Unified Approach to Enhanced Sampling

Michele Invernizzi[1,2,*] Pablo M. Piaggi[3] and Michele Parrinello[4,2,†]

[1]*Department of Physics, ETH Zurich, c/o Università della Svizzera italiana,*
6900 Lugano, Switzerland
[2]*Facoltà di Informatica, Institute of Computational Science,*
National Center for Computational Design and Discovery of Novel Materials (MARVEL),
Università della Svizzera italiana, 6900 Lugano, Switzerland
[3]*Department of Chemistry, Princeton University, Princeton, New Jersey 08544, USA*
[4]*Department of Chemistry and Applied Biosciences, ETH Zurich, c/o Università della Svizzera italiana,*
6900 Lugano, Switzerland
and Italian Institute of Technology, 16163 Genova, Italy

The sampling problem lies at the heart of atomistic simulations and over the years many different enhanced sampling methods have been suggested toward its solution. These methods are often grouped into two broad families. On the one hand, are methods such as umbrella sampling and metadynamics that build a bias potential based on few order parameters or collective variables. On the other hand, are tempering methods such as replica exchange that combine different thermodynamic ensembles in one single expanded ensemble. We instead adopt a unifying perspective, focusing on the target probability distribution sampled by the different methods. This allows us to introduce a new class of collective-variables-based bias potentials that can be used to sample any of the expanded ensembles normally sampled via replica exchange. We also provide a practical implementation by properly adapting the iterative scheme of the recently developed on-the-fly probability enhanced sampling method [M. Invernizzi and M. Parrinello, J. Phys. Chem. Lett. **11,** 2731 (2020)], which was originally introduced for metadynamicslike sampling. The resulting method is very general and can be used to achieve different types of enhanced sampling. It is also reliable and simple to use, since it presents only few and robust external parameters and has a straightforward reweighting scheme. Furthermore, it can be used with any number of parallel replicas. We show the versatility of our approach with applications to multicanonical and multithermal-multibaric simulations, thermodynamic integration, umbrella sampling, and combinations thereof.

Subject Areas: Chemical Physics,
Computational Physics

I. INTRODUCTION

Sampling is one of the main challenges in atomistic simulations. In fact, even the most accurate models cannot produce high-quality results if the phase space is not properly sampled. The sampling issue is due to the large gap between the physical macroscopic timescales and the actual time that can be explored in standard atomistic simulations. This results in an ergodicity problem that can be encountered in fields as varied as materials science, chemistry, and biology. One facet of this problem is the

existence of different metastable states separated by kinetic bottlenecks that make the transition from one state to another a rare event. Enhanced sampling methods are a possible solution to this problem. Instead of extracting configurations from the relevant physical ensemble, these methods create an *ad hoc* modified ensemble in which the probability of sampling rare events is greatly enhanced. One kind of such target ensembles is obtained by combining multiple subensembles that differ only by the temperature or some other quantity, a typical example being parallel tempering [1]. We refer to these ensembles as expanded ensembles [2].

In the present paper we formulate the problem of generating such expanded ensembles in a way that allows us to use collective-variables-based methods. We find that the recently developed on-the-fly probability enhanced sampling (OPES) [3] can be adapted to the scope and provides an efficient implementation. This method was introduced as an evolution of metadynamics [4], since it

---

*michele.invernizzi@phys.chem.ethz.ch
†parrinello@phys.chem.ethz.ch

can provide the same type of enhanced sampling, but presents in most cases a faster convergence and has only few and robust adjustable parameters. These properties of OPES are retained when it is applied to sample expanded ensembles. This provides us with a general and reliable method that can be easily applied to sample many different ensembles.

We accompany this paper with a number of general considerations (Secs. II and VII), but the reader mostly interested in the method itself and its practical implementation can go directly to Sec. IV. Section III briefly recalls OPES in its original formulation for metadynamicslike sampling. We also present a variety of simulations to show the versatility of the new scheme, in particular (Sec. V) multicanonical, multithermal-multibaric, thermodynamic integration, and also (Sec. VI) enhanced sampling based on an order parameter, both alone and in combination with the previous ensembles.

II. UNIFIED APPROACH

The most popular approaches to enhanced sampling follow mainly two strategies. A first one was proposed in a pioneering work by Torrie and Valleau and is referred to as umbrella sampling [5,6]. This method starts by identifying one or few order parameters, or collective variables (CVs), $\mathbf{s} = \mathbf{s}(\mathbf{x})$, that are a function of the microscopic configuration and encode some of the slow modes of the system. Then a bias potential that is function of the CVs is added to the energy of the system, so that the sampling of the slow modes encoded by the CVs is accelerated. Many have followed this approach, and nowadays one of the most popular methods in this class is metadynamics [4].

A different perspective to enhanced sampling is that of parallel or simulated tempering [7,8]. In this case the idea is to combine in the same generalized ensemble the configurations explored by the system at different temperatures. This can improve the sampling because at higher temperatures the exploration of the phase space is often more efficient, and the system is less likely to remain stuck in metastable states. Over the years this approach has been extended and implemented in a variety of different methods, among which replica exchange [9] is probably the most widely employed.

These two families of enhanced sampling methods often have been seen as distinct and complementary. Although there are some papers in which the two perspectives are combined [10–13], typically they have been perceived as hybrid approaches [14,15]. Here we want to take a closer look at these two families and show that it is possible to provide a unified perspective to the enhanced sampling problem.

For a start we must specify that we are not interested in looking at the specific computational technique the various enhanced sampling methods use, since according to this criterion there would be many more than two

families. There are methods that use a bias potential and others that use specific Monte Carlo moves [9], and also methods that introduce a fictitious dynamics [16], or that focus on directly modifying the atomic forces [17], to name just a few. This kind of classification is of course perfectly legitimate, but we find it of limited relevance for our purposes.

The distinction between the two families cannot be based on the fact that one uses system-specific CVs, while the other makes use of general thermodynamic properties. For instance, it is known that Hamiltonian replica exchange can be used to enhance the fluctuations of any chosen CV [18], and on the other hand, that it is possible with metadynamics to use the potential energy itself as CV and sample a multithermal ensemble [19].

Thus we prefer to focus on the target distribution $p^{\mathrm{tg}}(\mathbf{x})$ that the different methods sample. In fact, each enhanced sampling method explicitly or implicitly aims at sampling a specific probability distribution in the configuration space that is not the physical one, but assigns a higher probability to some rare event. Designing such target distributions so that they are effective is far from trivial, and we can relate the two families of methods to the type of target distribution they imply.

A first class of enhanced sampling methods defines the target distribution by setting a constraint on its marginal distribution along some chosen CVs, $p^{\mathrm{tg}}(\mathbf{s}) = \int \delta(s(\mathbf{x}) - s) p^{\mathrm{tg}}(\mathbf{x}) d\mathbf{x}$. The most common choice is to impose a uniform marginal, $p^{\mathrm{tg}}(\mathbf{s}) = \mathrm{const}$. Among the methods that adopt this strategy are adaptive umbrella sampling [20] and metadynamics in its original formulation [4]. Also, the Wang-Landau algorithm [21] and various multicanonical algorithms [22,23] chose to sample a flat marginal distribution, using the potential energy as CV. An interesting case is the one of well-tempered metadynamics [24] that aims at sampling an $\mathbf{s}$ distribution that is a smoothed version of the original one. Contrary to the uniform case and in general to the fixed target case [25], the well-tempered target explicitly depends on the unbiased probability, and is thus not completely known beforehand. Other kinds of targets are also used in the $1/k$ ensemble [26] and in nested sampling [27].

Another class of methods will be the main focus of this paper, and it is the one that aims at sampling the so-called expanded ensembles [2]. These targets are not defined explicitly as a function of some CVs, but rather consist in the sum of overlapping probability distributions. A typical enhanced sampling technique that targets expanded distributions is, for example, replica exchange [9]. Expanded ensembles can be obtained by combining the same system at different temperatures or, more in general, different Hamiltonians [18,28,29]. They can be sampled also with single replica approaches, such as simulated tempering [8] and integrated tempering [30,31]. Broadly speaking, one could consider as part of this expanded ensemble class also

methods like multiple-windows umbrella sampling [32], or thermodynamic integration [33], where multiple separated ensembles are simulated and then combined into one via some postprocessing procedure such as the weighted histogram analysis method (WHAM) [34].

It is important to note that by classifying enhanced sampling methods with respect to $p^{tg}(\mathbf{x})$ we are not implying that methods in the same class are equivalent. Different methods in fact can use very different strategies to reach their target, each having its own strengths and weaknesses. However, this target-distribution perspective suggests that there is not a fundamental difference between the two traditional families, and that a unified approach is possible.

From this point of view, a special place is occupied by variationally enhanced sampling (VES) [35], targeted metadynamics [25], and on-the-fly probability enhanced sampling [3], since in these methods one has to explicitly choose a target distribution. This makes them particularly suited for developing a unified approach, since they are in principle capable of sampling the targets of both of the two families of enhanced sampling, and also combine them in new ways. In VES the usefulness of various target distributions has already been explored [36–38]. In particular, a target distribution has been proposed for sampling multi-thermal-multibaric ensembles [39] and also for combining them with a CV that drives a phase transition [40]. It is this paper that inspired us to try a generalized unified approach.

Our goal here is to introduce explicitly the expanded ensemble target distribution and show that it can be sampled by using a CV-based bias potential method such as VES or OPES. In doing so we will introduce the concept of expansion CVs, that allows us to define both the target expanded distribution and the bias needed to sample it. The method we propose is thus capable of sampling both kinds of target distributions, those typical of replica exchange, but also the uniform and well-tempered distributions typical of metadynamics. In this sense it provides a unified approach to enhanced sampling.

III. ON-THE-FLY PROBABILITY ENHANCED SAMPLING

The recently developed on-the-fly probability enhanced sampling [3] is a collective-variables-based method. Collective variables are a function of the microscopic configuration, $\mathbf{s} = \mathbf{s}(\mathbf{x})$, that provide a low-dimensional description of the system. In OPES we aim at modifying the physical probability distribution of $\mathbf{s}$, $P(\mathbf{s})$, in order to reach a given target probability distribution, $p^{tg}(\mathbf{s})$. To achieve this we must add the following bias potential,

$$V(\mathbf{s}) = -\frac{1}{\beta} \log \frac{p^{tg}(\mathbf{s})}{P(\mathbf{s})}, \qquad (1)$$

where β is the inverse temperature. OPES has been introduced as an evolution of metadynamics and in this spirit we first have used the well-tempered target distribution, defined as $p^{WT}(\mathbf{s}) \propto [P(\mathbf{s})]^{1/\gamma}$, where $\gamma > 1$ is known as the bias factor. This target distribution aims at increasing the transition rate between metastable states of the system, by lowering of a factor γ the free-energy barriers along the CVs. In the limit of $\gamma = \infty$ it is equivalent to choosing a uniform target.

Since $P(\mathbf{s})$ is not known *a priori*, we resort to an iterative scheme [20,41]. The core idea in OPES is to estimate the probability distribution at each step n, $P_n(\mathbf{s})$, by reweighting on the fly a simulation that is biased with $V_n(\mathbf{s})$, which is itself constructed from such $P_n(\mathbf{s})$ estimate according to Eq. (1). The $P_n(\mathbf{s})$ is obtained via a weighted kernel density estimation, adding a new kernel at a fixed small interval, similarly to metadynamics.

We refer the interested reader to Ref. [3], where the OPES iterative equations for the case of a well-tempered and a uniform target are presented in detail, and to Refs. [42,43] for some initial applications. In the present paper we introduce a class of target distributions that allows sampling any expanded ensemble. We also present in detail the OPES iterative scheme for this class of targets. While the core ingredients of OPES remain the same, the resulting method will look quite different from the one presented in Ref. [3]. In particular, when targeting expanded ensembles we will not need to use the kernel density estimation that plays instead a crucial role in the well-tempered case.

In applying OPES to sample expanded ensembles, we find a method that is similar in spirit to that of Ref. [2] and to other more recent methods, such as integrated tempering sampling [30], infinite switch simulated tempering [44], and variationally derived intermediates [45].

IV. TARGETING EXPANDED ENSEMBLES

Let us call $u(\mathbf{x})$ the adimensional reduced potential that contains all the terms depending on the thermodynamic constraints, such as temperature, pressure, or others. With $\mathbf{x}$ we concisely indicate the atomic coordinates and any other configurational variable that the reduced potential might depend on, such as the volume or the box tensor. As an example, in the case of the canonical ensemble one has $u(\mathbf{x}) = \beta U(\mathbf{x})$, where β is the inverse temperature and $U(\mathbf{x})$ is the potential energy of the system. Let us consider a system with a reduced potential $u_\lambda(\mathbf{x})$ that is a function of λ, where λ could be either a single parameter or a set of parameters, and might indicate, e.g., a thermodynamic property such as the temperature. At equilibrium its probability distribution follows Boltzmann statistics:

$$P_\lambda(\mathbf{x}) = \frac{e^{-u_\lambda(\mathbf{x})}}{Z_\lambda}, \qquad (2)$$

where Z_λ is the partition function, $Z_\lambda = \int e^{-u_\lambda(\mathbf{x})} d\mathbf{x}$.

We are interested in sampling configurations in a range $\Delta\lambda$ of λ values. Instead of running multiple independent

simulations at different values of λ, we can sample a generalized ensemble which contains all the relevant microscopic configurations, and then reweight them to retrieve the correct statistics for any $\lambda \in \Delta\lambda$. Sampling such an ensemble over $\Delta\lambda$ instead of the separate single λ ensembles is more efficient when different λ ensembles overlap in the coordinate space, and it can also help in solving ergodicity problems.

We must choose as target a distribution that covers all the microscopic configurations relevant to the chosen $\Delta\lambda$ range. Similarly to what is done in replica exchange, we choose a set $\{\lambda\}$ of $N_{\{\lambda\}}$ values $\lambda \in \Delta\lambda$ such that there is a good overlap between contiguous $P_\lambda(\mathbf{x})$. We then define our target distribution as

$$p_{\{\lambda\}}(\mathbf{x}) = \frac{1}{N_{\{\lambda\}}}\sum_\lambda P_\lambda(\mathbf{x}). \qquad (3)$$

We refer to this class of target probability distributions as expanded ensemble target distributions. In the present paper we limit ourselves to considering nonweighted expanded targets, that assign the same $1/N_{\{\lambda\}}$ weight to all the subensembles, but it is also possible to add some λ-dependent weights and give different importance to different $P_\lambda(\mathbf{x})$.

Without loss of generality, one can consider $\lambda = 0$ to be inside the desired interval $\Delta\lambda$. It is then possible to run a simulation at $\lambda = 0$ and use the OPES scheme to iteratively optimize a bias that allows one to sample $p_{\{\lambda\}}(\mathbf{x})$. Before proceeding to explicitly write the target distribution and the bias potential, we express $P_\lambda(\mathbf{x})$ as

$$P_\lambda(\mathbf{x}) = P_0(\mathbf{x})e^{-u_\lambda(\mathbf{x})+u_0(\mathbf{x})}\frac{Z_0}{Z_\lambda} = P_0(\mathbf{x})e^{-\Delta u_\lambda(\mathbf{x})+\Delta F(\lambda)}, \qquad (4)$$

where $\Delta u_\lambda(\mathbf{x}) = u_\lambda(\mathbf{x}) - u_0(\mathbf{x})$ is the potential energy difference and

$$\Delta F(\lambda) = -\log\frac{Z_\lambda}{Z_0} = -\log\langle e^{-\Delta u_\lambda}\rangle_{u_0} \qquad (5)$$

is the dimensionless free-energy difference from the reference system u_0, that can be expressed also as an ensemble average, indicated with the notation $\langle\cdot\rangle_{u_0}$. Our expanded target thus becomes

$$p_{\{\lambda\}}(\mathbf{x}) = P_0(\mathbf{x})\frac{1}{N_{\{\lambda\}}}\sum_\lambda e^{-\Delta u_\lambda(\mathbf{x})+\Delta F(\lambda)}. \qquad (6)$$

In order to define the target bias, we first rewrite Eq. (1) as

$$v(\mathbf{x}) = -\log\frac{p^{\mathrm{tg}}(\mathbf{x})}{P_0(\mathbf{x})}. \qquad (7)$$

Finally, the adimensional bias needed to sample the expanded target $p_{\{\lambda\}}(\mathbf{x})$ is

$$v(\mathbf{x}) = -\log\left(\frac{1}{N_{\{\lambda\}}}\sum_\lambda e^{-\Delta u_\lambda(\mathbf{x})+\Delta F(\lambda)}\right). \qquad (8)$$

Note that in writing the bias in this way $P_0(\mathbf{x})$ cancels out. It follows that the bias potential $v(\mathbf{x})$ depends on the coordinates $\mathbf{x}$ only through the $N_{\{\lambda\}}$ quantities $\Delta u_\lambda(\mathbf{x})$. We refer to these Δu_λ as expansion collective variables. The expansion CVs completely characterize the expansion, since not only the bias, but also $\Delta F(\lambda)$, Eq. (5), and the expanded target distribution $p_{\{\lambda\}}(\mathbf{x})$, Eq. (6), are unambiguously defined through these quantities. We will see how, by properly choosing the expansion CVs $\Delta u_\lambda(\mathbf{x})$, it is possible to sample different kinds of expanded ensembles. For each of them we also highlight the connection between these expansion CVs and more traditional CVs that have a straightforward physical interpretation.

Our target bias, Eq. (8), depends on the free energy along the λ parameter, $\Delta F(\lambda)$, that is in general unknown. In the OPES spirit we will reach the target bias iteratively, by estimating on the fly $\Delta F(\lambda)$ via a reweighting procedure, and using such an estimate to define the applied bias.

A. Iterative OPES scheme

The free-energy difference $\Delta F(\lambda)$ defined in Eq. (5) can be written using an ensemble average over the reference unbiased system u_0 [46]. However, estimating $\langle e^{-\Delta u_\lambda}\rangle_{u_0}$ from an unbiased trajectory is practically impossible due to the fact that $e^{-\Delta u_\lambda}$ can be significantly large in regions where P_0 is extremely small, and that are thus not properly sampled. For this reason we use reweighting to write it as an average over the biased ensemble,

$$e^{-\Delta F(\lambda)} = \langle e^{-\Delta u_\lambda}\rangle_{u_0} = \frac{\langle e^{-\Delta u_\lambda+v}\rangle_{u_0+v}}{\langle e^v\rangle_{u_0+v}}, \qquad (9)$$

where the ensemble average $\langle\cdot\rangle_{u_0+v}$ is computed as a time average over a biased trajectory. In this way, one can obtain a much more accurate estimate of $\Delta F(\lambda)$.

The problem with this procedure is that the target bias v, Eq. (8), is itself a function of $\Delta F(\lambda)$. Therefore, we set up an iterative scheme that consists in running a biased simulation whose bias is based on the estimate of the free-energy difference that we obtain via on-the-fly reweighting. At step n the simulation runs with potential $u_0(\mathbf{x}) + v_n(\mathbf{x})$, where

$$v_n(\mathbf{x}) = -\log\left(\frac{1}{N_{\{\lambda\}}}\sum_\lambda e^{-\Delta u_\lambda(\mathbf{x})+\Delta F_n(\lambda)}\right). \qquad (10)$$

The reweighted estimate $\Delta F_n(\lambda)$ is updated at every iteration step n. In between the iteration steps there is a

fixed short stride where the simulation proceeds and both $\Delta F_n(\lambda)$ and the bias $v_n(\mathbf{x})$ are kept constant. The free-energy estimate at the nth step can be explicitly written as

$$\Delta F_n(\lambda) = -\log \left(\frac{\sum_{k=1}^{n} e^{-\Delta u_\lambda^{(k)} + v_{k-1}^{(k)}}}{\sum_{k=1}^{n} e^{v_{k-1}^{(k)}}} \right), \qquad (11)$$

where we use the notation $\Delta u_\lambda^{(k)} \equiv \Delta u_\lambda(\mathbf{x}_k)$ and $v_n^{(k)} \equiv v_n(\mathbf{x}_k)$, with $\mathbf{x}_k$ the configuration at the kth iteration step.

As the bias approaches convergence, the ensemble sampled approaches the target one, and the $\Delta F_n(\lambda)$ estimates become more and more accurate. Thus not only do we obtain the target bias, but we also have an estimate of the free energy as a function of the λ parameter, i.e., $\Delta F(\lambda)$. Our iterative scheme is similar in spirit to the one used in integrated tempering sampling [30], but the two differ both in their implementation and in their applications.

Equations (10) and (11) are the explicit OPES iterative equations used for sampling the expanded ensemble defined by the target distribution $p_{\{\lambda\}}(\mathbf{x})$, Eq. (3), and are at the heart of our method. In the following sections we show how these equations can be used to sample different expanded ensembles, simply by specifying different expansion CVs $\Delta u_\lambda(\mathbf{x})$. Once these are chosen, the only free parameter of the method is the stride between the iterations. This should be set so that consecutive steps are not too correlated, as it is the case for the on-the-fly Gaussians' deposition in metadynamics.

It is possible to parallelize the procedure using multiple replicas of the system, as is done in multiple walkers metadynamics [47], where each replica shares the same bias and all contribute to the ensemble average in Eq. (11). At variance with replica exchange, the number of parallel simulations does not have to be equal to the number $N_{\{\lambda\}}$ of λ points that define the target.

Finally, we note that one could consider expressing the free energy $\Delta F(\lambda)$ via a cumulant expansion [48,49]. This generally provides a very good estimate close to the reference $\lambda = 0$, but can be very inaccurate when the range is broad, requiring a great number of terms in the expansion. Furthermore, we found it can introduce artificial barriers that might stop the system from visiting all the relevant configurations, thus making the OPES self-consistent procedure much less efficient.

B. Reweighting

Until now we have seen how to sample expanded ensembles by applying a bias potential. We now need a reweighting procedure in order to retrieve statistics at any desired value of λ. To this effect one can use standard umbrella sampling reweighting [6]. Given any observable $O = O(\mathbf{x})$ that is a function of the atomic coordinates, we can calculate its average in the ensemble λ via the following reweighting equation:

$$\langle O \rangle_{u_\lambda} = \frac{\langle O e^{-\Delta u_\lambda + v} \rangle_{u_0 + v}}{\langle e^{-\Delta u_\lambda + v} \rangle_{u_0 + v}} \approx \frac{\sum_k^n O_k w_k(\lambda)}{\sum_k^n w_k(\lambda)}, \qquad (12)$$

where $O_k \equiv O(\mathbf{x}_k)$ and the weight $w_k(\lambda)$ is defined as $w_k(\lambda) \equiv e^{-\Delta u_\lambda^{(k)} + v_{k-1}^{(k)}}$.

This equation assumes that the applied bias is static or quasistatic, meaning that it is updated in an adiabatic fashion. It is thus good practice to discard an initial transient of the simulation, where the bias changes quickly, and not use it for reweighting. Determining the exact length to be discarded might not be intuitive; however, OPES generally assigns a very low weight to this initial transient, and thus the result will not be significantly affected by this choice [3].

A useful diagnostic tool when performing reweighting is the so-called effective sample size, defined as the number of sampled points n times the ratio between the variance of an observable in the unbiased ensemble and its variance in the reweighted ensemble [50]. We use a popular estimator for the effective sample size, defined as [50,51]

$$n_{\text{eff}}(\lambda) = \frac{[\sum_k w_k(\lambda)]^2}{\sum_k w_k^2(\lambda)}, \qquad (13)$$

where $w_k(\lambda)$ are the importance sampling weights. Intuitively, the effective sample size for a given λ will be smaller than the total number of samples, $n_{\text{eff}}(\lambda) < n$. One should expect the efficiency to be roughly $n_{\text{eff}}(\lambda)/n \propto 1/N_{\{\lambda\}}$, given a minimal choice of λ points that properly covers the target range. Plotting n_{eff}/n as a function of λ can be a good diagnostic tool to monitor the consistency of the iterative procedure.

Finally, we note that the estimate of uncertainties requires some extra care in the case of weighted samples [52]. In the Appendix B we describe in detail the weighted block averaging procedure we adopt, and show how the effective sample size plays a role.

V. LINEARLY EXPANDED ENSEMBLES

An important type of expanded ensemble is the one obtained by linearly perturbing the reduced potential of the system, $u_\lambda(\mathbf{x}) = u_0(\mathbf{x}) + \lambda \Delta u(\mathbf{x})$. It is defined by the following expansion CVs:

$$\Delta u_\lambda(\mathbf{x}) = \lambda \Delta u(\mathbf{x}). \qquad (14)$$

Various different ensembles can be obtained in this way, such as the multicanonical ensemble and the multibaric ensemble, and also alchemical transformations, and others. Recently an interesting "multiforce" ensemble that falls in

this category has been proposed [53]. We present in detail some of these ensembles in the following sections.

It can be useful to group together these linearly expanded ensembles because they share some interesting properties. In particular, for these ensembles we can propose a simple automatic way to choose the λ points that define the target $p_{\{\lambda\}}(\mathbf{x})$. The idea is to have the λ points uniformly distributed in the $\Delta\lambda$ interval with a spacing $\Delta\lambda/N_{\{\lambda\}}$ estimated from the effective sample size as a function of λ, $n_{\mathrm{eff}}(\lambda)$. In practice what we do is to run a short unbiased simulation of n steps at $\lambda = 0$ and use a root finding algorithm to determine the points $\lambda_+ > 0$ and $\lambda_- < 0$ such that $n_{\mathrm{eff}}(\lambda_\pm)/n \approx 0.5$. Then one can use a total of $N_{\{\lambda\}} = \Delta\lambda/(\lambda_+ + \lambda_-)$ equally spaced λ points to define the target $p_{\{\lambda\}}(\mathbf{x})$.

This heuristic way of choosing the λ points is not optimal, and more elaborate options have been explored in the replica-exchange literature [54]. However, in our case this choice is less critical, since within our scheme one can increase $N_{\{\lambda\}}$ without the need to simulate additional replicas of the system. Thus this procedure provides an easy and automatic guess for linearly expanded ensembles that can be practically useful in many scenarios (see Sec. VII for further considerations on the choice of the λ points).

A. Multicanonical ensemble

We start by considering as an example of linearly expanded ensemble the case of the multicanonical ensemble, which is probably the one with the longest history. The goal is to sample all the configurations relevant for canonical simulations in a given range of temperatures. In a canonical simulation the reduced potential is $u(\mathbf{x}) = \beta U(\mathbf{x})$, where $U(\mathbf{x})$ is the potential energy, $\beta = 1/(k_B T)$ is the inverse thermodynamic temperature, and k_B is Boltzmann constant. It is possible to define a multicanonical linearly expanded ensemble, by putting $\Delta u(\mathbf{x}) = u_0(\mathbf{x}) = \beta_0 U(\mathbf{x})$ and $\lambda = (\beta - \beta_0)/\beta_0$, where β_0 is the inverse temperature set by the thermostat of the simulation and β spans the target range $\beta_{\mathrm{min}} < \beta < \beta_{\mathrm{max}}$.

The expansion CVs that define such a target are

$$\Delta u_\lambda(\mathbf{x}) = \lambda\beta_0 U(\mathbf{x}) = (\beta - \beta_0)U(\mathbf{x}), \quad (15)$$

and by using them in the OPES iterative equations, Eqs. (10) and (11), we obtain our multicanonical simulation. Given the physical significance of the inverse temperature β, it is more natural to directly consider β as parameter instead of the dimensionless λ. We thus write Δu_β and $\Delta F = \Delta F(\beta)$, where we have set $\Delta F(\beta_0) = 0$. Similarly, it is natural to consider the potential energy $U(\mathbf{x})$ as a collective variable, and thus write the bias as

$$v(U) = -\log\left(\frac{1}{N_{\{\beta\}}}\sum_\beta e^{-(\beta-\beta_0)U+\Delta F(\beta)}\right). \quad (16)$$

It is important to notice that we did not require the bias to be a function of a single CV, but rather we find it to be the case when we set as target the temperature-expanded ensemble. This is in fact a general property of linearly expanded ensembles. When expanding according to a given λ, the resulting bias will be a function only of the thermodynamic conjugate variable Δu. To define the bias $v = v(\Delta u)$ we then need to estimate the free energy along λ, $\Delta F(\lambda)$.

Other multicanonical methods aim instead at sampling a flat energy distribution [22,23,39]. In order to do so, they need to estimate the free energy as a function of U (or equivalently the density of states along U), while in our method, as in other tempering approaches [44], we instead need to estimate the free energy as a function of temperature, $\Delta F(\beta)$.

Example: Alanine dipeptide As an example of multicanonical sampling, we consider alanine dipeptide in a vacuum, at temperature $T_0 = 300$ K. This is a typical toy model for testing enhanced sampling methods, since at room temperature it presents two metastable states with an extremely low transition probability. A possible way of enhancing the sampling is to bias the ϕ and ψ dihedral angles, using as target a flat uniform distribution or the well-tempered distribution [3]. Here instead we bias the potential energy U, and use as a target the multicanonical ensemble over a temperature range from 300 up to 1000 K.

Simulations are performed with the molecular dynamics software GROMACS [55], patched with the enhanced sampling library PLUMED [56] (see Supplemental Material for computational details [57]). The only input needed for OPES, besides the temperature range we are interested in sampling, is the bias update pace, that is taken to be 500 simulation steps (1 ps).

Figure 1(a) shows on the ϕ, ψ plane the configurations sampled during the 100 ns multicanonical run. It is interesting to notice that the potential energy U would be considered a bad CV in enhanced sampling methods such as metadynamics, since it cannot distinguish between the two basins that have roughly the same energy. However, when using the multicanonical ensemble as target, by biasing U we can enhance the probability of visiting the transition state (roughly the region where $\phi = 0$), and thus observe multiple transitions between the basins and converge the free-energy difference between them, ΔF_{AB} [Fig. 1(b)]. We can use the angle ϕ to define this free-energy difference between the two basins:

$$\Delta F_{AB} = -\log\left(\frac{\langle\chi_{\phi\in[0,\pi]}\rangle}{\langle\chi_{\phi\in[-\pi,0]}\rangle}\right), \quad (17)$$
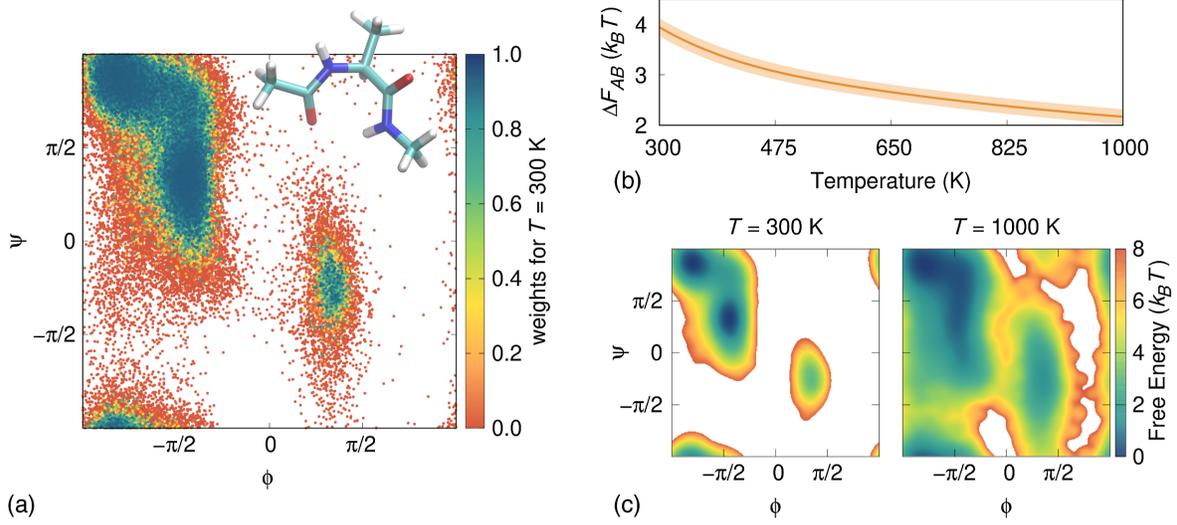
FIG. 1.   Alanine dipeptide in the multicanonical ensemble ($T_{\min} = 300$ K, $T_{\max} = 1000$ K). (a) Explored configurations as a function of the dihedral angles. Sampled points are colored according to their reweighting weight at $T = T_0 = 300$ K, $w_k(\beta) = e^{-(\beta-\beta_0)U_k + v_{k-1}^{(k)}}$. Notice how all the points in the transition state, close to $\phi = 0$, have extremely low probability of being sampled in an isothermal simulation at 300 K. (b) Free-energy difference between the two basins ΔF_{AB} as a function of temperature. (c) Reweighted free-energy surface at two different temperatures.

where χ is a characteristic function, equal to 1 if the variable is in the proper range and 0 otherwise.

In the Supplemental Material we discuss the differences with biasing the energy with metadynamics [12,19] and show a comparison between this multicanonical run and a well-tempered run biasing the two dihedral angles [57]. As expected, the latter is much more efficient (roughly 10 times) in converging the free-energy difference at 300 K, due to the fact that it focuses on the relevant degrees of freedom and on a single temperature.

B. Multithermal-multibaric ensemble

Within our scheme, combining different linearly expanded ensembles is straightforward. One simply has a two-dimensional parameter $\lambda = \{\lambda_1, \lambda_2\}$ and considers $u_\lambda(\mathbf{x}) = u_0 + \lambda_1 \Delta u_1(\mathbf{x}) + \lambda_2 \Delta u_2(\mathbf{x})$. This can be useful, for example, to sample multiple temperatures and multiple pressures in a single multithermal-multibaric simulation.

In this case we consider NPT simulations with a reference reduced potential $u_0(\mathbf{x}) = \beta_0 U(\mathbf{x}) + \beta_0 p_0 V(\mathbf{x})$, where p is the pressure and $V(\mathbf{x})$ the volume. Similarly to what was done before, it is more natural to use as λ parameters directly the temperature β and the pressure p, and write the expansion CVs $\Delta u_\lambda(\mathbf{x})$ as

$$\Delta u_{\beta,p}(\mathbf{x}) = (\beta - \beta_0)U(\mathbf{x}) + (\beta p - \beta_0 p_0)V(\mathbf{x}). \quad (18)$$

The target distribution is defined by a set of $N_{\{\beta\}}$ temperatures $\beta \in [\beta_{\min}, \beta_{\max}]$ and $N_{\{p\}}$ pressures $p \in [p_{\min}, p_{\max}]$,

for a total of $N_{\{\beta,p\}} = N_{\{\beta\}} \times N_{\{p\}}$ different $\Delta F(\beta, p)$ to be estimated. We will also express the bias, Eq. (8), as a function of the potential energy and the volume $v = v(U, V)$, which come as a natural CVs choice. As already discussed, the intermediate temperatures β and pressures p that define the target can be chosen automatically from a short unbiased simulation. We do this independently for the two parameters, despite the fact that the pressure term p is multiplied by β in Eq. (18).

Finally, we notice how the choice of β_0 and p_0 is completely free. As long as they lie inside the range of temperatures and pressures that we aim at sampling, no matter what thermodynamic conditions we start from at convergence we will sample the same configurational space. However, when the target range is very broad, choosing β_0 and p_0 roughly at the center can help to speed up convergence.

Example: Chignolin As an example of a multithermal-multibaric simulation we consider the miniprotein chignolin (CLN025) with CHARMM22* force field [58] and the three-site transferable intermolecular potential water model (about 2800 molecules), over a temperature range from $T_{\min} = 270$ K to $T_{\max} = 800$ K and a pressure range from $p_{\min} = 1$ bar to $p_{\max} = 4000$ bar. The velocity-rescaling thermostat [59] is set at $T_0 = 500$ K and the Parrinello-Rahman barostat [60] at $p_0 = 2000$ bar. The $\Delta F_n(\beta, p)$ estimates and the bias are updated every 500 simulation steps (1 ps). The $N_{\{\beta\}}$ temperature steps and $N_{\{p\}}$ pressure steps are chosen automatically based on a short 100 ps unbiased run. This results in 92 steps in temperature and 26 in pressure, for a total of $N_{\{\beta,p\}} = 2392$ points. In order to
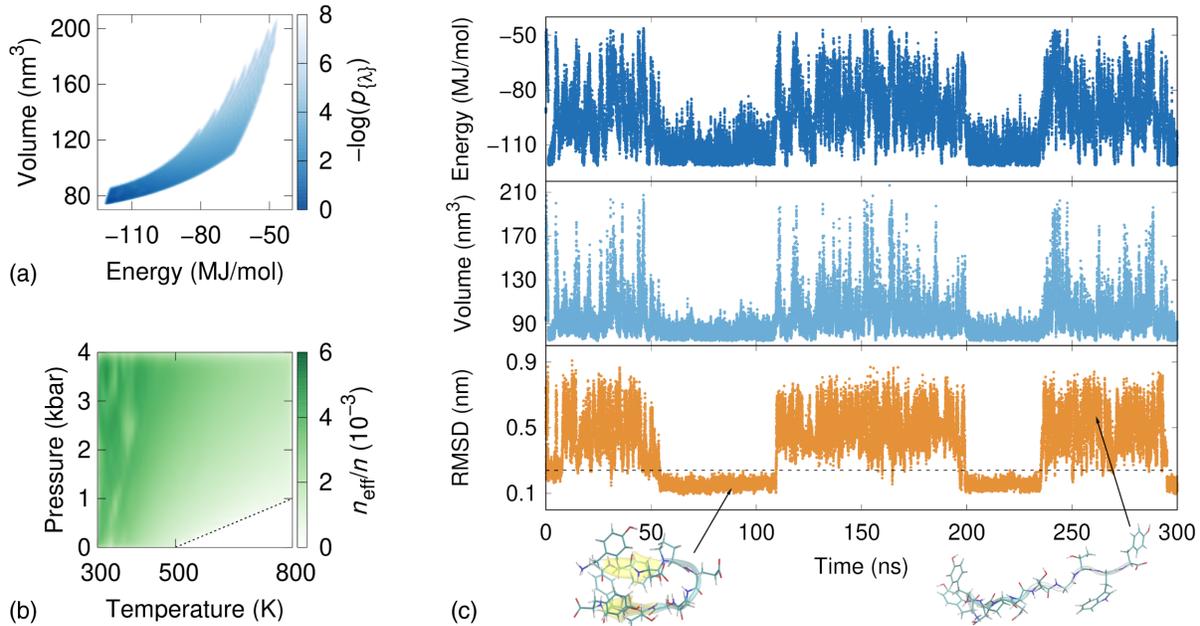
FIG. 2. Multithermal-multibaric simulation of chignolin. (a) Sampled target distribution in the CV space of potential energy and volume. (b) Relative effective sample size at different temperatures and pressures. The bottom corner of high temperatures and low pressures is excluded from the target to avoid vaporization of the system. (c) Time evolution of the two biased CVs and of the Cα RMSD for one of the 40 walkers. A RMSD threshold between folded and unfolded is highlighted with a dashed line.

avoid the region of low pressure and high temperature where water could evaporate, we discard any (β, p) point lying below the line from (500 K, 1 bar) and (800 K, 1000 bar). In this way, 91 (β, p) points are discarded.

The simulation is performed in parallel using 40 multiple walkers, and runs for a total of 300 ns per walker, of which roughly 10 ns are needed to converge the bias. Also in this case we use GROMACS patched with PLUMED.

In Fig. 2(a) we show the distribution sampled in the energy-volume space, while in Fig. 2(b) the corresponding effective sample size n_{eff} is plotted, as a function of temperature and pressure, and rescaled over the total number of samples n. The n_{eff}/n is not perfectly uniform, but it has the same order of magnitude over the whole target region. On the right, Fig. 2(c), we show for one of the 40 replicas the energy and volume trajectory, together with the trajectory of the Cα root-mean-square deviation (RMSD) to the experimental NMR structure [61].

In Fig. 3 we show the chignolin folded fraction at the different temperatures and pressures we targeted. The folded fraction is defined as in Ref. [62], using a dual cutoff on the Cα RMSD based on the CLN025 experimental NMR structure. A configuration is considered folded when the RMSD goes below 0.1 nm, and unfolded when it goes above 0.4 nm. In the inset we compare our results with those of Ref. [62] that considered a smaller temperature range at standard pressure. The confidence interval of our estimate is calculated with the block analysis described in Appendix B.

The stability diagram of chignolin (Fig. 3) does not present striking features, in qualitative agreement with Ref. [63]. However, it has been recently shown [64] that other miniproteins can have a nontrivial phase diagram, with unfolding both at low temperature and at high pressure.
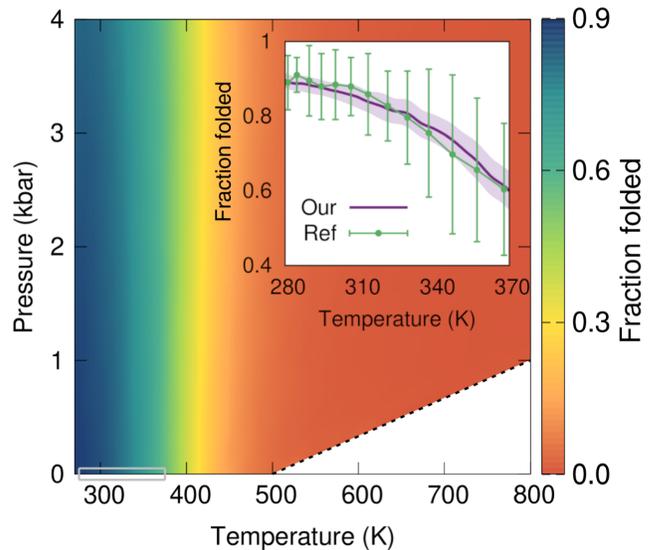
FIG. 3. The fraction folded of chignolin estimated from the multithermal-multibaric simulation. The inset shows the same quantity over a smaller range of temperatures at 1 bar (highlighted with a gray box), in order to compare it with the reference results from Lindorff-Larsen *et al.* [62].

C. Thermodynamic integration

Another interesting application of our method is its use for performing thermodynamic integration [65]. Let us consider a system with reduced potential energy $u_0(\mathbf{x})$ and free energy $F_0 = -\log Z_0$ and another similar system with potential $u_1(\mathbf{x})$ and free energy F_1. We are interested in calculating the free-energy difference $\Delta F_{0\to1} = F_1 - F_0$, for instance, because we know the free energy of one of the two systems and in this way we can retrieve the other one. The key idea of thermodynamic integration is to define a ladder of intermediate systems with reduced potentials $u_\lambda(\mathbf{x})$ and $0 < \lambda < 1$, to connect the two systems. The free-energy difference $\Delta F_{0\to1}$ to go from the u_0 system to the u_1 can be calculated via the following integral:

$$\Delta F_{0\to1} = \int_0^1 \left\langle \frac{\partial u_\lambda(\mathbf{x})}{\partial \lambda} \right\rangle_{u_\lambda} d\lambda. \qquad (19)$$

Typically individual simulations are run using $u_\lambda(\mathbf{x})$ for different values of λ and the ensemble average $\langle \partial u_\lambda(\mathbf{x})/\partial \lambda \rangle_{u_\lambda}$ is estimated for each of them. Then the integration in Eq. (19) can be carried out numerically, e.g., using the trapezoid rule or a Gaussian quadrature.

The most common way to define the intermediate states $u_\lambda(\mathbf{x})$ is via a linear interpolation,

$$u_\lambda(\mathbf{x}) = u_0(\mathbf{x}) + \lambda \Delta u(\mathbf{x}), \qquad (20)$$

where $\Delta u(\mathbf{x}) \equiv u_1(\mathbf{x}) - u_0(\mathbf{x})$. In this case we have $\partial u_\lambda/\partial \lambda = \Delta u$.

In the spirit of the present paper, we aim at performing a single simulation that samples all values of λ simultaneously, similarly to other methods [2,66,67]. It is then possible to reweight for any λ and calculate the integral in Eq. (19). Thus we simulate the system at $u_0(\mathbf{x})$ and build

a target $p_{\{\lambda\}}(\mathbf{x})$ as in Eq. (3) using $N_{\{\lambda\}}$ λ points in the interval $0 < \lambda < 1$. The OPES iterative equations, Eqs. (10) and (11), can be written using the expansion CVs $\Delta u_\lambda = \lambda \Delta u$ as defined in Eq. (20).

Finally, we note that thermodynamic integration can be performed using interpolation schemes different from the linear one, and our method is general and can be applied also in those scenarios, simply by properly defining the expansion CVs $\Delta u_\lambda(\mathbf{x})$.

Example: TIP4P water to Lennard-Jones fluid We now use the thermodynamic integration formalism described above to calculate the free energy of TIP4P water, relative to a reference Lennard-Jones system. The TIP4P potential energy (U_{TIP4P}) is made of an electrostatic energy term and a van der Waals-type interaction between the oxygens described by a Lenard-Jones potential (U_{LJ}). The free energy of the Lennard-Jones fluid has been fit to an equation of state and thus is a good reference [68]. For the simulations we use the LAMMPS [69] molecular dynamics software, patched with PLUMED. We perform an NVT canonical simulation at 443 K using the TIP4P water potential, thus $u_0(\mathbf{x}) = \beta U_{\text{TIP4P}}(\mathbf{x})$, with $N = 384$ molecules. The reference system is characterized by $u_1(\mathbf{x}) = \beta U_{\text{LJ}}(\mathbf{x})$.

With β a constant, we consider as collective variable $\Delta U(\mathbf{x}) \equiv U_{\text{LJ}}(\mathbf{x}) - U_{\text{TIP4P}}(\mathbf{x})$, and write the bias according to Eq. (8):

$$v(\Delta U) = -\log \left(\frac{1}{N_{\{\lambda\}}} \sum_\lambda e^{-\lambda \beta \Delta U + \Delta F(\lambda)} \right). \qquad (21)$$

From a short 20 ps unbiased run we obtain with the usual automatic procedure (Sec. V) 30 equispaced points in the interval $0 < \lambda < 1$, that define our target distribution. The evolution of ΔU as a function of simulation time is shown in Fig. 4(a). There is an initial transient of about 3 ns
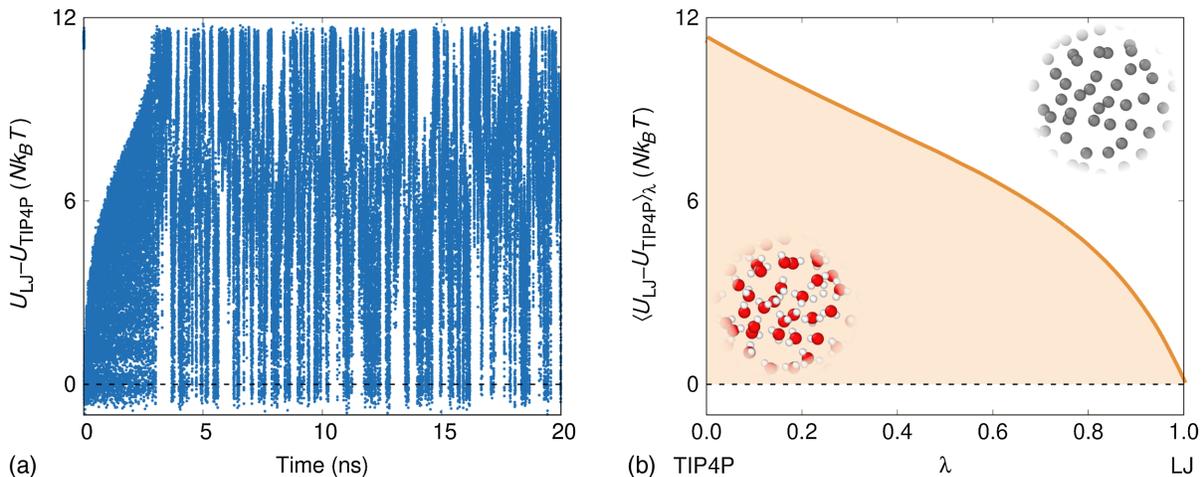
FIG. 4. Calculation of the free energy of liquid TIP4P water using a single-simulation thermodynamic integration. (a) Evolution of the collective variable ΔU as a function of simulation time. (b) Integrand for the thermodynamic integration obtained through reweighting.

until the bias potential is optimized and then the system diffuses freely. This has to be compared with a simulation for a given value of λ in which the fluctuations of ΔU would be very small. From this simulation the integrand $\langle \partial u_\lambda(\mathbf{x})/\partial \lambda \rangle_{u_\lambda} = \beta \langle \Delta U \rangle_\lambda$ can be calculated via reweighting, Eq. (12),

$$\langle \Delta U \rangle_\lambda = \frac{\sum_k^n \Delta U_k w_k(\lambda)}{\sum_k^n w_k(\lambda)}, \qquad (22)$$

where $\Delta U_k = U_{\text{LJ}}(\mathbf{x}_k) - U_{\text{TIP4P}}(\mathbf{x}_k)$ and $w_k(\lambda) = e^{-\lambda\beta\Delta U_k + v_{k-1}^{(k)}}$. The values of $\langle \Delta U \rangle_\lambda$ thus calculated are shown in Fig. 4(b). Using these results and Eq. (19) we find a free-energy difference $\Delta F_{\text{TIP4P}\to\text{LJ}} = F_{\text{LJ}} - F_{\text{TIP4P}} = 7.00(1)$ (Nk_BT units) in agreement with the result reported in Ref. [33].

VI. BEYOND LINEARITY: MULTIUMBRELLA ENSEMBLE

We consider now another important kind of expanded ensemble, namely the one obtained by combining all the different windows of a typical umbrella sampling simulation [18]. We refer to such an ensemble as a multiumbrella ensemble.

Multiple-windows umbrella sampling [32] allows for the free-energy surface (FES) reconstruction along some collective variable $s = s(\mathbf{x})$, that can be the reaction coordinate or some slow mode of the system. Typically one simulates multiple copies of the system, each one with a parabolic bias potential centered at a different s_λ point, in such a way that the resulting probability distributions have an overlap and cover the whole CV range. Postprocessing via WHAM [34] or other methods is then needed to estimate the relative free-energy differences and combine the data in a single FES estimate. Here, instead, we aim at sampling all the umbrella windows in the same simulation, by estimating on the fly these free-energy differences and building a single global potential. For this reason, the FES can then be obtained with the simple reweighting scheme described in Sec IV B. Both approaches have their own strengths and weaknesses, as we discuss while presenting the example below.

Given a system with reduced potential $u_0(\mathbf{x})$ and equilibrium Boltzmann distribution $P_0(\mathbf{x})$, we can write the reduced potential of each umbrella window as $u_\lambda(\mathbf{x}) = u_0(\mathbf{x}) + \Delta u_\lambda(\mathbf{x})$, with expansion CVs:

$$\Delta u_\lambda(\mathbf{x}) = \frac{(s(\mathbf{x}) - s_\lambda)^2}{2\sigma^2}. \qquad (23)$$

The associated probability distribution is $P_\lambda(\mathbf{x}) \propto P_0(\mathbf{x})G_\sigma(s(\mathbf{x}), s_\lambda)$, where $G_\sigma(s, s_\lambda)$ is a Gaussian of width σ centered in s_λ. The resulting expanded target $p_{\{\lambda\}}(\mathbf{x}) = (1/N_{\{\lambda\}}) \sum_\lambda P_\lambda(\mathbf{x})$ is clearly not linear in λ, and in fact

requires an extra parameter σ to be defined. The width σ can in principle vary with λ, but we consider here only the case of uniform umbrellas.

Since the expansion CVs, Eq. (23), depend on $\mathbf{x}$ only through $s = s(\mathbf{x})$, it is natural to write the bias as a function of the s CV:

$$v(s) = -\log\left(\frac{1}{N_{\{\lambda\}}}\sum_\lambda e^{-(s-s_\lambda)^2/2\sigma^2 + \Delta F(s_\lambda)}\right). \qquad (24)$$

Contrary to the linear case, in this multiumbrella case both the bias $v(s)$ and the free-energy differences $\Delta F(s) = -\log\langle G_\sigma(s(\mathbf{x}), s)\rangle_{u_0}$ are expressed as functions of the same CV.

The $N_{\{\lambda\}}$ s_λ points can be chosen to be uniformly distributed in the desired $\Delta s = s_{\max} - s_{\min}$ interval, in such a way to be at most at a distance of σ, ensuring overlap between contiguous P_λ. For a small enough σ, the estimate $\Delta F_n(s)$ converges precisely to the free-energy surface, while if σ is too broad there will be small artifacts, similarly to what happens when a too broad bandwidth is used in kernel density estimation.

It is instructive to consider the marginal of the target probability with respect to the CV, $p_{\{\lambda\}}(s)$. In the limit of infinitely small σ and thus infinitely large $N_{\{\lambda\}}$, the multiumbrella target $p_{\{\lambda\}}(s)$ is a uniform flat distribution over the Δs interval. In the opposite limit, of a very broad σ, the target distribution will look like the original, hard-to-sample P_0. As a rule of thumb σ should be as small as the smallest features of the FES we are interested in. We note that this is the same criterion used to choose the σ parameter in metadynamics [70], and it can typically be guessed from a short unbiased run. For this reason we prefer to use as parameter σ instead of the more commonly used strength of the harmonic umbrella potential $K = 1/\sigma^2$ [32].

In some cases it proved useful to introduce two small modifications to make the multiumbrella iterative optimization scheme more robust. We leave the explanation of them to Appendix C, since they have not been necessary for the examples presented in the paper.

For simplifying the exposition we presented the procedure in the case of a 1D CV, but it is straightforward to extend it to higher dimensions, by using multidimensional Gaussians and placing the s_λ points on an appropriate multidimensional grid. When dealing with higher dimensions it might be interesting to use some more elaborate shapes for the umbrellas, e.g., a Gaussian mixture in a similar but complementary way to Ref. [71], or to follow a specific path, as in Ref. [72].

Example: Double-well model As an example for the multiumbrella ensemble we consider a Langevin dynamics on a 2D model potential [73] using as CV the x coordinate only, Fig. 5(a). Such CV is suboptimal, in the sense that it
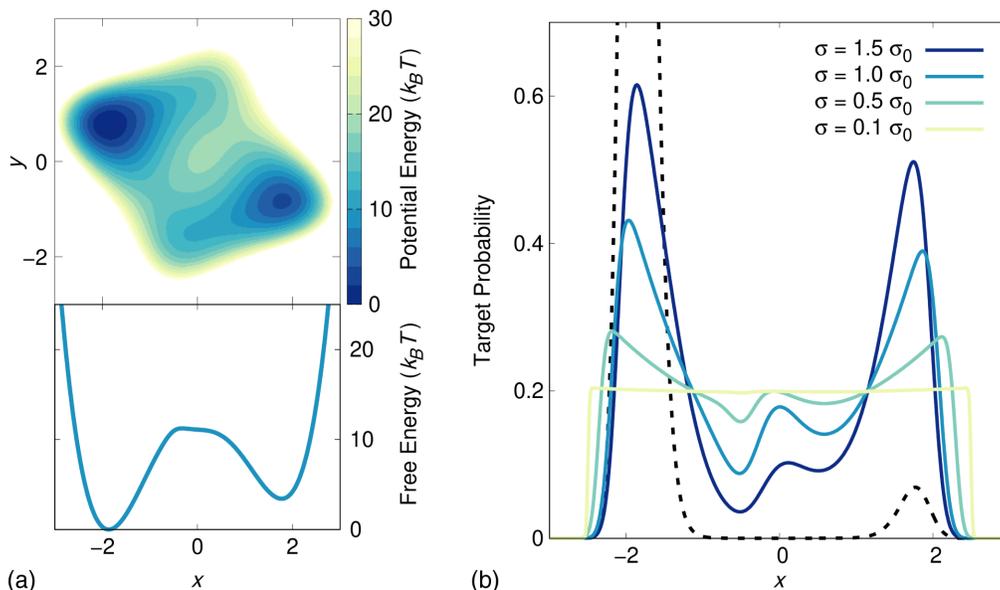
FIG. 5.   Double-well model in the multiumbrella ensemble. (a) Potential energy of the double-well 2D model system, and its free energy along the x coordinate. (b) Multiumbrella target distribution, for different values of umbrella width σ. The black dotted line is the unbiased probability distribution $P_0(x)$.

correctly describes the metastable basins, but does not include all the slow modes of the system [73]. One of the positive features of multiple-windows umbrella sampling is that it allows one to estimate the FES by running many short simulations in parallel. However, in this case some of the windows—for instance, the one centered at $x = 0$—cannot be efficiently sampled in a single short simulation due to the barrier along y. This is a known problem that can be diagnosed with dedicated consistency tests and is typically solved by performing more sampling [74]. With our approach, instead, we have a single combined simulation that in this example is long enough for the y mode to diffuse, and thus the CV suboptimality does not constitute a problem, and no extra care is required to handle it.

Figure 5(b) shows how the target distribution changes for different σ choices, expressed in units of the unbiased standard deviation in the basins, $\sigma_0 \approx 0.18$. The FES estimate could be directly obtained from the $\Delta F_n(s)$, but in the case of large σ this would lead to an estimate in which features are oversmoothed (see Supplemental Material [57]). Also, due to the CV suboptimality this estimate might not be very precise, but simply good enough for having that the main bottleneck for the transition is the unbiased slow mode y [73]. As a general rule, it is better to estimate the FES via the reweighting procedure.

In the Supplemental Material we provide all the simulation details and show the convergence of the free energy, comparing it to well-tempered OPES and metadynamics [57]. While in metadynamics and well-tempered OPES the bias is constructed in such a way to push the system

out of the visited areas, with multiumbrella OPES we are forcing the system to stay in a chosen CV region. Despite this difference in both cases we have similar target distributions and the resulting sampling allows us to reconstruct the FES.

A. Combining thermodynamic and order parameter expansions

An important characteristic of the present scheme is that it allows for a straightforward combination of different expanded ensembles. In particular, it allows for a rigorous and efficient combination of thermodynamic generalized ensembles with enhanced sampling along a system-specific order parameter.

To understand why this is important one can think about a first-order phase transition, where there is a kinetic bottleneck between the two phases that is responsible for an ergodicity problem. Increasing the temperature typically changes the relative stability of the two phases, but the free-energy barrier separating them might remain high along the whole coexistence line, thus making convergence very slow. A possible solution is to identify a suitable order parameter and biasing it to increase the transition probability. Combining the two approaches might actually outperform both [75,76]. This kind of combination can be useful not only for phase transitions; for instance, also in alchemical free-energy calculations an open problem is how to properly handle barriers orthogonal to the transformation [77].

We have already cited some hybrid methods that combine a replica-exchange approach with metadynamics, in order to enhance the sampling along both a

(a)   (b)
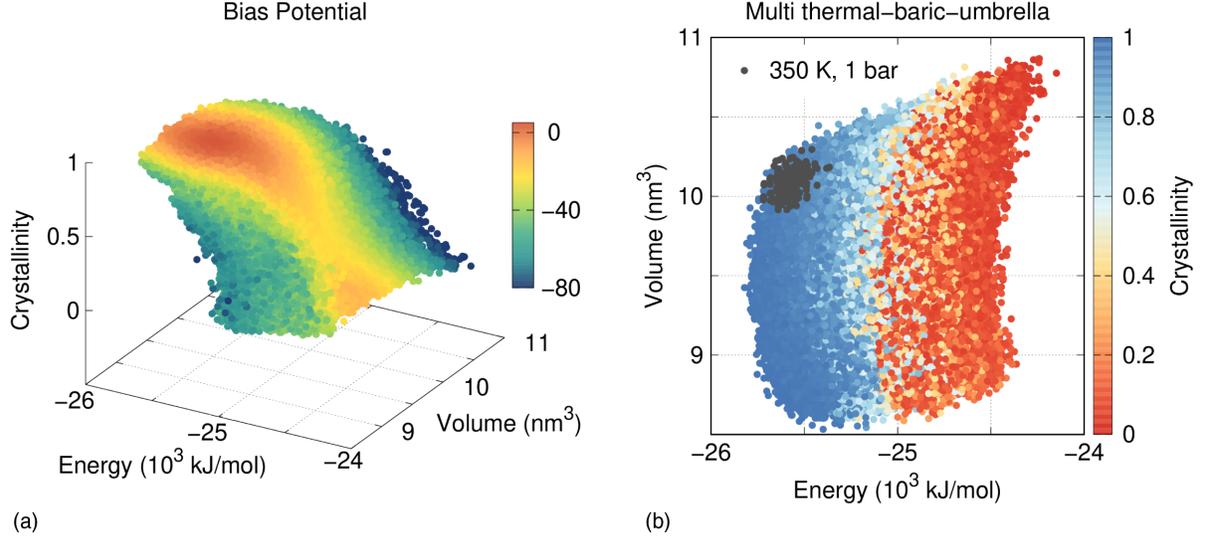
FIG. 6.   Configurations of sodium sampled during the multithermal-multibaric-multiumbrella simulation. (a) The points are colored accordingly to the value of the bias $v(U, V, s)$. (b) The points are shown in the energy-volume space and colored accordingly to the value of the crystallinity CV s. As reference we also show in gray the region sampled during an unbiased simulation in the bcc phase.

thermodynamic quantity and an order parameter [10–13, 75]. A nonhybrid approach has been first proposed with multidimensional replica exchange [18], but it has the drawback of requiring a sometimes impractical number of parallel replicas, due to the multidimensionality of the expansion. With OPES we can sample the same target distribution of multidimensional replica exchange, but using a bias potential and without requiring a minimum number of parallel replicas. In developing our method we followed the footsteps of another nonhybrid approach that has been recently proposed by our group, using the VES formalism and a custom target distribution [40,78]. Compared to the very flexible and customizable VES approach, OPES has the advantage of having much fewer free parameters and thus being simpler to set up and use.

Example: Sodium We consider here as an example the calculation of the liquid-bcc phase diagram of a model of sodium [79], the same studied in Ref. [40]. We sample the liquid and solid phase over a range of temperatures and pressures, using a recently proposed order parameter s, called environment similarity collective variable [40]. Such CV provides a measure of the crystallinity of the system, by comparing the local environment of the atoms to a reference one. For this reason we refer to it as crystallinity CV, but it is actually more general and can be used to describe a variety of phase transitions[78,80].

Using LAMMPS patched with PLUMED, we perform *NPT* simulations, $u_0(\mathbf{x}) = \beta_0 U(\mathbf{x}) + \beta_0 p_0 V(\mathbf{x})$. We can write the OPES equations, Eqs. (10) and (11), via the following expansion CVs:

$$\Delta u_{\beta,p,s}(\mathbf{x}) = (\beta - \beta_0)U(\mathbf{x}) + (\beta p - \beta_0 p_0)V + \frac{(s(\mathbf{x}) - s)^2}{2\sigma^2}.$$
(25)

The free-energy estimates $\Delta F_n(\beta, p, s)$ are expressed as a function of the inverse temperature β, the pressure p, and the crystallinity CV s. The bias $v = v(U, V, s)$ is expressed as a function of the potential energy U, the volume V, and the crystallinity CV s.

The simulation is performed with 250 atoms at $T_0 = 400$ K and $p_0 = 0.5$ GPa (5 kbar), using 4 multiple walkers that share the same bias and contribute to the same ensemble averages to update the $\Delta F_n(\beta, p, s)$ estimate. The aim is to sample liquid and solid configurations in the temperature range from 350 to 450 K and pressures from 0 to 1 GPa (10 kbar). The uniform grid over β and p to define the target distribution is automatically generated from a short 100 ps unbiased run, and consists of 4 temperature steps and 8 pressure steps. We chose as σ for the multiumbrella target a value of about 2.5 times the unbiased standard deviation in the basins, and it determines the presence of 26 umbrellas uniformly placed between $s_{\min} = 0$ (liquid) and $s_{\max} = 1$ (solid). In total the $\Delta F_n(\beta, p, s)$ to be estimated are $4 \times 8 \times 26 = 832$. After less than 3 ns the bias has clearly reached the adiabatic regime, and could be kept constant and used as static bias. However, since there is no drawback in doing so, we keep updating the bias and run 25 ns per walker, for a total combined simulation of 100 ns.

In Fig. 6(a) we show the points sampled in the CV space during the simulation, colored according to the value of the bias potential $v(U, V, s)$. We can clearly distinguish the
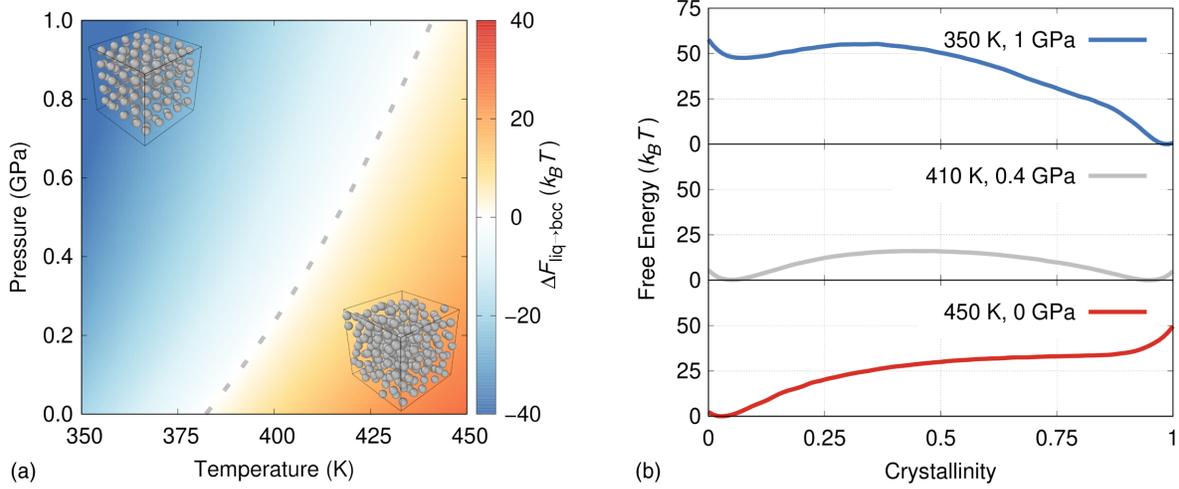
FIG. 7.   Phase equilibrium of liquid and solid (bcc) sodium using a combination of thermodynamic and order parameter expansions. (a) Free-energy difference between the phases, $\Delta F_{\text{liq}\rightarrow\text{bcc}}$, at different thermodynamic conditions. The coexistence line is shown as a gray dashed line. (b) Free-energy surfaces as a function of the crystallinity CV, for three representative thermodynamic conditions. Error bars are smaller than the linewidth.

hourglass shape described in Ref. [40]. In Fig. 6(b) we see the same trajectory plotted on the energy-volume plane. For comparison we show the configuration sampled in an unbiased simulation at a single temperature and pressure, in which the system remains all the time in the bcc crystal phase.

We can define the free-energy difference between the two phases as in Ref. [40]:

$$\Delta F_{\text{liq}\rightarrow\text{bcc}}(T,p) = -\log\left(\frac{\langle\chi_{s\in[0.5,1]}\rangle_{T,p}}{\langle\chi_{s\in[0,0.5]}\rangle_{T,p}}\right), \qquad (26)$$

where χ is a characteristic function, equal to 1 if the variable is in the proper range and 0 otherwise, and $\langle\cdot\rangle_{T,p}$ is the ensemble average at temperature T and pressure p.

In Fig. 7(a) we show $\Delta F_{\text{liq}\rightarrow\text{bcc}}(T,p)$ obtained by reweighting at different temperatures and pressures. The coexistence line $\Delta F_{\text{liq}\rightarrow\text{bcc}}(T,p) = 0$ is shown with a dotted gray line. On the right-hand side, Fig. 7(b), we provide the free-energy surface as a function of the CV, $F(s)$, at different thermodynamic conditions. Error bars are calculated with a weighted block average (Appendix B) and all the results are in agreement with Ref. [40]; see Supplemental Material [57].

It is important to notice that while the relative stability between liquid and solid changes across the range considered here, the probability of being in the transition state between the two is always extremely small, as can be seen from the high FES values around $s = 0.5$ in Fig. 7(b). By actively biasing the CV s we allow the system to efficiently sample the transition region as well, and this makes it possible to quickly converge the multithermal-multibaric simulation despite the presence of a first-order phase transition.

VII. ABOUT THE OPTIMAL TARGET DISTRIBUTION

Before reaching the conclusion of the paper we would like to add a final remark. At the beginning of Sec. IV we presented the nonweighted expanded ensemble target distribution, $p_{\{\lambda\}}(\mathbf{x}) = (1/N_{\{\lambda\}})\sum_\lambda P_\lambda(\mathbf{x})$. It is reasonable to wonder if this is an optimal target and in which sense. We argue here that the effective sample size can be used to define an optimality criterion.

Let us say our goal is to sample from a generalized ensemble that contains all the relevant microscopical configurations for a give range of the parameter λ. While the expanded target distribution $p_{\{\lambda\}}(\mathbf{x})$, Eq. (3), fulfills such a goal, there are in principle other possible choices.

It is useful to look at the special case of multicanonical ensembles, that has a long history (see also Sec. VA). In this context various different target distributions have been used, other than the nonweighted expanded ensembles one. One option is to have a uniform sampling in the energy [22,23,39], another one is to have a uniform sampling in the entropy [26], and a third one is to define the target by integrating the probability over the temperature, as in Refs. [30,44]. In this last approach one often approximates the integral with a sum, and effectively uses a target similar to our, Eq. (3), which is also the one used in temperature replica exchange. Another interesting perspective is presented in Ref. [81], where a Riemann metric is introduced to define optimality.

We believe that if our goal is to reweight at different temperatures, then the optimal target distribution is the one that yields the highest possible uniform effective sample size over the whole considered $\Delta\lambda$ range. Here we will not further explore such optimal target distribution nor dig

TABLE I. Some of the most common expanded ensembles, together with the expansion collective variables $\Delta u_\lambda(\mathbf{x})$ that define the OPES target bias, Eq. (8), and the free-energy differences $\Delta F(\lambda)$, Eq. (5). Each of the considered target biases can in turn be expressed as a function of one or two CVs. It is also possible to easily combine these ensembles to form new ones, as shown in Sec. VI A.

Target ensemble	Expansion CVs	Parameters	CVs
Linearly expanded	$\lambda \Delta u(\mathbf{x})$	$\{\lambda\}$	$\Delta u(\mathbf{x})$
Multicanonical	$(\beta - \beta_0)U(\mathbf{x})$	$\{\beta\}$	$U(\mathbf{x})$
Multibaric	$\beta_0(p - p_0)V(\mathbf{x})$	$\{p\}$	$V(\mathbf{x})$
Multithermal-multibaric	$(\beta - \beta_0)U(\mathbf{x}) + (\beta p - \beta_0 p_0)V(\mathbf{x})$	$\{\beta, p\}$	$U(\mathbf{x}), V(\mathbf{x})$
Multiumbrella	$(s(\mathbf{x}) - s_\lambda)^2/(2\sigma^2)$	$\{s_\lambda\}$	$s(\mathbf{x})$

deeper in the definition of effective sample size. However, simply by defining this criterion we can notice that one should not see the sum in Eq. (3) as an approximation to an integral. As a matter of fact, using as few as possible intermediate λ points brings us closer to this optimal target than having more, at least in the systems we studied (see Supplemental Material [57]).

It might also be the case that one is not interested in obtaining statistic for the whole $\Delta\lambda$ range, but only for a subset of λ states. In this case the optimal target would be the one that maximizes the effective sample size for those λ states while ensuring ergodicity. According to this criterion we argued in Ref. [3] that the well-tempered target is better than a uniform one, since it allows for an ergodic sampling while providing a higher n_{eff}/n ratio and avoids unimportant high free-energy regions.

VIII. CONCLUSION

In this paper we presented a general framework that provides a unified approach to enhanced sampling. To implement our method we leveraged the iterative scheme of OPES, an enhanced sampling method based on the construction of a bias potential along a set of collective variables, that was originally introduced for metadynamics-like sampling. We showed how this approach can be used to sample the same expanded ensembles typically sampled by a different family of enhanced sampling methods.

We also introduced the concept of expansion CVs, $\Delta u_\lambda(\mathbf{x})$, that can be used to fully characterize a non-weighted expanded target distribution $p_{\{\lambda\}}(\mathbf{x})$, Eq. (3), together with the free-energy differences to be iteratively estimated, Eq. (5), and the target bias, Eq. (8).

We then presented various examples of the application of the method to sample the most common expanded ensembles. These ensembles are summarized in Table I. In particular, we have shown how OPES can be used to enhance at the same time temperature-related fluctuations and a system-specific order parameter, Sec. VI A.

We notice that in defining the target distribution $p^{\text{tg}}(\mathbf{x})$ we consider only the positional degrees of freedom, and not the atomic velocities. Thus the ensembles sampled by our method are not identical to the ones sampled, for instance, by replica exchange, even though the target distribution is the same. In fact, the two methods sample the same configuration space, but a different velocity space. This does not have an effect on any statistical average of observables that are function of the coordinates only, but might be an interesting point for further research.

In the future it would be interesting to combine expanded target distributions with well-tempered-like distributions, which can scale better with higher dimensionality. Also weighted expanded targets might be of interest, where each subensemble λ has a specific different normalized weight. It might also be useful to implement the target distribution implicitly used by bias-exchange and parallel-bias metadynamics [11,82], that scales efficiently with the number of CVs. More generally, we believe that our perspective of focusing on the target distribution has further potential that should be explored.

The method is implemented in the open source PLUMED enhanced sampling library [56], and will be available as a contributed module called OPES. All the input files and postprocessing scripts used for this paper are openly available on the PLUMED-NEST [83,84] and in the Materials Cloud Archive [85], where the trajectories of the simulations are also stored.

APPENDIX A: ALGORITHMIC SCHEME

We schematically present the OPES algorithm used for sampling expanded ensembles.

(i) INITIALIZATION: each set of expansion CVs can be initialized independently, following one of two strategies

manual Choose the expansion CVs $\Delta u_\lambda(\mathbf{x})$ and initialize $\Delta F_0(\lambda) = \Delta u_\lambda(\mathbf{x}_0)$

automatic (linear expansions only) Choose $\lambda_{\min}$ and $\lambda_{\max}$. Run n_{init} unbiased simulation steps, and use $n_{\text{eff}}(\lambda)$ as described in Sec. V to define the $\{\lambda\}$ set. Initialize $\Delta F_0(\lambda)$ via Eq. (11), using the n_{init} points and $v = 0$.

(ii) SIMULATION: run the simulation for τ steps with potential $u(\mathbf{x}) + v_n(\mathbf{x})$. Add the point $\mathbf{x}_{n+1} = \mathbf{x}(n\tau)$ to the estimate $\Delta F_{n+1}(\lambda)$ and thus define $v_{n+1}(\mathbf{x})$, according to Eqs. (11) and (10). If multiple walkers are present, use each different configuration $\mathbf{x}_{n+1}$ to update the same $\Delta F_{n+1}(\lambda)$ estimate.

(iii) REWEIGHTING: to reweight for an observable $O(\mathbf{x})$ store its value during the simulation, together with the value of the potential v and use Eq. (12).

APPENDIX B: WEIGHTED BLOCK AVERAGE

When estimating the uncertainty related to an ensemble average obtained from molecular dynamics, one must take into account for the time correlation between the samples. Methods such as block averaging [86] are commonly used to properly handle this. The effect of such time correlation is to make the sample size effectively smaller; thus simply taking the square root of the variance divided by the number of samples would underestimate the actual uncertainty. When dealing with weighted samples, as it is the case when a bias potential is used, the effective sample size is further reduced by the presence of these weights. To account also for this effect we should slightly modify the block averaging procedure. We report here the protocol that we follow to estimate uncertainties, which is the same as the one presented in Ref [52], but here we highlight the role played by the effective sample size.

We are interested in estimating the ensemble average of an observable $O = O(\mathbf{x})$ from a biased ensemble,

$$\langle O \rangle \approx \hat{O} = \frac{\sum_{k=1}^{n} O_k w_k}{\sum_{k=1}^{n} w_k}, \tag{B1}$$

where w_k are the weights due to the bias potential, as in Eq. (12). In order to estimate the uncertainty we divide the data into M subsets or blocks, each containing an equal number of samples n/M. We then calculate the estimate $\hat{O}_i$ from each block, via a weighted average, and also the weight of the ith block,

$$W_i = \sum_{k=(i-1)(n/M)}^{i(n/M)} w_k. \tag{B2}$$

In this way the total estimate can be obtained as the weighted average of the blocks:

$$\hat{O} = \frac{\sum_{i=1}^{M} W_i \hat{O}_i}{\sum_{i=1}^{M} W_i}. \tag{B3}$$

Then according to the usual block average procedure we estimate the unbiased variance between the blocks, that in this case is a weighted variance:

$$\sigma_O^2 = \frac{M_{\text{eff}}}{M_{\text{eff}} - 1} \frac{\sum_{i=1}^{M} W_i (\hat{O}_i - \hat{O})^2}{\sum_{i=1}^{M} W_i}, \tag{B4}$$

where instead of the total number of blocks M, we use the effective block size $M_{\text{eff}} < M$:

$$M_{\text{eff}} = \frac{(\sum_{i=1}^{M} W_i)^2}{\sum_{i=1}^{M} W_i^2}, \tag{B5}$$

that is the same as Eq. (13). The statistical error on the $\hat{O}$ estimate is then $\sigma_O / \sqrt{M_{\text{eff}}}$. When the number of blocks M is small, or when the weight W_i are unbalanced, using M instead of M_{eff} can introduce a considerable underestimate of the real uncertainty.

The usual block average procedure can then be carried out, repeating the analysis using different number of blocks M and looking for a plateau in the error estimate.

APPENDIX C: IMPROVING ROBUSTNESS FOR THE MULTIUMBRELLA TARGET

In some cases, when targeting the multiumbrella ensemble it can be important to introduce two small modifications, to make the iterative optimization scheme more effective. We did not use these modifications in any of the examples presented, but we found them useful for dealing with more challenging systems and thus have been implemented in the code.

When the iterative scheme starts, the first guess of the $\Delta F_n(s_\lambda)$ comes from just one single point, and is thus very inaccurate for CV values far away from the visited one. In particular, it tends to become extremely large, which causes the bias to be very strong in pushing the system to the farthest s_λ value. This initial bias is stronger the smaller the σ, and might even cause the simulation to fail during the very first biased steps. To avoid this, we can limit the initial value of the $\Delta F_n(s_\lambda)$ estimates to always be smaller than a given value; thus $\Delta F_0(s_\lambda) \leq \Delta E$. This ΔE value can be set to be equal to an estimate of the free-energy barrier that has to be overcome. Thus, similarly to Ref. [3], we add an extra optional parameter called "barrier" that sets the value of ΔE. This barrier guess does not have to be perfect and a very rough estimate typically suffices. Also, we did not observe significant change in the convergence speed; thus we suggest to use this extra parameter only in case of an initial failure of the simulation, due to a too strong initial bias. The barrier parameter can in principle be used also

with types of expansion other than the multiumbrella one, even though in those cases it might be less useful.

The second modification comes from the observation that, contrary to the previously considered expanded ensembles, the multiumbrella one is not guaranteed to sample the full unbiased distribution $P_0(\mathbf{x})$. This can be a problem for the iterative scheme, because all the $\Delta F_n(s_\lambda)$ use as reference the unbiased free energy $F_0 = -\log Z_0$, whose estimate can vary substantially if P_0 is not properly sampled. Typically it should be easy to choose an interval Δs that contains all the s values relevant for P_0, but if this is not the case then the problem can be fixed by simply adding P_0 itself to the target distribution. To add P_0 to the target, it is sufficient to add an extra expansion CV $\Delta u_0(\mathbf{x}) \equiv 0$ that always returns zero.

---

[1] D. J. Earl and M. W. Deem, *Parallel Tempering: Theory, Applications, and New Perspectives*, Phys. Chem. Chem. Phys. **7,** 3910 (2005).

[2] A. P. Lyubartsev, A. A. Martsinovski, S. V. Shevkunov, and P. N. Vorontsov-Velyaminov, *New Approach to Monte Carlo Calculation of the Free Energy: Method of Expanded Ensembles*, J. Chem. Phys. **96,** 1776 (1992).

[3] M. Invernizzi and M. Parrinello, *Rethinking Metadynamics: From Bias Potentials to Probability Distributions*, J. Phys. Chem. Lett. **11,** 2731 (2020).

[4] A. Laio and M. Parrinello, *Escaping Free-Energy Minima*, Proc. Natl. Acad. Sci. U.S.A. **99,** 12562 (2002).

[5] G. M. Torrie and J. P. Valleau, *Monte Carlo Free Energy Estimates Using Non-Boltzmann Sampling: Application to the Sub-Critical Lennard-Jones Fluid*, Chem. Phys. Lett. **28,** 578 (1974).

[6] G. Torrie and J. Valleau, *Nonphysical Sampling Distributions in Monte Carlo Free-Energy Estimation: Umbrella Sampling*, J. Comput. Phys. **23,** 187 (1977).

[7] R. H. Swendsen and J.-S. Wang, *Replica Monte Carlo Simulation of Spin-Glasses*, Phys. Rev. Lett. **57,** 2607 (1986).

[8] E. Marinari and G. Parisi, *Simulated Tempering: A New Monte Carlo Scheme*, Europhys. Lett. **19,** 451 (1992).

[9] Y. Sugita and Y. Okamoto, *Replica-Exchange Molecular Dynamics Method for Protein Folding*, Chem. Phys. Lett. **314,** 141 (1999).

[10] G. Bussi, F. L. Gervasio, A. Laio, and M. Parrinello, *Free-Energy Landscape for β Hairpin Folding from Combined Parallel Tempering and Metadynamics*, J. Am. Chem. Soc. **128,** 13435 (2006).

[11] S. Piana and A. Laio, *A Bias-Exchange Approach to Protein Folding*, J. Phys. Chem. B **111,** 4553 (2007).

[12] M. Bonomi and M. Parrinello, *Enhanced Sampling in the Well-Tempered Ensemble*, Phys. Rev. Lett. **104,** 190601 (2010).

[13] M. Nava, R. Quhe, F. Palazzesi, P. Tiwary, and M. Parrinello, *de Broglie Swapping Metadynamics for Quantum and Classical Sampling*, J. Chem. Theory Comput. **11,** 5114 (2015).

[14] C. Abrams and G. Bussi, *Enhanced Sampling in Molecular Dynamics Using Metadynamics, Replica-Exchange, and Temperature-Acceleration*, Entropy **16,** 163 (2014).

[15] Y. I. Yang, Q. Shao, J. Zhang, L. Yang, and Y. Q. Gao, *Enhanced Sampling in Molecular Dynamics*, J. Chem. Phys. **151,** 070902 (2019).

[16] L. Rosso, P. Mináry, Z. Zhu, and M. E. Tuckerman, *On the Use of the Adiabatic Molecular Dynamics Technique in the Calculation of Free Energy Profiles*, J. Chem. Phys. **116,** 4389 (2002).

[17] E. Darve and A. Pohorille, *Calculating Free Energies Using Average Force*, J. Chem. Phys. **115,** 9169 (2001).

[18] Y. Sugita, A. Kitao, and Y. Okamoto, *Multidimensional Replica-Exchange Method for Free-Energy Calculations*, J. Chem. Phys. **113,** 6042 (2000).

[19] C. Micheletti, A. Laio, and M. Parrinello, *Reconstructing the Density of States by History-Dependent Metadynamics*, Phys. Rev. Lett. **92,** 170601 (2004).

[20] M. Mezei, *Adaptive Umbrella Sampling: Self-Consistent Determination of the Non-Boltzmann Bias*, J. Comput. Phys. **68,** 237 (1987).

[21] F. Wang and D. P. Landau, *Efficient, Multiple-Range Random Walk Algorithm to Calculate the Density of States*, Phys. Rev. Lett. **86,** 2050 (2001).

[22] B. A. Berg and T. Neuhaus, *Multicanonical Algorithms for First Order Phase Transitions*, Phys. Lett. B **267,** 249 (1991).

[23] J. Lee, *New Monte Carlo Algorithm: Entropic Sampling*, Phys. Rev. Lett. **71,** 211 (1993).

[24] A. Barducci, G. Bussi, and M. Parrinello, *Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method*, Phys. Rev. Lett. **100,** 020603 (2008).

[25] A. D. White, J. F. Dama, and G. A. Voth, *Designing Free Energy Surfaces that Match Experimental Data with Metadynamics*, J. Chem. Theory Comput. **11,** 2451 (2015).

[26] B. Hesselbo and R. B. Stinchcombe, *Monte Carlo Simulation and Global Optimization without Parameters*, Phys. Rev. Lett. **74,** 2151 (1995).

[27] L. B. Pártay, A. P. Bartók, and G. Csányi, *Efficient Sampling of Atomic Configurational Spaces*, J. Phys. Chem. B **114,** 10502 (2010).

[28] P. Liu, B. Kim, R. A. Friesner, and B. J. Berne, *Replica Exchange with Solute Tempering: A Method for Sampling Biological Systems in Explicit Water*, Proc. Natl. Acad. Sci. U.S.A. **102,** 13749 (2005).

[29] G. Bussi, *Hamiltonian Replica Exchange in GROMACS: A Flexible Implementation*, Mol. Phys. **112,** 379 (2014).

[30] Y. Q. Gao, *An Integrate-over-Temperature Approach for Enhanced Sampling*, J. Chem. Phys. **128,** 064105 (2008).

[31] H. Nymeyer, *Serial Tempering without Exchange*, J. Chem. Phys. **133,** 114113 (2010).

[32] J. Kästner, *Umbrella Sampling*, Comput. Mol. Sci. **1,** 932 (2011).

[33] C. Vega, E. Sanz, J. L. F. Abascal, and E. G. Noya, *Determination of Phase Diagrams via Computer Simulation: Methodology and Applications to Water, Electrolytes and Proteins*, J. Phys. Condens. Matter **20,** 153101 (2008).

[34] S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, and P. A. Kollman, *The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules. I. The Method*, J. Comput. Chem. **13,** 1011 (1992).

[35] O. Valsson and M. Parrinello, *Variational Approach to Enhanced Sampling and Free Energy Calculations*, Phys. Rev. Lett. **113**, 090601 (2014).

[36] P. Shaffer, O. Valsson, and M. Parrinello, *Enhanced, Targeted Sampling of High-Dimensional Free-Energy Landscapes Using Variationally Enhanced Sampling, with an Application to Chignolin*, Proc. Natl. Acad. Sci. U.S.A. **113**, 1150 (2016).

[37] M. Invernizzi, O. Valsson, and M. Parrinello, *Coarse Graining from Variationally Enhanced Sampling Applied to the Ginzburg-Landau Model*, Proc. Natl. Acad. Sci. U.S.A. **114**, 3370 (2017).

[38] J. Debnath, M. Invernizzi, and M. Parrinello, *Enhanced Sampling of Transition States*, J. Chem. Theory Comput. **15**, 2454 (2019).

[39] P. M. Piaggi and M. Parrinello, *Multithermal-Multibaric Molecular Simulations from a Variational Principle*, Phys. Rev. Lett. **122**, 050601 (2019).

[40] P. M. Piaggi and M. Parrinello, *Calculation of Phase Diagrams in the Multithermal-Multibaric Ensemble*, J. Chem. Phys. **150**, 244119 (2019).

[41] D. Reith, M. Pütz, and F. Müller-Plathe, *Deriving Effective Mesoscale Potentials from Atomistic Simulations*, J. Comput. Chem. **24**, 1624 (2003).

[42] L. Bonati, V. Rizzi, and M. Parrinello, *Data-Driven Collective Variables for Enhanced Sampling*, J. Phys. Chem. Lett. **11**, 2998 (2020).

[43] D. Mandelli, B. Hirshberg, and M. Parrinello, *Metadynamics of Paths*, Phys. Rev. Lett. **125**, 026001 (2020).

[44] A. Martinsson, J. Lu, B. Leimkuhler, and E. Vanden-Eijnden, *The Simulated Tempering Method in the Infinite Switch Limit with Adaptive Weight Learning*, J. Stat. Mech. (2019) 013207.

[45] M. Reinhardt and H. Grubmüller, *Determining Free Energy Differences through Variationally-Derived Intermediates*, J. Chem. Theory Comput. **16**, 3504 (2020).

[46] R. W. Zwanzig, *High Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases*, J. Chem. Phys. **22**, 1420 (1954).

[47] P. Raiteri, A. Laio, F. L. Gervasio, C. Micheletti, and M. Parrinello, *Efficient Reconstruction of Complex Free Energy Landscapes by Multiple Walkers Metadynamics*, J. Phys. Chem. B **110**, 3533 (2006).

[48] S. Park, *Comparison of the Serial and Parallel Algorithms of Generalized Ensemble Simulations: An Analytical Approach*, Phys. Rev. E **77**, 016709 (2008).

[49] B. K. Radak, D. Suh, and B. Roux, *A Generalized Linear Response Framework for Expanded Ensemble and Replica Exchange Simulations*, J. Chem. Phys. **149**, 072315 (2018).

[50] A. Kong, J. S. Liu, and W. H. Wong, *Sequential Imputations and Bayesian Missing Data Problems*, J. Am. Stat. Assoc. **89**, 278 (1994).

[51] L. Kish, *Sampling Organizations and Groups of Unequal Sizes*, Am. Sociol. Rev. **30**, 564 (1965).

[52] G. Bussi and G. A. Tribello, *Analyzing and Biasing Simulations with PLUMED*, in *Biomolecular Simulations: Methods and Protocols*, edited by M. Bonomi and C. Camilloni (Springer, New York, 2019), pp. 529–578.

[53] M. J. Hartmann, Y. Singh, E. Vanden-Eijnden, and G. M. Hocky, *Infinite Switch Simulated Tempering in Force (FISST)*, J. Chem. Phys. **152**, 244120 (2020).

[54] M. K. Prakash, A. Barducci, and M. Parrinello, *Replica Temperatures for Uniform Exchange and Efficient Round-trip Times in Explicit Solvent Parallel Tempering Simulations*, J. Chem. Theory Comput. **7**, 2025 (2011).

[55] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl, GROMACS: *High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers*, SoftwareX **1–2**, 19 (2015).

[56] G. A. Tribello, M. Bonomi, D. Branduardi, C. Camilloni, and G. Bussi, PLUMED 2: *New Feathers for an Old Bird*, Comput. Phys. Commun. **185**, 604 (2014).

[57] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevX.10.041034 for additional figures and computational details.

[58] S. Piana, K. Lindorff-Larsen, and D. E. Shaw, *How Robust Are Protein Folding Simulations with Respect to Force Field Parameterization?*, Biophys. J. **100**, L47 (2011).

[59] G. Bussi, D. Donadio, and M. Parrinello, *Canonical Sampling through Velocity Rescaling*, J. Chem. Phys. **126**, 014101 (2007).

[60] M. Parrinello and A. Rahman, *Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method*, J. Appl. Phys. **52**, 7182 (1981).

[61] S. Honda, T. Akiba, Y. S. Kato, Y. Sawada, M. Sekijima, M. Ishimura, A. Ooishi, H. Watanabe, T. Odahara, and K. Harata, *Crystal Structure of a Ten-Amino Acid Protein*, J. Am. Chem. Soc. **130**, 15327 (2008).

[62] K. Lindorff-Larsen, P. Maragakis, S. Piana, M. P. Eastwood, R. O. Dror, and D. E. Shaw, *Systematic Validation of Protein Force Fields against Experimental Data*, PLoS One **7**, e32131 (2012).

[63] H. Okumura, *Temperature and Pressure Denaturation of Chignolin: Folding and Unfolding Simulation by Multibaric-Multithermal Molecular Dynamics Method*, Proteins Struct. Funct. Bioinf. **80**, 2397 (2012).

[64] B. Uralcan and P. G. Debenedetti, *Computational Investigation of the Effect of Pressure on Protein Stability*, J. Phys. Chem. Lett. **10**, 1894 (2019).

[65] D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, Vol. 1 (Academic Press, New York, 2001).

[66] X. Kong and C. L. Brooks, *λ-Dynamics: A New Approach to Free Energy Calculations*, J. Chem. Phys. **105**, 2414 (1996).

[67] R. Bitetti-Putzer, W. Yang, and M. Karplus, *Generalized Ensembles Serve to Improve the Convergence of Free Energy Simulations*, Chem. Phys. Lett. **377**, 633 (2003).

[68] J. K. Johnson, J. A. Zollweg, and K. E. Gubbins, *The Lennard-Jones Equation of State Revisited*, Mol. Phys. **78**, 591 (1993).

[69] S. Plimpton, *Fast Parallel Algorithms for Short-Range Molecular Dynamics*, J. Comput. Phys. **117**, 1 (1995).

[70] O. Valsson, P. Tiwary, and M. Parrinello, *Enhancing Important Fluctuations: Rare Events and Metadynamics from a Conceptual Viewpoint*, Annu. Rev. Phys. Chem. **67**, 159 (2016).

[71] J. Debnath and M. Parrinello, *Gaussian Mixture -Based Enhanced Sampling for Statics and Dynamics*, J. Phys. Chem. Lett. **11,** 5076 (2020).

[72] D. Branduardi, F. L. Gervasio, and M. Parrinello, *From A to B in Free Energy Space*, J. Chem. Phys. **126,** 054103 (2007).

[73] M. Invernizzi and M. Parrinello, *Making the Best of a Bad Situation: A Multiscale Approach to Free Energy Calculation*, J. Chem. Theory Comput. **15,** 2187 (2019).

[74] F. Zhu and G. Hummer, *Convergence and Error Estimation in Free Energy Calculations Using the Weighted Histogram Analysis Method*, J. Comput. Chem. **33,** 453 (2012).

[75] Y. I. Yang, H. Niu, and M. Parrinello, *Combining Metadynamics and Integrated Tempering Sampling*, J. Phys. Chem. Lett. **9,** 6426 (2018).

[76] H. Niu, Y. I. Yang, and M. Parrinello, *Temperature Dependence of Homogeneous Nucleation in Ice*, Phys. Rev. Lett. **122,** 245501 (2019).

[77] D. F. Hahn, G. König, and P. H. Hünenberger, *Overcoming Orthogonal Barriers in Alchemical Free Energy Calculations: On the Relative Merits of λ-Variations, λ-Extrapolations, and Biasing*, J. Chem. Theory Comput. **16,** 1630 (2020).

[78] H. Niu, L. Bonati, P. M. Piaggi, and M. Parrinello, *Ab Initio Phase Diagram and Nucleation of Gallium*, Nat. Commun. **11,** 2654 (2020).

[79] S. R. Wilson, K. G. S. H. Gunawardana, and M. I. Mendelev, *Solid-Liquid Interface Free Energies of Pure bcc Metals and B2 Phases*, J. Chem. Phys. **142,** 134705 (2015).

[80] P. M. Piaggi and R. Car, *Phase Equilibrium of Liquid Water and Hexagonal Ice from Enhanced Sampling Molecular Dynamics Simulations*, J. Chem. Phys. **152,** 204116 (2020).

[81] V. Lindahl, J. Lidmar, and B. Hess, *Riemann Metric Approach to Optimal Sampling of Multidimensional Free-Energy Landscapes*, Phys. Rev. E **98,** 023312 (2018).

[82] J. Pfaendtner and M. Bonomi, *Efficient Sampling of High-Dimensional Free-Energy Landscapes with Parallel Bias Metadynamics*, J. Chem. Theory Comput. **11,** 5062 (2015).

[83] The PLUMED Consortium, *Promoting Transparency and Reproducibility in Enhanced Molecular Simulations*, Nat. Methods **16,** 670 (2019).

[84] https://www.plumed-nest.org, https://www.plumed-nest.org/eggs/20/022.

[85] https://www.materialscloud.org, https://archive.materialscloud.org/record/2020.81.

[86] H. Flyvbjerg and H. G. Petersen, *Error Estimates on Averages of Correlated Data*, J. Chem. Phys. **91,** 461 (1989),