

Estimating Free-Energy Surfaces and Their Convergence from Multiple, Independent Static and History-Dependent Biased Molecular-Dynamics Simulations with Mean Force Integration

Antoniu Bjola and Matteo Salvalaglio*





Cite This: *J. Chem. Theory Comput.* 2024, 20, 5418–5427



Read Online

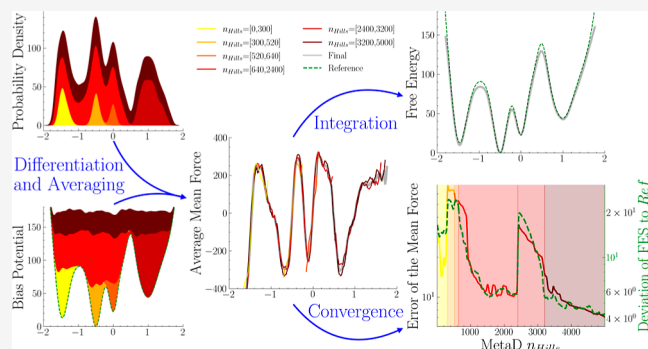
ACCESS |

 Metrics & More

 Article Recommendations

 Supporting Information

ABSTRACT: Addressing the *sampling problem* is central to obtaining quantitative insight from molecular dynamics simulations. Adaptive biased sampling methods, such as metadynamics, tackle this issue by perturbing the Hamiltonian of a system with a history-dependent bias potential, enhancing the exploration of the ensemble of configurations and estimating the corresponding free energy surface (FES). Nevertheless, efficiently assessing and systematically improving their convergence remains an open problem. Here, building on mean force integration (MFI), we develop and test a metric for estimating the convergence of FESs obtained by combining asynchronous, independent simulations subject to diverse biasing protocols, including static biases, different variants of metadynamics, and various combinations of static and history-dependent biases. The developed metric and the ability to combine independent simulations granted by MFI enable us to devise strategies to systematically improve the quality of FES estimates. We demonstrate our approach by computing FES and assessing the convergence of a range of systems of increasing complexity, including one- and two-dimensional analytical FESs, alanine dipeptide, a Lennard–Jones supersaturated vapor undergoing liquid droplet nucleation, and the model of a colloidal system crystallizing via a two-step mechanism. The methods presented here can be generally applied to biased simulations and are implemented in *pyMFI*, a publicly accessible, open-source Python library.



1. INTRODUCTION

Molecular dynamics (MD) simulations have become a powerful tool for studying and predicting the dynamics and thermodynamics of molecular systems. They allow scientists to develop insight into the collective behavior of complex systems at the atomistic scale. The quantitative assessment of equilibrium properties in molecular systems and their interpretation is associated with estimating free-energy surfaces (FES) as a function of a low-dimensional set of collective variables (CVs), s . An FES provides information on a molecular system of interest by quantifying the equilibrium probability of ensembles of configurations corresponding to relevant states and providing information on low-energy transition pathways. However, molecular systems are often characterized by multiple metastable states, separated by high free-energy barriers. This renders transitions between states rare and the sampling necessary to converge thermodynamic properties computationally inaccessible. Numerous methods have been proposed to overcome the sampling problem and enhance the exploration of configuration spaces despite the presence of high-energy barriers.^{1–5}

A subclass of these methods is based on perturbing the system Hamiltonian via an opportunistically defined bias potential,

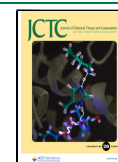
which allows for efficient exploration of the relevant configuration space. Multiple approaches belong to this class.^{3,4,6–9} Among these, two widely used methods are Umbrella sampling (US)^{4,10} and metadynamics (MetaD).^{3,7} These two methods exemplify two opposite and complementary ways of defining a potential to enhance the sampling of rare transitions. US introduces multiple, independent replicas—often referred to as *windows*—in which a time-independent, harmonic bias potential defined in CV space is introduced to localize the sampling on a specific set of configurations. On the other hand, MetaD, as well as other adaptive sampling methods,^{11,12} introduces a history-dependent bias potential that evolves dynamically with the system and, in the long-time limit, provides an estimate of the free energy in the CV space explored. While in US a global FES is obtained by merging the sampling obtained in each window with

Received: January 23, 2024

Revised: May 26, 2024

Accepted: May 28, 2024

Published: June 24, 2024



algorithms such as the weighted histogram analysis method^{6,13} or Umbrella integration (UI),^{14–16} MetaD provides a global FES directly as a function of the cumulative bias potential,^{3,7,8} or via reweighting.^{9,17–20} However, since static bias approaches such as US are based on independent windows, they are trivially parallel and enable a systematic sampling augmentation to reduce the uncertainty of the reconstructed FES. On the other hand, MetaD offers an autonomous exploration of the configuration space, essential when pathways connecting metastable states in CV space are unknown a priori.

Recently, we demonstrated that a single global FES can be obtained from multiple independent asynchronous MetaD replicas by mean force integration (MFI).²¹ Here, we build on this result to develop a systematic approach to combine the information obtained from sampling phase space with multiple independent MD simulations performed under the effect of various biases, both static and history-dependent. For instance, using MFI to reconstruct an FES from independent US and MetaD simulations opens up the possibility of systematically improving the uncertainty of FESs by sampling poorly converged regions. In addition, based on the MFI formalism, we develop a systematic estimation of local and global convergence of the target FES that can be computed on-the-fly and applied to multiple asynchronous replicas subject to different biasing methods. Such measures can be used to inform and systematically improve free-energy estimates by concentrating the sampling in poorly converged regions of CV space.

2. THEORETICAL BACKGROUND: MFI

The FES $F(\mathbf{s})$, as a function of a low-dimensional set of CVs \mathbf{s} , in the presence of a bias potential defined in CV space $V(\mathbf{s})$, can be expressed as^{6,14,21}

$$F(\mathbf{s}) = -k_{\text{B}}T \ln p(\mathbf{s}) - V(\mathbf{s}) - k_{\text{B}}T \ln \langle e^{-\beta V(\mathbf{s})} \rangle \quad (1)$$

where k_{B} is the Boltzmann constant, T is the temperature, and $p(\mathbf{s})$ is the biased probability density sampled under the effect of the bias potential $V(\mathbf{s})$. The term $-k_{\text{B}}T \ln \langle e^{-\beta V(\mathbf{s})} \rangle$ is the reversible work associated with the introduction of the bias potential $V(\mathbf{s})$ in the unperturbed ensemble.

Inspired by UI,^{1,14–16} MFI relies on the calculation of the mean thermodynamic force in \mathbf{s} , $-\nabla_{\mathbf{s}}F(\mathbf{s})$. The advantage of computing the gradient of the free energy in \mathbf{s} , instead of directly estimating $F(\mathbf{s})$, lies in the fact that the former does not require the estimate of the term $-k_{\text{B}}T \ln \langle e^{-\beta V(\mathbf{s})} \rangle$. In fact, such a term is independent of \mathbf{s} and represents a constant offset of the free energy in eq 1. As a consequence, the mean force $-\nabla_{\mathbf{s}}F(\mathbf{s})$ can be estimated from simulations performed under the effect of different bias potentials without evaluating alignment constants between samples obtained under the effect of different biases. We note that several techniques have been developed to obtain such constants for simulations evolving under the effect of static^{13,22} and MetaD^{9,17–20} biases.

By focusing on the derivative of the free energy, MFI²¹ is similar in spirit to force-based methods for the calculation of probability densities from atomistic data,^{23,24} which have been shown to converge to a smoother estimate of the probability density and have been recently used in conjunction with MetaD to accelerate the convergence of FES estimates from ab initio MetaD calculations.²⁵

Moreover, as shown by Awasthi et al.,²⁶ combining US and WTmetaD, that is, static and history-dependent biases, can be

beneficial in several use cases. MFI provides an alternative, flexible framework for combining multiple independent simulations carried out asynchronously with different biasing approaches, thus enabling the systematic refining of FES estimates by mixing data obtained with different sampling protocols.

2.1. MetaD with MFI. In this section, we summarize the main features of MFI applied to obtain an FES from MetaD in a monodimensional CV space s to provide a basis for introducing additional bias potentials and mean force convergence estimators. For history-dependent biasing methods such as MetaD, the simulation time is divided into time windows of constant bias denoted by subscript i . Without loss of generality, the mean force of time-window i , $\frac{dF_i(s)}{ds}$, in a monodimensional CV space s , can be written as^{10,14,21}

$$\frac{dF_i(s)}{ds} = -k_{\text{B}}T \frac{d \ln p_i(s)}{ds} - \frac{dV_i(s)}{ds} \quad (2)$$

where $p_i(s)$ is the bias probability density, sampled during a time-window i under the effect of a constant bias potential $V_i(s)$. In ref 21, it has been shown that this expression holds in the case of a MetaD bias potential updated in *discrete* time-steps.

The average mean force obtained after an arbitrary number of N time-windows can be obtained as a weighted average of $\frac{dF_i(s)}{ds}$, with weights proportional to $p_i(s)$

$$\left\langle \frac{dF(s)}{ds} \right\rangle = \frac{\sum_{i=1}^N \left[p_i(s) \frac{dF_i(s)}{ds} \right]}{\sum_{i=1}^N p_i(s)} \quad (3)$$

Both eq 3 and the first term in eq 2 require an estimate of the bias probability density, which is constructed from configurations sampled during some time-window i . The time-window i starts at time t_i when the system is first sampled under the effect of a new (updated) bias potential $V_i(s)$ and ends with the last sample under the effect of that bias at time $t_i + \tau$, where τ is the time between two consecutive updates of the bias potential. Thus, the bias probability density of time-window i is estimated as a sum of Gaussian kernels

$$p_i(s) = \frac{f_c}{h\sqrt{2\pi}} \sum_{t=t_i}^{t_i+\tau} \exp \left[-\frac{(s-s_t)^2}{2h^2} \right] \quad (4)$$

where h is the bandwidth of the Gaussian kernel and f_c is the height, which corresponds to the sampling rate of configurations. The term f_c/h serves as a scaling factor that facilitates the combination of bias probability densities that employ different h or f_c . This choice leads to the following expression of the first term in eq 2

$$\frac{d \ln p_i(s)}{ds} = \frac{\sum_{t=t_i}^{t_i+\tau} \frac{-(s-s_t)}{h^2} \exp \left[-\frac{(s-s_t)^2}{2h^2} \right]}{\sum_{t=t_i}^{t_i+\tau} \exp \left[-\frac{(s-s_t)^2}{2h^2} \right]} \quad (5)$$

The second term in eq 2 is straightforward. It represents the contribution to the mean force associated with the bias potential $V_i(s)$ and is computed as the derivative of the MetaD potential accumulated up to time-window i as a sum of Gaussian kernels centered in s_i , with height w_i , and width $\sigma_{M,i}$

$$\frac{dV_i(s)}{ds} = -\sum_{i=1}^N \frac{w_i(s-s_i)}{\sigma_{M,i}^2} \exp\left[-\frac{(s-s_i)^2}{2\sigma_{M,i}^2}\right] \quad (6)$$

By combining eqs 2–6, the unbiased average mean force for N consecutive time-windows is estimated in a closed analytical form.²¹ The resulting unbiased average mean force $\langle \frac{dF(s)}{ds} \rangle$ is then integrated numerically to obtain the FES $F(s)$. In analogy with UI, obtaining $F(s)$ by integrating the mean force, while general, is practically limited by the dimensionality of the CV space used to define the bias potential, and therefore, as UI, MFI is practically applicable for CV spaces with dimensionality ≤ 3 . Moreover, integration can introduce numerical errors in the calculation of $F(s)$. As discussed in detail by Kästner in ref 16 for UI, such errors are significantly smaller than inherent sampling errors. This is empirically observed in all the analytical potentials studied in this work, where the MFI estimate of $F(s)$ achieves vanishingly small deviations from the analytical free energies. Finally, integration accuracy and efficiency depend on the algorithm adopted and on the density of the chosen integration grid. In ref 21, we show convergence for increasingly dense grids using finite-differences integration methods. Here, we obtain analogous results by adopting either finite-differences²¹ or a fast Fourier transform integration.¹⁶

2.2. Combining Independent Simulations. As mentioned above, the strength of MFI is the possibility of combining the sampling obtained from *independent* and *asynchronous* simulations. This can be done by appropriately extending the weighted average given by eq 3 to include the mean forces of multiple independent simulations. The resulting equation combines the average mean forces $\langle \frac{dF(s)}{ds} \rangle_j$ of simulations j , with a weight corresponding to the biased probability density of that simulation $H_j(s) = \sum_{i=1}^N p_i(s)$. The combined average mean force over M independently biased trajectories can thus be determined using a weighted average analogous to eq 3, as

$$\left\langle \frac{dF(s)}{ds} \right\rangle = \frac{\sum_{j=1}^M [H_j(s) \langle \frac{dF(s)}{ds} \rangle_j]}{\sum_{j=1}^M H_j(s)} \quad (7)$$

Combining independent simulations in this fashion opens the door to asynchronous MetaD simulation campaigns that may optimize the use of computational resources but also allow for local refinement of the sampling of configurations.

In the following, we discuss how, within the framework of MFI, one can estimate the convergence of average mean force and the FES to identify regions of configuration space that require additional sampling. Later, we lay out methods to combine static and history-dependent bias potentials to systematically improve the sampling as well as parallel and serial simulation structures. We do so by demonstrating these features for Langevin models in mono- and bidimensional CV spaces, for the *ever-present* alanine dipeptide and for a more challenging collective process of liquid droplet nucleation from a supersaturated argon vapor and for a colloidal crystal nucleating from solution. The methods and analyses reported here are implemented in *pyMFI*, a publicly accessible Python library available at <https://github.com/mme-ucl/pyMFI>. Example scripts using *pyMFI* are reported in the [Supporting Information](#).

3. CONVERGENCE

With bias-enhanced sampling simulations, practitioners typically pursue two concurrent objectives. The first is the exploration of relevant configurations that are rarely sampled during standard MD simulations. The second is the estimate of the equilibrium probability of such configurations. History-dependent biased sampling methods combine these two objectives by evolving under the effect of a bias potential that encourages the autonomous exploration of configuration space and that enables the estimate of the equilibrium probability of appropriately defined sets of configuration. Therefore, a useful convergence metric for biased sampling simulations should acknowledge these two complementary aspects by accounting for the uncertainty in the FES calculation and for the extent of the sampling. Typically, the convergence of the FES is estimated using block averaging techniques,^{27,28} and error propagation.²⁹ This approach is carried out *a posteriori*, often requiring the assumption of a time-independent bias, even though time-dependent reweighting techniques can also be used.^{7,18–20}

3.1. On-the-Fly Assessment of Biased Sampling Convergence. With MFI, the time-independent average mean force $\langle \frac{dF(s)}{ds} \rangle$ is estimated as the running weighted average with eq 3. This implies that the uncertainty of the average can be estimated by computing the weighted variance^{30,31} of the mean forces of each time window, with weights proportional to the biased probability density $p_i(s)$ sampled in the respective time window.^{10,14} Employing the notation introduced so far, the variance of the average mean force can be expressed as

$$\sigma^2(s) = \text{BC} \times \left[\frac{\sum_{i=1}^N p_i(s) \left(\frac{dF_i(s)}{ds} \right)^2}{\sum_{i=1}^N p_i(s)} - \left\langle \frac{dF(s)}{ds} \right\rangle^2 \right] \quad (8)$$

where $\sigma(s)$ is the weighted standard deviation, and BC represents the Bessel correction for the variance of the weighted mean, defined as

$$\text{BC} = \frac{n_{\text{eff}}(s)}{n_{\text{eff}}(s) - 1} = \frac{[\sum_{i=1}^N p_i(s)]^2}{[\sum_{i=1}^N p_i(s)]^2 - \sum_{i=1}^N [p_i(s)]^2} \quad (9)$$

where $n_{\text{eff}} = [\sum_{i=1}^N p_i(s)]^2 / \sum_{i=1}^N [p_i(s)]^2$ is the effective sample size. The standard error of the weighted mean is expressed as

$$\sigma_E(s) = \sqrt{\frac{\sigma^2(s)}{n_{\text{eff}}(s)}} \quad (10)$$

All the terms appearing in eqs 8–10 can thus be computed based on the history of the simulation up to time-window i . As such, the local estimate of the variance obtained from eqs 8–10 provides an *on-the-fly, local* measure of convergence of the mean thermodynamic force. By averaging $\sigma_E(s)$ value over the *sampled* CV space, we can obtain a global convergence estimator of the mean force: $\overline{\sigma}_E$.

Furthermore, we note that $\overline{\sigma}_E$ can only be evaluated in *explored* regions of s since in unexplored regions, the force is not determined. To develop a convergence metric incorporating both the uncertainty in the FES calculation and the extent of the configurational volume explored, we divide σ_E by ν , a measure of the volume of CV space explored by the simulation. In this work, we express ν as a nondimensional quantity by

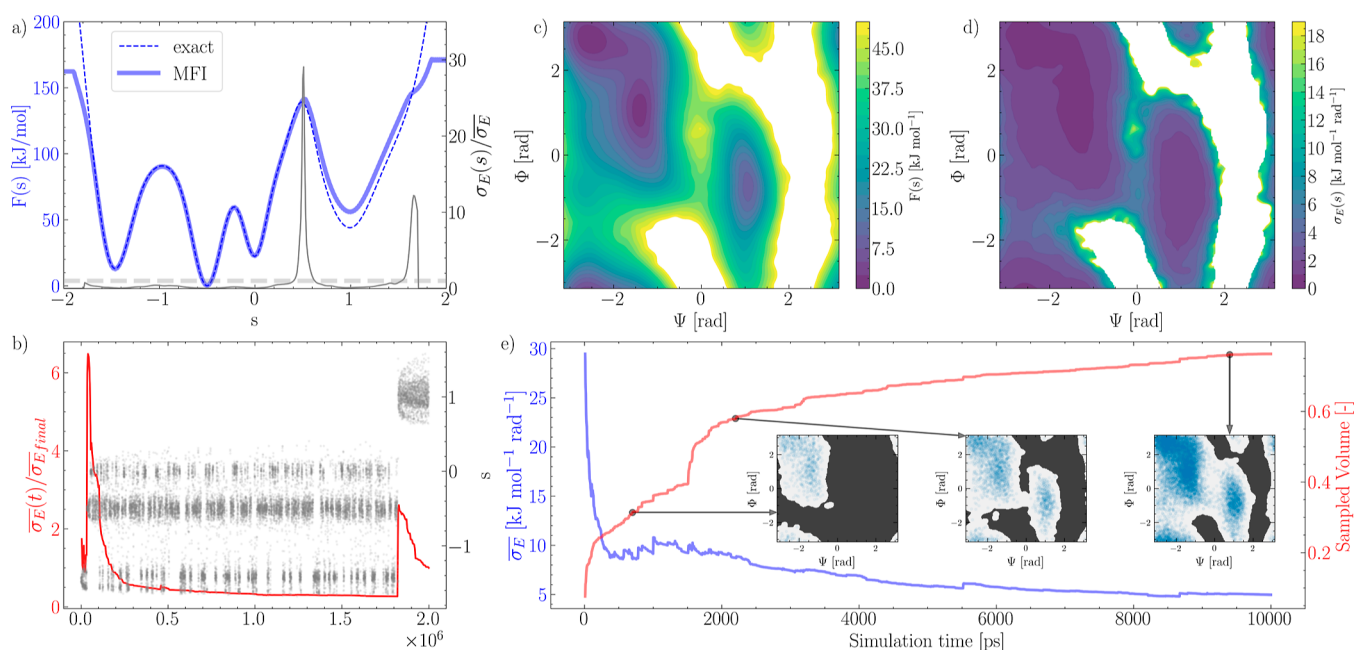


Figure 1. (a) Comparison between the MFI estimate of a monodimensional multiwell FES obtained postprocessing a single MetaD simulation (solid blue line) with the exact FES (dashed blue line), and the normalized standard error of the mean force, $\sigma_E(s)/\bar{\sigma}_E$ (solid gray line), normalized by its average value over the sampled CV space, $\bar{\sigma}_E$ (dashed gray line). (b) CV space exploration (gray symbols) and normalized average variance of the mean force (solid red line). (c) Alanine dipeptide FES, function of the Φ and Ψ dihedral angles. (d) Standard error of the mean force mapped in the Φ and Ψ torsional angles. (e) Average standard error of the mean force (blue) and ratio of configurational space explored (red) as a function of the simulation time.

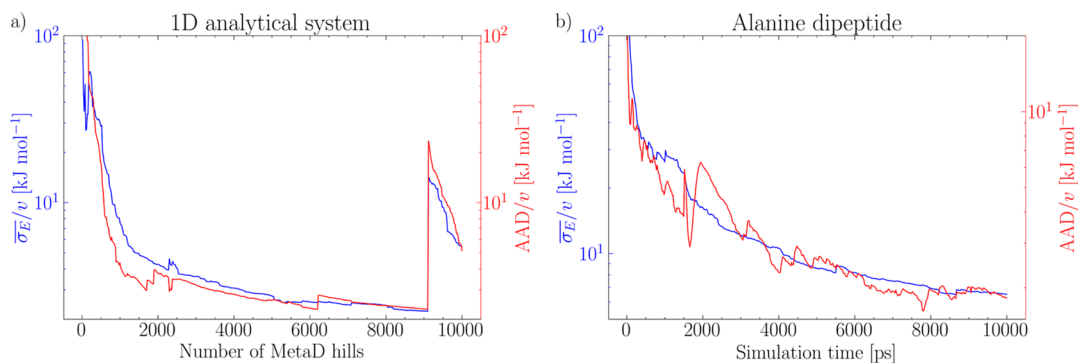


Figure 2. Comparing the dynamic evolution of the average error of the mean force normalized by the sampled volume (*left y-axis*) and the AAD normalized by the sampled volume (*right y-axis*). The former can be computed on-the-fly based on the history of the biased simulation without knowledge of an external reference FES. The latter can only be estimated only for toy examples with a known exact FES. (a) Multiwell one-dimensional FES (see Figure 1a). (b) Alanine dipeptide (see Figure 1c).

dividing the volume of the sampled configuration in CV Space by the total volume of the domain in CV space considered to be relevant for the problem studied. The former is evaluated as the volume of the CV domain where the biased histogram of configurations $H_i(s)$ is larger than a lower-bound arbitrary threshold.

The different components of the convergence estimator $\bar{\sigma}_E/v$ are illustrated in Figure 1: panels a and c report the FES computed by integrating the average mean thermodynamic force computed with eq 3 for a monodimensional Langevin dynamics model evolving on an analytical potential defined as

$$F_{\text{exact}}(s) = s^8 - 103e^{-(s+1.5)^2/0.07} - 93e^{-(s+0.5)^2/0.05} - 70e^{-s^2/0.03} + 50e^{-(s-0.5)^2/0.03} - 50e^{-(s-1)^2/0.1}$$

and for Alanine Dipeptide, a typical example of a two-dimensional FES.

In Figure 1a (in gray, right axis) and Figure 1d, we report for these two systems $\sigma_E(s)$, the standard error of the weighted mean mapped onto the CV space s . In Figure 1a, $\sigma_E(s)$ is normalized by its average value estimated over the entire sampled CV space, $\bar{\sigma}_E$.

In Figure 1b, it can be seen that the dynamic evolution of $\bar{\sigma}_E(t)$ captures the overall convergence of the sampling in s . In particular, it can be seen that $\bar{\sigma}_E(t)$ relaxes rapidly, with the system reversibly sampling numerous transitions between the multiple local basins characterizing the FES for $s < 0.5$. Such relaxation is representative of the increasing confidence in the estimate of the local mean force. When the system discovers a new metastable state, for $s > 0.5$ at $\approx 1.7 \times 10^6$ steps, the mean force in the newly discovered metastable state is characterized

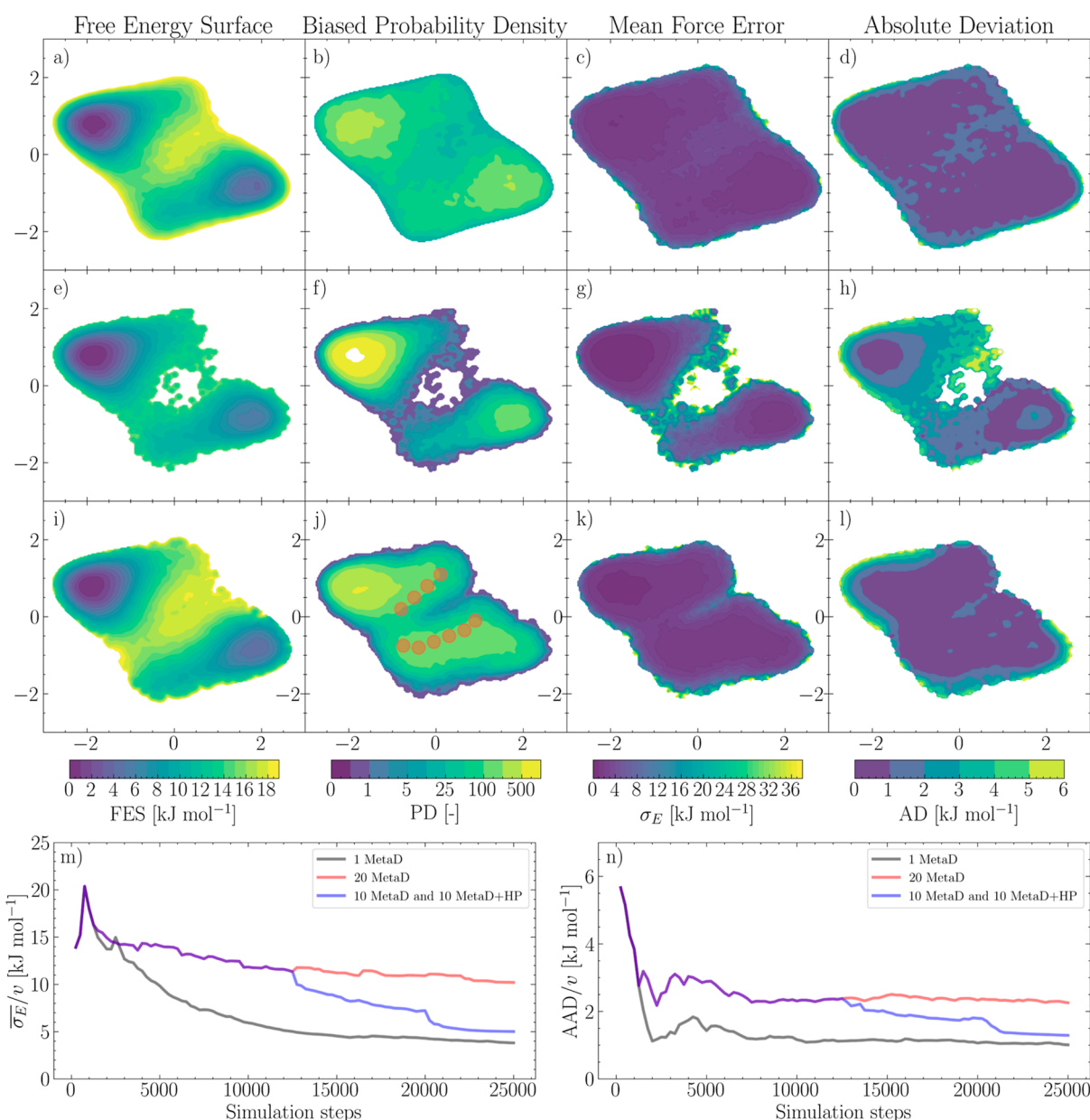


Figure 3. FESs from independent, biased simulations. FES, total biased probability density ($\sum_{j=1}^M H_j(s)$), CV map of the mean force error, and absolute deviation from the analytical FES for a two-dimensional double-well model potential. (a–d) Single, long WTmetaD simulation (e–h) twenty randomly initialized WTmetaD simulations (i–l) ten randomly initialized WTmetaD simulations and 10 WTmetaD simulations are subject to a two-dimensional harmonic potential localized in CV space [the harmonic potential centers are represented as circles in (j)]. (m) Error of the mean force normalized by the explored CV-space volume as a function of the number of simulation steps. (n) AAD from the analytical FES, normalized by the explored CV-space volume as a function of the number of simulation steps.

by a lower level of confidence, thus resulting in a sudden increase in $\overline{\sigma_E}(t)$.

A similar trend can be seen in Figure 1e, displaying the dynamic evolution of $\overline{\sigma_E}(t)$ of alanine dipeptide (blue, left axis) together with the total CV space sampled (red, right axis). At the start of the simulation, the low-energy region of the left basin is being sampled, causing an immediate decrease in the mean force error. As the height of the MetaD potential increases, higher energy regions are discovered, and $\overline{\sigma_E}(t)$ stops decreasing.

It should be noted that $\sigma_E(s)$, and $\overline{\sigma_E}$ enable to assess the convergent behavior of the simulations but provide only a qualitative measure of the real error in the mean force due to

correlations in the mean force that might be present at small sample sizes. Notably, however, the normalized convergence estimator $\overline{\sigma_E}/v$ correlates strongly with the normalized average absolute deviation (AAD/ v) of the FES from an independent reference, as shown in Figure 2a for the monodimensional FES and for alanine dipeptide in Figure 2b. This is an important observation as, while the AAD is arguably an objective and accurate measure of convergence, it cannot be computed for any realistic application, while $\overline{\sigma_E}/v$ can be computed on-the-fly and enables a systematic assessment of FES convergence.

3.2. Convergence for Sets of Independent Biased Simulations. The measure of convergence discussed in the previous section can be naturally extended to cases where the

FES is computed from multiple independent simulations by computing the variance of the time-averaged mean force over M independent simulations (see eq 7) as

$$\sigma^2(s) = \text{BC} \times \left[\frac{\sum_{j=1}^M H_j(s) \left\langle \frac{dF(s)}{ds} \right\rangle_j^2}{\sum_{j=1}^M H_j(s)} - \left\langle \frac{dF(s)}{ds} \right\rangle^2 \right] \quad (11)$$

where the Bessel correction now takes into account the total weight of each simulation via $n_{\text{eff}}(s)$ defined as

$$n_{\text{eff}}(s) = \frac{[\sum_{j=1}^M H_j(s)]^2}{\sum_{j=1}^M [H_j(s)]^2} \quad (12)$$

For such cases, the mean forces obtained from individual simulations are not correlated and can be used to estimate the FES error with a bootstrapping analysis.²⁸ This is demonstrated in the first example of Section 5.

4. COMBINING MULTIPLE BIASES AND INDEPENDENT SIMULATIONS

The ability to estimate the local convergence of the mean force in CV space *on-the-fly* combined with the possibility of locally refining the sampling in specific regions of phase space by merging the information obtained from independent simulations through eq 7 suggests that to efficiently achieve local refinement of free energy estimates one may aim to apply different biasing strategies at different stages of exploration/convergence. In particular, to locally enhance the accuracy of FES estimates, it might be useful to combine static and time-dependent biases. By introducing n_{B} bias potentials that simultaneously act on a system, the mean force of a generic time-window i becomes

$$\frac{dF_i(s)}{ds} = -k_{\text{B}}T \frac{d \ln p_i(s)}{ds} - \sum_k^{n_{\text{B}}} \frac{dV_k(s)}{ds} \quad (13)$$

where the term $\sum_k^{n_{\text{B}}} \frac{dV_k(s)}{ds}$ can include any differentiable static bias. Typical choices that enable to focus the sampling in specific regions of the CV space of interest include harmonic potentials upper and lower walls, but static biases are not restricted to these cases.

To demonstrate the feasibility of combining different biases and using an on-the-fly convergence metric to monitor the behavior of the calculations, we investigate a two-dimensional double-well analytical potential often used in the enhanced sampling literature.^{12,32} The exact free energy for such model is defined as $F_{\text{exact}}(s_1, s_2) = 1.35s_1^4 + 1.90s_1^3s_2 + 3.93s_1^2s_2^2 - 6.44s_1^2 - 1.90s_1s_2^3 + 5.59s_1s_2 + 1.33s_1 + 1.35s_2^4 - 5.56s_2^2 + 0.90s_2 + 18.59$.

4.1. Combining Multiple and Three Short MetaD Simulations. As discussed in the introductory section, an advantage of MFI is the ability to construct an FES from multiple asynchronous MetaD simulations. This can be exploited by running simulations in a trivially parallel manner and increasing the sampling for configurations that, while having been visited by previous trajectories, are far from convergence. This provides additional flexibility for running MetaD simulations and allows for a more efficient use of computational resources.

To present this feature, we compare a long MetaD simulation with 20 short simulations totaling the same number of steps. The short MetaD simulations are performed under a

set of parameters identical to the long simulation (see the Supporting Information for details) and are initialized in a random configuration. We note that, for atomistic examples, random reinitialization would only be possible within the set of configurations that have already been explored by existing trajectories. This is demonstrated and discussed in the Results section for the condensation of a LJ vapor, where many short simulations are initialized from the two known metastable states, that is, liquid droplet in vapor and supersaturated vapor, and in ref 33, where an analogous approach is used to study the free-energy landscape associated with reactive events.

The long simulation, depicted in the first row of Figure 3, shows a similar result to the alanine dipeptide example reported in Figure 1c–e. In comparison, the short simulations depicted in the second-row sample less of the configuration space. Given the same simulation and biasing parameters, the short simulations do not have the time to build up a MetaD potential that enables a reversible crossing of the free energy barriers separating two metastable states. However, crucially, at least some of these simulations can undergo a transition. As a result, even if no back and forth *recrossings* are observed in any individual simulation, the transition region is captured with a moderate error, while the relative stability of the metastable states is captured with accuracy comparable to that granted by the long simulation reference (top row of Figure 3). Nonetheless, for a more accurate estimate of the FES, further sampling of the transition region is required, which can be done efficiently by employing other biasing methods.

4.2. Combining MetaD with Static Harmonic Potentials. The results presented in the section above exemplify the fairly typical situation in which the convergence of an FES needs to be locally improved in specific regions of CV-space. In such cases, additional simulations can be performed using a combination of biases aimed to locally improve the accuracy of the free energy estimate for configurations mapped in those undersampled regions. To demonstrate this approach, we consider improving the FES calculation for the first ten short simulations from those employed in the section above. In Figure 3f, it can be seen that the region connecting the two basins is poorly sampled. To increase sampling in this region, we perform ten additional short MetaD simulations subject also to a two-dimensional harmonic potential centered in the red circles depicted in Figure 3j. The combination of ten exploratory MetaD and ten MetaD simulations localized in a specific region CV-space leads to the results reported in the third row of Figure 3. As can be observed, the reconstructed FES provides a better estimate of the equilibrium probability not only in the local minima corresponding to metastable states but also for high-free energy, low-probability configurations.

In Figure 3m,n, we report the on-the-fly convergence to measure $\overline{\sigma}_{\text{E}}/\nu$ side by side with the AAD/ ν . It can be seen that the global convergence of the different sets of simulations can be systematically monitored and compared, quantitatively capturing the improvement introduced by focusing on undersampled CV regions with time-independent harmonic potentials.

This analysis shows that the flexibility granted by using multiple biasing potentials together with independent short simulations enables to obtain, monitor and improve convergence with independent simulations subject to different biases.

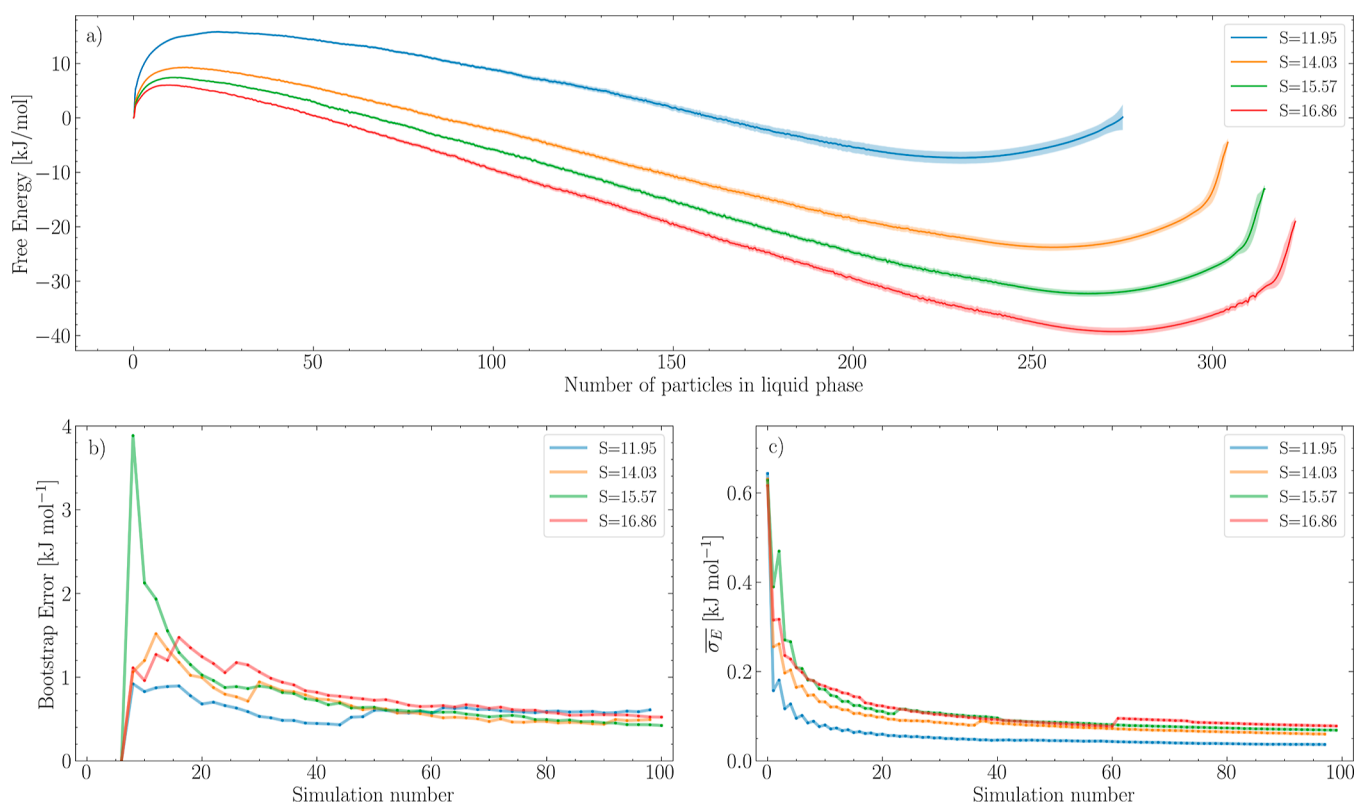


Figure 4. Combining multiple short simulations to estimate the FES associated with the nucleation of a liquid droplet from supersaturated vapor. (a) FES associated with the number of molecules in the liquid phase³⁴ (see the [Supporting Information](#) for the simulation details) at different supersaturation levels. The shaded region represents the standard deviation calculated with a bootstrap analysis. (b) Progression of the standard deviation of the FES as a function of bootstrap iterations for various levels of supersaturation. (c) Progression of the standard error of the mean force as a function of the number of simulations for various levels of supersaturation.

5. APPLICATIONS

In this section, we discuss how the approaches illustrated on simple model systems in the previous sections can be used to monitor and improve the convergence of FESs for complex processes. We focus our attention on modeling nucleation, a task where converging FESs are often limited by the inherent slow dynamics in CV space and where combining multiple simulations enables us to improve our ability to obtain accurate FESs.

5.1. Liquid Droplet Nucleation from a Supersaturated Vapor. The first application is the numerical calculation of FESs associated with the nucleation of a liquid droplet from a supersaturated vapor, a rare-event process initiating the condensation of a liquid phase. In this case, a system with 512 argon atoms was simulated under the canonical ensemble at four increasing supersaturation levels. Sampling this process with brute force simulations is extremely impractical and only possible for *billion* atom simulations.³⁶ However, even enhanced sampling techniques such as MetaD, while being instrumental in efficiently recovering the kinetics of nucleation,³⁴ are rather inefficient at determining the full FES. This is due to the fact that a large, asymmetric, free energy barrier separates the basins corresponding to the metastable parent phase and the stable state. Moreover, the characteristic fluctuations in CV space are orders of magnitude different in the two states. As such, different WTmetaD parameters (such as Gaussian width and bias factor) are desirable for efficiently sampling the forward (condensation) and backward (evaporation) transitions.

Here, we show that the biasing strategy can be tailored specifically to this problem by using MFI to postprocess the simulation results. For the forward transition, a WTmetaD bias is applied with narrow hills, whereas for the backward transition, a MetaD bias with wider hills is applied. Moreover, given the asymmetry of the barrier, forward transitions can take place with much smaller bias factors than backward transitions. Additionally, for the lowest supersaturation level, a higher energy barrier is expected, and thus, an additional static bias potential is added to the forward simulation, increasing the efficiency of the construction of the bias potential necessary to observe nucleation events. To adequately sample the whole FES and provide sufficient data sampling configurations that cross the free energy barrier, 50 forward simulations and 50 backward simulations were conducted until the other stable state was reached. This protocol was repeated for various levels of supersaturation. Additional information regarding the simulation setup is reported in the [Supporting Information](#). The forces from all trajectories were calculated and patched together with MFI to find the FES, depicted in [Figure 4a](#), and the convergence was monitored with the on-the-fly error of the mean force, illustrated in [Figure 4c](#). Additionally, a bootstrap analysis was performed on the independent forces, yielding an uncorrelated standard deviation of the FES. That was used as error bars for the FES, and the progression of the global average is presented in [Figure 4b](#). The shape of the FES of the nucleation event is captured well, and the bootstrap analysis indicates a low error in the transition region, whereas the tail of the FES has a larger error.

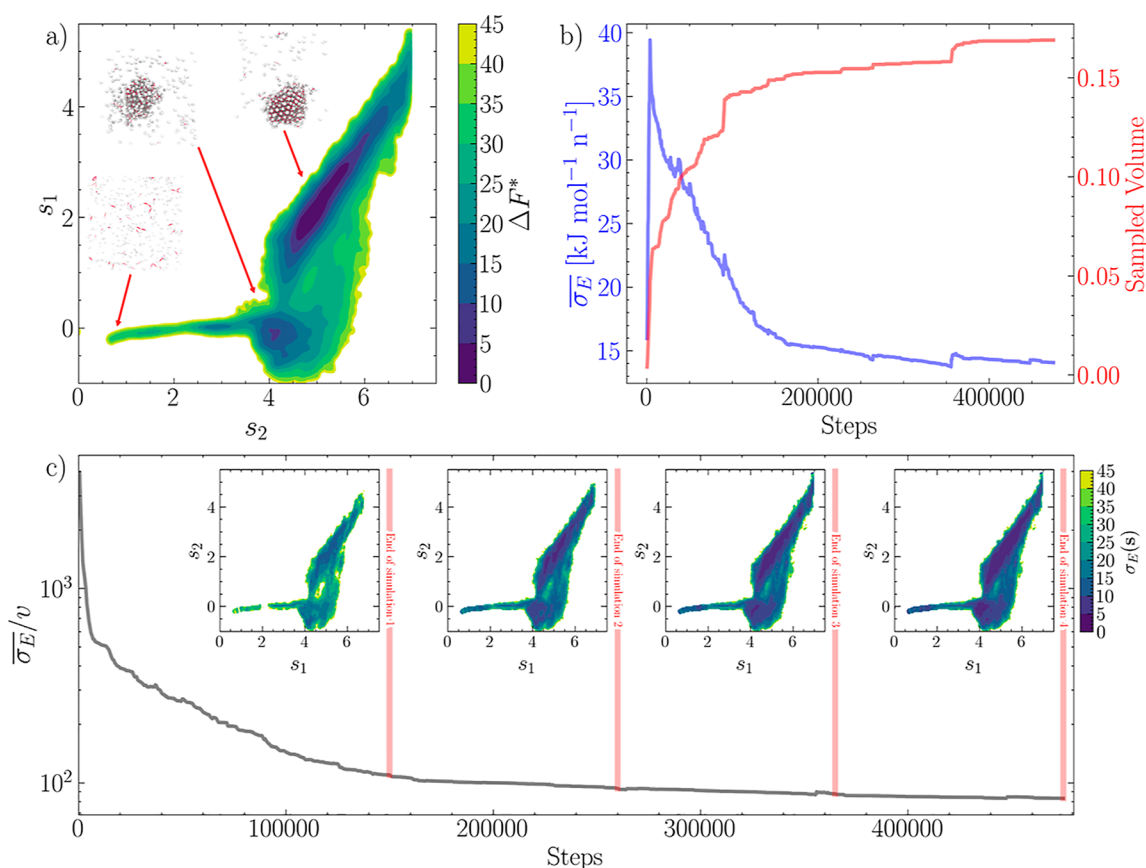


Figure 5. Improving and monitoring the convergence of FES by combining multiple independent WTmetaD simulations. (a) Converged FES (in reduced units) obtained merging four independent WTmetaD and MetaD simulations performed with different biasing parameters. The CVs s_1 and s_2 are GNN-based approximations of nucleation collective variables n and $n(Q6)$ discussed in detail in ref 35. (b) Time series of the configurational volume explored by the simulations (blue) and the CV-space averaged value of the standard error of the mean force in CV space. (c) Global $\overline{\sigma_E}/\nu$ and local convergence estimators, demonstrating a systematic convergence of the FES associated with providing additional data obtained from independent simulations performed with different biasing setups.

5.2. Two-Step Crystallization of a Colloidal System.

To demonstrate the application of the $\overline{\sigma_E}/\nu$ convergence estimator when deploying a set of independent simulations in a complex application, we analyze the convergence behavior of multiple MetaD calculations modeling a colloidal system undergoing a two-step nucleation process.³⁷ The CVs used to describe this process are n , counting the number of particles with a coordination number above some threshold, and $n(Q6)$, counting the number of particles with a local Steinhardt order parameter ($Q6$) above a representative threshold. In such simulations, the CVs are efficiently computed via a graph neural network (GNN) model, which offers orders-of-magnitude gains in computational efficiency in the on-the-fly evaluation of the CVs necessary when conducting biased sampling. Additional information about the GNN method and the simulation details can be found in ref 35 and in the [Supporting Information](#).

The system modeled consists of 421 particles in a cubic box of length 92.83σ , modeled via a Derjaguin–Landau–Verwey–Overbeek^{38–40} potential. All simulations were performed in the NVT ensemble, tempered at $2T^*$. This investigation entailed four independent MetaD simulations. Three simulations utilized WTmetaD with varying bias factors, while the fourth was a nontempered MetaD simulation. Additional simulation details^{41–48} are reported in the [Supporting Information](#), input files necessary to reproduce the relevant examples are available

on PLUMED-NEST.⁴⁹ The configurations sampled across all simulations were postprocessed and combined into a single FES using MFI. The combined FES is shown in [Figure 5a](#), where two main basins can be seen. The deeper basin at large values of both s_1 and s_2 represents configurations embedding a crystalline particle surrounded by a vapor of colloidal particles in dynamic equilibrium with each other. The second basin, at low s_2 , represents configurations where a dense liquid droplet is present in the simulation box as a metastable intermediate along the crystallization pathway. These two metastable states are separated by a free energy barrier around $n(Q6) \approx 1$. In [Figure 5b](#), the progression of the on-the-fly error of the mean force and the progression of the sampled volume is shown. While the error $\overline{\sigma_E}$ (blue line) shows a decreasing trend, there are occasional upward fluctuations corresponding to an increase in the explored space ν (red line).

The overall convergent behavior of the four combined simulations is exemplified by the convergence measure $\overline{\sigma_E}/\nu$, reported in [Figure 5c](#), clearly demonstrating the systematic improvement of the FES obtained by combining self-consistent data generated via independent MetaD simulations. A mapping of σ_E in CV space is reported in the insets of [Figure 5c](#), demonstrating that the overall, average convergent behavior of the set of simulations performed here is indeed accompanied by an overall convergent behavior across the entire CV space, and it is not dictated by a local reduction of the error $\sigma_E(s)$.

6. CONCLUSIONS

In this work, we developed a measure of the convergence of the sampling of configurational FESs based on MFI and biased dynamics. Recognizing that convergence, in the context of adaptive sampling, refers to the ability to visit new and relevant molecular configurations, as well as the accurate determination of their equilibrium probability, the convergence measure that we propose is the error of the *mean force* in CV space, normalized by a measure of the *volume* explored in CV space. We show that this measure of biased sampling convergence can be computed on-the-fly and correlates strongly with the FES error computed a posteriori via bootstrapping independent simulations or block averaging. In addition to providing a measure of convergence, we show with examples and applications that, postprocessing biased simulations with MFI, one can combine different static and dynamic biases, thereby targeting the convergence of FES in specific regions of CV space with significant flexibility. Combining the sampling obtained under the effect of different biases also enables us to systematically improve on simulations performed with suboptimal setups without discarding data, thus making the most of the often unreported computing time allocated to fine-tune bias simulation parameters. The convergence estimators, as well as different strategies for combining biases, are demonstrated with a range of examples of increasing complexity, including analytical models, conformational changes of alanine dipeptide, the nucleation of a liquid argon droplet from a supersaturated vapor and the nucleation of a colloidal crystal from solution. All examples are implemented via the pyMFI Python library, which is publicly accessible at <https://github.com/mme-ucl/pyMFI>. Use cases and simple examples of the use of pyMFI to postprocess biased simulations are provided online and in the [Supporting Information](#).

■ ASSOCIATED CONTENT

Data Availability Statement

The input files for the simulations using PLUMED are available at PLUMED NEST under the ID plumID:24.013 (<https://www.plumed-nest.org/eggs/24/013/>).⁴⁹ The MFI method, implemented in the “pyMFI” Python library, can be accessed at <https://github.com/mme-ucl/pyMFI>. Instructions for using this library to reproduce the results of this work are provided as Jupyter Notebooks at <https://github.com/mme-ucl/MFI>.

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jctc.4c00091>.

Scripts employing the pyMFI library, analysis of the convergence estimator dependence on the kernel bandwidth, and simulation setup details ([PDF](#))

■ AUTHOR INFORMATION

Corresponding Author

Matteo Salvalaglio – Thomas Young Centre and Department of Chemical Engineering, University College London, London WC1E 7JE, U.K.; orcid.org/0000-0003-3371-2090;
Email: m.salvalaglio@ucl.ac.uk

Author

Antoniù Bjola – Thomas Young Centre and Department of Chemical Engineering, University College London, London WC1E 7JE, U.K.

Complete contact information is available at:
<https://pubs.acs.org/10.1021/acs.jctc.4c00091>

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

M.S. acknowledges funding from the Crystallization in the Real World EPSRC Programme Grant (EP/R018820/1) and from the ht-MATTER UKRI Frontier Research Guarantee Grant (EP/X033139/1). M.S. and A.B. thank Florian Dietrich for the simulation data analyzed in Figure S, and Francesco Serse for his feedback and contribution to the pyMFI library.

■ REFERENCES

- (1) Kirkwood, J. Statistical mechanics of fluid mixtures. *J. Chem. Phys.* **1935**, *3*, 300–313.
- (2) Geissler, P.; Dellago, C.; Chandler, D. Kinetic pathways of ion pair dissociation in water. *J. Phys. Chem. B* **1999**, *103*, 3706–3710.
- (3) Laio, A.; Parrinello, M. Escaping free-energy minima. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12562–12566.
- (4) Torrie, G.; Valleau, J. Nonphysical sampling distributions in monte carlo free-energy estimation: Umbrella sampling. *J. Comput. Phys.* **1977**, *23*, 187–199.
- (5) Earl, D.; Deem, M. Parallel tempering: Theory, applications, and new perspectives. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3910–3916.
- (6) Roux, B. The calculation of the potential of mean force using computer simulations. *Comput. Phys. Commun.* **1995**, *91*, 275–282.
- (7) Barducci, A.; Bonomi, M.; Parrinello, M. Metadynamics. *WIREs Comput. Mol. Sci.* **2011**, *1*, 826–843.
- (8) Barducci, A.; Bussi, G.; Parrinello, M. Well-tempered metadynamics: A smoothly converging and tunable free-energy method. *Phys. Rev. Lett.* **2008**, *100* (2), 020603.
- (9) Bonomi, M.; Barducci, A.; Parrinello, M. Reconstructing the equilibrium Boltzmann distribution from well-tempered metadynamics. *J. Chem. Phys.* **2009**, *30*, 1615–1621.
- (10) Kästner, J. Umbrella sampling. *WIREs Comput. Mol. Sci.* **2011**, *1*, 932–942.
- (11) Valsson, O.; Parrinello, M. Variational approach to enhanced sampling and free energy calculations. *Phys. Rev. Lett.* **2014**, *113*, 090601.
- (12) Invernizzi, M.; Parrinello, M. Rethinking metadynamics: From bias potentials to probability distributions. *J. Phys. Chem. Lett.* **2020**, *11* (7), 2731–2736.
- (13) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. The weighted histogram analysis method for free-energy calculations on biomolecules. i. the method. *J. Comput. Chem.* **1992**, *13*, 1011–1021.
- (14) Kästner, J.; Thiel, W. Bridging the gap between thermodynamic integration and umbrella sampling provides a novel analysis method: “umbrella integration. *J. Chem. Phys.* **2005**, *123* (14), 144104.
- (15) Kästner, J.; Thiel, W. Analysis of the statistical error in umbrella sampling simulations by umbrella integration. *J. Chem. Phys.* **2006**, *124* (23), 234106.
- (16) Kästner, J. Umbrella integration in two or more reaction coordinates. *J. Chem. Phys.* **2009**, *131*, 034109.
- (17) Branduardi, D.; Bussi, G.; Parrinello, M. Metadynamics with adaptive gaussians. *J. Chem. Theory Comput.* **2012**, *8*, 2247–2254.
- (18) Tiwary, P.; Parrinello, M. A time-independent free energy estimator for metadynamics. *J. Phys. Chem. B* **2015**, *119*, 736–742.
- (19) Giberti, F.; Cheng, B.; Tribello, G.; Ceriotti, M. Iterative unbiasing of quasi-equilibrium sampling. *J. Chem. Theory Comput.* **2020**, *16*, 100–107.

- (20) Schäfer, T. M.; Settanni, G. Data reweighting in metadynamics simulations. *J. Chem. Theory Comput.* **2020**, *16*, 2042–2052.
- (21) Marinova, V.; Salvalaglio, M. Time-independent free energies from metadynamics via mean force integration. *J. Chem. Phys.* **2019**, *151*, 10.
- (22) Shirts, M. R.; Chodera, J. D. Statistically optimal analysis of samples from multiple equilibrium states. *J. Chem. Phys.* **2008**, *129* (12), 124105.
- (23) de Las Heras, D.; Schmidt, M. Better than counting: Density profiles from force sampling. *Phys. Rev. Lett.* **2018**, *120*, 218001.
- (24) Rotenberg, B. Use the force! reduced variance estimators for densities, radial distribution functions, and local mobilities in molecular simulations. *J. Chem. Phys.* **2020**, *153*, 150902.
- (25) Agosta, L.; Brandt, E. G.; Lyubartsev, A. Improved sampling in ab initio free energy calculations of biomolecules at solid–liquid interfaces: Tight-binding assessment of charged amino acids on tio2 anatase (101). *Computation* **2020**, *8* (1), 12.
- (26) Awasthi, S.; Kapil, V.; Nair, N. Sampling free energy surfaces as slices by combining umbrella sampling and metadynamics. *J. Comput. Chem.* **2016**, *37*, 1413–1424.
- (27) Frenkel, D.; Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications, Chapter D.3 Block Averages*; Academic Press, 1996.
- (28) Grossfield, A.; Patrone, P.; Roe, D.; Schultz, A. J.; Siderius, D.; Zuckerman, D. Best practices for quantification of uncertainty and sampling quality in molecular simulations [article v1.0]. *Living J. Comput. Mol. Sci.* **2019**, *1*, 5067.
- (29) Gimondi, I.; Tribello, G.; Salvalaglio, M. Building maps in collective variable space. *J. Chem. Phys.* **2018**, *149*, 104104.
- (30) Bevington, P.; Robinson, D. *Data Reduction and Error Analysis for the Physical Sciences*; McGraw-Hill, 1969.
- (31) Kirchner, J. *Weighted Averages and Their Uncertainties*, 2006.
- (32) Invernizzi, M.; Parrinello, M. Making the best of a bad situation: A multiscale approach to free energy calculation. *J. Chem. Theory Comput.* **2019**, *15*, 2187–2194.
- (33) Serse, F.; Bjola, A.; Salvalaglio, M.; Pelucchi, M. Unveiling solvent effects on β -scissions through metadynamics and mean force integration. *ChemRxiv* **2024**.
- (34) Salvalaglio, M.; Tiwary, P.; Maggioni, G.; Mazzotti, M.; Parrinello, M. Overcoming time scale and finite size limitations to compute nucleation rates from small scale well tempered metadynamics simulations. *J. Chem. Phys.* **2016**, *145*, 211925.
- (35) Dietrich, F. M.; Advincola, X. R.; Gobbo, G.; Bellucci, M. A.; Salvalaglio, M. Machine learning nucleation collective variables with graph neural networks. *J. Chem. Theory Comput.* **2024**, *20*, 1600–1611.
- (36) Diemand, J.; Angéil, R.; Tanaka, K.; Tanaka, H. Large scale molecular dynamics simulations of homogeneous nucleation. *J. Chem. Phys.* **2013**, *139*, 074309.
- (37) Finney, A.; Salvalaglio, M. A variational approach to assess reaction coordinates for two-step crystallization. *J. Chem. Phys.* **2023**, *158*, 094503.
- (38) Derjaguin, B.; Landau, L. The theory of stability of highly charged lyophobic sols and coalescence of highly charged particles in electrolyte solutions. *Acta Physicochimica U.R.S.S.* **1941**, *14*, 633–662.
- (39) Verwey, E.; Overbeek, J. *Theory of the Stability of Lyophobic Colloids: The Interaction of Sol Particles Having an Electric Double Layer*; Elsevier, 1962.
- (40) Loeb, A.; Overbeek, J.; Wiersema, P.; King, C. The electrical double layer around a spherical colloid particle. *J. Electrochem. Soc.* **1961**, *108*, 269.
- (41) Thompson, A.; Aktulga, H.; Berger, R.; Bolintineanu, D.; Brown, W.; Crozier, P.; in 't Veld, P. J.; Kohlmeyer, A.; Moore, S.; Nguyen, T. D.; et al. LAMMPS—a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Comput. Phys. Commun.* **2022**, *271* (2), 108171.
- (42) Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in Neural Information Processing Systems 32 (NeurIPS 2019)*; Wallach, H., Larochelle, H., Beygelzimer, A., d Alché-Buc, F., Fox, E., Garnett, R., Eds; Curran Associates, Inc., 2019; Vol. 32.
- (43) McGibbon, R.; Beauchamp, K.; Harrigan, M.; Klein, C.; Swails, J.; Hernández, C.; Schwantes, C.; Wang, L.; Lane, T.; Pande, V. Mdtraj: A modern open library for the analysis of molecular dynamics trajectories. *Biophys. J.* **2015**, *109*, 1528–1532.
- (44) Michaud-Agrawal, N.; Denning, E.; Woolf, T.; Beckstein, O. Mdanalysis: A toolkit for the analysis of molecular dynamics simulations. *J. Comput. Chem.* **2011**, *32*, 2319–2327.
- (45) *OSTI.GOV. Mdanalysis: A python Package for the Rapid Analysis of Molecular Dynamics Simulations*, 2016.
- (46) Giorgino, T. Pycv: a plumed 2 module enabling the rapid prototyping of collective variables in python. *J. Open Source Software* **2019**, *4*, 1773.
- (47) Bradbury, J.; Frostig, R.; Hawkins, P.; Johnson, M. J.; Leary, C.; Maclaurin, D.; Necula, G.; Paszke, A.; VanderPlas, J.; Wanderman-Milne, S.; et al. Jax: composable transformations of python + numpy programs, 2018, <http://github.com/google/jax> (accessed 18 Aug 2023).
- (48) Heek, J.; Levskaya, A.; Oliver, A.; Ritter, M.; Rondepierre, B.; Steiner, A.; van Zee, M. Flax: A neural network library and ecosystem for jax. 2020, <http://github.com/google/flax> (accessed Aug 18, 2023).
- (49) The PLUMED consortium. Promoting transparency and reproducibility in enhanced molecular simulations. *Nat. Methods* **2019**, *16*, 670–673.