. Henceforth we will refer to the former variables as “Hamiltonian parameters” or HPs and the latter variables as “order parameters” or OPs. The value of a HP or its change can be specified and enacted in a straightforward way and is associated with a change in either the thermodynamic state of the system or alchemical transformations. In contrast, OPs can only be indirectly controlled (e.g., by a biasing potential) and their changes typically occur within the same thermodynamic state. We will use symbol X to denote an OP and f for a HP. The boldfaced symbols $\mathbf{f}$ and $\mathbf{X}$ indicate “vectors” containing multiple parameters; i.e., $\mathbf{f} = \{f_1, f_2, \dots\}$ and $\mathbf{X} = \{X_1, X_2, \dots\}$.

We stress that our usage of the term *Hamiltonian* in HPs is broader than simply referring to the total energy of a system and pertains to a generalized Hamiltonian function appearing inside the exponential terms of the partition function [as per Eq. (1) below]. Examples of HPs include state parameters like temperature, pressure, chemical potentials, chemical potential differences between species, volume, and number of particles (changes in the latter often used to compute chemical potentials), but also parameters of the “force field” like energy well depths, site diameters, charges, force constants, etc., whose changes effectively enact alchemical transformations or mutations. Examples of OPs include the distances between atoms or multiatom groups (often used to compute binding free energies and potentials of mean force), torsional angles, potential energy or subcomponents thereof, root mean square deviations from reference

structures, and structural metrics of local order such the largest cluster of an incipient phase (appropriate for computing free-energy nucleation barriers).

The configurational part of the partition function for an isothermal ensemble can be written as:

$$Q(\mathbf{f}) = \sum \exp[-\mathcal{H}(\mathbf{X}|\mathbf{f})] \quad (1)$$

where the sum is over configurations, and $\mathcal{H}$ is a dimensionless Hamiltonian, e.g., with energy units reduced by kT (where T is temperature and k is Boltzmann's constant). The associated dimensionless free energy Φ (also reduced by kT) is related to the partition function via

$$\Phi(\mathbf{f}) = -\ln Q(\mathbf{f}) \quad . \quad (2)$$

The generalized Hamiltonian corresponding to a particular point p will be denoted thus:

$$\mathcal{H}_p(\mathbf{X}) = \mathcal{H}(\mathbf{X}|\mathbf{f}_p)$$

For future reference, we write the special but important case of a “linear” Hamiltonian corresponding to:

$$\mathcal{H}(\mathbf{f}, \mathbf{X}) = \mathcal{H}_0 + \mathbf{f} \cdot \mathbf{X} = \mathcal{H}_0 + \sum_k f_k X_k \quad (3)$$

where the sum is over all relevant coupling parameters. Here the f 's would typically represent reduced thermodynamic fields or TI coupling parameters as in Einstein integrations³ for solids [where the modified Hamiltonian is often expressed as $\mathcal{H} = \sum_k f_k \mathcal{H}_k$ so that $\mathcal{H}_k$ would essentially be X_k in Eq. (3)], the X 's correspond to extensive fluctuating quantities in the ensemble, and changes in Φ relate to changes in f_1 and f_2 through the fundamental thermodynamic equation:

$$d\Phi = \sum_i X_i df_i \quad (4)$$

where:

$$\left(\frac{\partial \Phi}{\partial f_i}\right)_{f_j} = X_i \quad (5)$$

$$\left(\frac{\partial^2 \Phi}{\partial f_i^2}\right)_{f_j} = -\text{cov}(X_i, X_i), \quad \frac{\partial^2 \Phi}{\partial f_i \partial f_j} = -\text{cov}(X_i, X_j) \quad (6)$$

$$\text{where:} \quad \langle X_i \rangle \equiv X_i, \quad \text{cov}(X_i, X_j) \equiv \sigma_{ij}^2 = \langle X_i X_j \rangle - \langle X_i \rangle \langle X_j \rangle \quad (7)$$

where $\langle \rangle$ denote ensemble averages.

In general (not just for linear Hamiltonians), a probability density function associated with point ‘ p ’ having field $\mathbf{f}_p = \{f_{1p}, f_{2p}, \dots\}$ (in general a vector) is

$$\Pi_p(\mathbf{X}) = \Pi(\mathbf{X}|\mathbf{f}_p) \quad (8)$$

Π_p is often plotted as a normalized histogram of collected $\mathbf{X}$ data, but any discretization is immaterial. Such a Π_p can be characterized by its moments; in particular, the first two moments are the averages X_{ip} and covariances σ_{ijp}^2 as defined in relations (7).

A “Landau” free energy can also be associated with microstates corresponding to macrostate X , so that for fixed $\mathbf{f}$ and within an additive constant:

$$\Phi(\mathbf{X}|\mathbf{f}) = -\ln \Pi(\mathbf{X}|\mathbf{f}) \quad (9)$$

Or simply $\Phi(\mathbf{X}) = -\ln \Pi(\mathbf{X})$ where $\Pi(\mathbf{X})$ is the probability of observing configurations where $\mathbf{X}$ has a particular value. Note that Φ from Eq. (9) is not the same as that in Eq. (2), being distinguishable by the variable they are function of ($\mathbf{f}$ or $\mathbf{X}$), and are related, e.g., as:

$$\Phi(\mathbf{f}_B) - \Phi(\mathbf{f}_A) = \Phi(\mathbf{X}|\mathbf{f}_B) - \Phi(\mathbf{X}|\mathbf{f}_A) + \mathcal{H}(\mathbf{X}|\mathbf{f}_B) - \mathcal{H}(\mathbf{X}|\mathbf{f}_A) \quad (10)$$

As stated in the Introduction, the formulas to be derived in the following sections exploit the fact that the information content of the Π functions, if assumed to be Gaussian-like, can be suitably distilled into its first two moments which are also the data that are most accurately estimated in simulations and circumvent the need for bookkeeping histogram or transition data. As equations

(5)-(6) illustrate, the first two moments typically translate into information about the first and second derivatives of $\Phi(\mathbf{f})$. If the Π functions are Gaussians, then for a single X :

$$\Pi_p(X) \approx \frac{1}{(2\pi\sigma_p^2)^{1/2}} \exp\left(-\frac{(X-X_p)^2}{2\sigma_p^2}\right) \quad (11)$$

where $\sigma_p^2 = \langle X^2 \rangle_p - \langle X \rangle_p^2$; and for a two-dimensional function $\Pi_p(\mathbf{X}) = \Pi(X_1, X_2 | f_{1p}, f_{2p})$ is a bivariate Gaussian, i.e.,

$$\Pi_p(\mathbf{X}) \approx \frac{1}{2\pi(\sigma_{11p}^2 \sigma_{22p}^2 \theta)^{1/2}} \exp\left(-\frac{1}{\theta} \left[\frac{(X_1 - X_{1p})^2}{2\sigma_{11p}^2} + \frac{(X_2 - X_{2p})^2}{2\sigma_{22p}^2} - \frac{(X_1 - X_{1p})(X_2 - X_{2p})}{\sigma_{11p}^2 \sigma_{22p}^2 / \sigma_{12p}^2} \right]\right) \quad (11')$$

where σ_{ijp}^2 are covariances at point p [as defined in Eq. (7)] and $\theta = 1 - (\sigma_{12p}^2)^2 / (\sigma_{11p}^2 \sigma_{22p}^2)$.

In the following sections we will consider formulas to obtain free-energy differences $\Delta\Phi_{AB}$ between two states A and B using FEP and TI methods when both states are simulated. In these two methods the integration step is conducted in different order: in FEP one finds the finite difference of Φ (between states A and B) by directly evaluating an integrated form which entails instantaneous “switching” between the A and B Hamiltonians, while in TI one first considers the infinitesimal difference in Φ which is then integrated to “gradually” get the finite difference between A and B . Their connection can be made even more apparent by describing them as limiting cases of Non-Equilibrium Work “NEW” methods.^{2,25-27} In FEP we will exploit knowledge of the first two moments of the underlying Π distributions while in TI we will exploit the (largely equivalent) knowledge of the first and second order derivatives of Φ .

Figure 1 illustrates some of the connections between Φ, f, X and $\Pi(X)$ for the case of a one-dimensional linear Hamiltonian.

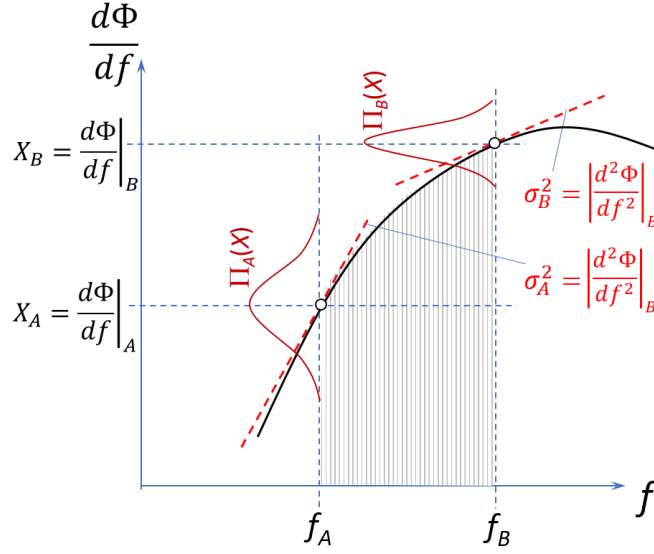


Fig. 1. Depiction of the connection between changes in free-energy Φ , Hamiltonian Parameter (HP) shown as the variable f (horizontal axis), and Order Parameter (OP) shown as the X variable (vertical axis) in the case of a linear Hamiltonian. $\Pi(X)$ distributions, whose variances are related to the slopes of the free-energy derivative curve, are also illustrated. $\Delta\Phi_{AB}$ appears here as the hatched area under the curve. Because of the different slopes marked by red dashed lines, the area of rectangle $(f_B - f_A) \times X_B$ is closer than $(f_B - f_A) \times X_A$ to the correct hatched area representing $\Delta\Phi_{AB}$.

III. FEP-BASED FORMULAS

3.1 Free energies over HPs

3.1.1 Optimized Simple Overlap Sampling (OSOS).

While there are multiple FEP-based working expressions for $\Delta\Phi_{AB} = \Phi_B - \Phi_A$, we consider first a basic formula that combines forward and reverse perturbations, related to the simple overlap sampling (SOS),^{1,2} wherein virtual perturbations do not occur between A and B but through an intermediate state with Hamiltonian:

$$\mathcal{H}_M = (1 - \alpha)\mathcal{H}_A + \alpha\mathcal{H}_B, \quad (12)$$

where α is an interpolating factor; in such a case it can be shown that the free-energy difference can be found from:

$$e^{-\Delta\Phi_{AB}} = \frac{Q_B}{Q_A} = \frac{\langle e^{-[\mathcal{H}_M - \mathcal{H}(\mathbf{X}, \mathbf{f}_A)]} \rangle_A}{\langle e^{-[\mathcal{H}_M - \mathcal{H}(\mathbf{X}, \mathbf{f}_B)]} \rangle_B} = \frac{\int d\mathbf{X} \Pi_A(\mathbf{X}) e^{-\alpha[\mathcal{H}(\mathbf{X}, \mathbf{f}_B) - \mathcal{H}(\mathbf{X}, \mathbf{f}_A)]}}{\int d\mathbf{X} \Pi_B(\mathbf{X}) e^{(1-\alpha)[\mathcal{H}(\mathbf{X}, \mathbf{f}_B) - \mathcal{H}(\mathbf{X}, \mathbf{f}_A)]}} \quad (13)$$

where the brackets denote ensemble average at the subscripted state. Equation (13) reduces to the standard SOS for a constant $\alpha = 1/2$. The variance in the free-energy estimation can be obtained from:^{3,15}

$$\sigma_{\Delta\Phi}^2 = \frac{1}{l_A} \left[\frac{\langle e^{-2\alpha[\mathcal{H}(\mathbf{X}, \mathbf{f}_B) - \mathcal{H}(\mathbf{X}, \mathbf{f}_A)]} \rangle_A}{\langle e^{-\alpha[\mathcal{H}(\mathbf{X}, \mathbf{f}_B) - \mathcal{H}(\mathbf{X}, \mathbf{f}_A)]} \rangle_A^2} - 1 \right] + \frac{1}{l_B} \left[\frac{\langle e^{2(1-\alpha)[\mathcal{H}(\mathbf{X}, \mathbf{f}_B) - \mathcal{H}(\mathbf{X}, \mathbf{f}_A)]} \rangle_B}{\langle e^{(1-\alpha)[\mathcal{H}(\mathbf{X}, \mathbf{f}_B) - \mathcal{H}(\mathbf{X}, \mathbf{f}_A)]} \rangle_B^2} - 1 \right] \quad (14)$$

where l_p is the number of statistically independent samples taken in state p ; we will henceforth assume that $l_A = l_B \equiv l$.

If α in Eqs. (13)-(14) is assumed to be a function of $\mathbf{X}$ [i.e., $\alpha = \alpha(\mathbf{X})$], then it can be shown that minimizing the variance with respect to α leads to the same result obtained by Bennett¹⁵ (the BAR formula shown later in Eq. (32) for $l_A = l_B$). We will consider instead the case that parameter α in Eqs. (12)-(14) is a constant (independent of $\mathbf{X}$); finding then the α value ($0 \leq \alpha \leq 1$) that minimizes $\sigma_{\Delta\Phi}^2$ in Eq. (14) is workable through numerical methods akin to those needed to implement BAR. The math simplifies greatly if the Π functions needed for the estimation of $\sigma_{\Delta\Phi}^2$ can be approximated by Gaussian distributions and $\mathcal{H}$ is a polynomial function of $\mathbf{X}$; in such a case the integrals in Eq. (14) are analytically solvable. Of course that Gaussian Π s will also simplify the evaluation of $\Delta\Phi_{AB}$ in Eq. (13), a trait that we also exploit. We will denote this approach as the “optimized SOS” or “OSOS” and will be exemplified in subsections 3.1.2 and 3.1.3.

3.1.2 Linear Hamiltonian

For the single-HP case Eq. (3) is $\mathcal{H}=\mathcal{H}_0+fX$ and the variance in the estimation of $\Delta\Phi_{AB}$ from Eq. (14) [with Gaussian Π as in Eq. (11)] simplifies to:

$$\sigma_{\Delta\Phi}^2 = \frac{1}{l} \left[e^{\sigma_A^2 \alpha^2 \Delta f^2} + e^{\sigma_B^2 (1-\alpha)^2 \Delta f^2} - 2 \right] \quad (15)$$

where $\Delta f = f_B - f_A$, and $\Delta f^2 = (\Delta f)^2$ and the optimal value of α found by setting $d\sigma_{\Delta\Phi}^2/d\alpha = 0$ satisfies:

$$(\sigma_A^2 \alpha^2 - \sigma_B^2 (1-\alpha)^2) \Delta f^2 = \ln \frac{(1-\alpha) \sigma_B^2}{\alpha \sigma_A^2} \quad (16)$$

which can be simply solved for α via a numerical procedure.

If the Gaussian Π approximation is also used in Eq. (13), then after taking logarithms the result can be expressed as:

$$\Delta\Phi_{AB} = [\alpha X_A + (1-\alpha) X_B] \Delta f + \frac{1}{2} [(1-\alpha)^2 \sigma_B^2 - \alpha^2 \sigma_A^2] \Delta f^2 \quad (17)$$

which can be interpreted as a TI polynomial formula, with the 1st and 2nd terms providing weighed contributions from the 1st and 2nd derivatives to Φ as per Eqs. (5) and (6), respectively. An alternative approach to find an optimal α value is to consider that the truncation error in $\Delta\Phi_{AB}$ can be (over) estimated as the second order term in Eq. (17). Indeed, based on the similarity of this result to the Euler-Maclaurin series expansion to be discussed later [see Eq. (58)], one can conjecture that more accurate estimates of $\Delta\Phi_{AB}$ would produce a polynomial with additional higher order terms of the form: prefactor $\times [(1-\alpha)^{2n} \sigma_B^{2n} - \alpha^{2n} \sigma_A^{2n}] \Delta f^{2n}$ with $n \geq 1$ (and Gaussian central moments proportional to powers of σ^2). In such a case, higher order corrections are rendered negligible by setting $(1-\alpha)^n \sigma_B^n = \alpha^n \sigma_A^n$, resulting in:

$$\alpha = \sigma_B / (\sigma_A + \sigma_B) \quad (18)$$

which, as a bonus, simplifies formula (17) by dropping the 2nd order term leading to the remarkably simple expression:

$$\Delta\Phi_{AB} = \frac{\sigma_B X_A + \sigma_A X_B}{\sigma_B + \sigma_A} \Delta f \quad (19)$$

Equations (16)-(19) embody the intuitive idea that since $\sigma_{\Delta\Phi}^2$ increases with the magnitude of the distribution variances, the intermediate Hamiltonian between points A and B should be apportioned weighing more conservatively the point with larger standard deviation/variance or, leveraging the relation (6), larger curvature. Figure 1 also illustrated this point geometrically, interpreting relations (17) and (19) as TI formulas. In our numerical tests, very little difference was found in the quality of the $\Delta\Phi_{AB}$ results when computing α from either (16)-(17) or (18)-(19), likely reflecting the presence of a broad minimum in the $\sigma_{\Delta\Phi}^2(\alpha)$ function. Our results are reminiscent but different from a formula suggested in Ref. [7] to optimally combine exponential (free-energy) averages from forward ($A \rightarrow B$) and backward ($B \rightarrow A$) FEP simulations.

The analysis above can be readily extended to the multidimensional linear Hamiltonian of Eq. (3); i.e., for the case of two X properties: $\mathcal{H} = \mathcal{H}_0 + f_1 X_1 + f_2 X_2$. Here we choose as intermediate Hamiltonian

$$\mathcal{H}_M = \mathcal{H}_0 + \sum_{i=1}^2 [(1 - \alpha_i) f_{iA} + \alpha_i f_{iB}] X_i, \quad (20)$$

where the α 's are adjustable factors (noting that we could choose $\alpha_1 = \alpha_2$). Hence,

$$e^{-\Delta\Phi_{AB}} = \frac{\langle e^{-(\mathcal{H}_M - \mathcal{H}_A)} \rangle_A}{\langle e^{-(\mathcal{H}_M - \mathcal{H}_B)} \rangle_B} = \frac{\int d\mathbf{X} \Pi_A(\mathbf{X}) e^{-\alpha_1 \Delta f_1 X_1 - \Delta f_2 X_2}}{\int d\mathbf{X} \Pi_B(\mathbf{X}) e^{(1-\alpha_1) \Delta f_1 X_1 + (1-\alpha_2) \Delta f_2 X_2}} \quad (21)$$

where $\Delta f_i = f_{iB} - f_{iA}$. Adopting the bivariate Gaussian of Eq. (11') for the Π 's, the integrals in Eq. (21) are solved analytically to yield:

$$\begin{aligned}\Delta\Phi_{AB} = & [\alpha_1 X_{1A} + (1 - \alpha_1) X_{1B}] \Delta f_1 + [\alpha_2 X_{2A} + (1 - \alpha_2) X_{2B}] \Delta f_2 + \frac{1}{2} [(1 - \alpha_1)^2 \sigma_{11B}^2 - \\ & \alpha_1^2 \sigma_{11A}^2] \Delta f_1^2 + \frac{1}{2} [(1 - \alpha_2)^2 \sigma_{22B}^2 - \alpha_2^2 \sigma_{22A}^2] \Delta f_2^2 + [(1 - \alpha_1)(1 - \alpha_2) \sigma_{12B}^2 - \alpha_1 \alpha_2 \sigma_{12A}^2] \Delta f_1 \Delta f_2\end{aligned}\quad (22)$$

Likewise, the associated variance in $\Delta\Phi_{AB}$ can be estimated from:

$$l\sigma_{\Delta\Phi}^2 = \exp(\alpha_1^2 \sigma_{11A}^2 \Delta f_1^2 + \alpha_2^2 \sigma_{22A}^2 \Delta f_2^2 + 2\alpha_1 \alpha_2 \sigma_{12A}^2 \Delta f_1 \Delta f_2) + \exp[(1 - \alpha_1)^2 \sigma_{11B}^2 \Delta f_1^2 + (1 - \alpha_2)^2 \sigma_{22B}^2 \Delta f_2^2 + 2(1 - \alpha_1)(1 - \alpha_2) \sigma_{12B}^2 \Delta f_1 \Delta f_2] - 2 \quad (23)$$

which could be minimized to find the optimal values of α_1 and α_2 using any suitable numerical procedure.³⁷ (This would simplify somewhat if $\alpha_1 = \alpha_2 = \alpha$ were appropriate, albeit having two parameters allows greater flexibility for minimizing $\sigma_{\Delta\Phi}^2$). As in the single- X case, an approximate solution for the α 's could be chosen so that the second order terms associated with Δf_1^2 and Δf_2^2 in Eq. (22) vanish:

$$\alpha_1 = \sigma_{11B} / (\sigma_{11A} + \sigma_{11B}), \quad \alpha_2 = \sigma_{22B} / (\sigma_{22A} + \sigma_{22B}) \quad (24)$$

$$\Delta\Phi_{AB} = \frac{\sigma_{11B} X_{1A} + \sigma_{11A} X_{1B}}{\sigma_{11B} + \sigma_{11A}} \Delta f_1 + \frac{\sigma_{22B} X_{2A} + \sigma_{22A} X_{2B}}{\sigma_{22B} + \sigma_{22A}} \Delta f_2 + \frac{\sigma_{11A} \sigma_{22A} \sigma_{12B}^2 + \sigma_{11B} \sigma_{22B} \sigma_{12A}^2}{(\sigma_{11B} + \sigma_{11A})(\sigma_{22B} + \sigma_{22A})} \Delta f_1 \Delta f_2 \quad (22')$$

The above equations can be straightforwardly extended to systems with more than 2 X variables.

3.1.3 Non-linear Hamiltonians

As illustration, we present results for a Hamiltonian of the form:

$$\mathcal{H}(X, f) = \mathcal{H}_0(X) + \frac{f_1}{2} (X - f_2)^2 \quad (25)$$

For which the exact solution to Eqs. (13)-(14) can be expressed as:

$$\Delta\Phi_{AB} = \frac{1}{2} \left(\frac{X_A^2}{\sigma_A^2} - \frac{X_B^2}{\sigma_B^2} + f_{1B} f_{2B}^2 - f_{1A} f_{2A}^2 + \ln \frac{\rho_A}{\rho_B} + w_B - w_A \right) \quad (26)$$

where:

$$w_A = \frac{1}{\rho_A} \left(\frac{X_A}{\sigma_A} + \alpha \sigma_A (f_{1B} f_{2B} - f_{1A} f_{2A}) \right)^2 \quad (27)$$

$$w_B = \frac{1}{\rho_B} \left(\frac{X_B}{\sigma_B} + (1 - \alpha) \sigma_B (f_{1A} f_{2A} - f_{1B} f_{2B}) \right)^2 \quad (28)$$

$$\rho_A = 1 + \alpha \sigma_A^2 \Delta f_1, \quad \rho_B = 1 - (1 - \alpha) \sigma_B^2 \Delta f_1 \quad (29)$$

Unlike the linear Hamiltonian case, a simple relation to find the optimal α is not generally accessible, but this simplifies greatly for two limiting cases:

- (i) If $f_{2A} = f_{2B} = f_2$, then this case could be mapped onto a linear Hamiltonian by redefining $(X - f_2)^2$ in Eq. (25) as X and henceforth using the solution of Sec. 3.1.2, and
- (ii) If $f_{1A} = f_{1B} = f_1$ then Eq. (26) simplifies to:

$$\Delta \Phi_{AB} = \left[\frac{f_{2A} + f_{2B}}{2} - \alpha X_A - (1 - \alpha) X_B \right] f_1 \Delta f_2 + \frac{1}{2} [(1 - \alpha)^2 \sigma_B^2 - \alpha^2 \sigma_A^2] f_1^2 \Delta f_2^2 \quad (30)$$

and expression (14) becomes:

$$\sigma_{\Delta \Phi}^2 = \frac{1}{l} \left[e^{\sigma_A^2 \alpha^2 f_1^2 \Delta f_2^2} + e^{\sigma_B^2 (1 - \alpha)^2 f_1^2 \Delta f_2^2} - 2 \right] \quad (31)$$

We should point out the similarity between Eqs. (17)+(15) (for linear Hamiltonian) and Eqs. (30)+(31) and that Eq. (30) would also be obtained for choices of the intermediate Hamiltonian other than Eq. (12), e.g., for $\mathcal{H}_M = \frac{1}{2} f_1 [X - (1 - \alpha) f_{2A} - \alpha f_{2B}]^2$. The optimal α can be found either as the value that minimizes the function in Eq. (31):

$$(\sigma_A^2 \alpha^2 - \sigma_B^2 (1 - \alpha)^2) f_1^2 \Delta f_2^2 = \ln \frac{(1 - \alpha) \sigma_B^2}{\alpha \sigma_A^2} \quad , \quad (31')$$

or, as in Sec. 3.1.2, by choosing α to vanish the 2nd order term in Eq. (30), which again results in formula (18) and simplifies Eq. (30) to:

$$\Delta\Phi_{AB} = \left[\frac{f_{2A}+f_{2B}}{2} - \frac{\sigma_B X_A + \sigma_A X_B}{\sigma_B + \sigma_A} \right] f_1 \Delta f_2 \quad (30')$$

3.1.4 Other perturbation methods

The Gaussian approximation for Π deployed for OSOS in 3.1.1-3.1.3 can also be deployed for other free-energy methods including WHAM and MBAR methods¹⁶⁻²⁰ for multiple points or for BAR¹⁵ for the two-point situation; in the latter case, this requires solution of

$$\int d\mathbf{X} \frac{\Pi_A(\mathbf{X})}{1+e^{\mathcal{H}(\mathbf{X},f_B)-\mathcal{H}(\mathbf{X},f_A)-\Delta\Phi_{AB}}} = \int d\mathbf{X} \frac{\Pi_B(\mathbf{X})}{1+e^{-\mathcal{H}(\mathbf{X},f_B)+\mathcal{H}(\mathbf{X},f_A)+\Delta\Phi_{AB}}} \quad (32)$$

Unfortunately, there is no analytical solution for the definite integral in Eq. (32) for Gaussian Π functions and hence finding $\Delta\Phi_{AB}$ requires the use of either: (i) numerical integration coupled to a root find method – a combination that will be henceforth referred as BAR-G = BAR-Gaussian method, or (ii) approximate expressions and series expansions that have been proposed for the related problem of Gauss-Fermi integrals for charge carrier problems⁵⁴ which also require numerical root finding. In our preliminary numerical tests, however, we found no distinct advantage of the latter formulas, in terms of precision or ease of use, relative to the equations presented in Sections 3.1.2-3.1.3 and henceforth will not be considered any further.

3.2 Free energies over OPs

Φ changes are significantly more involved when evaluated with respect to coordinate-dependent OPs. While different approaches can be followed, often involving the evaluation of forces,^{2,55} here we only consider umbrella sampling with harmonic potentials, a widely used and easy-to-implement method.^{18-20,33} The key idea is to introduce a (harmonic) bias in the Hamiltonian so that OP values are not fixed but targeted in an average sense. The formulas presented in Sec. 3.1.3 were precisely for a Hamiltonian having such a harmonic term, but now we make explicit that such a

term is an add-on to the original Hamiltonian and hence represents the *biasing function*. Formulas (25)-(31) still apply to obtain $\Delta\Phi$ for this case with 3 caveats:

- 1) The probability function Π to be used in FEP formulas like (13) should be the *biased* distributions since these are the ones expected to be approximately Gaussian, stemming from the strength of the harmonic bias. This is unlike what is customary with HP-based simulations discussed in Section 3.1 where the unbiased Π s were relevant and states A and B represented different thermodynamic states or chemistries, sampling of relevant microstates was assumed unconstrained, and the Gaussian nature of Π was associated with large system sizes. In the present case, however, we are typically simulating the same thermodynamic state, each simulation focusing the sampling over a particular region of the OP given that the unbiased simulation would not be able to sample all microstates of interest (e.g., due to the presence of regions with vastly different probability of occurrence).
- 2) Even though $\Delta\Phi_{AB}$ in formula (26) is associated with changes in HPs (i.e., f_1 and f_2), the physically relevant OP is the X variable in the harmonic function; hence $\Delta\Phi_{AB}$ must now be associated with changes of the average values of X (i.e., X_A and X_B).
- 3) Equation (26) must now be interpreted as a biased free-energy, $\Delta\Phi'$, the unbiased value $\Delta\Phi$ is obtained by subtracting the biasing functions associated with the *average* values of X_A and X_B , namely:

$$\Delta\Phi_{AB} \equiv \Phi(X_B) - \Phi(X_A) = \Delta\Phi'_{AB} - \frac{1}{2}f_{2,B}(X_B - f_{1,B})^2 + \frac{1}{2}f_{2,A}(X_A - f_{1,A})^2 \quad (33)$$

IV. Thermodynamic Integration

We will consider the possibility of integrating over multiple HPs or OPs and using λ as a general symbol for the integration variable (either f or X), the standard TI formula for an specified $A \rightarrow B$ integration path is:

$$\Delta\Phi_{AB} = \Phi_B - \Phi_A = \int_A^B \sum_i \frac{\partial\Phi}{\partial\lambda_i} d\lambda_i \quad (34)$$

where for simplicity of notation it is implicit that the partial derivative with respect to λ_i is while keeping constant all other f_j property and any other ensemble field, and for multiple λ s (as in Sec. V). Numerical integration of Eq. (34) requires knowledge of at least $\partial\Phi/\partial\lambda_i$ data but, following the guiding strategy of Section III, 2nd derivative data will be leveraged as well. In the following subsections we first describe how those derivatives can be obtained for HP and OP as TI variables.

4.1 TI: Derivatives with respect to Hamiltonian-dependent HP

From Eqs. (1) and (2), any first and second order partial derivative with respect to any HP can be written as:

$$\frac{\partial\Phi}{\partial f_i} = \left\langle \frac{\partial\mathcal{H}}{\partial f_i} \right\rangle \quad (35)$$

$$\frac{\partial^2\Phi}{\partial f_i^2} = \left\langle \frac{\partial^2\mathcal{H}}{\partial f_i^2} \right\rangle - \text{cov} \left(\frac{\partial\mathcal{H}}{\partial f_i}, \frac{\partial\mathcal{H}}{\partial f_i} \right) \quad (36)$$

$$\frac{\partial^2\Phi}{\partial f_i \partial f_j} = \left\langle \frac{\partial^2\mathcal{H}}{\partial f_i \partial f_j} \right\rangle - \text{cov} \left(\frac{\partial\mathcal{H}}{\partial f_i}, \frac{\partial\mathcal{H}}{\partial f_j} \right), \quad (37)$$

where it is implicit that the derivative with respect to any f_i is made for all other f_j kept constant with $j \neq i$, $\langle \rangle$ indicate ensemble averages, and covariances defined as in Eq. (7). Higher order derivatives can be similarly obtained. Equations (35)-(37) are indicated when the averages in their right-hand sides are measurable in simulation, i.e., the $\mathcal{H}$ derivatives are continuous functions, involving elementary operations over readily accessible microscopic quantities. Equations (35)-(37) simplify into formulas (5)-(6) in the case of the linear Hamiltonian of Eq. (3) where the f 's would typically represent dimensionless thermodynamic fields.

In some situations, however, the sought-after $\mathcal{H}$ derivative is not a continuous function, as is the case when the f_i parameter describes hard-core interactions⁶⁷ or it is difficult to evaluate analytically, as when rigid or Holonomic constraints are involved or if f_i is a property convoluted in the Hamiltonian (like volume) or an integer (like number of particles). In such cases, finite-size approximations are suitable, which can be derived by starting with the well-known expression for a simple FEP:

$$\Delta\Phi = -\ln\left\{\frac{Q(f_i+\Delta f_i)}{Q(f_i)}\right\} \approx -\ln\langle e^{-\Delta\mathcal{H}_i} \rangle \quad (38)$$

where $\Delta\mathcal{H}_i = \mathcal{H}(f_i + \Delta f_i) - \mathcal{H}(f_i)$ and the $\langle \rangle$ brackets indicate ensemble average evaluated at f_i .

For an infinitesimal change in f_i , i.e., $\Delta f_i \rightarrow \delta f_i$, then:

$$\frac{\partial\Phi}{\partial f_i} \approx \frac{\delta\Phi}{\delta f_i} = -\frac{1}{\delta f_i} \ln\langle e^{-\delta\mathcal{H}_i} \rangle \quad (39)$$

with $\delta\mathcal{H}_i = \mathcal{H}(f_i + \delta f_i) - \mathcal{H}(f_i)$. Likewise, we can derive:

$$\frac{\partial^2\Phi}{\partial f_i^2} \approx \frac{1}{\delta f_i} \left[\frac{\delta\Phi}{\delta f_i} \Big|_{f_i+\delta f_i} - \frac{\delta\Phi}{\delta f_i} \Big|_{f_i} \right] = -\frac{1}{(\delta f_i)^2} \ln \frac{\langle e^{-\delta^2\mathcal{H}_i} \rangle}{\langle e^{-\delta\mathcal{H}_i} \rangle^2} \quad (40)$$

where $\delta^2\mathcal{H}_i = \mathcal{H}(f_i + 2\delta f_i) - \mathcal{H}(f_i)$. Note that in principle one could estimate perturbations of $\mathcal{H}$ by either increasing or decreasing f_i ($\delta f_i > 0$ or < 0) or a combination of both, but we assume that one direction leads to more accurate estimation of $\delta\Phi$. Cross 2nd derivatives can be similarly evaluated, e.g., if finite differences are to be used for derivatives of f_i and f_j :

$$\frac{\partial^2\Phi}{\partial f_i \partial f_j} \approx -\frac{1}{\delta f_i \delta f_j} \ln \left\{ \frac{\langle e^{-\delta\mathcal{H}_{ij}} \rangle}{\langle e^{-\delta\mathcal{H}_i} \rangle \langle e^{-\delta\mathcal{H}_j} \rangle} \right\} \quad (41)$$

where $\delta\mathcal{H}_{ij} = \mathcal{H}(f_i + \delta f_i, f_j + \delta f_j) - \mathcal{H}(f_i, f_j)$, $\delta\mathcal{H}_i = \mathcal{H}(f_i + \delta f_i, f_j) - \mathcal{H}(f_i, f_j)$, and $\delta\mathcal{H}_j = \mathcal{H}(f_i, f_j + \delta f_j) - \mathcal{H}(f_i, f_j)$. If now we assume that the derivative of $\mathcal{H}$ with respect to f_j is analytical but that with respect to f_i should be done by finite difference, then it is easy to show that:

$$\frac{\partial^2 \Phi}{\partial f_i \partial f_j} \approx -\frac{1}{\delta f_i} \left[\frac{1}{\langle e^{-\delta \mathcal{H}_i} \rangle} \left\langle \frac{\partial \mathcal{H}(f_i + \delta f_i, f_j)}{\partial f_j} e^{-\delta \mathcal{H}_i} \right\rangle - \left\langle \frac{\partial \mathcal{H}(f_i, f_j)}{\partial f_j} \right\rangle \right] \quad (42)$$

while perturbation formulas like (39)-(42) entail an additional cost relative to evaluating simple averages and covariances like (5)-(6), they all can readily be computed on-the-fly during a given simulation. Note that chemical potentials, which could be evaluated using Eq. (38) for $\Delta f_i = \Delta N_i = 1$, more commonly require more gradual coupling and specialized approaches.^{3,11}

4.2 Derivatives with respect to Coordinate-dependent OP.

Evaluating first and second order derivatives of Φ with respect to coordinate-dependent OPs is more involved. Again, here we only consider the case when harmonic potentials are used to bias the sampling to allow targeting specific OP values on an average sense. In this way, the TI can be performed over the distinct average OP values. Recalling our definition of $\Phi(X)$ in Eq. (9), we can write:

$$\frac{d\Phi}{dX} = -\frac{1}{\Pi(X)} \frac{d}{dX} \Pi(X) \quad (43)$$

Assuming that the simulation was conducted by adding a biasing function $-\psi(X, \mathbf{f})$ to the Hamiltonian, then, the probability of observing different X -macrostates is, within a constant prefactor:

$$\Pi(X) \propto \Pi_\psi(X) e^{\psi(X, \mathbf{f})} \quad (44)$$

Substituting Eq. (44) into Eq. (43) and simplifying:

$$\frac{d\Phi}{dX} = -\frac{d\psi(X, \mathbf{f})}{dX} - \frac{1}{\Pi_\psi(X)} \frac{d}{dX} \Pi_\psi(X) \quad (45)$$

The second derivative is then:

$$\frac{d^2 \Phi}{dX^2} = -\frac{d^2 \psi(X, \mathbf{f})}{dX^2} - \frac{1}{\Pi_\psi(X)} \frac{d^2}{dX^2} \Pi_\psi(X) + \left[\frac{1}{\Pi_\psi(X)} \frac{d}{dX} \Pi_\psi(X) \right]^2 \quad (46)$$

Equations (45) and (46) are henceforth specialized for the choice:

$$\psi(X, \mathbf{f}) = \frac{1}{2}f_{1p}(X-f_{2p})^2 \quad (47)$$

where f_{1p} and f_{2p} are specified HPs, designed to bias the sampling of X to values near f_{2p} . For large enough f_{1p} , the Π_ψ distribution will be unimodal with a well-defined peak. Assuming Π_ψ to be approximately Gaussian, introducing Eq. (47) into Eqs. (45) and (46) evaluating them at $X = \langle X \rangle_p = X_p$:

$$\left. \frac{d\Phi}{dX} \right|_{X_p} = -f_{1p}(X_p - f_{2p}) \quad (48)$$

$$\left. \frac{d^2\Phi}{dX^2} \right|_{X_p} \approx -f_{1p} + \frac{1}{\sigma_{X_p}^2} = -f_{1p} + (\langle X^2 \rangle_p - \langle X \rangle_p^2)^{-1} \quad (49)$$

Equations (48) and (49) show how the required derivatives of Φ at a given point can be estimated from a biased simulation. The same results can be alternatively obtained if one assumes that within the relevant sampling region, the local free energy profile can be approximated by a 2nd order polynomial in X , say, $\Phi = a + bX + \frac{1}{2}cX^2$. In such a case and with both Φ and ψ being quadratic functions of X , we have that $\Pi_\psi(X) \propto \exp(-\Phi(X) - \psi(X, \mathbf{f}))$ will be Gaussian whose mean $\langle X \rangle$ and variance σ_X^2 as measured in the biased simulation can be related to the 1st and 2nd derivatives of Φ (i.e., to $b+cX$ and c according to the local model) to give Eqs. (48) and (49), respectively. Note that while one cannot precisely specify the value of X at which the derivatives are evaluated since $\langle X \rangle$ is not known a-priori, it will be rather close to the chosen f_{2p} value for large f_{1p} .

Formulas (48) and (49) are consistent with the so-called umbrella integration approach,³³⁻³⁶ but in such a case the free energy is reconstructed differently; i.e., by (i) describing first the free-energy derivatives as linear functions around each simulated point, (ii) combining these results using a weighing scheme that takes into account the sampling frequencies of X values (registered in binned histograms) from all simulated points, and (iii) using a suitable quadrature method to integrate the

weighed derivative averages from (ii). In contrast, the approach outlined here and in Section 4.3 directly uses first and second derivative data at each point to form piece-wise polynomial fits to extract free-energy differences.

4.3 Polynomial Approximations for TI

Polynomial approximations are particularly convenient to use for interpolations and extrapolations of free-energies. While data from numerous simulation points can be considered, in the following we restrict the analysis to computing the free-energy difference between points A and B given solely data from those two points, so that the formulas thus derived are a TI counterpart to those presented in Section III for FEP methods.

In the following we present the case of a single TI variable which can be a HP (f) or an OP (X) and we will hence use the symbol λ as a generalized integration variable. For brevity of notation, we define:

$$\phi_1 \equiv \frac{\partial \Phi}{\partial \lambda}, \quad \phi_2 \equiv \frac{\partial^2 \Phi}{\partial \lambda^2} \quad (50)$$

where derivatives are evaluated for any (other) ensemble field held constant. Figure 1 can be used as a guide letting $\lambda \rightarrow f$.

Matching polynomial.

Enforcing that values of ϕ_1 and ϕ_2 match at both ends of the integration step $\{\phi_{1A}, \phi_{1B}, \phi_{2A}, \phi_{2B}\}$ requires a polynomial with 4 coefficients:

$$\phi_1 = c_1 + 2c_2\lambda + 3c_3\lambda^2 + 4c_4\lambda^3 \quad (51)$$

where, consistent with our convention in a related paper,⁴² the coefficients are written as “ $n \times c_n$ ” so that the prefactors “ n ” go away when Eq. (51) is integrated to get $\Delta\Phi_{AB}$. The 4 coefficients are readily found as:

$$c_1 = \phi_{1,A}, \quad c_2 = \frac{\phi_{2A}}{2}, \quad c_3 = \frac{\Delta\phi_1}{\Delta\lambda^2} - \frac{\Delta\phi_2 + 3\phi_{2A}}{3\Delta\lambda}, \quad c_4 = -\frac{\Delta\phi_1}{2\Delta\lambda^3} + \frac{\Delta\phi_2 + 2\phi_{2A}}{4\Delta\lambda^2}, \quad (52)$$

where $\Delta\lambda = \lambda_B - \lambda_A$.

Best-Fitting Polynomial.

Two disadvantages of the exact polynomial match are that: (i) it disregards that ϕ_2 data typically have significantly larger statistical error than ϕ_1 data, and (ii) any extrapolation becomes more unwieldy the higher the polynomial order. While fewer coefficients imply fewer assumptions about the true model, a linear extrapolation based on ϕ_1 data at A and B could be too conservative and would disregard the effect of any curvature in the ϕ_1 function (see Fig. 1). We hence propose a compromise where a ϕ_1 quadratic model is constructed (the minimal polynomial able to capture curvature in ϕ_1), so that its 3-coefficients are found by matching exactly the more accurate ϕ_1 data and by “fitting”, in a least-square deviation sense, the ϕ_2 data. In such a case

$$\Phi = c_0 + c_1\lambda + c_2\lambda^2 + c_3\lambda^3 \quad (53)$$

so that:

$$\phi_1 = c_1 + 2c_2\lambda + 3c_3\lambda^2, \quad \phi_2 = 2c_2 + 6c_3\lambda \quad (54)$$

and besides matching the ϕ_{1A} and ϕ_{1B} data, we minimize the function:

$$R = b_A(\phi_{2A} - \phi_{2A}^*)^2 + b_B(\phi_{2B} - \phi_{2B}^*)^2 \quad (55)$$

where ϕ_2 and ϕ_2^* represent the actual and model (54) predicted values, respectively, and the b 's are error-dependent coefficients here assumed equal $b_A \approx b_B$. The solution is:

$$c_1 = \phi_{1A}, \quad c_2 = \frac{\Delta\phi_1}{2\Delta\lambda} - \frac{\Delta\phi_2}{4}, \quad c_3 = \frac{\Delta\phi_2}{6\Delta\lambda}, \quad (56)$$

The free-energy difference between point A and B is then given by:

$$\Delta\Phi_{AB} = \frac{1}{2}(\phi_{1A} + \phi_{1B})\Delta\lambda - \frac{1}{12}(\phi_{2B} - \phi_{2A})\Delta\lambda^2 \quad (57)$$

If we evaluate $\Delta\Phi_{AB}$ for model (51)-(52), we surprisingly find that the resulting expression is again Eq. (57), despite the differences in polynomial order and parameters. This suggests that Eq. (57) is a robust formula to estimate $\Delta\Phi$ based on data of ϕ_1 and ϕ_2 at points A and B , somewhat independent of the underlying Φ model assumed, and is rooted in a more general mathematical framework. Indeed, the Euler-Maclaurin summation formula^{37,59} can be expressed as:

$$\Delta\Phi_{AB} = \int_{\lambda_A}^{\lambda_B} \phi_1(\lambda) d\lambda = \frac{\Delta\lambda}{2}(\phi_{1A} + \phi_{1B}) - \frac{(\Delta\lambda)^2}{12}(\phi_{2B} - \phi_{2A}) + \frac{(\Delta\lambda)^4}{720}(\phi_{4B} - \phi_{4A}) + \dots - \frac{B_{2k}}{(2k)!}(\Delta\lambda)^{2k}(\phi_{(2k-1)B} - \phi_{(2k-1)A}) - \dots \quad (58)$$

where B_{2k} is a Bernoulli number of order $2k$. Our proposed polynomials are consistent with this series truncated up to the 2nd order term and, with vanishing 3rd (and higher odd) order terms, showing that the error in our numerical integrations is less than twice the 4th order term. Equation (58) also shows how simulation data on 3rd order derivatives would not improve the free-energy estimation, but data on 4th order derivatives would, albeit with harder-to-get and less accurate data. Due to this connection, henceforth will refer to Eq. (57) as the TI Euler Maclaurin 2nd order formula or “TI-EM2”.

It is also informative to rewrite the linear Hamiltonian OSOS, Eq. (17), for the SOS case, i.e., $\alpha = \frac{1}{2}$, and upon replacing the moments of the distribution by the 1st and 2nd derivatives of Φ [as per Eq. (5) and (6)] and letting $\lambda \leftrightarrow f$:

$$\Delta\Phi_{AB} = \frac{1}{2}(\phi_{1A} + \phi_{1B})\Delta\lambda - \frac{1}{8}(\phi_{2B} - \phi_{2A})\Delta\lambda^2 \quad (59)$$

Equation (59) is very similar to Eq. (57), except that the 2nd order term prefactor in the former is 50% larger [1/8 vs. 1/12]. Equations (57)-(59) highlight the connection between FEP and TI methods (at least for linear Hamiltonians), albeit it is also meaningful that Eq. (57) cannot be recovered from Eq. (17) for any choice of α , showing that *the Gaussian approximation in the OSOS formulation is not equivalent to using 2nd derivative data in constructing polynomials* for $\Delta\Phi_{AB}$. To first order, Eqs. (57) and (59) do coincide, becoming the trapezoidal rule which, we write down for future reference:

$$\Delta\Phi_{AB} = \frac{1}{2}(\phi_{1A} + \phi_{1B})\Delta\lambda \quad (60)$$

In Sec. A of the Appendix we further leverage Eq. (58) to outline an error analysis when using the TI-EM2 formula for TI.

Similar formulas can be developed for the multiple TI variables; the case of two HPs is described in Sec. B of the Appendix. Free energy changes over multiple HPs are in order when mapping the properties of phases over various thermodynamic fields or alchemical variables, and particularly when mapping phase coexistence conditions as shown next.

V. Tracing coexistence lines via TI: FENEX method

A TI intended to trace a coexistence curve is typically denoted as a Gibbs-Duhem integration (GDI) method.^{3,38-42} The formulas described in Sections III and IV give free-energy differences between discrete points, but extrapolations of free-energy are also needed (or implied) in a number of methods for advancing GDI along a phase coexistence line where the conditions defining the next coexistence state on the line, i.e., the integration path, are not known a-priori. Furthermore,

when extrapolating, it is uncertain when GDI will break down; e.g., brought about by one of the original phases becoming unstable or by the proximity of a critical point.

Approaches such as WHAM^{16-19,56-58} and MBAR²⁰ have the advantage that the same formulas used to obtain free-energy differences between simulated states can be used to interpolate or even extrapolate free-energies to other conditions, so long as those conditions are within the scope of the simulated Π s. For a typical GDI stepwise extrapolation along a coexistence line, however, it is unclear if approaches (like WHAM and MBAR) that leverage multistate simulation data provides any distinct advantage over methods that only leverage the nearest two data points. Likewise, higher order polynomial forms or Gaussian process regression^{46,48} models that include/fit many points are not needed to capture the curvature of the free-energy function as it is projected into new territory, as long as information about 2nd order derivatives in the nearest points is available. In this context, the simple polynomial forms described in Sec. B of the Appendix are particularly well-suited to the extrapolations needed in GDI, a strategy denoted as FENEX in Ref. [42] and is outlined in the Appendix, Sec. C.

VI. RESULTS

5.1 Independent Harmonic Oscillator(s).

A model system that has been used to test and understand phase space overlap and biases in FEP methods,^{4,7-9} entails two Hamiltonians A and B defined by

$$\mathcal{H}_A = \sum_{i=1}^N \omega_A x_i^2, \quad \mathcal{H}_B = \sum_{i=1}^N \omega_B (x_i - x_0)^2 \quad (61)$$

where x_i is the coordinate of particle i , N is number of particles, and ω_B/ω_A and x_0 are preset parameters. The free-energy change is analytically solvable:

$$\frac{1}{N}\Delta\Phi_{AB} = \frac{1}{2}kT \ln(\omega_B/\omega_A) \quad (62)$$

Increasing the ratio ω_B/ω_A narrows the phase space of system B relative to A , which makes less likely to sample in system B configurations relevant to system A (asymmetric bias). Increasing parameter x_0 shifts the phase space of both systems apart and increasing N exacerbates this effect and the asymmetric bias.⁸⁻¹⁰ Limited sampling of the Π distribution tails in finite-length simulation runs is detrimental to the accuracy of FEP methods. For the FEP and TI considered in this work, however, since only the first two moments of Π are used, we simplify the analysis by assuming they are well sampled regardless of N , and hence we will just consider the limiting case that those first two moments and associated Φ derivatives *are known exactly* and set $N = 1$. In such a case, we no longer need to conduct actual simulations and can probe the intrinsic accuracy of the methods on this system, decoupled from statistical errors and associated overlap sampling. We further set $x_0 = 0$ (which has no effect on $\Delta\Phi$ as calculated by the analytical formulas to be compared here) and simply consider changes in the Hamiltonian

$$\mathcal{H} = \omega x^2 \quad (63)$$

where ω goes from ω_A to ω_B . Taking ω to be the HP variable λ ($=f$) it follows that:

$$\phi_1 = \sigma^2 = 1/2\omega, \quad \phi_2 = -2\sigma^4 = -1/2\omega^2 \quad (64)$$

which can be readily used with the TI formula (57). To implement OSOS, however, one must first decide what the X variable is in the Hamiltonian:

(a) If we assume that $X = x$ in Eq. (63), then we are dealing with a non-linear, harmonic Hamiltonian akin to that described in Sec. 3.1.3 with $f_1 = \omega$, $f_2 = 0$ and $\mathcal{H}_0 = 0$. In such a case, Eq. (26) gives the exact solution for $\Delta\Phi$ (regardless of α which drops out). This is to be expected since Eq. (26) was derived assuming $\Pi(X)$ to be Gaussian which is precisely the case here.

(b) If we assume that $X = x^2$ in Eq. (63), then we are dealing with a single- X linear Hamiltonian akin to that described in Sec. 3.1.2 with $f = \omega$ and $\mathcal{H}_0 = 0$. Unlike case (a), this is a non-trivial and more challenging case, since the $\Pi(X)$ is not Gaussian at all (as assumed to derive the OSOS formulas) but rather an exponential decay function and hence this case can test the ability of different methods to handle underlying Π s which strongly depart from normality. To apply Eq. (17) for this case, for a point p :

$$X_p = 1/4\omega_p, \quad \sigma_p^2 = 1/8\omega_p^2 \quad (65)$$

Such that the mean values (in $X = x^2$) in the Π distribution for Hamiltonians A and B no longer coincide (even though they did in x). Figure 2 shows a comparison of the $\Delta\Phi_{AB}$ predicted by different methods using a single step from A to B . Up to $\omega_B/\omega_A = 1.5$ all methods perform similarly well, with the TI and SOS equations, being identical when truncated to first order, always overestimate $\Delta\Phi_{AB}$ while most other methods underestimate it. The TI-EM2 Eq. (57) is the most accurate and significantly better than the closely related $\alpha = 0.5$ SOS formula (59), but only up to about $\omega_B/\omega_A = 3.6$, at which point it starts to diverge drastically. The OSOS formula (19) is the 2nd best performer, closely followed by OSOS (16)-(17), with both being significantly better than SOS Eq. (59), and less divergent than the TI-EM2 (57) formula. Interestingly, the BAR-G (11), (32) and OSOS formulas (19), (16) give comparable absolute deviations from $\Delta\Phi_{AB}$, the former always overestimating and the latter underestimating it, which suggest that the main source of error here is not the optimization recipe but the incorrect assumption of Gaussian Π s. While these relative performances are system-dependent, some of these trends could extend to other systems as well.

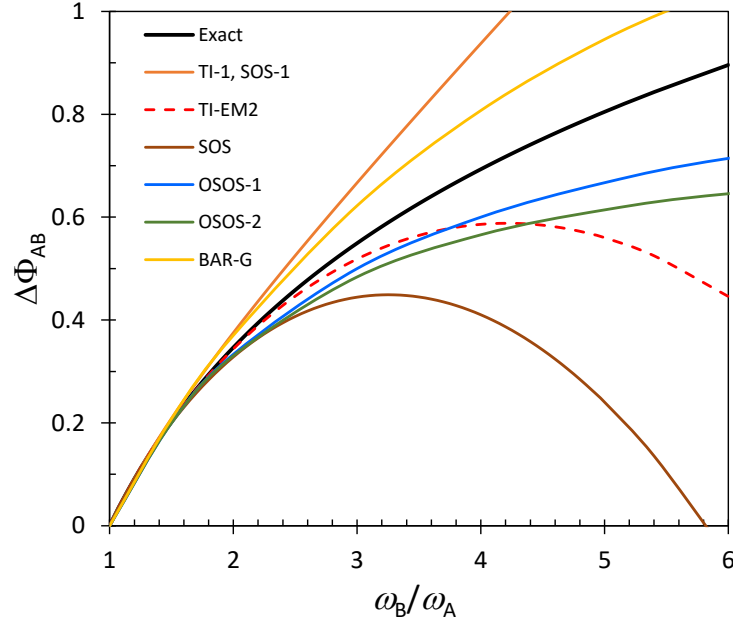


Fig. 2. Single step free energy change for the harmonic potential Hamiltonian model for different methods as marked in the legend: TI-1 = SOS-1 is Eq. (60), TI-EM2 is Eq. (57), SOS is Eq. (59), OSOS-1 is Eq. (19), OSOS-2 is Eq. (17)+(16), and BAR-G is Eq. (32)+(11).

In the context of TI, an integration step that doubles the coupling parameter can be considered as quite large (e.g., envision doubling the temperature) and it is significant that the TI-EM2 (57) and OSOS (19), for example, incur in less than 1% and 4% deviations, respectively, rather small errors, despite the very non-Gaussian nature of $\Pi(X)$.

5.2 Potential of mean force of cation-anion dissociation

We consider a fully atomistic polymeric system where a salt is dissolved to allow ionic conduction. The force field is based on those of Refs. [60]-[63]. The specific polymer is poly(3-(butoxybutyl)thiophene) [P3APAT] which is assumed to assemble into a bilayer ordered structure where the thiophenes crystallize into planes sandwiching amorphous layers of side chains whose oxygen groups solvate ionic species. As described in Ref. [63], molecular dynamic simulations⁶⁴ for different side-chain chemistries were performed using 16 decamers; here we only reproduce

simulation data for the dissociation free energy, often referred to as potential of mean force (PMF) of lithium bistrifluoromethanesulfonimide (LiTFSI) in the specific P3APAT polymer. The procedure consists of first equilibrating a salt molecule into the side-chain regions of the polymer and then separating the Li^+ and TFSI $^-$ ions step wisely using as OP the distance $X = r_{\text{sep}}$ between Li^+ and the center of mass of TFSI $^-$. The range of interest of r_{sep} was sampled in a series of umbrella sampling windows, each having the harmonic bias potential of Eq. (47) with $f_1 = 16 \text{ kcal/mol} \cdot \text{\AA}^2$, and changing the separation setpoint f_2 from $f_2 = 3.0 \text{ \AA}$ to $15.0 \text{ \AA}$ with $0.5 \text{ \AA}$ increments. The resulting histograms from each window were combined with WHAM¹⁶⁻¹⁹ and the PMF extracted therein.⁶⁵ In each window, four random initial configurations were used to get better sampling, and each simulated for 1 ns in a canonical ensemble at 400 K to collect the $\Pi(X)$ data.

Since in this case the coupling parameter is an OP (as opposed to a HP), we need to use the formulas we developed having a harmonic biasing function to indirectly constrain simulations around target OP values. Specifically, for TI we use Eqs. (48) and (49) to get required first and second derivatives of the unbiased Φ [for application in formulas (57), (59) or (60)]. For OSOS we use Eq. (30') or Eqs. (30)-(31') to get the biased $\Delta\Phi'_{AB}$ values, which we then unbiased using the correction of Eq. (33).

To implement the OSOS and TI formulas described in Sections III and IV we only used average and variance data from every-other (non-successive windows; i.e., for f_2 values separated by $\sim 1 \text{ \AA}$ rather than $0.5 \text{ \AA}$; in this way we also test the ability of those formulas to estimate free energy differences when minimal or no overlap between histograms exists. These selected data are depicted as Gaussians in Fig. 3. We note that using data from all windows would lead to negligible differences of results across methods relative to statistical error bars. The inset in Fig. 3 shows the complete set of simulated histograms collected at the different US windows, confirming that successive histograms duly overlap as required by WHAM.

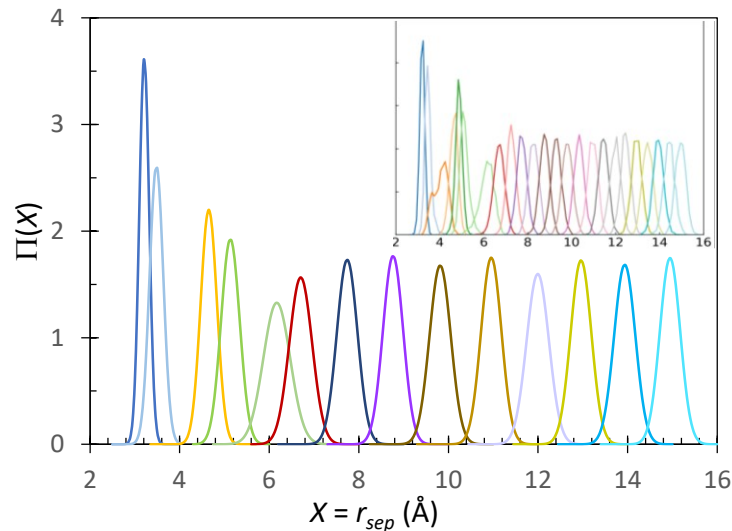


Fig. 3. Depiction of the implied Gaussians that match the mean and variance of selected simulated histograms from different umbrella sampling windows for various target separations between Li^+ and TFSI^- ions in P3APAT polymer at 400 K. Inset shows the complete set of actual simulated histograms, colored as in main figure. Inset adapted from data generated for Ref. [63].

Figure 4 shows that the PMF curve changes quickly across the ion separation range, with one minimum and 4 inflexion points and a large overall free-energy change spanning a barrier of over $56 kT$. It hence constitutes a challenging curve to describe via global fit models given the multiple localized curvature changes and stiffer regions.

The inset of Fig. 4 shows the deviations, relative to the reference results from WHAM for $0.5 \text{ \AA } f_2$ -intervals, for free-energy difference over the $1 \text{ \AA } f_2$ -intervals at which the TI and OSOS calculations were performed. These mean square deviations are about $\pm 0.31 kT$ per $\text{\AA}$ which, being comparable to the corresponding errorbar for the WHAM results, indicate that any of the proposed methods produces the PMF with largely comparable accuracy. Deviations with respect to the WHAM results are expectedly not uniform over the X range, reflecting the characteristics of the PMF curve; i.e., deviations tend to be larger around inflexion points. The dissociation free energies thus computed (i.e., the difference between the minimum and maximum in the curves) are

consistent within errorbars: $56.5 \pm 2.5 \text{ kT}$. This comparison illustrates that the proposed methods can get free-energy results with similar accuracy but higher computational efficiency than conventional WHAM as we can use half the number of US windows (i.e., have them more spaced out).

While this comparison does not favor a particular method amongst those tested, given that they all use the same input data, it is suggested that at least two of them be implemented to check for consistency in their predictions, the lack of which could signal deeper issues with the sampling or staging. In particular, we suggest using TI-EM2 which does not assume Gaussian Π s, and OSOS-1 which does, as these do not require any iterative calculation of weights and are hence the simplest to implement.

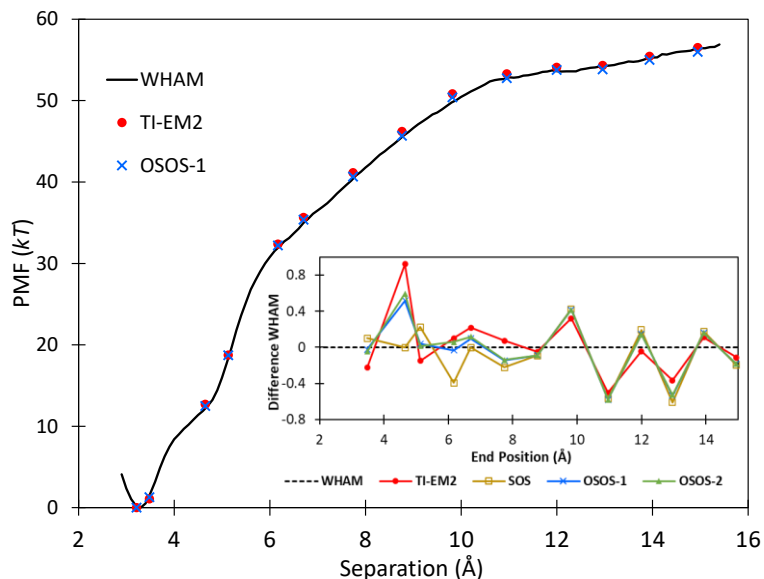


Fig. 4. Potential of mean force (PMF) calculated through 3 different methods. Inset shows the difference between the free-energy difference computed via different methods in successive intervals and the WHAM results. TI-EM2 refers to Eq. (57), SOS to Eq. (59), OSOS-1 to Eq. (30') and OSOS-2 to Eq. (30)+(31'), with Eq. (33) correction used for SOS and OSOS formulas.

5.3 TI of phase diagram for mixture of spheres and cubes

A coarse-grained representation of selective inter-species interactions was used where like-species interactions are hard-core and the cube (1) and sphere (2) potential energy is given by:^{66,67}

$$U_{12} = \begin{cases} \infty, & \text{if overlap} \\ -\varepsilon, & r \leq r_c, \delta \leq \delta_c \\ 0, & \text{otherwise} \end{cases} \quad (66)$$

where r is the distance between the particle centers of mass, and δ is the distance between the sphere center and the vector normal from the center of the cube's facet closest to the sphere. Unless otherwise stated, a particle size-ratio of $\zeta = \sigma/a = 1.23$ was adopted, where σ is the sphere core diameter and a is the cube side edge. The attractive well for cube-facet/sphere alignment has a depth ε , a radial length of $0.15a$, and a (lateral) δ -width of $0.4a$. Accordingly, the relevant cutoff distances are $r_c = 0.5\sigma + 0.65a$, and $\delta_c = 0.4a$.

Monte Carlo (MC) simulations were performed in a semigrand isobaric ensemble^{39,41-42} ($N\Delta\mu PT$) to obtain the equilibrium compositions of coexisting solid-liquid phases for given pressure (P), temperature (T), chemical potential difference between components ($\Delta\mu$), and fixed total number of particles (N). In reporting simulation results, the following dimensionless quantities are used: $\beta = \varepsilon/kT$, $p = \beta Pa^3$, $\beta(\mu_2 - \mu_1) = \beta\Delta\mu$ (chemical potential difference between species), and x is the mole fraction of cubes.

Simulations were performed in a cubic box with $N = 1728$ for cube-rich cubic phases and $N = 864$ for sphere-rich fcc phases, with any isotropic phase having the same N as its coexisting ordered phase. At or near each coexistence state of interest, both phases are simulated for at least 10^6 MC cycles for equilibration and 2×10^6 MC cycles for production. Each MC cycle consisted of N translational, N rotational (for the cubes), $N/5$ swap, 3 volume move attempts, and $N/2$ mutations. The mutations change one species into the other and cause composition fluctuations in accord to the specified $\beta\Delta\mu$. All attempted moves were accepted using the Metropolis criterion.³ Overlap

detection involved the separating axes theorem⁶⁸ for two cubes and Arvo's algorithm⁶⁹ for cube-sphere pairs.

To map out the pressure composition phase diagram for this system with the TI formulas described in Sec. B of the Appendix, the $N\Delta\mu PT$ ensemble is first cast into a two-field *linear* Hamiltonian:

$$N\beta\mu_2 = -\ln Q = -\ln\{\sum \exp[-\beta U - \beta PV - \beta N_1\Delta\mu]\} \quad (67)$$

where N_1 is the number of cubes in the system. By comparing to Eqs. (1)-(3), we can let $H_0 = \beta U$, $\Phi = N\beta\mu_2$ and either choose: (a) $f_1 = p$, $X_1 = V$ and $f_2 = \beta\Delta\mu$, $X_2 = N_1$ or (b) $f_1 = \beta\Delta\mu$, $X_1 = N_1$ and $f_2 = p$, $X_2 = V$. Choice (a)/(b) is indicated if driving steps in $p/\beta\Delta\mu$ are prescribed. For most points either choice gives consistent results, with the preferred choice being the one making the differential equation less stiff, i.e., leading to smaller changes of f_2 . Note that the first and second derivatives of Φ are simply given by averages and covariances as per Eqs. (5) and (6). The integrations were started from both ends of the composition spectrum: the pure sphere and pure cube systems whose hard-core isotropic-solid phase coexistence conditions are known^{70,71} (noting that thermal attractions only ensue across species). Step sizes were chosen, one at a time, such that anticipated changes in the coexistence composition of any one phase did not change by more than 0.05. The estimated errors in the estimation of free-energy steps based on Eq. (A6) was $\sim 0.05 kT$.

Figure 5(a) is the simulated pressure-composition phase diagram which shows that for near equimolar compositions a compound state (C*) is favored where cubes and spheres form a NaCl lattice. Sphere-rich states (with cube compositions < 0.5) exhibit a first order phase transition to the C* phase from either the isotropic (I) phase or the fcc solid phase (S). Indeed, the spheres and the compound form a eutectic at $x \approx 0.2$ and $p \approx 7.6$. It is at this eutectic point that integrations started from the left (pure spheres) and from the right (pure cubes) intersected, with the former process using prescribed steps in $f_1 = p$ (with $f_2 = \beta\Delta\mu$ computed to target phase coexistence)

and the latter using prescribed steps in $f_1 = \beta\Delta\mu$ (with $f_2 = p$ computed to stay at phase coexistence). In contrast, the C* compound is approached continuously from the cube solid solution (C), at least for the range of pressures near the simulated phase-coexistence conditions. In the C solid solution, however, while the spheres are spatially distributed at random, the 6 nearest neighbors of any given sphere are exclusively cubes, as is the case in the compound (see Fig. 5(b)). The continuous nature of the (C-to-C*) solution-to-compound transition depends on the system's characteristics, e.g., for a size ratio of $\sigma/a = 1$ (results not shown) a discontinuous transition is observed instead.

As shown in Fig. 6, there is very little difference in the $\Delta\Phi$ s obtained via either the TI-EM2 formula (B5) or the OSOS formula (22') (which is appropriate here given the linear Hamiltonian). The absolute average deviations in $\Delta\Phi$ per molecule per integration step $\Delta(\Delta\phi)$ between these 2 methods was $\sim 0.0002kT$ which is much smaller than the statistical errorbars of 0.065 of Eq. (B6) which reflect the neglected 2nd order terms in Eq. (B5) and, as can be seen in Fig. 6, can vary widely over the integration range. By accounting for the ϕ_1 curvature and integrand stiffness effects, the 2nd order terms provide a particularly significant correction for the I-S phase coexistence results, as can be seen by the large deviations in the 1st-order formula (open red circles in Fig. 6). Either set of free-energies (from TI-EM2 or OSOS) essentially generated indistinguishable phase coexistence curves.

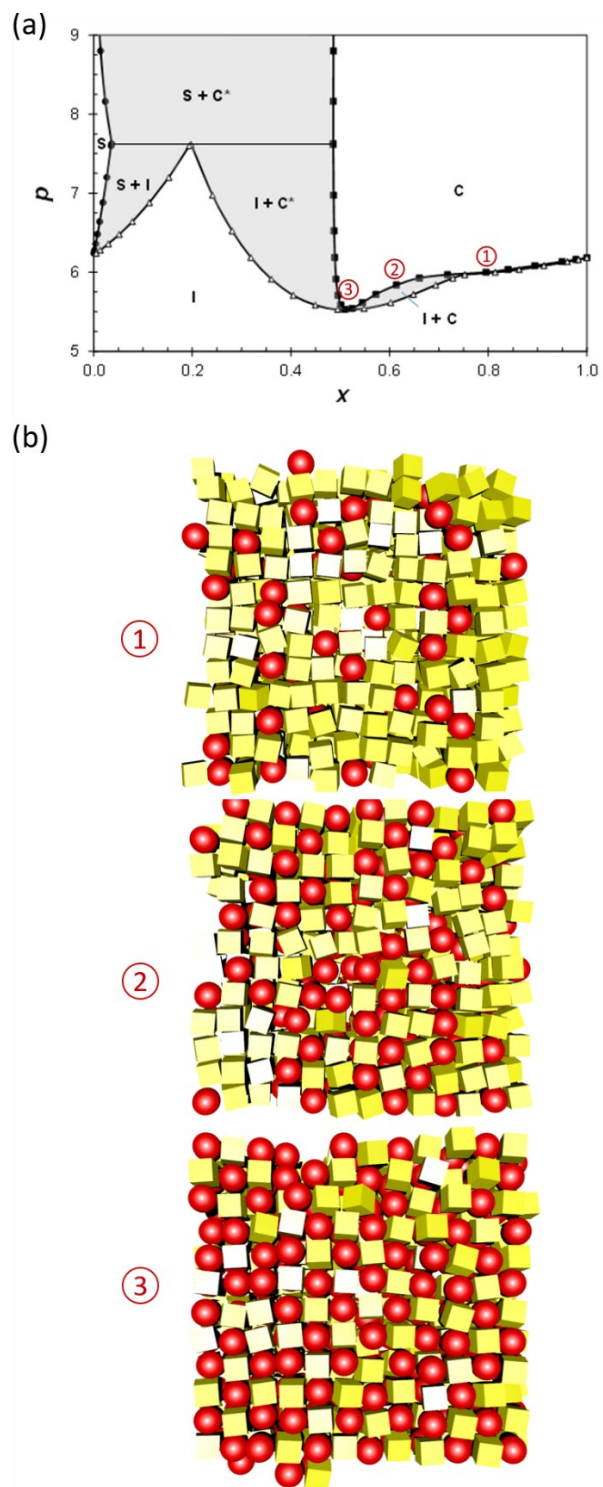


Fig. 5. (a) Pressure (p) vs. composition (x = cube mole fraction) phase diagram for the sphere+cube system with size ratio $\sigma/a = 1.23$ and $\beta = 1.0$. Gray regions indicate two-phase domains. C = Cubic phase, S = fcc solid, I = Isotropic phase, C^* = cubic compound. (b) Snapshots of C phase for 3 points along the I-C coexistence line as marked inside of diagram (a).

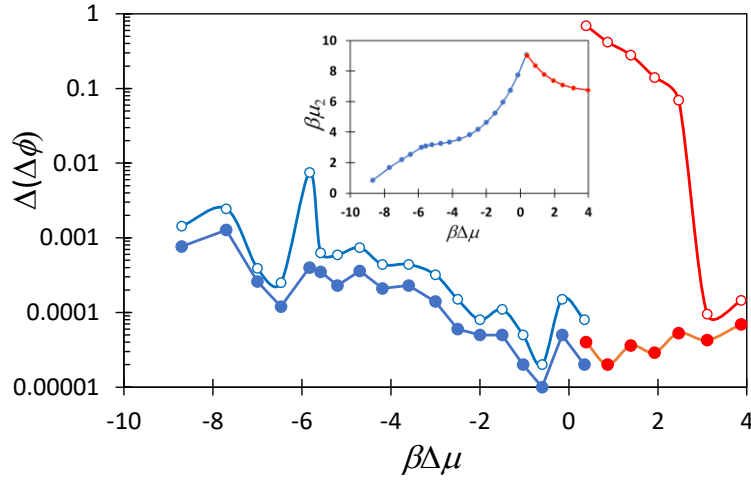


Fig. 6. Absolute departure in free-energy change per integration step (and per molecule) of OSOS Eq. (22') (filled circles) and of Eq. (B6) (open circles), relative to the full TI-EM2 formula (B5) results. Blue/red symbols correspond to the I-C/I-S coexistence points in Fig. 5(a). Inset shows values of $\Delta\phi = \Delta\Phi/N = \beta\mu_2$ of I phase along the integration path.

One of the advantages of accessing free-energy data associated with a phase diagram is that it allows to assess the relative thermodynamic stability of different systems and conditions. From the $\Delta\Phi/N = \beta\mu_2$ and $\beta\Delta\mu$ data we computed the corresponding $\beta\mu_1$ and $\beta\mu_2$ values at any point and therefrom find, within an additive constant, the Gibbs free energy for any coexistence phase as $\beta G = x \beta\mu_1 + (1-x)\beta\mu_2$ where x is the mole fraction of cubes. Figure 7 shows the results for the range of conditions at which the isotropic phase (I) coexists with the cubic phase (C) for two cases, the $\sigma/a = 1.23$ system discussed in Fig. 5, and an additional $\sigma/a = 1.0$ case. Starting from the pure cube state ($x = 1$), the addition of spheres lowers G due to the energy reduction afforded by the inter-species thermal attractions; this trend continues until the I phase encounters the compound solid C^* (at $x = 0.5$) where favorable energetic interactions in the C phase are maximized. For $\sigma/a = 1.23$, G of the C^* phase coexisting with the I phase with $x = 0.5$ is also the lowest achievable G , but for $\sigma/a = 1.0$ the I phase can lower its G even further (until $x \approx 0.38$). However, for congruent

crystallization (i.e., from I phase at $x = 0.5$) the $\sigma/a = 1.0$ C* phase has a lower G than the $\sigma/a = 1.23$ C* phase, indicating that the former system has a more optimal geometry for C* phase stability, as formerly ascertained with a different TI over σ/a values in Ref. [67].

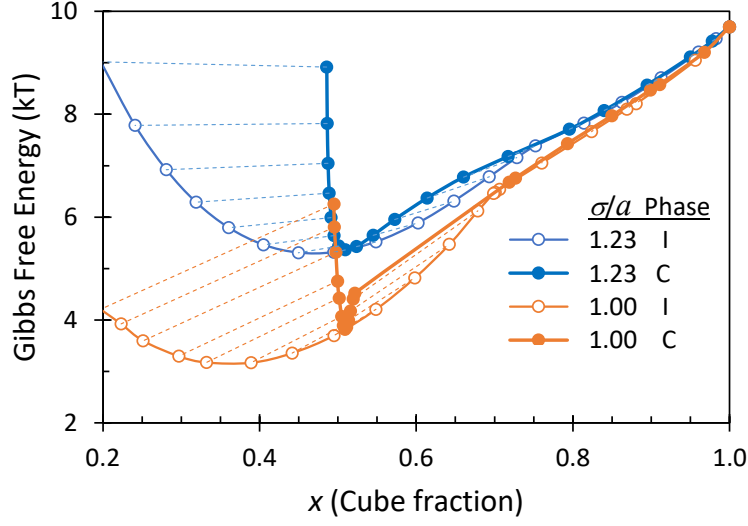


Fig. 7. Gibbs free energy of the I phase (open circles) and C phase (filled circles) at coexistence for systems at $\beta = 1$ with $\sigma/a = 1.23$ (blue) and $\sigma/a = 1.0$ (orange). Dashed traces are tie lines. Phase diagram for $\sigma/a = 1.23$ is shown in Fig. 5(a).

VI. FINAL REMARKS AND OUTLOOK

In this work a general framework is outlined to cast FEP and TI methods in a unified manner and derive specialized, optimized formulas to obtain free-energy changes between two states $\Delta\Phi_{AB}$, that make use of simulation data from first and second moments of the order parameter distribution Π in FEP, or equivalently, first and second derivatives of the free-energy in TI. Using such simulation data only up to 2nd order, is seen as a sensible compromise between typical TI, that only use first order derivatives, and FEP overlap methods which use complete Π data (whose

higher moments become rapidly more inaccurate). The resulting formulas are found to have some similarities in form (e.g., for linear Hamiltonians) and in performance with our testbeds comprising an analytically solvable harmonic Hamiltonian (for assessing error progression with step size), an atomistic system (for computing a potential of mean force with coordinate-dependent OP), and a binary-component coarse-grained model (for computing a phase diagram in system whose compositional space is sampled through two HPs driving alchemical transformations).

Table 1 summarizing the main formulas developed in this work. Equations (12)-(14) provide the general FEP template for the advocated OSOS approach, applicable regardless of Π being Gaussian or the Hamiltonian being linear. The OSOS free-energy formulas adopt particularly simple, analytical forms for Gaussian Π and linear Hamiltonians, but examples are shown of how non-linear Hamiltonians can also be processed that potentially lead to simple analytical expressions like Eqs. (30)-(31'). The OSOS optimization approach could also be deployed when $\Delta\Phi_{AB}$ is computed from the simulated $\Pi(X)$ data (i.e., using Eq. (13) rather than Eqs. like (17) or (26) derived from leveraging the Gaussian approximation), which would represent a simpler alternative to the conventional BAR method.¹⁵ Such considerations do not apply to the TI formulas presented, as they do not make any explicit assumption on the form of Π (Gaussian or not) or of the Hamiltonian (linear or not), notwithstanding the similarities with OSOS formulas that use such assumptions [like between Eq. (57) and (17)→(59)]. The key requirement in TI is the ability to find the needed derivatives of the free-energy with respect of the integration HP or OP, which is detailed in Secs. 4.1-4.2 through Eqs. (35)-(42) and (48)-(49).

Table 1. List of main free-energy formulas and key features from this work. SOS = Simple Overlap Sampling, OSOS = Optimized Simple Overlap Sampling, TI-EM2 = Euler-Maclaurin 2nd order, BAR-G = Bennett’s Acceptance Ration Method with Gaussian Π .

Type of method/ collective variable	Formula & key features	Equation number(s)
FEP/HP	OSOS (general, non Gaussian Π)	(12)-(14)
	OSOS-1 (Gaussian Π , linear $\mathcal{H}$)	(19)
	OSOS-2 (Gaussian Π , linear $\mathcal{H}$)	(16)-(17)
	OSOS-2 two HPs (Gaussian Π , linear $\mathcal{H}$)	(22')
	BAR-G (Gaussian Π)	(32),(11)
	SOS (Gaussian Π , linear $\mathcal{H}$)	(59)
	SOS-1 (delta function Π , linear $\mathcal{H}$)	(60)
FEP/HP or OP	OSOS-1 (Gaussian Π , non-linear $\mathcal{H}$)	(30')+(33) if OP
	OSOS-2 (Gaussian Π , non-linear $\mathcal{H}$)	(30), (31')+(33) if OP
TI/HP or OP	TI-EM2	(57) [(53) & (56)]
	Trapezoidal rule (= SOS-1)	(60)
	TI-EM2, two HPs	(B5) [(B3) & Table 2]
	Trapezoidal rule, 2HPs	(B6)
TI/HP	Derivatives	(35)-(37), (39)-(42)
TI/OP	Derivatives (harmonic bias)	(48)-(49)

In the case of FEP, the closed-form formulas presented to optimize the simple overlap sampling (OSOS) are based on the strategy of sampling (from states A and B) to a virtual intermediate state defined by a parameter α . The resulting prescriptions for α embody the physically intuitive idea that the optimal intermediate Hamiltonian need not be exactly halfway between A and B as in SOS but depend on the intrinsic variances of the Π distributions at those points. This was best illustrated by casting the problem from the perspective of TI: step sizes should be shorter around points where the integrand is fast varying or “stiff”. Since larger Π variances imply steeper slopes [$d^2\Phi/d\mathbf{f}^2$ as per Eqs. (6), (36) and (37)], it follows that given two points of different slopes, the intermediate state should be closer to that having the steeper slope, as also illustrated in Fig. 1.

A Gaussian Π embodies a 2nd order approximation for how X varies around its average value, and hence is not expected to be generally applicable. While the quality of this approximation will strongly depend on the property X , it is expected to improve the more degrees of freedom contribute to X , e.g., for an OP associated with numerous molecular interactions. The example of Sec. 5.1 illustrates, however, that even for a single degree of freedom and a very non-Gaussian Π , the Gaussian OSOS formulas can still be accurate for a fairly wide range of step sizes.

Among the GDI variants that have been used to trace phase coexistence lines of single and multicomponent systems, the free-energy extrapolation method⁴² (FENEX) had been put forward as a means to combine the ability histogram reweighting methods to interpolate/extrapolate free energy values and operate without the need to collecting overlapping histograms between successive simulation step points. While polynomial extrapolation formulas have been presented in previous studies,^{42,67} the new formulas proposed here (Appendix) take into account the larger statistical errors of 2nd order derivative data to give more robust local extrapolations for estimating free-energy differences.

Although the main purpose of this work was to propose new, easy-to-use formulas to compute $\Delta\Phi_{AB}$, it is also hoped that the framework provided also has some pedagogical value. The new formulas were developed in the context of a systematic organization and categorization of FEP and TI methods, showing how these methods are connected and how they can be used with either HPs or OPs as coupling parameters.

In developing the proposed formulas, we have only optimized the methods at the most basic level, e.g., to minimize the variance in the estimation of $\Delta\Phi_{AB}$ or the contribution of higher order terms relative to 1st order contributions. There are many other well-known tricks that can be used to optimize the integration processes underlying FEP and TI. For example, one common strategy to reduce integration error is to transform the coupling (integration path) variable $\lambda \rightarrow \lambda^*(\lambda)$ so that the integrand is a slowly varying function of λ^* . Indeed “stiff” integrals are difficult

to resolve accurately and often require a judicious on-the-fly reduction of stepsizes. For the case of sufficiently smooth integrands, choosing integration points/steps a priori could also be used to maximize the order (and often) the precision of the underlying integration polynomial, as in the case of Gaussian quadrature schemes.³⁷ Also, we only considered a single step and the case of equal sampling of points A and B . For multiple steps one could consider optimizing their sampling using, e.g., serial or parallel expanded ensemble wherein the frequency or biasing weight applied to each state be tunable.^{11,72-74}

Finally, the approaches presented in this work could also be extended to other types of free-energy calculations, for example to non-isothermal ensembles^{75,76} and to non-equilibrium work methods,²⁵⁻³⁰ which share a fundamental structure with FEP and TI methods described here, with some variants already exploiting Gaussian approximations for the work distributions.²⁷⁻²⁹ We also expect some of the formulas presented here to be beneficially deployed in specialized TI implementations⁷⁷ and approaches where specific values of free-energies differences or derivatives thereof are sought-after, e.g., in methods targeting phase coexistence conditions and the critical nucleus size in interfacial pinning methods.^{78,79}

Acknowledgments

Funding support from NSF awards CBET-1907369 and CHE-2101829 is gratefully acknowledged. The author thanks Y. Sun for providing the data for inset of Fig. 3 and WHAM results for Fig. 4 pertaining to Ref. [63].

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

APPENDIX

A. Model for TI Error Estimation

Previous studies have already addressed the question of estimating errors from FEP and TI calculations (see, e.g., Ref. [7]). Here we provide a simple analysis for TI tailored to the EM2 formula (57)-(58). For simplicity we assume a single integration path variable λ . Let us assume that L denotes the total range of integration, $\Delta\lambda$ be the step length (assumed to be uniform), t the length (in MD steps of MC cycles) used at each λ point, ε be the statistical error associated the evaluation of ϕ_1 at any simulation point. The total error can be seen as arising primarily from: (i) the statistical uncertainties in the calculation of f_1 values, and (ii) the discretization of the integral. It is known that the former error is inversely proportional to the square root of the simulation length,^{3,4} while the latter can be estimated, according to the Euler-Maclaurin formula (58), as a fourth order power of $\Delta\lambda$ for the integration formulas we adopt. Hence the error in free-energy in a particular step i of size $\Delta\lambda$ can be approximately expressed as:

$$\varepsilon_i \approx \varepsilon_{\phi_1,i} \Delta\lambda + \frac{|\phi_{4,i+1} - \phi_{4,i}|}{720} (\Delta\lambda)^4 \quad (\text{A1})$$

where ε_{ϕ_1} is the error in the estimation of ϕ_1 average values in a simulation run. If we assume an average ε_i value as representative of the error for each step and that step sizes are equal, then the total error in the estimation of $\Delta\Phi$ over the entire integration path $\Delta\lambda_{tot}$ would be:

$$\varepsilon_{tot} = \sum_i^{N_{steps}} \varepsilon_i \approx \frac{\Delta\lambda_{tot}}{\Delta\lambda} \varepsilon_i = \Delta\lambda_{tot} \left[\varepsilon_{\phi_1} + \frac{|\Delta\phi_4|}{720} (\Delta\lambda)^3 \right] \quad (\text{A2})$$

where $\Delta\phi_n$ is an average value of $\phi_{n,i+1} - \phi_{n,i}$ over steps. In practice, it will be difficult to have accurate estimates of $\Delta\phi_4$ and so, alternatively we overestimate the integration error as half of the 2nd order term in Eq. (58), thus:

$$\varepsilon_i \leq \varepsilon_{\phi_1} |\Delta\lambda| + \frac{|\Delta\phi_2|}{24} (\Delta\lambda)^2, \quad \varepsilon_{tot} \leq \frac{\Delta\lambda_{tot}}{\Delta\lambda} \varepsilon_i \quad (\text{A3})$$

Formulas (A2) or (A3) could form the basis for optimizing the stepsize $\Delta\lambda$; e.g., by maximizing the statistical efficiency. The latter is often taken to be inversely proportional to the product of the computational cost and the statistical error ε_{tot} . The optimization can be performed under different constraints; here we illustrate just one such scenario. If the length per simulation run is preset to guarantee suitable sampling, then ε_{ϕ_1} will be approximately constant and the computational cost will be proportional to the number of steps, so that the function to minimize is $(\Delta\lambda_{tot}/\Delta\lambda) \times \varepsilon_{tot}$. Using ε_{tot} from Eq. (A2), the final result can be written as:

$$\varepsilon_{\phi_1} |\Delta\lambda_{opt}| = 2 \frac{|\Delta\phi_4|}{720} \Delta\lambda_{opt}^4 \quad (A4)$$

which simply states that the optimal $\Delta\lambda$ occurs when the error associated with the uncertainty in ϕ_1 values (left hand side) is about twice the error associated with the quadrature discretization (right hand side). This shows that it would not make sense to increase the computational cost by reducing $\Delta\lambda$ and the integration error when the total error and ε_{ϕ_1} are unaffected. If our means to estimate the quadrature error is via a second order term as in relation (A3), then (A4) would translate to:

$$\varepsilon_{\phi_1} |\Delta\lambda_{opt}| \approx \frac{|\Delta\phi_2|}{12} \Delta\lambda_{opt}^2 \quad (A5)$$

Thus, if the 2nd order quadrature correction is larger (smaller) than the ϕ_1 -statistical error, then $\Delta\lambda$ could be made smaller (larger). A similar analysis could be put forward for TI over multiple variables; e.g., ε_i from relation (A3) for the two-variable TI-EM2 of Eq. (B5) could be written as:

$$\varepsilon_i \leq \varepsilon_{\phi_{10}} |\Delta f_1| + \varepsilon_{\phi_{01}} |\Delta f_2| + \left| -\frac{\Delta\phi_{20}}{24} (\Delta f_1)^2 - \frac{\Delta\phi_{02}}{24} (\Delta f_2)^2 + \frac{\Delta\phi_{11}}{12} \Delta f_1 \Delta f_2 \right| \quad (A6)$$

where $\varepsilon_{\phi_{10}}, \varepsilon_{\phi_{01}}$ are the errors in the estimation of ϕ_{10}, ϕ_{01} average values in a simulation run of a preset length.

B. TI with Best-Fit Polynomials for Multiple HPs

For concreteness, only formulas for the two-variable case are presented. While the formulas to be derived here can be applied to integrations over either HPs or OPs, they were only applied (in Sec. 5.3) to the case of HPs, and hence we specialized them using the symbol f (rather than the generic λ used in Sec. 4.3.1). For brevity of notation, we define:

$$\phi_{10} \equiv \left(\frac{\partial \Phi}{\partial f_1} \right)_{f_2}, \quad \phi_{01} \equiv \left(\frac{\partial \Phi}{\partial f_2} \right)_{f_1}, \quad (\text{B1})$$

$$\phi_{20} \equiv \left(\frac{\partial^2 \Phi}{\partial f_1^2} \right)_{f_2}, \quad \phi_{02} \equiv \left(\frac{\partial^2 \Phi}{\partial f_2^2} \right)_{f_1}, \quad \phi_{11} \equiv \frac{\partial^2 \Phi}{\partial f_1 \partial f_2} \quad (\text{B2})$$

so that ϕ_{ij} denotes a partial derivative that is order i with respect to f_1 and order j with respect to f_2 (other ensemble properties are implicitly held constant). These derivatives can be obtained at any simulation point " p " (i.e., ϕ_{ijp}) using the formulas derived in Sec. 4.1.

Following the strategy of Section 4.3, we favor the lowest order polynomials capable of capturing the variations of 1st and 2nd derivatives of Φ ; i.e., a maximum order of 3 over individual fields. A key difference between the one-field Φ -model described before and multi-field models is that cross-interaction terms should now be included as the effect of each field on Φ need not be independent of other fields. Indeed, capturing such inter-field coupling is essential to model Φ . Here we propose:

$$\Phi = c_{00} + c_{10}f_1 + c_{20}f_1^2 + c_{30}f_1^3 + c_{01}f_2 + c_{02}f_2^2 + c_{03}f_2^3 + c_{11}f_1f_2 + c_{22}f_1^2f_2^2 \quad (\text{B3})$$

The $c_{11}f_1f_2$ cross term embodies the leading (first-order) effect of the cross interaction of f_1 and f_2 on Φ . By itself however, it implies that ϕ_{11} is a constant which need not be true. To allow a greater flexibility, at least an additional higher-order cross term is needed. While 3rd order terms

$c_{21}f_1^2f_2$ and $c_{12}f_1f_2^2$ could be used, the three coefficients c_{11} , c_{21} , and c_{12} would be underspecified by only two ϕ_{11} data (ϕ_{11A} and ϕ_{11B}). We hence favor instead a single $f_1^2f_2^2$ term, which can be seen as the quadratic version of the basic cross variable (f_1f_2) to help describe positive and inverse correlation effects.

In model (B3) we have 8 c -parameters (disregarding c_{00} which is irrelevant since only free-energy differences are needed), while the points A and B provide 10 simulation data. Following Sec. 4.3, we can find 4 of those c -parameters by matching the most accurate data ϕ_{10A} , ϕ_{01A} , ϕ_{10B} and ϕ_{01B} . Since the ϕ_{11} data is crucial in capturing the coupling of λ_1 and λ_2 that shape the free-energy surface, we also choose to match the ϕ_{11A} and ϕ_{11B} data. This leaves 2 c -parameters that can be found by fitting the 4 data ϕ_{20A} , ϕ_{02A} , ϕ_{20B} and ϕ_{02B} , and minimizing the function:

$$R = (\phi_{20A} - \phi_{20A}^*)^2 + (\phi_{20B} - \phi_{20B}^*)^2 + (\phi_{02A} - \phi_{02A}^*)^2 + (\phi_{02B} - \phi_{02B}^*)^2 \quad (\text{B4})$$

The resulting coefficients are listed in Table 2.

Table 2. Constants of model (B3). $\Delta f_i = f_{iB} - f_{iA}$, $\Delta\phi_{ij} = \phi_{ijB} - \phi_{ijA}$

$c_{10} = \phi_{10B}$
$c_{01} = \phi_{01B}$
$c_{11} = \phi_{11B}$
$c_{22} = \frac{\Delta\phi_{11}}{4\Delta f_1\Delta f_2}$
$c_{20} = \frac{1}{8\Delta f_1}(4\Delta\phi_{10} - 2\Delta f_1\Delta\phi_{20} - \Delta f_2\Delta\phi_{11} - 4\Delta f_2\phi_{11A})$
$c_{02} = \frac{1}{8\Delta f_2}(4\Delta\phi_{01} - 2\Delta f_2\Delta\phi_{02} - \Delta f_1\Delta\phi_{11} - 4\Delta f_1\phi_{11A})$
$c_{30} = \frac{1}{12\Delta f_1^2}(2\Delta f_1\Delta\phi_{20} - \Delta f_2\Delta\phi_{11})$
$c_{03} = \frac{1}{12\Delta f_2^2}(2\Delta f_2\Delta\phi_{02} - \Delta f_1\Delta\phi_{11})$

With these c 's, the free-energy difference can be expressed as:

$$\Delta\Phi_{AB} = \frac{1}{2}(\phi_{10A} + \phi_{10B})\Delta f_1 + \frac{1}{2}(\phi_{01A} + \phi_{01B})\Delta f_2 - \frac{\Delta\phi_{20}}{12}\Delta f_1^2 - \frac{\Delta\phi_{02}}{12}\Delta f_2^2 + \frac{\Delta\phi_{11}}{6}\Delta f_1\Delta f_2 \quad (\text{B5})$$

It can be shown that this expression is the 2-variable, 2nd order generalization of the TI-EM2 formulas (57)-(58). A fourth-order model akin to Eq. (B3) but including two extra terms $c_{40}\Delta f_1^4$ and $c_{04}\Delta f_2^4$ was presented in Ref. [42] whose 10-parameters allowed to exactly match all 10 simulation data at points A and B and, remarkably, gives an expression for $\Delta\Phi_{AB}$ which is identical to Eq. (B5). This outcome is similar to what we observed in comparing the best-fit 3rd order model and the exact-match 4th order model for the single-field case described in Sec. 4.3. It should be pointed out that while multiple matching-polynomials and best-fitting-polynomials were proposed in Ref. [42], model (B3) was not considered.

For reference, we also list the trapezoidal rule by neglecting the 2nd order terms in Eq. (B5):

$$\Delta\Phi_{AB} = \frac{1}{2}(\phi_{10A} + \phi_{10B})\Delta f_1 + \frac{1}{2}(\phi_{01A} + \phi_{01B})\Delta f_2 \quad (\text{B6})$$

In the context of extrapolation, Eq. (B3) is most indicated for “lateral” extrapolation; i.e., when trying to estimate free-energies beyond one of the boundaries (say point B). One can also consider cases when “central” extrapolations are needed; e.g., to explore a new point (f_1, f_2) such that $f_{1A} < f_1 < f_{1B}$ (say) but f_2 is not bracketed between f_{2A} and f_{2B} but $f_{2B} < f_2 < f_{2A'}$ where A' is another simulated point proximal to B . In such a case, rather than only using one point (B) or two points to construct the polynomial (say A and B with the former being the closest to B in f_1 space) one could advantageously use data from A , B and A' . In particular, it would be sensible to use all 1st and 2nd derivative data at point B while only 1st order derivative data from points A and A' : those 9 total data could then be fitted to match the 9 c -constants in a full 3rd order polynomial (i.e., one similar to Eq. (B3) but replacing the single $c_{22}f_1^2f_2^2$ term by terms $c_{21}f_1^2f_2$ and $c_{12}f_1f_2^2$); this case was

in fact already considered in Ref. [42] and expressions for the c -constants given therein [Eqs. (A12)-(A17)].

C. FENEX IMPLEMENTATION

For concreteness, we consider here a common situation where GDI is used to map the coexistence between two phases to be denoted by superscripts I and II, and where two fields f_1 and f_2 are used to determine the thermodynamic state. Any Hamiltonian parameters constraining the system size of each phase are also assumed fixed throughout and be the same for both phases. The phase diagram in the f_1 - f_2 plane is mapped by stepping over prescribed values of $f_1 = f_{1,new}$ and calculating the corresponding coexistence values of $f_2 = f_{2,new}$ which should satisfy the equation:

$$\Gamma(f_{2,new}) = \Phi^I(f_{2,new}) - \Phi^{II}(f_{2,new}) = 0 \quad (C1)$$

whose solution implies that at $f_{2,new}$: $\Phi^I = \Phi^{II} = \Phi_{new}$.

The stepwise mapping proceeds as follows:

1) *Initialization.* At the outset it is assumed that at point “A” the values of f_{1A}^I and f_{2A}^I that lead to (near) coexistence of phases J = I and II are known (i.e., the free-energy difference $\Phi_A^{II} - \Phi_A^I$ is zero or known), and simulations were conducted so that ϕ_{10A} , ϕ_{01A} , ϕ_{20A} , ϕ_{02A} , and ϕ_{11A} are known for each phase. Note that the value of, Φ_A^I say, can be set arbitrarily to zero. Given $\Delta f_1^J = f_{1,new} - f_{1A}^J$, the value of $f_{2,new}$ that complies with phase coexistence is found by solving Eq. (C1). Since only point A has been simulated at this point, one can use a 2nd order Taylor expansion to estimate the free-energy of any phase; i.e.:

$$\Phi_{new} - \Phi_A = \phi_{10A}\Delta f_1 + \phi_{01A}\Delta f_2 + \frac{1}{2}\phi_{20A}\Delta f_1^2 + \frac{1}{2}\phi_{02A}\Delta f_2^2 + \phi_{11A}\Delta f_1\Delta f_2 \quad (C2)$$

where $\Delta f_i = f_{i,new} - f_{iA}$, and is applied separately to each phase (I and II) to get Φ^I and Φ^{II} to use in Eq. (C1) to solve for $f_{2,new}$. Specialized formulas may be needed if point A represent an infinitely dilute state in a mixture.^{39,42}

2) Reset & simulate. Redefine point B by letting (f_{1B}, f_{2B}, Φ_B) be $(f_{1,new}, f_{2,new}, \Phi_{new})$ from the previous step, perform simulations at these conditions for the two phases I and II and obtain the corresponding values for the 1st and 2nd derivatives (ϕ_{10B} , ϕ_{01B} , ϕ_{20B} , ϕ_{02B} , and ϕ_{11B}) at this state B .

3) Free-energy calculation. Calculate the free-energy differences between the last two states simulated for each phase, and hence estimate Φ_B^I and Φ_B^{II} for $\Delta f_i^J = f_{iB}^J - f_{iA}^J$. For these calculations we can use either relevant polynomial fits like (B5) or, depending on the Hamiltonian, OSOS formulas like (22)-(22') by leveraging the connection between Π distribution moments and derivatives' data [i.e., Eqs. (5)-(6)].

4) Step forward/extrapolation. At this stage, all relevant properties of (at least) the previous two points A and B and for both phases I and II are known. If $\Delta f_1^J = f_{1,new} - f_{1A}^J$ is given for a new point $(f_{1,new}, f_{2,new})$, then the value of $f_{2,new}$ for which coexistence is expected is estimated by solving Eq. (C1) where now $\Phi^J(f_{2,new})$ for each phase J is found by via Eq. (B3) using properties appropriate to the corresponding phase and $\Delta f_1 \rightarrow \Delta f_1^J = f_{1,new} - f_{1A}^J$, and $\Delta f_2 \rightarrow \Delta f_2^J = f_{2,new} - f_{2A}^J$.

At this point relabel point B of each phase (and all properties thereof) as point A , go back to step 2 and iterate steps 2-4 until a final target state has been reached.

5) Post processing. Coexistence properties are refined from the “near coexistence” simulated data; improved estimates (marked with the superscript “coex”) for any state A can be obtained by extrapolating the results obtained at (f_{1A}^J, f_{2A}^J) to $(f_{1A}^{coex} = f_{1A}^J, f_{2A}^{coex})$ by first finding f_{2A}^{coex} as the root of $\Gamma(f_{2A}^{coex}) = 0$ from Eq. (C1); the extrapolating polynomial model for Φ^J (as a function of

$\Delta f_1^I = f_{1A}^{coex} - f_{1A}^I$ and $\Delta f_2^I = f_{2A}^{coex} - f_{2A}^I$) could be based on single-point A data as in Eq. (C2), or on points A and B data as with Eq. (B3) (point B being the closest to A in f_1 space), or on three-point A - B - A' data ($f_{1A} < f_{1A'} < f_{1B}$) as with the *central* extrapolation discussed at the end of Sec. B of the Appendix. The Φ^I values found this way (with $f_{2A}^I = f_{2A}^{coex}$) correspond to Φ^{coex} (at point A). For steps 1, 4, and 5, the sought-after root of Eq. (C1) can be found via Newton's method using as first guess the first-order solution:

$$f_{2,new} = (\Phi_A^I - \Phi_A^{II} + \phi_{10A}^I \Delta f_1^I - \phi_{10A}^{II} \Delta f_1^{II} - \phi_{01A}^I f_{2A}^I + \phi_{01A}^{II} f_{2A}^{II}) / (\phi_{01A}^{II} - \phi_{01A}^I) \quad (C3)$$

REFERENCES

1. D. A. Kofke, Fluid Phase Equilibria **228-229**, 41 (2005).
2. C. Chipot, A. Pohorille editors, *Free Energy Calculations. Theory and application in Chemistry and Biology*, Study Edition, Springer-Verlag Berlin Heidelberg, 2007.
3. D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*. Academic, Sand Diego, 2002.
4. M. Allen and D. Tildesley, *Molecular Simulation of Liquids*, 2nd Ed. Oxford, 2017.
5. A. Pohorille, C. Jarzynski, and C Chipot, J. Phys. Chem. B **114**, 10235 (2010).
6. J. D. Chodera, D. L. Mobley, M. R. Shirts, R. W. Dixon, K. Branson, and V. S. Pande, Curr. Opinion Struct. Bio. **21**, 150 (2011).
7. M. R. Shirts and V. S. Pande, J. Chem. Phys. **122**, 144107 (2005).
8. D. Wu and D. A. Kofke, Phys. Rev. E **70**, 066702 (2004).
9. D. Wu and D. A. Kofke, J. Chem. Phys. **123**, 054103 (2005).
10. D. Wu and D. A. Kofke, J. Chem. Phys. **123**, 084109 (2005).
11. F.J. Martinez-Veracoechea and F.A. Escobedo, J. Phys. Chem. B. **112**, 8120 (2008).

12. P. V. Klimovich, M. R. Shirts, and D. L. Mobley, *J. Computer-Aided Molec. Design* **29**, 397 (2015).
13. T. Boon, A. J. Schultz, and D. A. Kofke, *J. Chem. Phys.* **132**, 214103 (2010).
14. S. G. Moustafa, A. J. Schultz, and D. A. Kofke, *J. Chem. Phys.* **139**, 084105 (2013).
15. C. H. Bennett, *J. Comput. Phys.* **22**, 245 (1976).
16. M. Ferrenberg and R. H. Swendsen, *Phys. Rev. Lett.* **63**, 1195 (1989).
17. M. Ferrenberg and R. H. Swendsen, *Phys Rev. Lett.* **61**, 2635 (1988).
18. S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, P. A. Kollman, *J. Comput. Chem.* **13**, 1011 (1992).
19. M. Souaille and B. Roux, *Computer Phys. Comm.* **135**, 40 (2001).
20. M. R. Shirts and J. D. Chodera, *J. Chem. Phys.* **129**, 129105 (2008).
21. J. R. Errington, *J. Chem. Phys.* **118**, 9915 (2003).
22. M. Fenwick and F.A. Escobedo, *J. Chem. Phys.*, **120**, 3066 (2004).
23. F.A. Escobedo and C.R.A. Abreu, *J. Chem. Phys.* **124**, 104110 (2006).
24. H. Paliwal and M. R. Shirts, *J. Chem. Theory Comput.* **7**, 4115 (2011).
25. C. Jarzynski, *Phys. Rev. Lett.* **78**, 2690 (1997).
26. G. E. Crooks, *Phys. Rev. E* **60**, 2721 (1999).
27. S. Park and K. Schulten, *J. Chem. Phys.* **120**, 5946 (2004).
28. O. Perisic and H. Lu, *PLOS ONE* **9**, e101810 (2014).
29. P. Procacci, *J. Chem. Phys.* **142**, 154117 (2015).
30. C. Dellago and G. Hummer, *Entropy* **16**, 41 (2014).
31. G. M. Torrie and J. P. Valleau, *Chem. Phys. Lett.* **28**, 578 (1974).
32. N. D. Lu, T. B. Woolf, D. A. Kofke, *J. Comput. Chem.* **25**, 29 (2004).
33. J. Kastner and W. Thiel, *J. Chem. Phys.* **123**, 144104 (2005).
34. J. Kastner and W. Thiel, *J. Chem. Phys.* **124**, 234106 (2006).
35. J. Kastner, *J. Chem. Phys.* **131**, 034109 (2009).
36. L. Maragliano and E. Vanden-Eijnden, *J. Chem. Phys.* **128**, 184110 (2008).

37. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran 90. The Art of Scientific computing*, 2nd Ed. Cambridge Univ. press, New York 1996.
38. D. Kofke, J. Chem. Phys. **98**, 4149 (1993).
39. M. Mehta and D Kofke, Chem. Eng. Sci. **49**, 2633 (1994).
40. A. van 't Hof, C. J. Peters, and S. W. de Leeuw, J. Chem. Phys. **124**, 054906 (2006).
41. F.A. Escobedo, J. Chem. Phys. **110**, 11999 (1999).
42. F.A. Escobedo, J. Chem. Phys. **140**, 094102 (2014).
43. F.A. Escobedo, J. Chem. Phys. **108**, 8761 (1998).
44. Y. Meng and B. Roux, J. Chem. Theory Comput. **11**, 3523 (2015).
45. G. Hummer, L. Pratt, and A. E. Garcia, J. Phys. Chem. **100**, 1206 (1996).
46. T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd Ed. Springer -Verlag, Berlin, 2009.
47. T. Stecher, N Bernstein, and G. Csanyi, J. Chem. Theory Comput. **10**, 4079 (2014).
48. L. Mones, N. Bernstein, and G. Csanyi, J. Chem. Theory Comput. **12**, 5100 (2016).
49. A. L. Ferguson, J. Phys.: Condens. Matter **30**, 043002 (2017).
50. C. Dai and S. C. Glotzer, J. Phys. Chem. B **124**, 1275 (2020).
51. H. Sidky and J. K. Whitmer, J. Chem. Phys. **148**, 104111 (2018).
52. C. Desgranges and J. Delhommelle, Mol. Syst. Des. Eng. **6**, 52 (2021).
53. N. A. Mahynski, H. W. Hatch, M. Witman, D. A. Sheen, J. R. Errington, and V. K. Shen, Molecular Simulation **47**, 395 (2021).
54. J. P. Selvaggi, J. Comput. Electron **17**, 61 (2018).
55. E. Darve and A. Pohorille, J. Chem. Phys. **115**, 9169 (2001).
56. N. B. Wilding, Phys. Rev. E **52**, 602 (1995).
57. J. J. Potoff, A. Z. Panagiotopoulos, J. Chem. Phys. **109**, 10914 (1998).
58. N. B. Wilding, Amer. J. of Phys. **69**, 1147 (2001).
59. T. M. Apostol, The American Mathematical Monthly. Mathematical Association of America. **106**, 409 (1999). doi:10.2307/2589145.

60. B. X. Dong, C. Nowak, J. W. Onorato, J. Strzalka, F. A. Escobedo, C. K. Luscombe, P. F. Nealey and S. N. Patel, *Chem. Mater.* **31**, 1418 (2019).
61. D. M. Huang, R. Faller, K. Do and A. J. Moulé, *J. Chem. Theory Comput.* **6**, 526 (2010).
62. W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives, *J. Am. Chem. Soc.* **118**, 11225 (1996).
63. J. W. Onorato, Z. Wang, Y. Sun, C. Nowak, L. Q. Flagg, R. Li, B. X. Dong, L. J. Richter, F. A. Escobedo, P. F. Nealey, S. N. Patel, and C. K. Luscombe. Side Chain Engineering Control of Mixed Conduction in Oligoethylene Glycol-Substituted Polythiophenes, *J. Mat. Chem. A.*, under review (2021).
64. LAMMPS web page: <https://www.lammps.org/>
65. A. Grossfield, WHAM: the weighted histogram analysis method, version 2.0.10, <http://membrane.urmc.rochester.edu/wordpress/>
66. F.A. Escobedo, *J. Chem. Phys.* **146**, 134508 (2017).
67. F. A. Escobedo, *J. Chem. Phys.* **147**, 214501 (2017).
68. S. Gottschalk, M.C. Lin, and D. Manocha, *Computer Graphics* **30**, 171 (1996).
69. J. Arvo, *Graphic Gems*, Acad. Press, San Diego, CA, 1990.
70. D. Frenkel and A. J. C. Ladd, *J. Chem. Phys.* **81**, 3188 (1984).
71. F. Smallenburg, L. Fillion, M. Marechal, and M. Dijkstra, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 17886 (2012).
72. N. Lu, D. A. Kofke, and T. B. Woolf, *J. Phys. Chem. B* **107**, 5598 (2003).
73. F.A. Escobedo and F.J. Martinez-Veracoechea, *J. Chem. Phys.* **129**, 154107 (2008).
74. M. Lundborg, J. Lidmar, and B. Hess, *J. Chem. Phys.* **154**, 204103 (2021).
75. F.A. Escobedo, *J. Chem. Phys.* **123**, 044110 (2005).
76. F.A. Escobedo, *Phys. Rev. E* **73**, 056701 (2006).
77. C. Nowak and F. A. Escobedo, *J. Chem. Theor. & Comput.* **14**, 5984 (2018).
78. V. Thapar and F.A. Escobedo, *J. Chem. Phys.* **141**, 124117 (2014).
79. A. K. Sharma and F. A. Escobedo, *J. Chem. Phys.* **148**, 184104 (2018).

