THE JOURNAL OF PHYSICAL CHEMISTRY Letters

pubs.acs.org/JPCL Letter

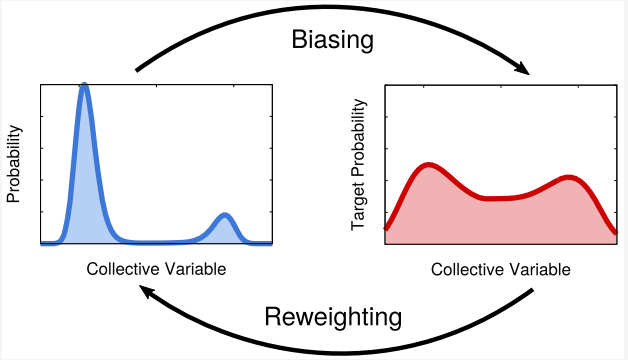
Rethinking Metadynamics: From Bias Potentials to Probability Distributions

Michele Invernizzi and Michele Parrinello*

Cite This: J. Phys. Chem. Lett. 2020, 11, 2731−2736 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Metadynamics is an enhanced sampling method of great popularity, based on the on-the-fly construction of a bias potential that is a function of a selected number of collective variables. We propose here a change in perspective that shifts the focus from the bias to the probability distribution reconstruction while retaining some of the key characteristics of metadynamics, such as flexible on-the-fly adjustments to the free energy estimate. The result is an enhanced sampling method that presents a drastic improvement in convergence speed, especially when dealing with suboptimal and/or multidimensional sets of collective variables. The method is especially robust and easy to use and in fact requires only a few simple parameters to be set, and it has a straightforward reweighting scheme to recover the statistics of the unbiased ensemble. Furthermore, it gives more control of the desired exploration of the phase space since the deposited bias is not allowed to grow indefinitely and it does not push the simulation to uninteresting high free energy regions. We demonstrate the performance of the method in a number of representative examples.



Enhanced sampling plays a crucial role in modern simulation techniques and is a very active area of research.[1] Of particular historical importance has been the work of Torrie and Valleau.[2] They consider a system with an interaction potential $U(\mathbf{R})$, where $\mathbf{R}$ denotes the atomic coordinates. Sampling is accelerated by adding a bias potential $V(\mathbf{s})$ that depends on $\mathbf{R}$ via a set of collective variables (CVs), $\mathbf{s} = \mathbf{s}(\mathbf{R})$. The CVs are chosen so as to describe the modes of the system that are more difficult to sample. The choice of a proper set of CVs is critical because it determines the efficiency of the method. The properties of the unbiased system are then calculated by using a reweighting procedure. In fact, the unbiased probability density $P(\mathbf{s}) = \langle \delta[\mathbf{s} - \mathbf{s}(\mathbf{R})] \rangle \propto \int d\mathbf{R}\, e^{-\beta U(\mathbf{R})} \delta[\mathbf{s} - \mathbf{s}(\mathbf{R})]$ can be written as an average over the biased ensemble

$$P(\mathbf{s}) = \frac{\langle \delta[\mathbf{s} - \mathbf{s}(\mathbf{R})] e^{\beta V(\mathbf{s})} \rangle_V}{\langle e^{\beta V(\mathbf{s})} \rangle_V} \tag{1}$$

where β is the inverse temperature. In this way, it is also possible to reconstruct the free energy surface (FES), defined as $F(\mathbf{s}) = -\frac{1}{\beta} \log P(\mathbf{s})$.

Since the work of Torrie and Valleau, a large number of CV-based enhanced sampling methods have been proposed. Among them is metadynamics[3,4] (MetaD), which builds the bias $V(\mathbf{s})$ by adding at fixed intervals repulsive Gaussians centered at the instantaneous point sampled. At the nth iteration, the bias is given by

$$V_n(\mathbf{s}) = \sum_k^n e^{-\beta V_{k-1}(\mathbf{s}_k)/(\gamma-1)} G(\mathbf{s}, \mathbf{s}_k) \tag{2}$$

where the parameter $\gamma > 1$ is called the bias factor, and the Gaussian function is defined as $G(\mathbf{s}, \mathbf{s}') = h \exp\left[-\frac{1}{2}(\mathbf{s} - \mathbf{s}')^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{s} - \mathbf{s}')\right]$, with height h and variance $\boldsymbol{\Sigma}$ set by the user. Typically only diagonal variances $\Sigma_{ij} = \sigma_i^2 \delta_{ij}$ are employed, but more general choices have also been suggested.[5] It has been proven[6] that at convergence there is a simple relationship between the bias and the free energy, $V(\mathbf{s}) = -(1 - 1/\gamma) F(\mathbf{s})$, and the sampled ensemble is a smoothed version of the unbiased one, with FES barriers lowered by a factor of γ.

Arguably, a major developments of MetaD has been its well-tempered variant.[4] With only a simple change to the original MetaD equations, it brought about many improvements, especially regarding the following points. (1) By damping the

Received: February 14, 2020
Accepted: March 19, 2020
Published: March 19, 2020

https://dx.doi.org/10.1021/acs.jpclett.0c00497
J. Phys. Chem. Lett. 2020, 11, 2731−2736

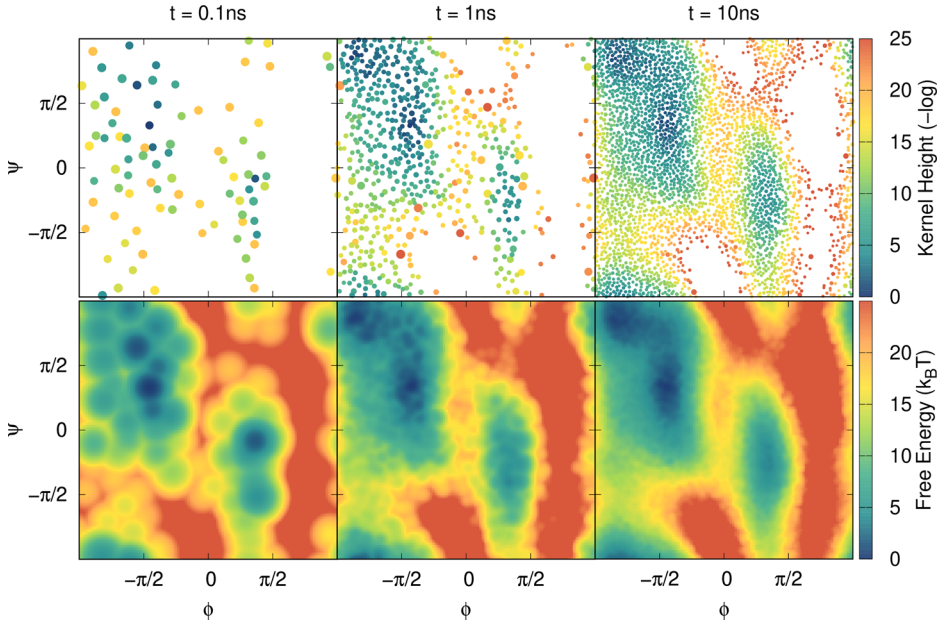


Figure 1. Time evolution of a typical OPES simulation of alanine dipeptide in vacuum, using as CVs dihedral angles ϕ and ψ. On the left, the compressed kernels forming $\tilde{P}_n(\phi, \psi)$ are shown, with the point size indicating the bandwidth, while on the right is the corresponding free energy estimate $F_n(\phi, \psi) = -\frac{1}{\beta} \log \tilde{P}_n(\phi, \psi)$, shifted to have a zero minimum. See the SI for the computational details and a performance comparison with MetaD.

bias oscillations, it allows for better handling of suboptimal CVs, that is, CVs that do not include all of the slow modes of the system. This is a crucial issue since finding a good CV for a complex system is nontrivial and even a good CV is usually suboptimal.[7] (2) It opens up the possibility of performing reweighting, which is a fundamental aspect of any enhanced sampling method, since it allows the retrieval of unbiased statistics of any quantity of interest. (3) It gives more control over the regions explored since the bias does not push the system to extremely high free energy regions. (4) Thanks to this property, it also improves the handling of multiple CVs by reducing the volume of CV space that is sampled at convergence.

Despite the success of MetaD, there is certainly room for further improvement. In fact, over the years many new MetaD variants have been proposed, which put particular emphasis on one of the above-mentioned issues.[5,8−12] Particular attention has been paid to reweighting, and many different solutions have also been proposed in recent years.[5,13−18] With this letter, we want to take a step back and propose a new perspective on MetaD in order to provide a general improvement to all of these issues, as has been the case for well-tempered MetaD.

We start from the observation that in the case of suboptimal CVs, the FES estimate obtained via reweighting can converge faster than the bias itself.[19] In particular, it is more robust and does not present the strong oscillations typical of MetaD. Furthermore, a more static bias can help with the reweighting procedure itself, giving rise to a positive feedback loop. Thus, we develop a method that is based on the reconstruction of the probability distribution via reweighting and uses this estimate to define the bias potential rather than directly building it as in eq 2.

Enhanced sampling based on the probability reconstruction is not a new idea. It was first proposed in the adaptive umbrella sampling method[20] and later by many others.[21−23] Typically, in such methods the bias at the nth iteration is defined as

$$V_n(\mathbf{s}) = \frac{1}{\beta} \log \hat{P}_n(\mathbf{s}) \tag{3}$$

where $\hat{P}_n(\mathbf{s})$ is an estimate of the probability obtained via a weighted histogram or some more elaborate method[21] and is updated iteratively or on the fly.[22] In building our method, we will introduce few key differences that come from the long experience with MetaD, which allow us to overcome some of the limitations of previous probability-based methods.

First, we explicitly introduce a target distribution $p^{\mathrm{tg}}(\mathbf{s})$ that will be sampled once the method reaches convergence. This can be obtained with the following bias:

$$V(\mathbf{s}) = \frac{1}{\beta} \log \frac{P(\mathbf{s})}{p^{\mathrm{tg}}(\mathbf{s})} \tag{4}$$

In adaptive umbrella sampling, the target distribution is uniform, $p^{\mathrm{tg}}(\mathbf{s}) \propto 1$, while in MetaD, it is the well-tempered distribution, $p^{\mathrm{tg}}(\mathbf{s}) \propto [P(\mathbf{s})]^{1/\gamma}$. It is possible to modify MetaD in order to reach any arbitrary target,[11] and in general the concept of a target distribution has proven to be very useful, especially in the contest of variationally enhanced sampling.[19,24−28] In the present work, we will limit ourselves to a well-tempered target (or a flat target in the $\gamma \to \infty$ limit), leaving other interesting possibilities for future work. We notice here that a well-tempered target leads to more efficient importance sampling compared to the common choice of a flat target, and despite lowering the FES barriers by γ instead of flattening them, it generally does not give rise to a slower transition rate between the metastable states. In fact, in most applications suboptimal CVs are employed and the transition rate is limited by the slow modes not accelerated by $V(\mathbf{s})$ rather than by the small FES barriers left along $\mathbf{s}$ (Supporting Information (SI)).[29]

Since we can express the target distribution as a function of the unbiased one, $p^{\mathrm{tg}}(\mathbf{s}) \propto [P(\mathbf{s})]^{1/\gamma}$, we only need to estimate $P(\mathbf{s})$ via reweighting in order to calculate the bias. We build

https://dx.doi.org/10.1021/acs.jpclett.0c00497
J. Phys. Chem. Lett. 2020, 11, 2731−2736

our probability distribution estimate on the fly by periodically depositing Gaussians, similarly to how MetaD builds the bias potential. This is indeed a common way of reconstructing a probability, known as kernel density estimation (KDE), and we shall draw from the vast literature on the subject.[30] Each new Gaussian is weighted according to the previously deposited bias potential

$$\tilde{P}_n(\mathbf{s}) = \frac{\sum_k^n w_k G(\mathbf{s}, \mathbf{s}_k)}{\sum_k^n w_k} \tag{5}$$

where the weights w_k are given by $w_k = e^{\beta V_{k-1}(\mathbf{s}_k)}$.

We write the estimator in eq 5 with a tilde, $\tilde{P}_n(\mathbf{s})$, to indicate that it is not properly normalized, and we will take care of the normalization separately. $G(\mathbf{s}, \mathbf{s}_k)$ represents Gaussians such as those defined previously for MetaD, with diagonal variance $\Sigma_{ij} = \sigma_i^2 \delta_{ij}$ and fixed height $h = \prod_i (\sigma_i \sqrt{2\pi})^{-1}$. Contrary to MetaD, here the height of the Gaussians is not a free parameter, and changing it simply corresponds to changing the overall normalization.

It has been shown[30] that in KDE the most relevant parameter is the bandwidth (i.e., the width of the Gaussians). A good choice of the bandwidth should depend on the amount of available data: the larger the sampling, the smaller the bandwidth. Thus, we choose to shrink the bandwidth as the simulation proceeds according to the popular Silverman's rule of thumb.[30] At the nth iteration

$$\sigma_i^{(n)} = \sigma_i^{(0)} [N_{\text{eff}}^{(n)} (d + 2)/4]^{-1/(d+4)} \tag{6}$$

where $\sigma_i^{(0)}$ is the initial standard deviation estimated from a short unbiased simulation, d is the dimensionality of the CV space, and $N_{\text{eff}}^{(n)} = (\sum_k^n w_k)^2 / \sum_k^n w_k^2$ is the effective sample size.[31] The KDE literature presents many other promising alternatives for the bandwidth selection, but we leave their study to future investigation.

The number of kernels accumulated during the simulation quickly becomes very large, and summing all of them at each time step is prohibitive. To avoid this problem, we adapt to our needs a simple on-the-fly kernel compression algorithm[32] that allows the insertion of new kernels only in newly explored regions and otherwise merges them with existing ones. In the Supporting Information, we discuss this choice in detail, and we show the advantages over the more common approach of storing the bias on a grid.[33]

In Figure 1 we show how the FES estimate evolves during a typical simulation with our new method. Our choice of the probability estimator aims at quickly obtaining a coarse representation of the FES and then slowly converging the finer details, and it is one of the key novelties of our method.

We can now discuss the normalization problem. Any constant overall normalization of the probability estimate $\tilde{P}_n(\mathbf{s})$ would simply result in a global shift of the bias and thus would not have any influence over the simulation. However, $\tilde{P}_{nn}(\mathbf{s})$ should be normalized not with respect to the full CV space but only over the CV space actually explored up to step n, which we call Ω_n. Thus, we introduce the normalization factor

$$Z_n = \frac{1}{|\Omega_n|} \tilde{P}_n(\mathbf{s})\, d\mathbf{s} \tag{7}$$

that will change over time as the system explores new regions of the CV space, and it will have an impact on the biasing scheme. This impact becomes particularly relevant in CV spaces of dimension $d \gg 1$ since the volume explored Ω_n grows with a power of d. Adding such a normalization, together with the chosen probability estimator, helps us overcome the limitations in exploration speed that have affected some previously proposed on-the-fly probability-based methods.[34] To estimate Z_n, we take advantage of our compressed kernels representation and substitute the integral in eq 7 with a sum over the positions of the compressed kernels (SI).

Finally, we can explicitly write the bias at the nth step as

$$V_n(\mathbf{s}) = (1 - 1/\gamma)\frac{1}{\beta} \log\left(\frac{\tilde{P}_n(\mathbf{s})}{Z_n} + \epsilon\right) \tag{8}$$

where $\epsilon \ll 1$ can be seen as a regularization term that ensures that the argument of the logarithm is always greater than zero. We notice that the addition of this term is not merely a technicality to solve a numerical issue but rather it allows one to set a limit on the bias, thus providing better control over the desired exploration. It can be chosen to be $\epsilon = e^{-\beta \Delta E/(1-1/\gamma)}$, where ΔE is the height of the free energy barriers one wishes to overcome during the enhanced sampling (SI). We could have obtained the same effect of controlling the exploration by properly modifying the target distribution $p^{\text{tg}}(\mathbf{s})$, but we believe that introducing ϵ as a separate term makes for cleaner equations. By comparing eqs 8 and 3, it should be clear that our method distinguishes itself from previous adaptive umbrella sampling methods not only for the employed probability estimator but also for some other novel key components.

An important feature of our method is that it allows for a simple and straightforward reweighting scheme. In fact, reweighting can be performed in the usual umbrella sampling way (eq 1) without the need for further postprocessing analysis. The method has a rapid initial exploration phase, after which a quasi-static regime is reached, but it is by construction robust with respect to the initial nonadiabatic part of the trajectory so that the reweighting can in practice be performed without cropping out the initial transient (SI).

We implemented the new method, called on-the-fly probability-enhanced sampling (OPES), in enhanced sampling library PLUMED[33] and tested it on a variety of different systems. Here we provide only a quick overview of these tests, but the full results are presented in detail in the Supporting Information. The code and all of the files needed to reproduce the simulations are openly available on the PLUMED-NEST website[35] as plumID:19.068.

A full comparison of different enhanced sampling methods is a nontrivial task and is not the goal of the present letter. However, in order to give a better idea of our method, we compare it with standard well-tempered metadynamics whose performances might already be familiar to many readers.

We want to test the methods in an agnostic fashion, using very standard input parameters rather than running multiple different simulations and choosing the best-performing ones. One strength of OPES is that it is very simple to set up and needs just three main parameters: the pace at which the bias is updated, the initial bandwidth of the Gaussian kernels, and the approximate height of the barriers one wishes to cross. In our tests, we always keep the deposition pace equal to the one used in MetaD, typically 500 simulation steps. The initial bandwidth

https://dx.doi.org/10.1021/acs.jpclett.0c00497
J. Phys. Chem. Lett. 2020, 11, 2731−2736

is simply chosen to be equal to the smaller standard deviation of the CVs in the minima, which can be measured in a short unbiased run. The choice of the barrier parameter requires a minimal knowledge of the system, but only a vague idea is usually enough. This barrier parameter is used to set a reasonable default of both the regularization factor ϵ and the bias factor γ (SI). It should be noticed that the choice of γ is not as critical as in MetaD since here it does not directly influence the convergence speed but only the shape of the target distribution. In fact, in OPES the limit $\gamma \to \infty$ is not problematic, as in MetaD,[36] and OPES can also converge to the flat target distribution. In our tests, we always used the same value of γ for both OPES and MetaD.

In order to test the convergence speed in the case of suboptimal CVs, we consider Langevin dynamics on a 2D model potential,[29] Figure 2, and bias only the x coordinate. In

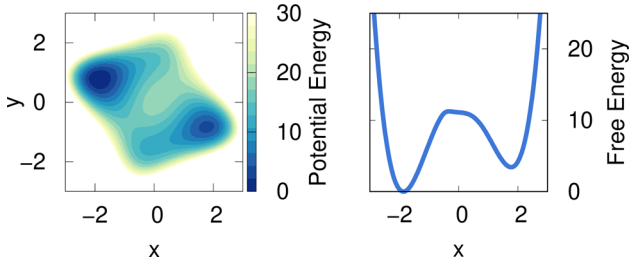


Figure 2. Potential energy of the suboptimal double-well 2D model system and its free energy along the x coordinate.

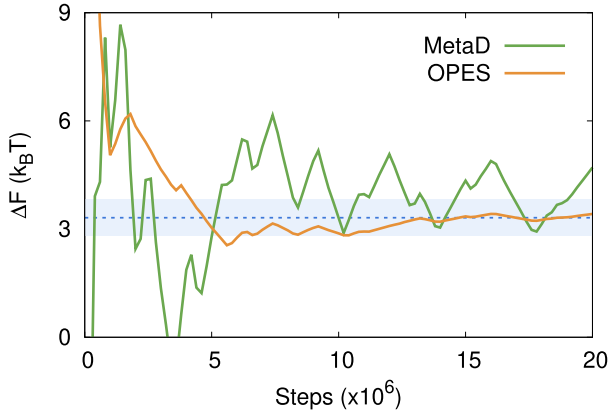


Figure 3. Typical time evolution of the free energy difference between the two basins of the model in Figure 2. We run the same simulation with 100 different initial conditions, and in the SI we show the average and uncertainty obtained from these estimates. The reference blue stripe is $1k_BT$ thick.

Figure 3, we compare MetaD and OPES by plotting the estimate of the free energy difference between the two basins as a function of time. Such estimates are obtained directly from the applied bias, as $F_n(x) = V_n(x)/(1/\gamma - 1)$. From Figure 3, one can see that as OPES converges it does not present the strong bias oscillations typical of MetaD.

We also run the typical benchmark system for novel enhanced sampling methods for alanine dipeptide, biasing dihedral angles ϕ and ψ. The results are in Figure 1 and the SI.

As an example of multidimensional bias, we run simulations of alanine tetrapeptide in vacuum, biasing all six dihedral angles. In Figure 4, we show how OPES is able to explore this high-dimensional CV space much more efficiently than MetaD.

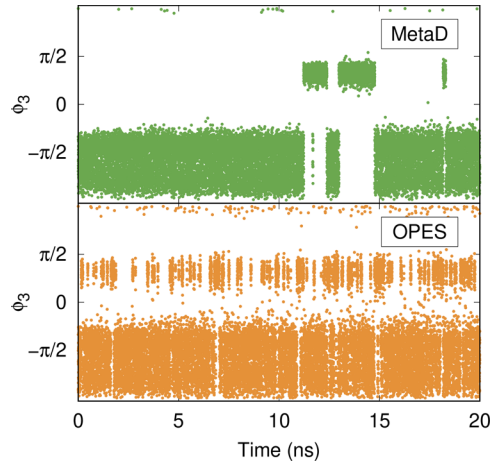


Figure 4. Φ_3 trajectory for alanine tetrapeptide in vacuum, obtained by biasing all six dihedral angles (ϕ_1, ϕ_2, ϕ_3, ψ_1, ψ_2, and ψ_3) using MetaD and OPES, with the same input parameters as the standard ones used for alanine dipeptide. We show the ϕ_3 angle because it is the hardest one to sample, but in the SI, all of them are presented, together with the reconstructed FES.

It is important to notice that we use the same input parameters for alanine dipeptide and alanine tetrapeptide. This should be a reasonable choice since the two systems are very similar from the point of view of the physics involved, the main difference being the increased dimensionality of the CV space. However, many CV-based enhanced sampling methods would require further tuning or a different set of inputs in order to perform well in both systems. This is not the case for OPES, thanks to its robustness to the choice of the input parameters.

In conclusion, in this letter we present a new enhanced sampling method, OPES, based on an on-the-fly reconstruction of the probability distribution. It performs such reconstruction via a weighted kernel density estimation (eq 5) with an on-the-fly compression algorithm that allows it to be most effective, starting from a coarse-grained guess of the free energy surface and then converging the finer details. Thanks to this strategy and to the introduction of a normalization over the explored CV space (eq 7), the method provides an extremely fast exploration, also at relatively high dimensions. Another peculiarity of the method is the presence of an upper limit to the applied bias (eq 8), which can be useful for avoiding the sampling of unphysical states. Most importantly, the proposed method requires few simple and robust input parameters, has very good convergence performance, and presents a straightforward reweighting scheme. Finally, we considered here only the case of a well-tempered target, but OPES provides a general framework in which different targets can also be considered.

We believe that this new method can become a handy tool in addressing enhanced sampling problems and has the potential for further interesting developments.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpclett.0c00497.

https://dx.doi.org/10.1021/acs.jpclett.0c00497
J. Phys. Chem. Lett. 2020, 11, 2731−2736

Full description of the kernel compression algorithm; further details on the bandwidth rescaling; estimate of the normalization factor Z_n; notes on the barrier parameter; and detailed results of all of the test systems (double-well model, alanine dipeptide, and alanine tetrapeptide) (PDF)

AUTHOR INFORMATION

Corresponding Author

Michele Parrinello − *Department of Chemistry and Applied Biosciences, ETH Zurich, 6900 Lugano, Switzerland; Facoltà di Informatica, Institute of Computational Science, Università della Svizzera Italiana, 6900 Lugano, Switzerland; Italian Institute of Technology, 16163 Genova, Italy*; Email: parrinello@phys.chem.ethz.ch

Author

Michele Invernizzi − *Department of Physics, ETH Zurich, 6900 Lugano, Switzerland; 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;* orcid.org/0000-0002-3328-6557

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpclett.0c00497

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This research was supported by the NCCR MARVEL, funded by the Swiss National Science Foundation, and European Union grant no. ERC-2014-AdG-670227/VARMET. Calculations were carried out on the Euler cluster at ETH Zurich and are openly available in the Materials Cloud Archive (www.materialscloud.org) with ID materialscloud:2019.0063. The authors thank Sandro Bottaro and Barak Hirshberg for useful feedback on the manuscript. M.I. thanks Dario Azzimonti and Manuel Schuerch for drawing his attention to kernel compression algorithms.

REFERENCES

(1) Peters, B. *Reaction Rate Theory and Rare Events*; Elsevier, 2017; p 619.

(2) Torrie, G.; Valleau, J. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J. Comput. Phys.* **1977**, *23*, 187−199.

(3) Laio, A.; Parrinello, M. Escaping free-energy minima. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 12562−12566.

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

(5) Branduardi, D.; Bussi, G.; Parrinello, M. Metadynamics with adaptive gaussians. *J. Chem. Theory Comput.* **2012**, *8*, 2247−2254.

(6) Dama, J. F.; Parrinello, M.; Voth, G. A. Well-tempered metadynamics converges asymptotically. *Phys. Rev. Lett.* **2014**, *112*, 240602.

(7) Pietrucci, F. Strategies for the exploration of free energy landscapes: Unity in diversity and challenges ahead. *Reviews in Physics* **2017**, *2*, 32−45.

(8) Raiteri, P.; Laio, A.; Gervasio, F. L.; Micheletti, C.; Parrinello, M. Efficient Reconstruction of Complex Free Energy Landscapes by Multiple Walkers Metadynamics †. *J. Phys. Chem. B* **2006**, *110*, 3533−3539.

(9) Piana, S.; Laio, A. A Bias-Exchange Approach to Protein Folding. *J. Phys. Chem. B* **2007**, *111*, 4553−4559.

(10) Dama, J. F.; Rotskoff, G.; Parrinello, M.; Voth, G. A. Transition-tempered metadynamics: Robust, convergent metadynamics via on-the-fly transition barrier estimation. *J. Chem. Theory Comput.* **2014**, *10*, 3626−3633.

(11) White, A. D.; Dama, J. F.; Voth, G. A. Designing Free Energy Surfaces That Match Experimental Data with Metadynamics. *J. Chem. Theory Comput.* **2015**, *11*, 2451−2460.

(12) Pfaendtner, J.; Bonomi, M. Efficient Sampling of High-Dimensional Free-Energy Landscapes with Parallel Bias Metadynamics. *J. Chem. Theory Comput.* **2015**, *11*, 5062−5067.

(13) Bonomi, M.; Barducci, A.; Parrinello, M. Reconstructing the equilibrium Boltzmann distribution from well-tempered metadynamics. *J. Comput. Chem.* **2009**, *30*, 1615−1621.

(14) Tiwary, P.; Parrinello, M. A time-independent free energy estimator for metadynamics. *J. Phys. Chem. B* **2015**, *119*, 736−742.

(15) Mones, L.; Bernstein, N.; Csányi, G. Exploration, Sampling, And Reconstruction of Free Energy Surfaces with Gaussian Process Regression. *J. Chem. Theory Comput.* **2016**, *12*, 5100−5110.

(16) Donati, L.; Keller, B. G. Girsanov reweighting for metadynamics simulations. *J. Chem. Phys.* **2018**, *149*, 072335.

(17) Marinova, V.; Salvalaglio, M. Time-independent free energies from metadynamics via mean force integration. *J. Chem. Phys.* **2019**, *151*, 164115.

(18) Giberti, F.; Cheng, B.; Tribello, G. A.; Ceriotti, M. Iterative Unbiasing of Quasi-Equilibrium Sampling. *arXiv* 2019, 1911.01140.

(19) Invernizzi, M.; Valsson, O.; Parrinello, M. Coarse graining from variationally enhanced sampling applied to the Ginzburg-Landau model. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, 3370−3374.

(20) Mezei, M. Adaptive umbrella sampling: Self-consistent determination of the non-Boltzmann bias. *J. Comput. Phys.* **1987**, *68*, 237−248.

(21) Maragakis, P.; van der Vaart, A.; Karplus, M. Gaussian-Mixture Umbrella Sampling. *J. Phys. Chem. B* **2009**, *113*, 4664−4673.

(22) Marsili, S.; Barducci, A.; Chelli, R.; Procacci, P.; Schettino, V. Self-healing Umbrella Sampling: A Non-equilibrium Approach for Quantitative Free Energy Calculations. *J. Phys. Chem. B* **2006**, *110*, 14011−14013.

(23) Dickson, B. M.; Legoll, F.; Lelièvre, T.; Stoltz, G.; Fleurat-Lessard, P. Free Energy Calculations: An Efficient Adaptive Biasing Potential Method. *J. Phys. Chem. B* **2010**, *114*, 5823−5830.

(24) Valsson, O.; Parrinello, M. Variational approach to enhanced sampling and free energy calculations. *Phys. Rev. Lett.* **2014**, *113*, 1−5.

(25) Valsson, O.; Parrinello, M. Well-tempered variational approach to enhanced sampling. *J. Chem. Theory Comput.* **2015**, *11*, 1996−2002.

(26) Shaffer, P.; Valsson, O.; Parrinello, M. 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.* **2016**, *113*, 1150−1155.

(27) Debnath, J.; Invernizzi, M.; Parrinello, M. Enhanced Sampling of Transition States. *J. Chem. Theory Comput.* **2019**, *15*, 2454−2459.

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

(29) 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.

(30) Silverman, B. *Density Estimation for Statistics and Data Analysis*; Routledge: New York, 1998.

(31) Kish, L. *Sampling Organizations and Groups of Unequal Sizes*; 1965; Vol. *30*, pp 564−572.

(32) Sodkomkham, D.; Ciliberti, D.; Wilson, M. A.; Fukui, K.-I.; Moriyama, K.; Numao, M.; Kloosterman, F. Kernel density compression for real-time Bayesian encoding/decoding of unsorted hippocampal spikes. *Knowledge-Based Systems* **2016**, *94*, 1−12.

(33) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED 2: New feathers for an old bird. *Comput. Phys. Commun.* **2014**, *185*, 604−613.
(34) Fort, G.; Jourdain, B.; Lelièvre, T.; Stoltz, G. Self-healing umbrella sampling: convergence and efficiency. *Statistics and Computing* **2017**, *27*, 147−168.
(35) The PLUMED Consortium. Promoting transparency and reproducibility in enhanced molecular simulations. *Nat. Methods* **2019**, *16*, 670−673.
(36) Lelièvre, T.; Rousset, M.; Stoltz, G. *Free Energy Computations*; Imperial College Press, 2010.

https://dx.doi.org/10.1021/acs.jpclett.0c00497
J. Phys. Chem. Lett. 2020, 11, 2731−2736