Using metadynamics to explore complex free-energy landscapes

Exploring complex free-energy landscapes by metadynamics

Giovanni Bussi* and Alessandro Laio*,+

* SISSA, Via Bonomea 265, I-34136 Trieste, Italy

+ ICTP, Strada Costiera 11, 34151 Trieste, Italy

July 3, 2020

Abstract

Metadynamics is an atomistic simulation technique that allows, within the same framework, accelerating rare events and estimating the free energy of complex molecular systems. The key idea at its basis is modifying iteratively the potential energy of the system by a sum of Gaussians centered along the trajectory followed by a suitably chosen set of collective variables (CVs). These Gaussians iteratively "fill" the free energy landscape as a function of the CVs, forcing the system to migrate from one minimum to the next. The potentiality of this idea is demonstrated by the large number of extensions and variants of the original approach that were developed during the years. The first scope of this review is presenting a critical comparison of these variants, discussing their advantages and disadvantages. The efficaciousness of metadynamics, as well as that of the numerous alternative methods that enhance the sampling by biasing selected CVs, is strongly influenced by the choice of the CVs: if an important variable is forgotten, the approach does not provide a reliable estimate of the free energy, and can predict qualitatively wrong transition mechanisms. The second scope of this review is discussing how the CVs should be selected, how one can verify if the CVs that are chosen are sufficient or redundant, and how one can iteratively improve the CVs using machine-learning approaches.

Many systems of interest in material science and biophysics are characterized by metastability. Technically speaking, this means that the probability distribution as a function of the atomic coordinates is characterized by the presence of at least two peaks (the metastable states) separated by a region in which the probability is many orders of magnitude lower. The prototypical example of this situation is a molecule which can undergo a chemical reaction: the two probability peaks correspond to the reactant and the product states. Other secondary probability peaks, when present, correspond to the intermediate states of the reaction. In practical terms, metastability means that a molecular dynamics (MD) or a Monte Carlo (MC) simulation is likely to remain stuck in only one probability maximum (typically an energy minimum) for all the duration of the run. Since the very early days of molecular simulations, researchers have attempted developing approaches to fight this problem, and observe all the relevant metastable states in the limited time that can be afforded in a simulation.

The simplest root to achieve this goal requires choosing, based on chemical or physical intuition, a collective variable, namely a function of the coordinates which takes a different value in all the relevant metastable states. Let us denote by x the coordinates on the system, and by $P(x)$ the probability distribution. Denote by $S(x)$ the collective variable (CV). The reduced probability distribution as a function of the CV is obtained by integrating over all the coordinates x under the constraint $s = S(x)$:

$$P(s) = \int dx P(x) \delta(s - S(x)). \tag{1}$$

We assume that $P(x)$ is the canonical distribution associated to a potential energy function $V(x)$: $P(x) \propto \exp(-V(x)/T)$, where T is the temperature (to simplify the notation we assume that the Boltzmann constant is one). The free energy as a function of s is then given by

$$F(s) = -T \log (P(s)) . \tag{2}$$

If the CV is well chosen, see next section, the metastable states will appear as separate and well-defined peaks in $P(s)$. Correspondingly, the free energy as a function of a good CV, for a system with metastable states, has (at least) two minima.

If in a molecular system one knows a good collective variable, and an approximation $B(s)$ of the negative of the free energy, the metastability problem in that system can be considered as solved. Indeed, one can run MD or MC with a modified potential $\tilde{V}(x) = V(x) + B(S(x))$. The probability distribution as a function of the CV becomes

$$\tilde{P}(s) = C \int dx e^{-\frac{V(x)+B(S(x))}{T}} \delta(s - S(x)) = C' e^{-\frac{F(s)+B(s)}{T}} . \tag{3}$$

where C and C' are normalization constants. Therefore, if $B(s) \sim -F(s)$ the probability distribution as a function of s is *approximately flat*. The simulation is not confined any more in a metastable state, and can freely diffuse across the barrier. Even if the simulation is performed under the action of a potential which is modified by an external bias, one can easily estimate the unbiased free energy from the biased probability distribution of s. Taking the logarithm of both members of Eq. 3 gives $F(s) = -B(s) - T \log \left(\tilde{P}(s) \right) + C$, where C is a constant.

This "trick" is well known since the very first days of molecular simulations [1], but its practical applicability is hindered by three problems:

1. Before performing the simulation one does not know how the free energy looks like, and a good choice for $B(s)$ is in general unavailable.

2. In many cases finding a good CV is non-trivial. One can build, based on intuition, a CV capable of distinguishing the metastable states, but this variable is not necessarily good for describing the transition. This topic is discussed in detail in Section 1.

3. In other cases one does not even know where the relevant metastable states are. For example, one may want to study the conformational transition of a complex biomolecule, knowing only the structure of the molecule in one state. This situation is possibly the most relevant for practical applications.

Metadynamics[2] is an algorithm which can satisfactorily solve problem 1 by building B in an iterative process. Instead, it does not offer a solution to problem 2, even if, as we will see, it allows *verifying* if a CV is good, and *improving* it for a successive simulation. Moreover, and possibly most importantly, metadynamics allows using at the same time multiple CVs. This, as we will see, allows being less greedy in the choice of the CVs and, in some special cases, even addressing problem 3.

Several reviews discussing the theory of metadynamics and its applications in a number of different fields are available [3, 4, 5, 6, 7, 8]. This review is mostly focused on the technical decisions one has to take before performing a metadynamics simulation. In particular, we will discuss the advantages and disadvantages of the different variants of this approach, the proper assessment of errors, the detection of critical cases where metadynamics is difficult to apply, the recently introduced methods to determine CVs using machine-learning techniques, and the available implementations of the method.

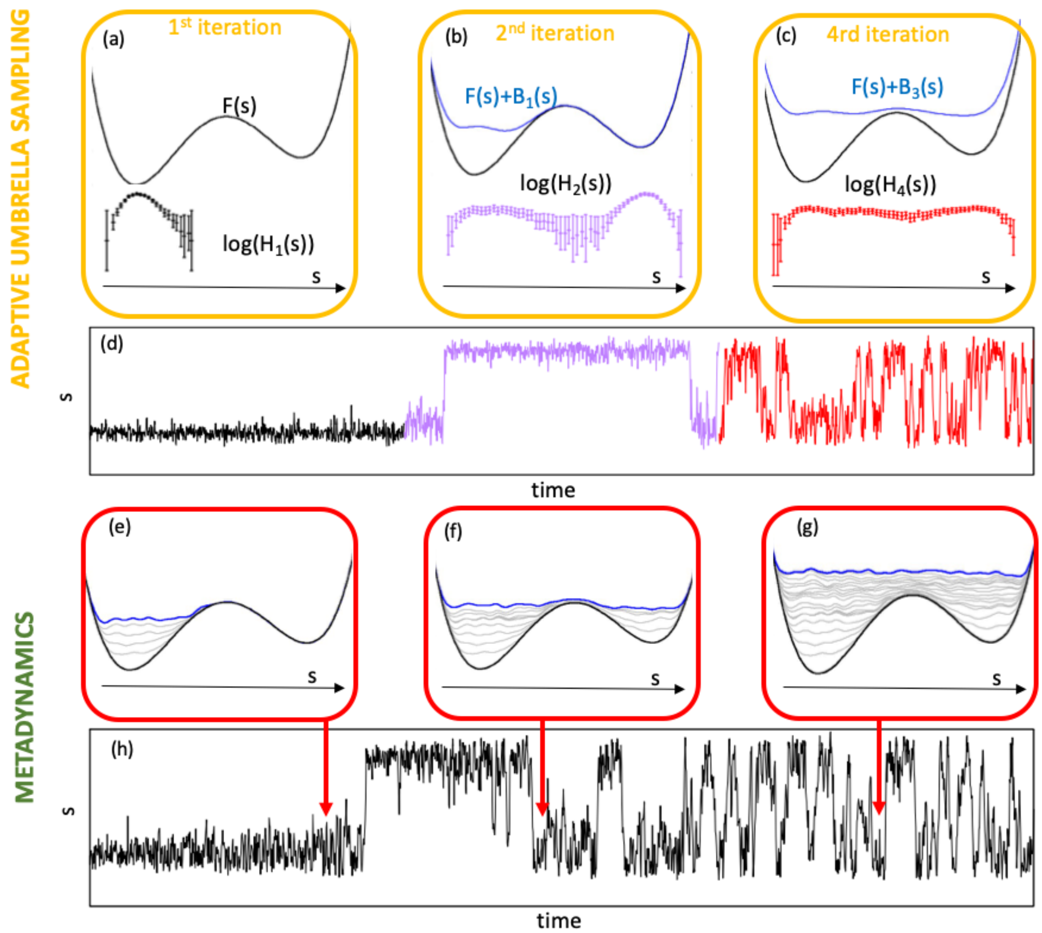

Figure 1: Illustration of the working principle of adaptive umbrella sampling, AUS (top) and metadynamics (bottom). Panel a, b and c illustrate three successive iterations of AUS. The points with error bars represent the histogram of the CV s estimated in each iteration. The purple lines in panel b and c represent the sum of the free energy and of the bias. Panel d shows the CV s as a function of time in the three iterations, represented with different colors. Panel e, f and g show the sum of the free energy and of the metadynamics bias potential (Eq. 7) at three different times, marked by arrows in panel h. Panel h shows the CV s as a function of time.

The algorithm was originally developed following the spirit of the time-stepper based approaches[9], embedding in this framework the idea of filling the free energy minima, like proposed in ref. [10], and biasing the dynamics by a history-dependent potential, like it was previously done in Taboo Search[11, 12], local elevation[13] and Wang-Landau sampling[14]. However, in order to understand more clearly the working principle of metadynamics, it is convenient heuristically introducing it as a limiting case of adaptive umbrella sampling (AUS)[15]. AUS requires running a sequence of relatively short simulations, here labeled by an index r. Each simulation is biased by a different external potential $B_r(s)$, built according to the following rules:

- Run a simulation under the action of the potential $V(x) + B_r(S(x))$.
- Compute the histogram of the collective variable $H_r(s)$; the first part of the simulation should be discarded, in order to allow an appropriate equilibration.
- Update the bias:

$$B_{r+1}(s) = B_r(s) + T\log(H_r(s)) \tag{4}$$

  In practice, if one wants to use B_r as an external bias in a MD simulation, one should take care to represent the logarithm of the histogram by a smooth function, for example a spline. Defining appropriately this function can be non-trivial, since the histogram is affected by non-uniform errors.

The first three iterations of the algorithm are illustrated in Figure 1. At the beginning, the bias is zero. Since the simulation is short, the system remains stuck in the first metastable state. The histogram $H_1(s)$ spans only that minimum (black points with error bars in panel a). The key idea of AUS is that the logarithm of this histogram is an estimate of the free energy, restricted to the region that has been explored so far. In other words, $F(s) + T\log(H_r(s))$ is approximately constant in the region of the minimum. In the second run, the system, due to the effect of the bias, will explore a wider range of the CV (purple line), performing a transition to the second minimum. The new histogram (purple points with error bars in panel b) will be approximately flat in the region already explored in the first run, and it will provide information on the shape of the free energy on a wider range. In the example in Figure, the bias potential has already "filled" the free energy landscape after three iterations (top-right panel, red curves). At this point, $B_r(s) \sim -F(s)$.

Metadynamics can be viewed as a limiting case of AUS: imagine to make the simulation time between two updates of the bias so short that the value of the CV does not change significantly anymore during the run. In this case, the label r designing the different runs in AUS can be replaced by a label t, labeling simulation time. The histogram $H_t(s)$ becomes a single peak, localized in the close neighbourhood of $s_t = S(x_t)$. The key idea of metadynamics is approximating the logarithm of this histogram with a simple, differentiable function. Typically one uses a Gaussian of width σ and height w:

$$\frac{1}{\beta}\log(H_t(s)) \sim w\exp\left(-\frac{(s_t - s)^2}{2\sigma^2}\right) \tag{5}$$

This turns Eq. 4 in

$$B_{t+1}(s) = B_t(s) + w\exp\left(-\frac{(s_t - s)^2}{2\sigma^2}\right) \tag{6}$$

or, equivalently

$$B_t(s) = w\sum_{t'<t}\exp\left(-\frac{(s_{t'} - s)^2}{2\sigma^2}\right) \tag{7}$$

The behaviour of the algorithm is illustrated in Fig. 1, bottom. At the beginning, the Gaussians are all localized in the first free energy minimum. These Gaussians induce larger and larger fluctuations to the CV (panel h). After some time, the first free energy minimum is almost completely filled by Gaussian (panel e) and the system performs a transition to the second minimum. This is also filled with Gaussians (panel f). After that moment, the CV starts diffusing freely between the two minima (panel h). The sum of the Gaussians now compensate almost exactly the free energy (panel g). This sum can be therefore used to estimate the $F(s)$. The two parameters w and σ can be tuned to control the speed at which the free energy landscape is filled and, thus, flattened. If one uses larger Gaussians, the bias will grow quickly, but the system will be strongly out of equilibrium. If, instead, the Gaussians are small, metadynamics becomes a quasi-equilibrium process, very similar to AUS. The precise role of these parameter is discussed in Section 2.

Replacing the logarithm of the histogram in Eq. 4 with a Gaussian in Eq. 6 can be viewed as a convenient manner of smoothing the former, in a spirit similar to kernel density estimators [16]. If, for example, one wants to compute the free energy as a simultaneous function of three different CVs, in AUS one should first compute a histogram as a function of three coordinates, and then represent its logarithm by a regular and differentiable function. In metadynamics, this function is built as a sum of three-dimensional Gaussians localized along the trajectory followed by the system in CV space. As it has been shown in many applications, and rigorously demonstrated for model dynamics, Eq. 7 provides a good approximation to the negative free energy in three dimensions or even more. This is the most important practical advantage of metadynamics with respect to other methods.

In the formulation of Eq. 7 it is apparent that in metadynamics the coordinates of the system evolve under the action of a non-Markovian process, in which the forces are history-dependent: indeed, the dynamics at time t is biased by an external potential defined by a sum of Gaussians localized on the sequence of values taken by the collective variable up to that moment. Conjecturing that the history-dependent potential in Eq. 7 can be used to estimate the free energy is the most important contribution of the original work where metadynamics was introduced[2] with respect to local elevation[13], where the idea of enhancing the sampling by using Gaussians is introduced, but no attempt to estimate the free energy from the sampled states is done.

The non-Markovianity of the dynamics makes its theoretical description more complex. However, as shown in Ref. [17], by explicitly *considering the external bias as a dynamic variable* the resulting dynamics is fully-Markovian in an extended space, the coordinates x and the bias $B(s)$ In these variables, the evolution of the system at time t depends only on its state at that time. When compared with the heuristic derivation above, the demonstration of Ref. [17] makes the assumption of adiabatic separation between the biased CV and the other degrees of freedom of the system but allows for an exact demonstration that is valid also in the case of a finite Gaussian deposition rate. This is the most important conceptual difference with respect to AUS. In this formulation it becomes natural treating the bias potential as one would treat an ordinary observable in a finite temperature MD or MC run: its instantaneous shape is not particularly meaningful, since it is affected by fluctuations. Instead, the relevant free energy estimator is not the bias itself, but its time average, in which the fluctuations become progressively smaller and smaller. These topics are treated extensively in Section 2.

Since the amplitude of the fluctuations in the bias potential is proportional to the height of the added Gaussian w, it is possible to reduce these fluctuations during the simulation by suitably reducing w. A possible way to do it is to employ the *well-tempered* variant of metadynamics [18], where the height of the Gaussians is chosen proportional to a decaying exponential function of the potential deposited in the currently visited point of the CV space. This turns Eq. 6 into

$$B_{t+1}(s) = B_t(s) + \exp\left(-\frac{B_t(s)}{\Delta T}\right) \exp\left(-\frac{(s_t - s)^2}{2\sigma^2}\right) \tag{8}$$

Here ΔT is a parameter that controls how quickly the Gaussian height is decreased. Often this is written in term of a so-called bias factor $\gamma = \frac{T+\Delta T}{T}$. It can be shown that with such a choice the height of the Gaussian deposited in a given point will decrease proportionally to the inverse of the time the simulation spent in that point [18]. This one-over-time relationship is a commonly used schedule for the learning rate in stochastic minimizations in machine-learning approaches [19], since it is guaranteed to converge [20]. However, by using a height that implicitly depends on the position in the CV space, in the long-time limit the system will not sample a flat distribution. It can be shown that the bias potential does not converge to $-F(s)$ but rather to an a priori known fraction of the free energy $-\frac{\Delta T}{T+\Delta T} F(s)$, and that the system will thus sample the distribution $P \propto \exp\left(-\frac{F(s)}{(T+\Delta T)}\right)$. This means that the ΔT parameters has both the role of damping the fluctuations of the estimator and that of controlling the effective temperature at which the chosen CV is sampled. In non-well-tempered metadynamics, this effective temperature is infinite. The scheme can be further generalized to independently control the fluctuations of the estimator and the effective temperature of the CV [20].

An important difference between well-tempered and non-well-tempered metadynamics is that, if boundary conditions are properly treated, the latter is guaranteed to reach a stationary state where the bias potential fluctuates as long as the relevant free-energy wells have been filled. The former, instead, will reach a quasi-equilibrium state, where the bias potential provides an exact estimator for the free-energy [18, 20]. Since this is guaranteed only for an infinitely long trajectory, it is heuristically possible to use the time-average of the bias potential as a better estimator for the free-energy surface also in well-tempered metadynamics (see Section 2).

Finally we notice that whereas metadynamics has been initially introduced as a method to explore and reconstruct the free-energy along biased variables, by suitable reweighting techniques [21, 22, 23, 24] it can be used also to reconstruct the free-energy along non-biased variables, although in these cases one should pay particular attention to make sure that analyzed variables are sufficiently sampled.

1 The key ingredient: choosing the collective variables

The capability of metadynamics to accelerate rare-event sampling and to reconstruct free-energy landscapes crucially depends on the employed collective variables (CVs). This dependence is common to all methods based on adding a bias potential that only depends on selected CVs.

CVs are arbitrary functions of the atomic coordinates and, since their are usually much less than the number of atomic coordinates, provide a low-dimensional projection of the conformational space. For a multistable system, a minimum criterion for this low-dimensional projection is that different metastable states should correspond to different values of the CVs. If this condition is not satisfied, any bias potential added to one state will equally disfavor all the other states that correspond to the same value of the CVs. An example of such a case is shown in Figure 2, panel d. Even if the potential energy landscape (panel a) has two minima, the free energy as a function of x has a single minimum. In this condition metadynamics is not able to accelerate in any manner the transitions between the two minima (panel g). A second requirement is that the CVs should be able to distinguish transition states. Indeed, metadynamics tends to work similarly to biological enzymes, namely it accelerates transitions by stabilizing the transition state relatively to reactants and product states. If the CV distinguishes the metastable states, but not the transition state, the transition will also not be enhanced. An example is provided by the landscape in panel b. The corresponding free energy as a function of x (panel e) has two minima, but the value of the CV at the transition state approximately coincides with the value of the CV in states with lower free energy that are part of the basin of attraction of reactants and products. In this case, under the action of metadynamics the CV reaches a perfectly diffusive dynamics (panel h). However, this behaviour is not an indication of convergence: after the first transition observed after approximately 5000 steps, the dynamics explores only the product state and the secondary minimum in the top-left corner. Therefore, the bias potential estimates the free energy without taking into account the state R. Indeed, the transitions between R and P are not enhanced at all by a bias acting on s_1, and transitions between R and P can happen only due to thermal fluctuations. As we will discuss in Section 2, in such a situation the bias potential in Eq. 7 cannot be used to estimate the free energy. Instead, the WT version of the bias, defined in Eq. 8, asymptotically provides a correct estimate. However, the bias does not accelerate the transition between the two minima, and therefore convergence is not significantly enhanced with respect to unbiased molecular dynamics.

The approach effectively works only if the CV takes different values in the metastable states and in the transition state between them. In other words, from the value of the CV one should be able to deduce with certainty whether the system is in a metastable state or in the other or in the transition state. An example is shown in panel c and f. In this situation, metadynamics induces several transitions between the two metastable states, and the free energy can be reliably estimated from the bias in Eq. 7 or in Eq. 8.

A further requirement is that the number of employed CVs should not be too large. Filling a multidimensional space becomes more expensive as the dimensionality of the space grows. Since the overall idea of metadynamics is to disfavor the conformations that have been already visited, if the number of CVs is too large the system will never return to exactly the same value of all the CVs. Approaches that employ multiple replicas to allow a large number of CVs to be biased alternately [25] or simultaneously [26] can be used to alleviate this requirement. In the first case, bias-exchange metadynamics, each replica biases a single CV, so that a large number of CVs can be simultaneously probed. In the second case, multiple independent biases are constructed in order to flatten the distribution of multiple CVs using well-tempered metadynamics. All CVs are biased in all replicas, but the ΔT parameter is modulated across the replica ladder, so that one replica provides unbiased sampling and the other replicas provide the capability to easily cross barriers. In both cases, coordinates are exchanged between replicas using an acceptance ratio that depends on the value of the biased CV of the different replicas. PBMETAD [27] allows to reproduce the character of bias-exchange metadynamics by using a single replica where the weight (or probability) for each of the variables to be biased is computed on the fly, thus with the practical advantage of allowing simulating a single replica.

Over the years, many other variants of metadynamics have been developed with the scope of addressing the problem of reconstructing the free energy in high dimensions. A very popular strategy is using a variable describing a *path* in a multidimensional CV space [28, 29]. This approach will be described and discussed in detail in Section 3. In ref. [30] Metadynamics is combined with standard Umbrella Sampling to sample orthogonal collective variables in a simultaneous way. In

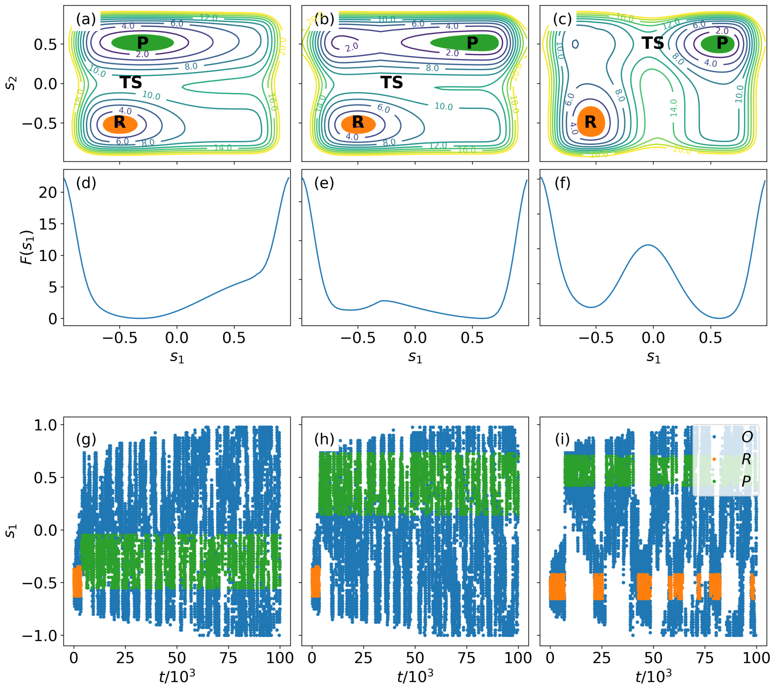

Figure 2: Three potential energy landscapes, the corresponding free energies, and metadynamics trajectories. The three landscapes are representative of cases where the chosen collective variable (CV) cannot distinguish reactant (R) and product (P) (panel (a)), cannot distinguish the transition state (TS) from R (panel (b)), and can distinguish R, P and TS (panel (c)). (a), (b) and (c): The landscapes $V(s_1, s_2)$ are shown as functions of two coordinates (s_1 and s_2) using contour lines. Regions within k_BT from the two mestastable minima are colored in orange (R) and green (P). s_1 represents the chosen (biased) CV. The corresponding free energies $F(s_1)$ are shown in panels (d), (e), and (f). Panels (g), (h), and (i) report metadynamics trajectories for the biased CV s_1. The color of the points depends on where they are located in the (s_1,s_2) space. The points are green (resp, orange) if they are located in the green (resp, orange) region in the top panel. Points that do not belong to these regions are shown in blue.

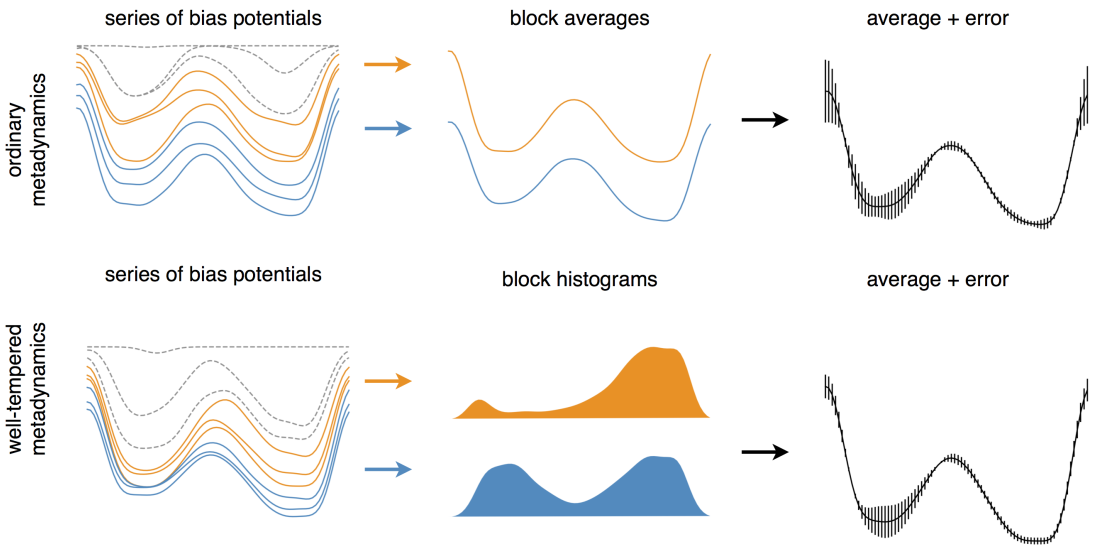

Figure 3: Schematic representation of error calculation in ordinary (upper) and well-tempered (lower) metadynamics. In ordinary metadynamics, the negative bias potential is used to estimate the free energy as a function of the biased CV. As the simulation proceeds, the time series of the potential changes and fluctuates around the correct profile. The initial part of the simulation should be discarded (dashed profiles). After that, a time average of the potential can be computed and block analysis can be used to compute the error. In well-tempered metadynamics a similar procedure can be used to analyze the histogram of the biased CV. In particular, one computes the histogram over a series of blocks and use their standard deviation to obtain the error on the free energy profile. Also in this case it is convenient to discard the initial part of the simulation, where the bias potential has not filled yet the relevant free-energy basins.

ref. [31] it is proposed to perform metadynamics on a one-dimensional variable embedded in multidimensional CV space, whose direction is learned on-the-fly during the simulation. In altruistic metadynamics [32], the computational cost is reduced by simulating simultaneously multiple different molecular systems, predicting simultaneously their free energy surfaces. In Ref. [33] the free energy estimator of metadynamics and of adaptive force bias[34] are combined in a single history-dependent bias, significantly boosting convergence speed. Another variant of metadynamics has been developed to deal with situation in which the presence of large free energy basins stabilized by the entropy hinder convergence. This is the example for instance of protein-ligand binding, where one entropic bottleneck arises from the need of the ligand to find its way to the binding pocket. In cases like this one, where these bottlenecks are a priori known, their effect can be moderated by setting proper restraints [35].

2 Computing the free energy and controlling the error

The manner of estimating the free energy and controlling the error is different in ordinary Metadynamics and in its well-tempered variant. Therefore, we consider these two cases separately.

In ordinary Metadynamics, if the CVs are well chosen, after a time t_{fill} all the free energy minima are filled with Gaussians, and the dynamics becomes diffusive in CV space (see for example Figure 2i). After this time, the bias potential keeps on changing, since new Gaussians are added again and again. However, $B_t(s)$ becomes *stationary*, namely its shape remains qualitatively the same at different times, as shown in Figure 3. $B_t(s)$ behaves like an ordinary observable in a finite temperature molecular simulation: after a transient time (the equilibration time) the observable does not remain constant, but keeps on fluctuating. A meaningful estimator of its thermodynamic average is the *time average*. Similarly, in metadynamics one does not consider meaningful the instantaneous value of the bias potential, but, rather, its time average after t_{fill}, which plays the role of an equilibration time. Therefore, at time t the best estimator of the negative of the free energy is

$$\overline{B}_t(s) = \frac{1}{t - t_{fill}} \sum_{t'=t_{fill}}^{t} B_{t'}(s) \tag{9}$$

This average has been shown to converge exactly to $-F(s)$ for $t \to \infty$ if the dynamics in the CV is adiabatically separated from the dynamics in the other variables [36, 17]. Even if this condition is violated, if transitions between reactants and products are observed on a time scale compatible with diffusion on a flat landscape, the average in Eq. 9 converges to the free energy within numerical accuracy. This is illustrated in the example in figure, and has been verified numerically in a Metadynamics simulation of a lattice model in which the free energy is known analytically [37]. If under the action of metadynamics transitions between reactants and products happen only rarely, like in the example in Figure 2h, the time average is not guaranteed to converge to the negative of the free energy. In particular, transitions should be considered as rare if they happen on a timescale that is compatible with that of an unbiased plain MD simulation, which implies that they are not accelerated by the metadynamics bias. It is important to remark that if in a Metadynamics simulation of a real system one observes a behavior like the one in figure 2-h, one should conclude that the CV is not appropriate, stop the simulation, and try to find a more appropriate variable. The statistical accuracy of the estimator in Eq. 9 is controlled by the same techniques one would use to monitor the accuracy of the average value of an observable in a finite temperature run. For example, one can perform a block analysis: one first splits the simulation time after t_{fill} in blocks and estimate the error by looking at how different are the average bias potentials in the used blocks. One can then monitor how the error estimator depends on the number of blocks. If the total averaging time $t - t_{fill}$ is significantly larger than the correlation time, the error estimate will be approximately independent on the number of blocks (see schematic representation in Fig. 3).

A similar analysis can be performed when running well-tempered metadynamics. In this case, however, one should take into account the fact that since the increments in the bias potential become by construction smaller and smaller as the simulation progresses, the changes in the bias potential itself are not a good indication of the convergence of the simulation. Qualitatively, one should make sure that, even if the bias potential becomes quasi-constant, the system still undergoes transitions between the relevant free-energy minima. Quantitatively, this can be observed by performing a block analysis in the histogram of the biased CV. Indeed, as shown in [22] the standard free-energy estimator used in well-tempered metaydnamics can be replaced with the standard umbrella sampling formula $F(s) = -T \log H(s) + V(s)$, where H is the histogram of the visited points in the CV space, that must be computed with a proper smoothing. This procedure is equivalent to reweighting each visited frame of coordinates q with a factor proportional to $\exp\left(\frac{V(s(q), t_{final}}{T}\right)$, where t_{final} is the simulation length. The following procedure can be used to estimate the error on the free-energy:

- Discard a suitable initial part of the simulation, where the main free-energy wells are filled
- Break the following part of the simulation in blocks and compute the histogram of the CVs in each block, as well as its error using block analysis
- Convert the error on the histogram to an error on the free-energy estimator.

In particular, when computing the error on the histogram one can exploit the relationship between the bias potential and the histogram itself [18], namely $N(s) \propto C + \exp\left(\frac{V(s)}{\Delta T}\right)$, where C is an arbitrary constant. The histogram accumulated in the i-th block is thus proportional to $e^{\frac{B(s, t_{fill}+(i+1)L_b)}{\Delta T}} - e^{\frac{B(s, t_{fill}+iL_b)}{\Delta T}}$, where L_b is the length of the block. The error on the free energy can thus be estimated as

$$\frac{1}{\sqrt{N_b}} \sqrt{\mathrm{Var}_i \left[\log \left(e^{\frac{B(s, t_{fill}+(i+1)L_b)}{\Delta T}} - e^{\frac{B(s, t_{fill}+iL_b)}{\Delta T}} \right) \right]} \tag{10}$$

where N_b is the number of blocks and the variance should be computed across all the blocks. A schematic representation of this procedure is shown in Figure 3.

The free-energy difference between reactants and products evaluated for two of the model systems reported in Fig. 2 is reported in Table 1. Here one can appreciate that in case 2, where

	Reference	Ordinary (short)	Ordinary (long)	WT (short) Final/Average	WT (long) Final/Average
case 1	1.07	1.40 ± 0.34	1.44 ± 0.02	1.20 / 1.47 ± 0.32	1.03 / 1.14 ± 0.09
case 2	1.57	2.45 ± 0.56	1.50 ± 0.02	1.65 / 0.88 ± 0.78	1.57 / 1.59 ± 0.03

Table 1: Estimated free-energy differences between the two minima and statistical errors on the two model landscapes shown in Figure 1b (case 1) and 1c (case 2). The system is evolved for 10^5 steps, as in Figures 1h and 1i (short simulations) or for 10^8 steps (long simulations). Long simulations present a large number of recrossings also for the difficult case 1, and are thus representative of an infinitely long simulation or of a simulation where a different enhanced sampling method was used to enhance the probability of crossing a barrier not seen by the biased CVs. The reference value is obtained from the exact projections shown in Fig. 1e and 1f. When using ordinary metadynamics, the free-energy difference is estimated taking the time average of the bias potential and the statistical error is estimated using block analysis on the bias potential. When using well-tempered metadynamics, the difference is estimated either using the final bias potential or the average bias potential, and the error is computed using block analysis on the histogram.

the biased CV can identify the transition state correctly, both ordinary and well-tempered metadynamics provide results that are compatible with the reference value for a short simulation and converge to a virtually exact estimate when running an infinitely long simulation. On the other hand, in case 1, where the biased CV cannot identify the transition state correctly and transitions are not actually enhanced by the metadynamics potential, ordinary metadynamics reports a biased result. These cases are easy to detect and directly indicate that a different CV should be tried or an additional CV should be added.

3 Automatic determination of the collective variables

As we have seen, cases in which metadynamics does not converge or converges to an incorrect result can be often ascribed to a common problem: the chosen CVs do not correctly describe the relevant barriers. As we discussed in detail, this problem is easy to detect, if the simulation is analyzed properly. Practitioners using metadynamics, when they meet such a problem, are thus immediately exposed to the need to search a "better" CV capable to describe correctly the conformational change of interest. This is perhaps the reason why, especially in the community using metadynamics, a lot of different CVs have been developed in the last years. In this section we review recently developed approaches for “learning” the correct CVs automatically.

We first describe the so-called path CVs [28, 29]. These variables are based on the definition of a series of reference structures for the system under investigation. If a transition from a state A to a state B is to be studied, these reference structures are ideally located in between A and B (see Fig. 4, top). A progression CV is then defined using the following exponential average

$$s(r) = \frac{\sum_i i e^{-\lambda d_i}}{\sum_i e^{-\lambda d_i}} \tag{11}$$

Here d_i is the squared distance between the current atomic configuration and the i-th reference structure, and λ is a smoothing parameter. This exponential average identifies which among the reference structures are closest to the current one and assigns to the CV a value that interpolates between the indexes of those structures. A metadynamics simulation can then be used to enforce transitions between A and B and viceversa biasing this CV. The procedure can be generalized to vector indexes rather than scalar ones in order to allow for a higher dimensional embedding of the configurational space [38]. Importantly, it is possible to optimize the location of the landmark structures through an iterative series of simulations. In Ref. [28], a procedure inspired to nudged-elastic-band [39] was introduced, where at every new simulation the reference structures are changed in order to become more similar to the real intermediates observed in the MD simulation. The actual definition of the CV is thus different at every new iteration. Ideally, the intermediate structures will relax towards a path passing through the transition state and, after a sufficient number of iterations, a metadynamics simulations biasing this CV will be able to make the system diffuse from A to B and viceversa. In Ref. [29], a progression CV with a definition similar to Eq. 11 has been introduced. In this case, however, the definition of the path CV was evolved during the metadynamics simulation, potentially speeding up the search for the optimal path.

The function defined in 11 allows for a highly dimensional representation (coordinates of the atoms) to be reduced to a lower dimensional one (a single CV). More generally, arbitrary features can be used as a starting point to perform a dimensional reduction of this type. Clearly, if some previous knowledge is available, it is possible to exploit it in order to define the distances d_i in an already reduced space. All the automatic methods used to construct CVs are based on the idea of constructing a small number of linear or non-linear functions of a larger set of a priori chosen features.

One of the possible criteria used in dimensional reduction algorithms is that of preserving the distance between structures computed using the full set of coordinates. This is what is done for instance in classic multidimensional scaling [40]. However, one should consider that distances between structures are typically informative only for a narrow range of values. Diffusion maps can be used to tackle this issue [41]. Here, a fictitiuous random walk connecting microstates that are close to each other in the initial feature space is constructed, and the slow modes of this random walk are assigned the larger distances in the low-dimensional representation. The sketch-map algorithm instead constructs a low-dimensional embedding where only distances in a selected range are preserved in the procedure. Specifically, sigmoid functions of the distances in the initial feature space are used in order to tune the distance range that is considered as relevant for the dimensional reduction [42].

		ADVANTAGES	**DISADVANTAGES**
Path CV [28, 38, 29]		It allows describing complex reaction pathways. It can be iteratively optimized	It requires the knowledge of the initial and the final states
Committor parametrization [43]		It provides the best possible reaction coordinate	It requires the knowledge of the initial and the final states; it requires running many short MD trajectories and can be very costly
Spectral gap optimization [44]		Simple and computationally cheap.	It requires the knowledge of the initial and the final states. It finds the best variable in a set, but does not allow parametrizing a new one.
Machine learning	**Distribution based** [41, 42, 45, 46, 47, 48]	It only requires ensemble averages.	Resulting CVs might suboptimally describe barriers.
	Dynamics based [49, 50, 51, 52, 53]	It takes into account explicitly the kinetics of the system.	It requires long unbiased trajectories or reweigthing methods.

Table 2

Recently, machine learning inspired procedures have been used in order to design variables the are capable to distinguish reactants and products. Indeed, CVs or this type are closely related to representations provided by supervised learning algorithms, where a parametric representation is optimized in order to be able to correctly distinguish *a priori* provided labeled examples. In Ref. [45] for instance, support vector machines and logistic regressions where used in order to classify folded and unfolded states of a protein, and the classifier was then used as a biased CV. A similar approach based on linear-discriminant analysis was used in Ref. [46]. This latter approach was applied for instance to the characterization of chemical reactions [47, 48]. The methods discussed so far aim at finding a low-dimensional embedding that correctly represents the structures

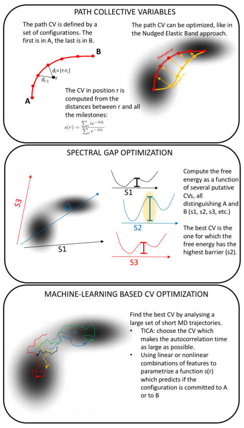

Figure 4: Sketch of three possible approaches for finding automatically the best CV.

seen in a preliminary run or to distinguish preassigned basins in the energetic landscape of the system. However, they are not explicitly taking into account how good these variables would be in representing the kinetics of the system. The ideal CV for describing a transition between two metastable states R and P is the so-called committor function [54]. The committor of a configuration x to the metastable state R is the probability that a trajectory starting from x reaches the state R before P. Finding the committor explicitly is possible only in simple model systems, but several approaches were developed to estimate its value, and parametrize it [43]. These approaches can in principle be used to find an appropriate CV to perform metadynamics.

As we have seen, typically good CVs exhibit large free-energy barriers, that are instead apparently removed when an incorrect dimensional reduction is done mixing the true transition states with other more stable states. The idea of spectral gap optimization [44] is to select linear combinations of putative CVs in order to choose the one that exhibits the slowest transition rate between two minima. Under the assumption that different CVs have comparable diffusion constants, this would be the one with the larger barrier separating reactants and products (see Fig. 4, central panel, for a pictorial representation of this procedure). Performing a proper reweighting is crucial in this approach to recover the free energy along putative linear combinations. With a similar goal, it is possible to use time-independent components analysis (TICA), that construct a linear combination of pre-selected features that is "as slow as possible" (i.e., with the largest possible autocorrelation time). The first few components of a precomputed TICA can be used as biased CVs for metadynamics [49]. Reference [50] also introduced a TICA-based approach, where however the TICA are directly computed during the *biased* simulation, thus allowing conformational changes that are only visible in biased sampling to be studied. We notice that if one performs metadynamics using the correct CV, one changes significantly the relaxation dynamics of the system, since, ideally, the lowest dynamics will take place in the hyperplane of constant CV. Therefore, in order to compute the correct TICAs one should apply a reweighting technique. This method was used then to identify slow molecular motions in complex chemical reactions [55]. Following a similar idea, it was recently shown that variational autoencoders can be used in a similar fashion in order to construct non-linear functions that optimally represent the kinetics of the system [51, 52]. A related approach was also used in [53], where a linear encoder was combined with a non-linear decoder. Limiting the encoder step to linear combinations in principle allows for the generated CVs to be easier to interpret.

4 Implementation

The CVs are often defined by very complex functional forms, but they normally depend on the coordinates of a limited number of atoms. Moreover, the same CVs and the same variants of metadynamics can be used across different applications (e.g., ab initio and classical molecular dynamics). For this reason metadynamics is optimally implemented in separate library (e.g. PLUMED[56], COLVARS[57], or SSAGES[58]), which can then be used in combination with any molecular dynamics code (see Table 3). These libraries typically have their own input files that are read during

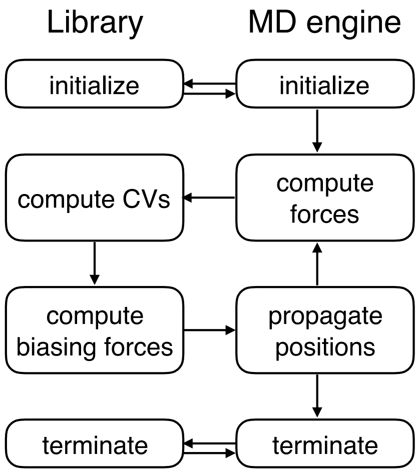

Figure 5: Typical architecture of a library to perform metadynamics simulations. The library is initialized at the beginning of the simulation and usually needs a separate input file specifying the options needed to perform metadynamics. At every step, coordinates are passed to the library and extra forces are received back and added to the physical forces computed by the MD engine. The library is then finalized when the simulation ends.

initialization and are then called at every iteration of the MD simulation (see Fig. 5). Coordinates should be passed to the library. In some cases, this might lead to a slow down of the simulation, in particular if the MD engine stores the coordinates on a graphical processing units whereas the library requires coordinates on the central processing unit. The library then computes the requested CVs and bias potentials resulting forces that should be added to those computed by the MD engine. In principle metadynamics can also be used with Monte Carlo, although we are not aware of Monte Carlo codes interfaced with the mentioned libraries.

The typical aim of these libraries is to allow a user to add arbitrary bias potentials on chosen CVs. In particular, a significant flexibility is usually given to the user in the choice of the CVs since tuning their definition, either manually or automatically as discussed above, is a crucial step in the application of any biasing technique. The code should then compute the derivatives of the i-th CV s_i with respect to $\mathbf{q}_j$, that is the position of the atom j: $\frac{\partial s_i}{\partial \mathbf{q}_j}$. An arbitrarily constructed bias potential can then be applied to these CVs and results in a force $\mathbf{f}_j$ acting on the j-th atom computed via the usual chain rule

$$\mathbf{f}_j = \sum_i \frac{\partial B(s)}{\partial s_i} \frac{\partial s_i}{\partial \mathbf{q}_j} \tag{12}$$

Arbitrary combinations of CVs can also be used, the gradient being propagated in the same manner. At least two of the mentioned packages indeed allow users to specify arbitrary algebraic functions in their input that are then automatically differentiated [57, 56]. The possibility of using arbitrary combinations of CVs makes it possible to implement some of the automatically determined CVs discussed in Section 3 directly in the input script. Whereas this option is often suboptimal from the performance point of view, it clearly speeds up the development of new ideas.

It must be observed that any method based on the idea of adding a bias potential or a force to a set of chosen CVs can be implemented in the same manner. This is the reason why these packages typically provide the user with many other enhanced sampling methods based on biasing CVs, such as umbrella sampling [1], steered MD [75], or other more recently developed techniques.

We also notice that whereas the calculation of metadynamics forces can be implemented by explicitly summing on the history of visited conformation (as in Eq. 7), usually the computational cost is significantly decreased by accumulating the sum of Gaussian functions on a grid so that cost of the calculation of the forces is independent of the simulation length. The update of the bias potential stored on the grid is done using Eq. 6 and its cost grows exponentially with the number of biased CVs. For this reason, it is convenient to use interpolation schemes that allow less dense grids to be employed, thus accelerating the update of the potential.

The availability of the discussed features in the three libraries is summarized in Table 4. The PLUMED package is the most complete in term of support for metadynamics variants, since it was

	native	PLUMED	COLVARS	SSAGES
ACEMD [59]	no	yes	no	no
AMBER [60]	no	yes	no	no
CP2K [61]	yes	yes	no	no
DLPOLY [62]	no	yes	no	no
DESMOND [63]	yes	no	no	no
GROMACS [64]	no	yes	yes	yes
i-Pi [65]	no	yes	no	no
HOOMD [66]	no	no	no	yes
LAMMPS [67]	yes[a]	yes	yes	yes
NAMD [68]	yes[b]	yes	yes	no
OPENMD [69]	no	no	no	yes
OPENMM [70]	yes	yes	no	no
ORAC [71]	yes	no	no	no
PINY-MD [72]	no	yes	no	no
QUANTUM-ESPRESSO [73]	no	yes	no	no
QBOX [74]	no	no	no	yes

Table 3: Availability of metadynamics in a list of commonly used MD codes. The 'native' column refers to implementations of metadynamics that do not require any additional libraries. Compatibility with the three libraries discussed in this review is also indicated. Notice that these three libraries are currently under development and that this table reflects the respective documentation in October 2019.

[a] A copy of the COLVARS code is included in the official LAMMPS repository.
[b] A copy of the COLVARS code is included in the official NAMD repository.

	PLUMED	COLVARS	SSAGES
Ordinary MTD	yes	yes	yes
WT-MTD	yes	yes	no
Grids	yes	yes	yes
Multiple walkers	yes	yes	yes
Bias exchange	yes[a]	no	no
Arbitrary CVs	yes	yes	no

Table 4: Feature comparison of three software that implement metadynamics and can be used as libraries in other MD codes. These three libraries are currently under development. This table reflects the respective documentation at the time of this writing.

[a] only in combination with GROMACS

developed by a community centered around the developers of the method, can be load at runtime as a plugin, and also implements a number of analysis and post-processing tools that might be crucial for dealing with advanced methods. The COLVARS package can be extended using TCL language and some of the non supported features might be implemented using scripts that have been made available by the community. Finally, the SSAGES package is still at a pre-release stage and thus its support is relatively limited. The documentation of the three mentioned packages can also be considered as a starting point to explore the wide range of CVs used in the literature. The PLUMED community has recently made available a public repository of data needed to reproduce enhanced sampling simulations named PLUMED-NEST, that can also be of great use to new users [76].

5 Discussion and perspectives

Metadynamics provides a framework for studying molecular systems affected by metastability, namely characterized by the presence of two or more probability maxima. The guiding principle of this approach is making the dynamics diffusive as a function of a set of suitably chosen collective variables. This is achieved by biasing the dynamics by a history-dependent potential consisting of Gaussians localized along the trajectory in CV space. This working principle proved to be very effective, stimulating the development of several variants and extensions over the years. The most important one is well tempered Metadynamics, in which the height of the Gaussians is progressively reduced during the dynamics, following a protocol which ensures asymptotic convergence. Other important extensions are the ones combining replica exchange and Metadynamics, or weighted histogram analysis and Metadynamics. A question that naturally arises in applications is which variant of the approach should one use. We are now going to discuss this issue.

A preliminary observation is that the convergence of well tempered Metadynamics has been proved rigorously in any condition and for any possible choice of the CV. Indeed, in this method the bias potential asymptotically does not change any more, making the approach technically equivalent to ordinary umbrella sampling. The approach requires choosing an extra parameter, the efficacious temperature ΔT in Eq. 8, whose optimal value depends on the knowledge of the height of the relevant barriers. A possible manner for overcoming the problem of choosing this parameter has been proposed in ref.[77]. However, as we already underlined, if the CV is not correctly chosen, the convergence speed is not significantly enhanced with respect to ordinary molecular dynamics. Ordinary Metadynamics is instead a process in which the dynamics happens in an extended space, including not only the coordinates but also the bias potential. This dynamics has the advantage to enforce diffusion in the CV even when this is not correct. The properties of this dynamics have not yet been fully understood, and its convergence was rigorously proved only even in conditions of adiabatic separation of the CV dynamics[17]. If this condition is violated, systematic errors may arise, like those observed in the examples described in this review, and as rigorously proven in Ref. [78]. However, if the CVs are appropriately chosen these errors are in practice well below the statistical accuracy of the free energy estimator[37]. It is also important to recall that in practical applications, where adiabatic separation is often violated, it can be convenient reconstructing the free energy *a posteriori*, not as an average of the bias potential, but using estimators based on the mean force observed during the biased run[57, 79] or Gaussian process regression[80].

To understand better the differences between these different approaches it is important to discuss what happens if the CV are not appropriately chosen. In Figure 2 we presented three examples of two-dimensional potential energy landscape associated to a system with two metastable states. In two cases, in particular, the free energy as a function of the x coordinate is very similar (two minima separated by a barrier). However, we have seen that the capability of the method of estimating the free energy is strikingly different in the two cases. In one of the cases, estimating the free energy is practically impossible: the collective variable identifies the two metastable states, but not the transition state between the two, and therefore the bias potential does not accelerate the transitions. In figure 6, left, we show a potential energy landscape characterized by the presence of 4 minima. The free energy landscapes as a function of x and y, also shown in same Figure, are approximately flat, despite the presence of very significant barriers in the two-dimensional landscape. In this case, if one uses the x as a CV, well tempered Metadynamics is by construction unable to generate a bias potential capable of enhancing the sampling in the x direction, and the same happens if one uses y as a CV. Indeed, at convergence the bias potential is a constant, and this bias is not affected anymore by the new Gaussians, whose height becomes smaller and smaller at the end of the run. Therefore, the system will get stuck in one of the minima. In short, the problem is that a bias potential compensating exactly the free energy does not necessarily make the landscape barrierless. The behaviour of ordinary Metadynamics on this system is rather different: the approach by construction enforces transitions even when the collective variable is not correct, as long as the CV takes different values in the different minima. The system will go on performing transitions but the bias potential will be affected by large fluctuations, and the free energy estimate will not converge. Indeed, since the transition states are not identified correctly by the CV, forward and backward transitions might follow different paths, determining hysteresis in the estimated free energy.

A possible way out to address this problem is offered by replica exchange methods. One can combine metadynamics with parallel tempering [81], collective variable tempering[26] or solute tempering [82] in order to enhance sampling in directions that are not directly biased. These approaches address to some extent the problem of metastability in degrees of freedom which are orthogonal to the biased collective variable. Another manner for addressing the same problem is by using Bias Exchange metadynamics[25], an approach in which several metadynamics are run in parallel and at the same temperature, each biasing a different CV. Exchanges of the coordinates between different replicas are attempted at regular time intervals, and accepted according to the Metropolis criterion. This approach allows using at the same time a very large number of CVs, and dramatically reduces the hysteresis if one biases in at least one replica all the relevant CVs.

Using replica Bias Exchange in combination with well-tempered Metadynamics is delicate, since in this approach the simplest manner of checking if the simulation is meaningful requires verifying if the system at the end of the simulation is able to diffuse through all the region in which one wants to estimate the free energy. In a situation like the one depicted in Figure 6, if one uses a single replica one will immediately notice that the CV tends to freeze in one of the local minima. If one uses any replica exchange method, an accepted exchange move can induce a jump of the CV from one minimum to another one. However, these jumps are not sufficient to ensure that the free energy estimator is correct. One should rather check that continuous trajectories, traveling across the space of available replicas, display transitions between the relevant states.

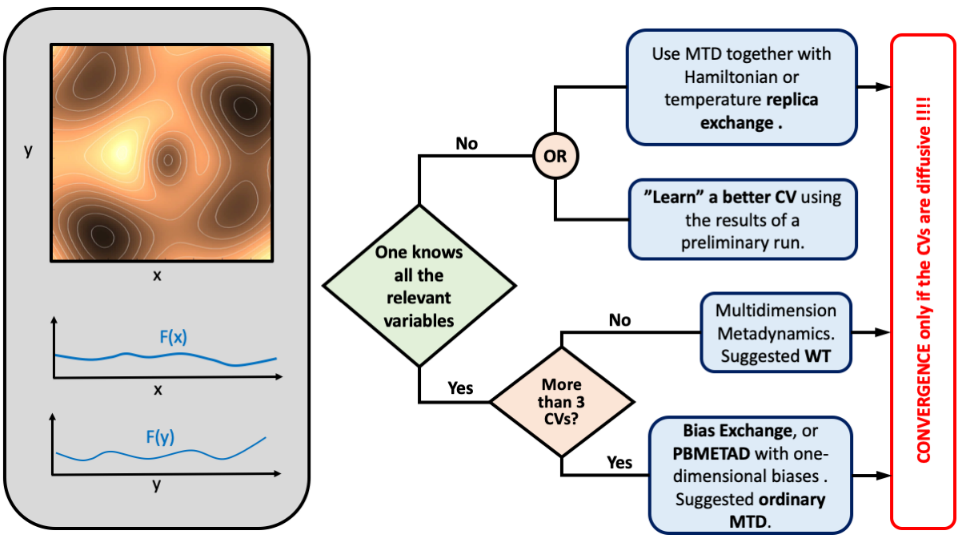

Figure 6: Left: an example of a potential energy landscape characterized by the presence of deep minima, but whose free energy as a function of two variables (x and y) is approximately flat. This situation is rather common in practical applications. Right: a decision tree for choosing the most appropriate version of metadynamics.

These considerations allow drawing some guidelines for choosing among the different versions of Metadynamics, summarized in the decision tree in figure 6, right. The well-tempered version of metadynamics in the opinion of the authors offers some advantages for estimating the free energy as a simultaneous function of two or more collective variables. Indeed, this approach has been proved to converge rigorously, and allows performing the calculation without defining explicitly the region in which the free energy should be estimated: one simply chooses the maximum value of the free energy that one considers interesting, and the approach automatically fills the free energy minima approximately up to that level. If one uses ordinary Metadynamics to estimate the free energy in a multidimensional domain, one should fix the appropriate boundary conditions, restricting the dynamics in the domain[83, 37, 8].

If one does not know the correct CV one can consider using a method combining Metadynamics and replica exchange. In Bias Exchange Metadynamics one can bias simultaneously an arbitrarily large number of CV. Ordinary Metadynamics has the advantage to force transitions even when the CV is not correct, like in the example of Fig 6. This features allows converging a low-dimensional free energy estimate even in the condition of Fig 6, if one uses the approach in combination with replica exchange. Well-tempered metadynamics instead might remain stuck unless the CV describe very well the transition state. Therefore, the authors see some advantages in using ordinary metadynamics with respect to the WT version in bias-exchange methods. It, is also important underlining that in any version of Metadynamics the free energy estimator is practically meaningful only if at the end of the simulation the dynamics in CV space is approximately diffusive, namely if the CV is able to explore in a short time all the relevant regions of the CV space. This can be considered as the true general guiding principle of this approach, in any of its variants.

As we hope it was clear in this review, one should always bear in mind that the usefulness of metadynamics (and of any enhanced sampling method based on biasing CVs) is largely determined by the capability to identify a proper CV able to describe the relevant transitions. Historically, CVs have been searched by trial and error, and their discovery has been part of the process of understanding the system under investigation. Going beyond this protocol is, in our opinion, the frontier of enhanced sampling methods. The community is well aware of the importance of this problem: recently, a number of approaches for the automatic search of CVs have been proposed. Some of these approaches, inspired by the quickly developing field of machine learning, are described in Section 3. Automatically training procedures allow using functions of arbitrary complexity, such as artificial neural networks. Whereas this can bring to important steps forward, one should also consider that complex functional forms might be difficult to interpret and, in the end, could teach less about the investigated system. In addition, highly flexible functional forms might easily lead to situations where data are overfitted. We thus believe that finding a solution to the problem of automatically finding the CVs ensuring the appropriate balance between accuracy, generality and interpretability, is still an open problem, that will likely attract a lot of interest in the near future.

References

[1] Torrie, G. M. & Valleau, J. P. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J. Comput. Phys.* **23**, 187–199 (1977).

[2] Laio, A. & Parrinello, M. Escaping free energy minima. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 12562–12566 (2002).

[3] Laio, A. & Gervasio, F. L. Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. *Rep. Prog. Phys.* **71**, 126601 (2008).

[4] Barducci, A., Bonomi, M. & Parrinello, M. Metadynamics. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **1**, 826–843 (2011).

[5] Sutto, L., Marsili, S. & Gervasio, F. L. New advances in metadynamics. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2**, 771–779 (2012).

[6] Bussi, G. & Branduardi, D. Free-energy calculations with metadynamics: Theory and practice. *Rev. Comput. Chem.* **28**, 1–49 (2015).

[7] Valsson, O., Tiwary, P. & Parrinello, M. Enhancing important fluctuations: Rare events and metadynamics from a conceptual viewpoint. *Annu. Rev. Phys. Chem.* **67**, 159–184 (2016).

[8] Baftizadeh, F., Cossio, P., Pietrucci, F. & A., L. Protein folding and ligand-enzyme binding from bias-exchange metadynamics simulations. *Curr. Phys. Chem.* **2**, 79–91 (2012).

[9] Theodoropoulos, C., Qian, Y.-H. & Kevrekidis, I. G. Coarse stability and bifurcation analysis using time-steppers: A reaction-diffusion example. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 9840–9843 (2000).

[10] Grubmüller, H. Predicting slow structural transitions in macromolecular systems: Conformational flooding. *Phys. Rev. E* **52**, 2893–2906 (1995).

[11] Glover, F. Future paths for integer programming and links to artificial intelligence. *Comput. Oper. Res.* **13**, 533–549 (1986).

[12] Cvijović, D. & Klinowski, J. Taboo search - an approach to the multiple minima problem. *Science* **267**, 664–666 (1995).

[13] Huber, T., Torda, A. E. & Van Gunsteren, W. F. Local elevation: a method for improving the searching properties of molecular dynamics simulation. *J. Comput. Aided Mol. Des.* **8**, 695–708 (1994).

[14] Wang, F. & Landau, D. Efficient, multiple-range random walk algorithm to calculate the density of states. *Phys. Rev. Lett.* **86**, 2050 (2001).

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

[16] Rosenblatt, M. Remarks on some nonparametric estimates of a density function. *Ann. Math. Stat.* **27**, 832–837 (1956).

[17] Bussi, G., Laio, A. & Parrinello, M. Equilibrium free energies from nonequilibrium metadynamics. *Phys. Rev. Lett.* **96**, 090601 (2006).

[18] Barducci, A., Bussi, G. & Parrinello, M. Well-tempered metadynamics: A smoothly converging and tunable free-energy method. *Phys. Rev. Lett.* **100**, 020603 (2008).

[19] Spall, J. C. *Stochastic Gradient Form of Stochastic Approximation* (2005).

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

[21] Bonomi, M., Barducci, A. & Parrinello, M. Reconstructing the equilibrium boltzmann distribution from well-tempered metadynamics. *J. Comput. Chem.* **30**, 1615–1621 (2009).

[22] Branduardi, D., Bussi, G. & Parrinello, M. Metadynamics with adaptive Gaussians. *J. Chem. Theory Comput.* **8**, 2247–2254 (2012).

[23] Tiwary, P. & Parrinello, M. A time-independent free energy estimator for metadynamics. *J. Phys. Chem. B* **119**, 736–742 (2014).

[24] Donati, L. & Keller, B. G. Girsanov reweighting for metadynamics simulations. *J. Chem. Phys.* **149**, 072335 (2018).

[25] Piana, S. & Laio, A. A bias-exchange approach to protein folding. *J. Phys. Chem. B* **111**, 4553–4559 (2007).

[26] Gil-Ley, A. & Bussi, G. Enhanced conformational sampling using replica exchange with collective-variable tempering. *J. Chem. Theory Comput.* **11**, 1077–1085 (2014).

[27] Pfaendtner, J. & Bonomi, M. Efficient sampling of high-dimensional free-energy landscapes with parallel bias metadynamics. *J. Chem. Theory Comput.* **11**, 5062–5067 (2015).

[28] Branduardi, D., Gervasio, F. L. & Parrinello, M. From A to B in free energy space. *J. Chem. Phys.* **126**, 054103 (2007).

[29] Leines, G. D. & Ensing, B. Path finding on high-dimensional free energy landscapes. *Phys. Rev. Lett.* **109**, 020601 (2012).

[30] Awasthi, S., Kapil, V. & Nair, N. Sampling free energy surfaces as slices by combining umbrella sampling and metadynamics. *J. Comp. Chem.* **37**, 1413–1424 (2016).

[31] Marinelli, F. Following easy slope paths on a free energy landscape: The case study of the trp-cage folding mechanism. *Biophysical journal* **105**, 1236–1247 (2013).

[32] Hošek, P., Toulcová, D., Bortolato, A. & Spiwok, V. Altruistic metadynamics: Multisystem biased simulation. *J Phys Chem B* **120**, 2209–2215 (2016).

[33] Fu, H. *et al.* Zooming across the free-energy landscape: Shaving barriers, and flooding valleys. *J. Phys. Chem. Lett.* **9**, 4738–4745 (2018).

[34] Darve, E. & Pohorille, A. Calculating free energies using average force. *J. Chem. Phys.* **115**, 9169–9183 (2001).

[35] Limongelli, V., Bonomi, M. & Parrinello, M. Funnel metadynamics as accurate binding free-energy method. *Procs. Natl. Acad. Sci. U.S.A.* **110**, 6358–6363 (2013).

[36] Iannuzzi, M., Laio, A. & Parrinello, M. Efficient exploration of reactive potential energy surfaces using car-parrinello molecular dynamics. *Phys. Rev. Lett.* **90**, 238302 (2003).

[37] Crespo, Y., Marinelli, F., Pietrucci, F. & Laio, A. Metadynamics convergence law in a multidimensional system. *Phys. Rev. E* **81**, 055701 (2010).

[38] Spiwok, V. & Králová, B. Metadynamics in the conformational space nonlinearly dimensionally reduced by Isomap. *J. Chem. Phys.* **135**, 224504 (2011).

[39] Henkelman, G., Uberuaga, B. P. & Jónsson, H. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *J. Chem. Phys.* **113**, 9901–9904 (2000).

[40] Cox, T. F. & Cox, M. A. *Multidimensional scaling* (Chapman and hall/CRC, 2000).

[41] Rohrdanz, M. A., Zheng, W., Maggioni, M. & Clementi, C. Determination of reaction coordinates via locally scaled diffusion map. *J. Chem. Phys.* **134**, 124116 (2011).

[42] Tribello, G. A., Ceriotti, M. & Parrinello, M. Using sketch-map coordinates to analyze and bias molecular dynamics simulations. *Proc. Natl. Acad. Sci. USA* **109**, 5196–5201 (2012).

[43] Peters, B., Beckham, G. T. & Trout, B. L. Extensions to the likelihood maximization approach for finding reaction coordinates. *J. Chem. Phys.* **127**, 034109 (2007).

[44] Tiwary, P. & Berne, B. Spectral gap optimization of order parameters for sampling complex molecular systems. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 2839–2844 (2016).

[45] Sultan, M. M. & Pande, V. S. Automated design of collective variables using supervised machine learning. *J. Chem. Phys.* **149**, 094106 (2018).

[46] Mendels, D., Piccini, G., Brotzakis, Z. F., Yang, Y. I. & Parrinello, M. Folding a small protein using harmonic linear discriminant analysis. *J. Chem. Phys.* **149**, 194113 (2018).

[47] Piccini, G. & Parrinello, M. Accurate quantum chemical free energies at affordable cost. *J. Phys. Chem. Lett.* **10**, 3727–3731 (2019).

[48] Rizzi, V., Mendels, D., Sicilia, E. & Parrinello, M. Blind search for complex chemical pathways using harmonic linear discriminant analysis. *J. Chem. Theory Comput.* **15**, 4507–4515 (2019).

[49] M. Sultan, M. & Pande, V. S. TICA-metadynamics: Accelerating metadynamics by using kinetically selected collective variables. *J. Chem. Theory Comput.* **13**, 2440–2447 (2017).

[50] McCarty, J. & Parrinello, M. A variational conformational dynamics approach to the selection of collective variables in metadynamics. *J. Chem. Phys.* **147**, 204109 (2017).

[51] Chen, W. & Ferguson, A. L. Molecular enhanced sampling with autoencoders: On-the-fly collective variable discovery and accelerated free energy landscape exploration. *J. Comput. Chem.* **39**, 2079–2102 (2018).

[52] Wehmeyer, C. & Noé, F. Time-lagged autoencoders: Deep learning of slow collective variables for molecular kinetics. *J. Chem. Phys.* **148**, 241703 (2018).

[53] Wang, Y., Ribeiro, J. M. L. & Tiwary, P. Past–future information bottleneck for sampling molecular reaction coordinate simultaneously with thermodynamics and kinetics. *Nat. Commun.* **10**, 1–8 (2019).

[54] Vanden-Eijnden, E. Transition-path theory and path-finding algorithms for the study of rare events. *Annu. Rev. Phys. Chem.* **61**, 391–420 (2010).

[55] Piccini, G., Polino, D. & Parrinello, M. Identifying slow molecular motions in complex chemical reactions. *J. Phys. Chem. Lett.* **8**, 4197–4200 (2017).

[56] Tribello, G. A., Bonomi, M., Branduardi, D., Camilloni, C. & Bussi, G. PLUMED2: New feathers for an old bird. *Comput. Phys. Comm.* **185**, 604–613 (2014).

[57] Fiorin, G., Klein, M. L. & Hénin, J. Using collective variables to drive molecular dynamics simulations. *Mol. Phys.* **111**, 3345–3362 (2013).

[58] Sidky, H. *et al.* SSAGES: Software suite for advanced general ensemble simulations. *J. Chem. Phys.* **148**, 044104 (2018).

[59] Harvey, M. J., Giupponi, G. & Fabritiis, G. D. Acemd: accelerating biomolecular dynamics in the microsecond time scale. *J. Chem. Theory Comput.* **5**, 1632–1639 (2009).

[60] Case, D. A. *et al.* The Amber biomolecular simulation programs. *J. Comput. Chem.* **26**, 1668–1688 (2005).

[61] Hutter, J., Iannuzzi, M., Schiffmann, F. & VandeVondele, J. cp2k: atomistic simulations of condensed matter systems. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **4**, 15–25 (2014).

[62] Todorov, I. T., Smith, W., Trachenko, K. & Dove, M. T. Dl_poly_3: new dimensions in molecular dynamics simulations via massive parallelism. *J. Mater. Chem.* **16**, 1911–1918 (2006).

[63] Bowers, K. J. *et al.* Scalable algorithms for molecular dynamics simulations on commodity clusters. In *SC'06: Proceedings of the 2006 ACM/IEEE Conference on Supercomputing*, 43–43 (IEEE, 2006).

[64] Abraham, M. J. *et al.* Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1**, 19–25 (2015).

[65] Kapil, V. *et al.* i-pi 2.0: A universal force engine for advanced molecular simulations. *Comput. Phys. Commun.* **236**, 214–223 (2019).

[66] Anderson, J. A., Lorenz, C. D. & Travesset, A. General purpose molecular dynamics simulations fully implemented on graphics processing units. *J. Comput. Phys.* **227**, 5342–5359 (2008).

[67] Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* **117**, 1–19 (1995).

[68] Phillips, J. C. *et al.* Scalable molecular dynamics with namd. *J. Comput. Chem.* **26**, 1781–1802 (2005).

[69] Louden, P. *et al.* OPENMD-2.5: Molecular dynamics in the open (2017).

[70] Eastman, P. *et al.* OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS Comput. Biol.* **13**, e1005659 (2017).

[71] Procacci, P. Hybrid mpi/openmp implementation of the orac molecular dynamics program for generalized ensemble and fast switching alchemical simulations. *J. Chem. Inf. Model.* **56**, 1117–1121 (2016).

[72] Tuckerman, M. E., Yarne, D., Samuelson, S. O., Hughes, A. L. & Martyna, G. J. Exploiting multiple levels of parallelism in molecular dynamics based calculations via modern techniques and software paradigms on distributed memory computers. *Comput. Phys. Commun.* **128**, 333–376 (2000).

[73] Giannozzi, P. *et al.* QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Phys. Condens. Matter* **21**, 395502 (2009).

[74] Gygi, F. Architecture of Qbox: A scalable first-principles molecular dynamics code. *IBM J. Res. Dev.* **52**, 137–144 (2008).

[75] Grubmüller, H., Heymann, B. & Tavan, P. Ligand binding: molecular mechanics calculation of the streptavidin-biotin rupture force. *Science* **271**, 997–999 (1996).

[76] The PLUMED consortium. Promoting transparency and reproducibility in enhanced molecular simulations. *Nat. Methods* **16**, 670–673 (2019).

[77] Dama, J., Rotskoff, G., Parrinello, M. & Voth, G. Transition-tempered metadynamics: Robust, convergent metadynamics via on-the-fly transition barrier estimation. *J. Chem. Theory Comput.* **10**, 3626–3633 (2014).

[78] Jourdain, B., Lelièvre, T. & Zitt, P.-A. Convergence of metadynamics: discussion of the adiabatic hypothesis arXiv:1904.08667 (2019).

[79] Cuendet, M. & Tuckerman, M. Free energy reconstruction from metadynamics or adiabatic free energy dynamics simulations. *J. Chem. Theory Comput* **10**, 2975–2986 (2014).

[80] Mones, L., Bernstein, N. & Csányi, G. Exploration, sampling, and reconstruction of free energy surfaces with gaussian process regression. *J. Chem. Theory Comput* **12**, 5100–5110 (2016).

[81] Bussi, G., Gervasio, F. L., Laio, A. & Parrinello, M. Free-energy landscape for beta hairpin folding from combined parallel tempering and metadynamics. *J. Am. Chem. Soc.* **128**, 13435–13441 (2006).

[82] Camilloni, C., Provasi, D., Tiana, G. & Broglia, R. A. Exploring the protein G helix free-energy surface by solute tempering metadynamics. *Proteins* **71**, 1647–1654 (2008).

[83] McGovern, M. & De Pablo, J. A boundary correction algorithm for metadynamics in multiple dimensions. *J. Chem. Phys.* **140**, 229901 (2013).