

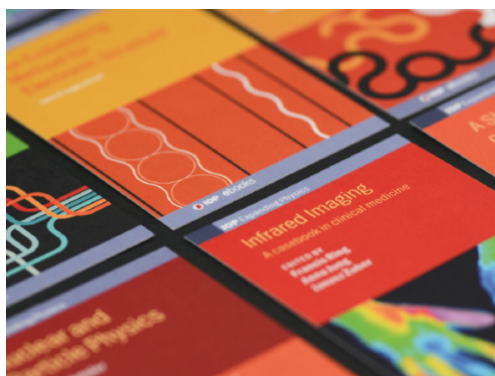
Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science

To cite this article: Alessandro Laio and Francesco L Gervasio 2008 *Rep. Prog. Phys.* **71** 126601

View the [article online](#) for updates and enhancements.

You may also like

- [An investigation of the tensile deformation and failure of an epoxy/Cu interface using coarse-grained molecular dynamics simulations](#)
Shaorui Yang and Jianmin Qu
- [Using first-principles metadynamics simulation to predict new phases and probe the phase transition of \$\text{NaAlH}_4\$](#)
Jianjun Liu and Qingfeng Ge
- [Metadynamics investigations of the AlN/GaN superlattice](#)
Yifeng Duan, Lixia Qin and Hanyu Liu



IOP | ebooks™

Bringing together innovative digital publishing with leading authors from the global scientific community.

Start exploring the collection—download the first chapter of every title for free.

Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science

Alessandro Laio¹ and Francesco L Gervasio²

¹ SISSA/ISAS, Statistical and biological physics, Via Beirut 2-4 Trieste, Italy

² Computational Science, Department of Chemistry and Applied Biosciences, ETH Zürich, USI Campus, Via Giuseppe Buffi 13, CH-6900 Lugano, Switzerland

Received 28 August 2008, in final form 27 October 2008

Published 26 November 2008

Online at stacks.iop.org/RoPP/71/126601

Abstract

Metadynamics is a powerful algorithm that can be used both for reconstructing the free energy and for accelerating rare events in systems described by complex Hamiltonians, at the classical or at the quantum level. In the algorithm the normal evolution of the system is biased by a history-dependent potential constructed as a sum of Gaussians centered along the trajectory followed by a suitably chosen set of collective variables. The sum of Gaussians is exploited for reconstructing iteratively an estimator of the free energy and forcing the system to escape from local minima. This review is intended to provide a comprehensive description of the algorithm, with a focus on the practical aspects that need to be addressed when one attempts to apply metadynamics to a new system: (i) the choice of the appropriate set of collective variables; (ii) the optimal choice of the metadynamics parameters and (iii) how to control the error and ensure convergence of the algorithm.

(Some figures in this article are in colour only in the electronic version)

This article was invited by Professor M Finnis

Contents

1. The algorithm	3	4.2. A practical example: alanine dipeptide in vacuum	15
1.1. Lagrangian metadynamics	5		
1.2. Discrete metadynamics	6	5. Extensions	17
2. The choice of CVs	7	5.1. Multiple walkers	17
2.1. Examples of collective variables	7	5.2. Parallel tempering metadynamics	17
3. Estimating the error	10	5.3. Bias exchange	18
3.1. Reducing the error by averaging several profiles	14	6. Conclusions and outlook	19
4. The algorithm in practice	14	Acknowledgments	19
4.1. Implementation	14	References	20

List of Abbreviations

CV	collective variables
FES	free energy surface
MD	molecular dynamics
FPMD	first-principle molecular dynamics
V_G	metadynamics potential
TMA	tetramethylammonium

AChE	acetylcholinesterase
PCV	path-like variables
RMSD	root mean square displacement
CMA	contact map
Nhb	number of backbone hydrogen bonds
NC_α	number of carbon alpha contacts
δs	width of the deposited Gaussian
w	height of the deposited Gaussian

τ_G	frequency of Gaussians deposition
WHAM	weighted histogram analysis method
PTMetaD	Parallel tempering metadynamics
BE	bias exchange
cHP	C-terminal headpiece (of Villin and Advillin)

Atomistic simulations of complex systems are nowadays routinely exploited in solid state physics, biophysics and chemistry and are becoming increasingly useful, as they have the potential to investigate processes with a high resolution in time and space. However in several important cases simulations are still not competitive if compared with more empirical methods. This is mainly due to the fact that using atomistic models is computationally expensive, as sufficiently realistic potential energy functions are intrinsically complex. Moreover, the dynamics of realistic systems spans a wide range of characteristic times and the integration time step that is used for evolving the system has to be commensurate with the fastest dynamics, namely of the order of 1 fs. As a result, the time scales that can be currently simulated are in the range of hundreds of nanoseconds for classical molecular dynamics and in the range of hundreds of picoseconds for first-principles molecular dynamics (FPMD). Most phenomena of interest take place on time scales that are orders of magnitude larger, and are, therefore, *rare events* on the currently accessible simulation time. Examples of rare events are chemical reactions and structural phase transitions and, in biophysics, protein folding, protein–protein interactions and molecular recognition. Given the spectacular increase in computer efficiency that we are witnessing, it is very probable that in the near future it will be possible to observe by direct simulation the reactive trajectories of more and more complex systems. Industry and academia are building larger and larger machines designed at this scope. Prominent examples are Blue Gene, the massively parallel supercomputer developed at IBM, and Desmond, the purpose-built machine developed at DESRES. Both supercomputers are designed to make the time scale necessary for folding an average size protein accessible to simulation, and the enormous sum of money that is invested in the projects is justified by the striking practical and theoretical importance of this goal. Still, the brute-force approach will probably remain for a long time available only to a few research groups. Moreover, observing one or a few reactive trajectories might not be sufficient for converging statistical averages and computing observables that can be compared with experiments.

A different way to overcome these difficulties is to renounce the all-atom description and use instead a coarse-grained model. This would retain only those characteristics which are essential but would require a detailed knowledge of the systems that is often unavailable.

If one does not want to renounce the atomistic description, one can exploit a methodology aimed at *accelerating rare events* using the available computer time with improved efficiency. Using these approaches, notable success has been achieved in several fields, ranging from solid state physics to quantum chemistry. Broadly speaking these methods can be

classified in four categories, according to their scope and range of applicability:

- (i) Methods aimed at reconstructing the probability distribution or enhancing the sampling as a function of one or a few predefined collective variables (CVs). For instance, in a chemical reaction one would choose the distance between two atoms that have to form a bond or, in the study of nucleation, the size of the nucleus and enhance the sampling as a function of these coordinates [1]. Examples of these methods include thermodynamic integration [2, 3], free energy perturbation [4], umbrella sampling [5], conformational flooding [6], weighted histogram techniques [7–9], Jarzynski's identity-based methods [10, 11], adaptive force bias [12, 13], steered MD [14] and adiabatic molecular dynamics [15]. These approaches are very powerful but require a careful choice of the CVs that must provide a satisfactory description of the reaction coordinate. If an important variable is forgotten they suffer from hysteresis and lack of convergence. Moreover, when more than a few CVs are used, the computational performance rapidly degrades as a function of the number of variables.
- (ii) Methods aimed at exploring the transition mechanism and constructing reactive trajectories [16], such as nudged elastic band [17], finite-temperature string method [18, 19], transition path sampling [20–22], transition interface sampling [23], milestoning [24] and forward flux method [25]. These methods do not require in most of the cases the explicit definition of a reaction coordinate, but require an *a priori* knowledge of the initial and final states of the process that has to be simulated. For instance, if applied to the study of folding, these methods require a knowledge of the folded and 'unfolded' state [26].
- (iii) Methods for exploring the potential energy surface and localizing the saddle points that correspond to the transition states such as eigenvalue following [27], the dimer method [28], hyperdynamics [29], multiple-time scale accelerated molecular dynamics [30] event-based relaxation [31]. These approaches are extremely powerful for exploring potential energy surfaces of low dimensionality, but their reliability degrades with the complexity of the system. Indeed, for very large or complex systems the number of possible transition states surrounding a minimum becomes rapidly too large for a deterministic search. Even if strategies have been designed to alleviate this problem that are effective in some special cases [31], in solvated systems the concept of saddle point on the potential energy surface becomes fuzzy, and these approaches cannot easily be applied.
- (iv) Methods in which the phase space is explored simultaneously at different values of the temperature, such as parallel tempering [32] and replica exchange [33], or as a function of the potential energy, such as multicanonical MD [34] and Wang–Landau [35]. These approaches are very general and powerful; however, they are not immune from some of the limitations listed in point 1. Indeed, these methods exploit more or less explicitly the potential energy as a generalized CV. In several cases, ordered and disordered states may correspond to the same value of

potential energy, or be present in the thermal ensemble at the same temperature. This may lead to hysteresis and convergence problems [36].

The metadynamics method [37] encompasses several features of other techniques and provides in many cases a unified framework for computing free energies and accelerating rare events. The algorithm is based on a dimensional reduction, in the spirit of the work by Kevrekidis [38, 39]. Like the approaches listed in point (i) above, metadynamics requires the preliminary identification of a set of CVs which are assumed to be able to describe the process of interest. Its power, as will be shown in the following, lies in its ability to treat several CVs simultaneously and in its flexibility: the method can be proficiently used both for reconstructing the free energy and for accelerating rare events. The dynamics in the space of the chosen CVs is enhanced by a history-dependent potential constructed as a sum of Gaussians centered along the trajectory followed by the CVs. This manner of biasing the evolution was first used by the taboo search method [40] and, in the context of MD, by the local elevation method [41]. A similar approach is also found in the Wang and Landau algorithm [35], adaptive force bias [12] and self-healing umbrella sampling [42]. In metadynamics the sum of Gaussians is exploited to reconstruct iteratively an estimator of the free energy.

The working principle of the algorithm can be qualitatively understood by a simple example. Imagine a walker who, during the night, falls into an empty swimming pool. The walls of the swimming pool are too steep for the walker to climb and the complete darkness makes it difficult for him to localize the shallowest point (lowest saddle). In these conditions the walker will move more easily downhill, and it is rather unlikely that he will find by chance the lowest saddle. His walk in these conditions resembles that performed by microscopic systems in normal molecular dynamics or Monte Carlo: a random walk with a bias in the direction of lower free energy, with a very small probability to explore transition regions (climb out of the pool). In metadynamics, the walker has access to a large source of sand that he can deposit in his current position. The sand will slowly fill the pool. Thus, even if at the beginning he visits more often the region at the bottom of the pool, little by little he fills the pool with sand (Gaussians), and he almost deterministically starts exploring regions that are higher and higher. Sooner or later, the walker is destined to fill sufficiently the pool to be able to climb out of it. And most probably he will climb out from the shallowest point of the pool. The novel idea that differentiates metadynamics from pre-existing methods is that if the walker is able to keep memory of all the positions in which he has deposited sand (the Gaussians), he will be able to reconstruct a negative image of the underlying pool (the free energy). More precisely, one assumes that the time dependent potential defined by the sum of Gaussians deposited up to time t provides an unbiased estimate of the free energy in the region explored during the dynamics. This property, that does not follow from any ordinary thermodynamic identity, such as umbrella sampling [5], was postulated on a heuristic basis in [37], and afterwards verified empirically in several systems of increasing complexity. Successively [43], it was shown

that this property derives from rather general principles, and can be demonstrated rigorously for a system evolving under the action of a Langevin dynamics. This will be discussed in detail in section 3.

Since the history-dependent potential iteratively compensates the underlying free energy, a system evolved with metadynamics tends to escape from any free energy minimum via the lowest free energy saddle point. This makes metadynamics a rather flexible tool that can be used not only to compute efficiently the free energy but also to explore new reaction pathways and accelerate the observation of rare events. If the CVs are chosen sensibly the system will quickly find its way over the lowest free energy saddle point and evolve over the next minimum as it would eventually do in a very long MD simulation. This flexibility is reflected in the disparate fields to which the method has been applied so far, including solid state and material science [44–53], crystal structure prediction [54–60], biophysics [61–74] and chemistry [75–105].

This review is organized as follows. In section 1 the metadynamics algorithm is described along with its main variants. In section 2 the manner in which the CVs have to be chosen is described and the most useful CVs used in the literature are listed. In section 3 the algorithm is analyzed in more detail, proving that the sum of the Gaussians that bias the dynamics provides an unbiased estimate of the free energy. Moreover, an explicit expression of the error as a function of the metadynamics parameter is provided. Section 4 contains practical examples of implementation and application of metadynamics. Finally, section 5 reports on the most recent extensions of the method, aimed at improving the capability of the algorithm to treat systems in which it is not easy to select *a priori* a small yet complete set of CVs.

1. The algorithm

Consider a system described by a set of coordinates x and a potential $V(x)$ evolving under the action of a dynamics, which could be for instance Langevin, Newtonian (under the action of a thermostat) or Monte Carlo, whose equilibrium distribution is canonical at a temperature T . The set of coordinates x may include ordinary atomic positions, but also electronic coordinates, as in Car–Parrinello molecular dynamics [106], box shape, as in Parrinello–Rahman [107], or any other auxiliary variables. We are interested in exploring the properties of the system as a function of a finite number of collective variables (CVs) $S_\alpha(x)$, $\alpha = 1, d$ where d is a small number, assuming that they provide a good coarse-grained description. The CVs can be any explicit function of x such as an angle, a distance, a coordination number or the potential energy (see section 2). The equilibrium behavior of these variables is completely defined by the probability distribution:

$$P(s) = \frac{\exp(-(1/T)F(s))}{\int ds \exp(-(1/T)F(s))}, \quad (1)$$

where s denotes the d dimensional vector (s_1, \dots, s_d) and the free energy $F(s)$ is given by

$$F(s) = -T \ln \left(\int dx \exp \left(-\frac{1}{T} V(x) \right) \delta(s - S(x)) \right). \quad (2)$$

In equation (2) (and in the following) capital S is used for denoting the function of the coordinates $S(x)$, while lower case s is used for denoting the value of the CVs.

Consider now a trajectory $x(t)$ at a temperature T . If this trajectory could be computed for a very long time, $P(s)$ could be obtained by taking the histogram of the CV s along this trajectory, i.e. at time t , $P(s) \sim \frac{1}{t} \int_0^t dt' \delta(S(x(t')) - s)$. If the system displays metastability, the motion of s will be bound in some local minimum of the free energy $F(s)$ (i.e. in a local maximum of $P(s)$) and it will escape from this minimum with a very low probability. In thermodynamic integration [2, 3] and weighted histogram analysis method [8, 9], the sampling of the phase space is enhanced by adding a constant external potential or a constraint. If, for example, the dynamics is restrained by an external potential of the form $\frac{k}{2}(s - S(x))^2$ [8], the system will explore preferentially the region around s even if it is not at the bottom of a free energy well. The free energy is then reconstructed using the combined information obtained from simulations restrained or constrained at several values of s . In this way it is possible to collect sufficient statistics for every value of s . In these methods the CV space is explored sequentially and systematically following a predefined scheme. In metadynamics, instead, the free energy is reconstructed recursively, starting from the bottom of the well by a history-dependent random walk that explores a larger and larger portion of configuration space. In the simplest molecular dynamics implementation of the algorithm the mathematical equivalent of the sand deposited by the walker, in the example of the introduction, is a small repulsive Gaussian potential added every τ_G MD steps. This manner of biasing the dynamics is usually referred to as ‘direct metadynamics’. The external (‘metadynamics’) potential acting on the system at time t is given by

$$V_G(S(x), t) = w \sum_{\substack{t' = \tau_G, 2\tau_G, \dots \\ t' < t}} \exp\left(-\frac{(S(x) - s(t'))^2}{2\delta s^2}\right), \quad (3)$$

where $s(t) = S(x(t))$ is the value taken by the CV at time t . Three parameters enter the definition of the V_G :

- (i) the Gaussian height w
- (ii) the Gaussian width δs
- (iii) the frequency τ_G at which the Gaussians are added.

These parameters influence the accuracy and efficiency of the free energy reconstruction as will be discussed in detail later. Qualitatively, they define the volume of the ‘sand’ the walker is depositing. If the Gaussians are large, the free energy surface will be explored at a fast pace, but the reconstructed profile will be affected by large errors. Instead, if the Gaussians are small or are placed infrequently the reconstruction will be accurate, but it will take a longer time.

As a first example, consider the system depicted in figure 1. Metadynamics is performed on the one-dimensional potential with three minima represented in the lower panel with a thick black line. The system evolves through an overdamped Langevin equation [108] with time step 1, diffusion coefficient

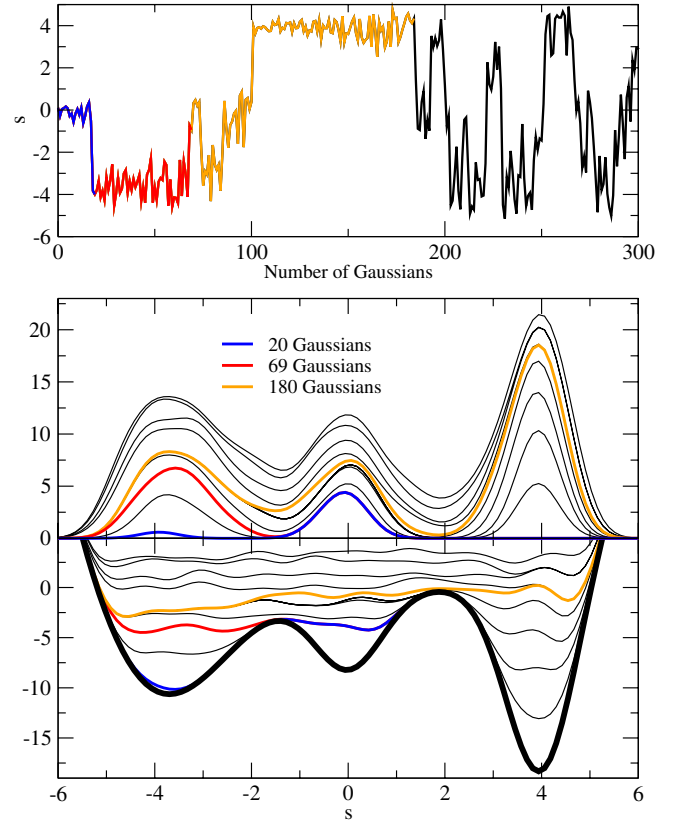


Figure 1. (Colour online.) Upper panel: trajectory of a one-dimensional system evolved by a Langevin equation on the 3-minima potential represented in the lower panel. The dynamics is biased with a metadynamics potential V_G as defined by equation (3). The parameters are $\delta s = 0.4$, $w = 0.3$ and $\tau_G = 300$. Middle panel: time evolution of the metadynamics bias potential V_G . Blue line: V_G as when the first minimum is filled and the system ‘escapes’ to the second minimum; red line: V_G as when also the second minimum is filled; orange line: V_G when the entire profile is filled and the dynamics becomes diffusive. Lower panel: time evolution of the sum of the metadynamics potential V_G and of the external potential, represented as a thick black line.

$D = 0.005$ and at a temperature of 1. The dynamics is started from the central minimum. Without any bias the system would escape from this minimum with very low probability, since the barrier separating it from the minimum on the left is approximately four times the thermal energy. The upper panel shows the trajectory followed by the system under the action of a metadynamics bias. A Gaussian of width $\delta s = 0.4$ and height $w = 0.3$ is added every 300 steps. After ~ 20 Gaussians (corresponding to 6000 steps) the central minimum is filled and the system escapes from the well through the lowest saddle point (blue lines in figure). The second well is filled after ~ 70 Gaussians (red lines). The second highest saddle point is reached after ~ 100 Gaussians, and the full free energy surface is filled with a total of ~ 180 Gaussians (orange lines). After that time, the motion of the system becomes diffusive and unbound in the region of CV space between ~ -5 and 5 . The metadynamics potential V_G (equation (3)) is represented at different times in figure 1, middle panel. The lower panel of figure 1 reports the sum of V_G and of the external potential

(thick black line). Clearly, as the simulation proceeds, V_G iteratively compensates the underlying potential.

This example provides a hint of the two different manners in which metadynamics can be used:

- It can be used to ‘escape free energy minima’ [37], i.e. to find the lowest free energy saddle point out of a local minimum. In this case the metadynamics should be stopped as soon as the walker exits from the minimum and starts exploring a new region of space. In figure 1, this happens after ~ 20 Gaussians are placed.
- It can be used to exhaustively explore a predefined region in the CV space and reconstruct the free energy surface. In this case the simulation should be stopped when the motion of the walker becomes diffusive in this region. In figure 1, this happens after ~ 180 Gaussians are placed.

The basic assumption of metadynamics is that $V_G(s, t)$ defined in equation (3) after a sufficiently long time provides an estimate of the underlying free energy:

$$\lim_{t \rightarrow \infty} V_G(s, t) \sim -F(s). \quad (4)$$

This equation states that an equilibrium quantity, namely the free energy, can be estimated by a non-equilibrium dynamics in which the underlying potential is changed every time a new Gaussian is added. This relation does not derive from any standard identity for the free energy, such as the umbrella sampling equality or the perturbation free energy formula. In [37], equation (4) was postulated heuristically, observing the effect of the history-dependent potential on the dynamics of the CVs on free energy surfaces of known functional form. For instance, in the example of figure 1, it is clear that the sum of F and V_G after ~ 180 Gaussians is approximately a constant, except for small ripples that would be in different positions in a statistically independent run. For an atomistic system in which the potential depends on the position of several atoms and the free energy is the result of a complex dimensional reduction, equation (4) can be qualitatively understood in the limit of slow ‘deposition’ (i.e. $w \rightarrow 0$). In this limit, $V_G(s, t)$ varies very slowly and the probability to observe s is always approximately proportional to $\exp[-\frac{1}{T}(F(s) + V_G(s, t))]$. If the function $F(s) + V_G(s, t)$ has some local minimum, s will preferentially be localized in the neighborhood of this minimum and increasing numbers of Gaussians will be added there until this minimum is completely filled. Let us consider instead the case in which $F(s) \sim -V_G(s, t)$ in a region $\Omega(s)$. The probability distribution will be approximately flat in this region, and the location of the new Gaussians will not be affected by the bias deriving from the difference $F(s) + V_G(s, t)$. Hence, if $w \rightarrow 0$, the only corrugations in the free energy that are not flattened by the dynamics will be of the order of the size of the newly added Gaussians. The validity of equation (4) for finite w will be discussed in detail in section 3.

If the CV is a d -dimensional vector, namely two or more CVs are used at the same time, the metadynamics potential is

given by

$$V_G(S(x), t) = w \sum_{\substack{t' = \tau_G, 2\tau_G, \dots \\ t' < t}} \exp \left(- \sum_{\alpha=1}^d \frac{(S_\alpha(x) - s_\alpha(t'))^2}{2\delta s_\alpha^2} \right) \quad (5)$$

and it is necessary to choose a width δs_α for each CV. The time required to escape from a local minimum in the free energy surface is determined by the number of Gaussians that are needed to fill the well. This number is proportional to $(1/\delta s)^d$, where d is the number of CVs used in the system. Hence, the efficiency of the method scales exponentially with the number of dimensions involved. If d is large, the only way to obtain a reasonable efficiency is to use Gaussians with a size comparable to that of the well. On the other hand, a sum of Gaussians can only reproduce features of the FES on a scale larger than $\sim \delta s$. Already from these simple considerations it is clear that the metadynamics works properly only if d is small, and that the quality of the reconstructed free energy is strongly influenced by the parameters w and δs . These parameters have to be carefully chosen to strike the best balance between accuracy and sampling efficiency. Large values for w and δs allow for a fast sampling of the CV space at the price of a low accuracy. However, if a large volume of Gaussians is placed in a short time, the reconstructed free energy can be totally wrong (see [79]).

1.1. Lagrangian metadynamics

If the metadynamics method is used for simulating chemical reactions by FPMD [75, 78, 106], the history-dependent potential has to force the system to cross barriers of several tenths of kcal mol⁻¹ in a short time, usually a few picoseconds. This implies that a lot of energy has to be injected in the degrees of freedom associated with the CVs. This might lead to a significant inhomogeneity in the temperature distribution of the system, and possibly to instabilities in the dynamics.

To address this problem, in the spirit of the extended Lagrangian approach [106, 107, 109], reference [75] introduces auxiliary variables \tilde{s} coupled to the system by harmonic restraining potentials of the form $\frac{1}{2}k(\tilde{s} - S(x))^2$. A fictitious kinetic energy $\frac{1}{2}M\dot{\tilde{s}}^2$ is also assigned to the auxiliary variables. The dynamics of these extra degrees of freedom can be explicitly controlled by suitable thermostats and the trajectory of \tilde{s} can be smoothened so as to control the stability of the algorithm.

The modified potential for the system is

$$\tilde{V}(x, \tilde{s}) = V(x) + \frac{1}{2}k(\tilde{s} - S(x))^2. \quad (6)$$

The free energy as a function of the \tilde{s} variables is given by

$$\begin{aligned} \tilde{F}(\tilde{s}) &= -T \ln \left(\int dx d\tilde{s}' \exp \left(-\frac{1}{T} \left[V(x) + \frac{1}{2}k(\tilde{s}' - S(x))^2 \right] \right) \delta(\tilde{s}' - \tilde{s}) \right) \\ &= -T \ln \left(\int dx \exp \left(-\frac{1}{T} \left[V(x) + \frac{1}{2}k(\tilde{s} - S(x))^2 \right] \right) \right). \end{aligned}$$

Since $\lim_{k \rightarrow \infty} \exp(-\frac{1}{T} \frac{1}{2} k (\tilde{s} - S(x))^2) \propto \delta(\tilde{s} - S(x))$, $\lim_{k \rightarrow \infty} \tilde{F}(\tilde{s}) = F(\tilde{s})$, modulus an additive constant. Hence, the free energy of a system of potential $V(x)$ can be obtained by performing metadynamics on the extended system of potential given by equation (6) if k is large enough. In practice, the value of this parameter is assigned by performing an ordinary MD run on the extended system. k must be chosen in such a way that the typical value of the difference $\tilde{s} - S(x)$ is smaller than the length on which the free energy varies of approximately T . This leads to the condition

$$\langle (\tilde{s} - S(x))^2 \rangle \sim \frac{T}{k} (\langle \tilde{s}^2 \rangle - \langle \tilde{s} \rangle^2), \quad (7)$$

where the averages are taken at a temperature T and in the absence of the metadynamics bias.

The value of the mass is a free parameter, that can be tuned in order to obtain a smooth evolution of the \tilde{s} . Within a Car–Parrinello scheme [106] an important requirement is the adiabatic separation from the electronic degrees of freedom. Since the extra term in the Hamiltonian introduces frequencies of the order of $\sqrt{k/M}$, and since k is fixed by equation (7), this defines a lower bound for M . On the other hand, if M is very large, the collective variables will relax to the equilibrium distribution slowly, reducing the efficiency of the method. For a detailed review about the proper way to tune these parameters see [79].

An advantage of taking a large M is also that, as will be shown in section 3, the metadynamics potential is an *exact* unbiased estimator of the free energy only if the unbiased dynamics along s can be described by a Markovian stochastic process. Markovianity in the s evolution can be imposed, at least in principle, by taking a sufficiently large M [110]. In fact, for large M all the coordinates of the system relax to their equilibrium distribution before s moves significantly, and this ensures a memory-less dynamics [108, 111]. However, in practical applications, if the CVs are properly chosen, the systematic error deriving from memory effects on the s dynamics is negligible (see sections 3 and 4.2). The situation is dramatically different if an important variable is not included in the CVs that determine the metadynamics bias. In this case, as will be discussed in section 2, the method does not converge in a finite simulation time, and the bias potential evolves unpredictably, in a way that is determined by the transitions in the hidden variables. This can be viewed as an extreme case of memory effect and one can imagine to cure the problem by choosing M so large that the system has time to perform several transitions in the hidden degrees of freedom before s moves significantly [109]. Unfortunately, this approach is impractical if the barriers in the hidden degrees of freedom are higher than a few $k_B T$.

1.2. Discrete metadynamics

Metadynamics was originally formulated [37] so as to enforce, at least in principle, the exact separation between the dynamics of the CVs and the dynamics of the normal microscopic variables. This variant of the algorithm, that will be referred to as *discrete metadynamics*, requires a stepwise evolution of

the CVs. The true dynamics of the system is used only for computing the derivative of the free energy $F(s)$ at fixed values of the CVs, $f_\alpha(s) = -\partial F / \partial s_\alpha$. This is done by performing short finite temperature molecular dynamics runs in which the normal Lagrangian of the system is modified with a term $\sum_{\alpha=1,d} \lambda_\alpha (S_\alpha(x) - s_\alpha)$, where λ_α are Lagrange multipliers. The derivative of the free energy with respect to the s_α s is calculated as the time-average of the Lagrange multipliers: $f_\alpha(s) = \langle \lambda_\alpha \rangle$ [2, 3]. Kinematic corrections due to the influence of inertial terms on the constraints can also be included [3]. The f s determined in this way are then used together with the forces coming from a history-dependent potential to update the collective variables as

$$s_\alpha(t+1) = s_\alpha(t) + \varepsilon \delta s_\alpha \frac{f_\alpha(s(t)) + f_\alpha^{\text{meta}}(s(t))}{|f(s(t)) + f^{\text{meta}}(s(t))|}, \quad (8)$$

where ε is a stepping parameter of the order of one. The history-dependent forces f_α^{meta} are derived by a potential that is identical to the one defined in equation (5), but depends directly on s and not on $S(x)$:

$$\begin{aligned} f_\alpha^{\text{meta}}(s) &= - \frac{\partial}{\partial s_\alpha} V_G(s, t) \\ &= - \frac{\partial}{\partial s_\alpha} w \sum_{t' \leq t} \exp \left(- \sum_{\alpha} \frac{(s_\alpha - s_\alpha(t'))^2}{2 \delta s_\alpha^2} \right). \end{aligned}$$

After the collective coordinates have been updated using equation (8), an ensemble with values $s(t+1)$ is prepared, and the new forces $f(t+1)$ are evaluated in another MD run. Similarly to what happens in the continuous algorithm, the forces due to the bias potential discourage the walker from revisiting the same spot and encourages an efficient sampling of the free energy surface. As the walker diffuses through the CV manifold, the Gaussian potentials accumulate and fill the free energy basins, allowing the system to explore all the minima. Eventually the sum of Gaussians, V_G , will approximately compensate the underlying free energy, i.e. $F(s) + V_G(s)$ will become approximately constant.

In [112] the quality of V_G as a statistical estimator of the free energy F is analyzed in detail, and a few tricks for improving the performance of the algorithm are described. Since in the discrete algorithm the force f is estimated by a molecular dynamics or a Monte Carlo simulation, its error will scale with the square root of the total computational time, and this will ultimately determine the cost of the free energy reconstruction. The recipe proposed in [112] is to stop sampling when the estimated uncertainty on the force is equal to the maximum force introduced by a single Gaussian, $w \exp(-1/2)/\delta s$. This choice ensures that for large values of t the typical force is of the order of $w/\delta s$. Improving the accuracy on the force calculation would lead to the same uncertainty because of the repeated superposition of the Gaussians. If this choice is made, the error at fixed δs becomes in practice linear in w and can be controlled. This allows merging, by weighted histogram techniques [8], free energy estimates obtained in two or more independent metadynamics [112].

As already mentioned, discrete metadynamics builds on the separation of time scales between the dynamics of the CVs

and the dynamics of the normal coordinates of the system. This feature makes it, at least in principle, more rigorous than the continuous version of the algorithm described in section 1. Nevertheless, as already underlined, systematic error deriving from memory effects on the s dynamics are not observed in practical applications, except when an important variable is left apart. If this is the case, also the discrete version of the algorithm does not converge, as the derivatives of the free energy are affected by large systematic errors. This has led to a wider diffusion of the continuous version of the algorithm, that is easier to implement in molecular dynamics codes (see section 4).

2. The choice of CVs

Similarly to other methods that reconstruct the free energy in a set of generalized coordinates, the reliability of metadynamics is strongly influenced by the choice of the CVs. Experience has demonstrated that choosing the right set of CVs is difficult but not impossible also in complex cases. If this is done, metadynamics provides very reasonable transition pathways and, more importantly, is able to discover new unpredicted stable and metastable states. Ideally the CVs should satisfy three properties:

- They should clearly distinguish between the initial state, the final state and the intermediates.
- They should describe all the slow events that are relevant to the process of interest.
- Their number should not be too large, otherwise it will take a very long time to fill the free energy surface.

Clearly, the second and third conditions might be mutually exclusive, and in many cases it can be very difficult to find a ‘good’ set of CVs. This happens, for instance, in the study of the folding of β -hairpin, a small polypeptide that forms a 20 residues-long β -turn [113]. In this case no combination of two or three variables was able to describe all the slow degrees of freedom. In similar cases the combination of metadynamics with parallel tempering or bias exchange or the use of path-like CVs provide good alternatives. These ‘special’ methods will be described in section 5.

Why is it so important to bias explicitly all the ‘slow’ variables and what happens if a relevant CV is neglected? In this respect, a simple metadynamics run on an idealized model can be enlightening. Consider the Z-shaped two-dimensional free energy depicted in figure 2. If a metadynamics simulation is performed biasing only CV1 and neglecting CV2 the simulation, that is started in basin B, is not able to perform in due time a transition toward A, and metadynamics goes on overfilling this minimum. A transition is finally observed only when the height of the accumulated Gaussians will largely exceed the true barrier height. This behavior will continue indefinitely without ever reaching a situation in which the free energy grows evenly as in the example of figure 1.

A similar behavior is often observed in real cases and is a strong indication that an important CV is missing. For instance, it was observed in the metadynamics simulation of the translocation of tetramethylammonium (TMA) in the

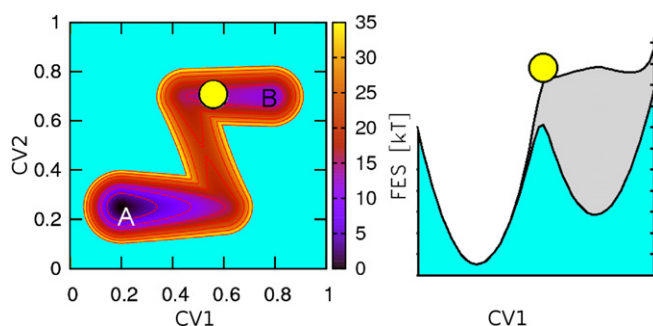


Figure 2. The effect of neglecting a relevant degree of freedom. Left side: 2D Z-shaped potential. Right side: the trajectory of a metadynamics simulation generated using only s_1 as CV. Transitions from A to B are properly described by CV1, causing strong hysteresis in the reconstructed free energy.

acetylcholinesterase (AChE) gorge [63]. The study was aimed at describing the mechanism of penetration of TMA in the gorge. A first metadynamics run was attempted using only one CV: the distance of the TMA center of mass from the active site of AChE. This resulted in a typical hysteretic behavior of the reconstructed FES. By an in-depth analysis of the trajectories it was then found that some aromatic residues block the gorge of AChE and act as a gate for the TMA translocation. If a second variable describing the opening of the gate is added, hysteresis is not observed anymore [63].

From the above discussion it should be clear that there is no *a priori* recipe for finding the correct set of CVs, and in many cases it is necessary to proceed by trial and error, attempting several metadynamics simulations with different combinations of variables. Of course, one can check *a posteriori* if the description provided by the chosen set is accurate, for instance by using transition path sampling techniques [20, 21] or by performing a committor test (see [20, 21]). Even without performing a specific check, an hysteretic behavior in the free energy reconstruction always signals the lack of a relevant CV. If, instead, the free energy grows ‘smoothly’ it is likely that the set of variables is complete [37, 114] (see also the example in section 4.2).

2.1. Examples of collective variables

In this section some of the variables that have been used in applications are briefly described.

Geometry-related variables. The simplest type of CVs are geometry related, such as distances, angles and dihedrals formed by atoms or groups of atoms. These variables are frequently used in the study of chemical reactions and biophysical systems. For example, to study protein-ligand recognition, metadynamics can be performed with the ‘standard’ variables used by many docking programs, namely the distance between the ligand and the cavity and one or more angles defining the orientation of the ligand [62].

Coordination numbers. The coordination number is probably the most general and useful collective variable. It is defined as

$$S(r) = \sum_{i,j} f(r_{ij}) \quad (9)$$

with

$$f(r_{ij}) = \frac{1 - (r_{ij}/r_0)^n}{1 - (r_{ij}/r_0)^m}, \quad (10)$$

where the sums on i and j run on two sets of atoms. The function $f(r)$ is approximately 1 for $r < r_0$ and goes to zero for large r . The parameters n and m can be used for tuning the smoothness of the function and its asymptotic behavior as, for large r , $f(r) \sim 1/r^{m-n}$. The coordination number can be used, for instance, to detect the presence of a bond between two atoms or for counting the bonds between two different atomic species. It is very commonly exploited for exploring and discriminating different pathways in chemical reactions [75–78, 115–117]. In biophysical systems, it can be used to count the hydrogen bonds or the hydrophobic contacts [65, 113, 118].

Potential energy. The potential energy of the system can be used as a CV and it is particularly useful for studying phase transitions [49, 112]. As shown previously [7], computing the free energy as a function of V allows estimating the density of states $\Omega(V)$ and extracting the temperature dependence of the observables, the specific heat, etc from a run performed at a single temperature. In fact

$$\begin{aligned} F_T(V) &= -T \log \int d\mathbf{r} e^{-\frac{1}{T} V(\mathbf{r})} \delta(V - V(\mathbf{r})) \\ &= V - T \log \int d\mathbf{r} \delta(V - V(\mathbf{r})) \\ &= V - T \log \Omega(V), \end{aligned}$$

where the last equation defines the density of states $\Omega(V)$. The free energy at a temperature T' different from T is given by

$$F_{T'}(V) = V - T' \log \Omega(V) = V - \frac{T'}{T} (V - F_T(V))$$

and the average value of V at T' is given by

$$\langle V \rangle = \frac{\int dV V \exp(-(1/T') F_{T'}(V))}{\int dV \exp(-(1/T') F_{T'}(V))}.$$

These relations are at the basis of the Ferrenberg and Swendsen approach [7] for combining energy histograms estimated at different temperatures, and, as shown in [112], can be used for combining the results of metadynamics runs performed at different T .

As also discussed in [112] the Wang and Landau sampling [35], a very popular approach aimed at reconstructing the density of states of complex systems, can be mimicked by performing metadynamics at a very high temperature and reducing iteratively the height of the Gaussians every time the system explores all the energy range. In fact, at large T the free energy is largely dominated by the density of states contribution, namely $T \log \Omega(V)$.

Box shape. In [54] metadynamics was combined with the Parrinello and Rahman approach [107] for studying crystal structure transformations. The three super-cell edges $\mathbf{a}, \mathbf{b}, \mathbf{c}$, arranged as a 3×3 matrix $\mathbf{h} = (\mathbf{a}, \mathbf{b}, \mathbf{c})$ [54, 55], are used as CVs. For relatively small systems, where the creation of defects is too expensive, the box matrix \mathbf{h} is likely to be simply

related to the unit cell \mathbf{u} via the relation $\mathbf{h} = \mathbf{u}\mathbf{m}$, where \mathbf{m} is an integer matrix. The matrix \mathbf{h} can therefore distinguish between different unit cells and crystal structures. As the matrix \mathbf{h} can always be chosen to be upper triangular, the independent CVs are 6. At pressure P and temperature T the appropriate thermodynamic potential for studying crystal transformations is the Gibbs free energy $\mathcal{G}(\mathbf{h}) = \mathcal{F}(\mathbf{h}) + PV$ where $\mathcal{F}(\mathbf{h})$ is the Helmholtz free energy of the system at fixed box and $V = \det(\mathbf{h})$ is the volume. During metadynamics the box shape changes, the wells in $\mathcal{G}(\mathbf{h})$ get filled and the system evolves from the initial structure toward novel structures, which correspond to new minima of the Gibbs free energy at the thermodynamic conditions of interest [56–60].

For studying crystal structure transformations it is convenient to use the algorithm in its discrete version. This requires calculating the derivatives of the free energy with respect to the collective variables

$$-\frac{\partial \mathcal{G}}{\partial h_{ij}} = V[\mathbf{h}^{-1}(\mathbf{p} - P)]_{ji}, \quad (11)$$

where \mathbf{p} is the internal pressure tensor, which can be evaluated in MD or Monte Carlo simulations at constant \mathbf{h} from the averaged microscopic virial tensor [119]. Further technical details on how to perform the simulation, choose the metadynamics parameters and analyze the results can be found in a dedicated review [55]. For the scope of the present review it is sufficient to note that the number of independent CVs used in crystal structure prediction, 6, is larger than that used in most other applications and precludes an accurate reconstruction of the free energy surface, as the complete exploration of the six-dimensional space would require a prohibitively long simulation time. This is a case where metadynamics is only used as a tool for exploring efficiently the configuration space and finding new structures. Once the structures are found, their free energy difference can be efficiently computed with other techniques [120].

Path variables. The ansatz of approximating the intrinsic reaction coordinate with a path connecting two stable basins in energy or free energy space has been used in several powerful approaches aimed at elucidating the reaction mechanism [16–19]. Path variables can also be used in metadynamics and effectively help alleviating the ‘curse of dimensionality’, namely the rapid loss of efficiency of the approach that takes place when more and more variables are biased. In fact, a path can be defined in a very high-dimensional space, in principle even in the full phase space [17–19], but the position of a system along a path is an intrinsically one-dimensional quantity.

Ensing *et al* [79] propose a multi-stage approach in which first a ‘coarse’ free energy surface is built with metadynamics in several dimensions (up to six). In the second step a path connecting different stable free energy basins is built on the coarse multidimensional FES. The free energy along the optimized path is then refined with umbrella sampling or one-dimensional metadynamics. This approach is very powerful as it exploits one of the most important features of metadynamics,

namely its capability of providing a quick and efficient coarse-grained map of the free energy, exploring rapidly the relevant metastable states.

In [121] two path-like variables (PCV) are introduced that are able to describe the position of a point in configurational space relative to a pre-assigned path:

$$s(x) = \lim_{\lambda \rightarrow \infty} \frac{\int_0^1 t e^{-\lambda \|S(x) - S(t)\|^2} dt}{\int_0^1 e^{-\lambda \|S(x) - S(t)\|^2} dt}, \quad (12)$$

$$z(x) = \lim_{\lambda \rightarrow \infty} -\frac{1}{\lambda} \ln \int_0^1 e^{-\lambda \|S(x) - S(t)\|^2} dt \quad (13)$$

where t parametrizes a path $S(t)$ in a high-dimensional CV space and $\|\dots\|$ indicates the distance in this space. For any microscopic configuration x , $s(x)$ and $z(x)$ measure, respectively, the progression along the path and the distance from the path. In practical applications, a first guess for the path is discretized with a discrete number of frames $S(l)$, $l = 1, P$ with $S(1) = S_A$ and $S(P) = S_B$, and equations (12) and (13) are approximated by finite sums over l . The distance $\|\dots\|$ in equations (12) and (13) can be defined in different

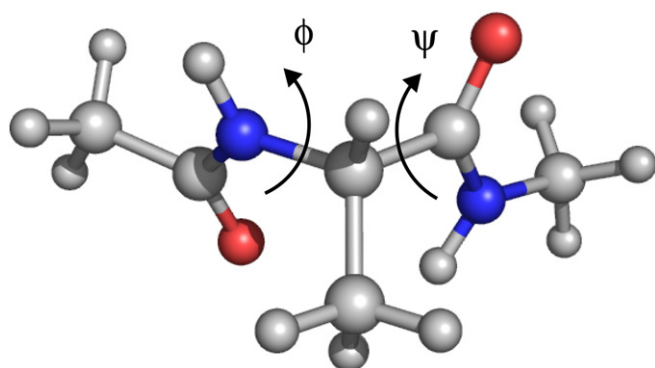
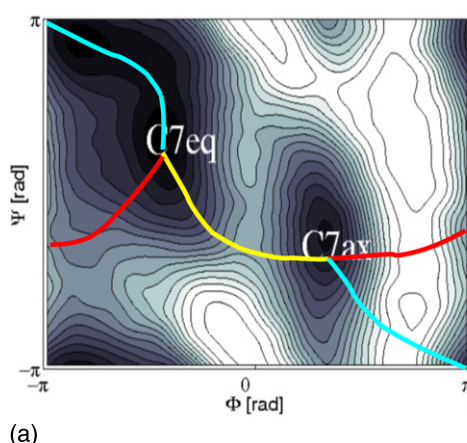


Figure 3. Schematic representation of the alanine dipeptide molecule. The backbone dihedral angles ϕ and ψ are shown. The C_{7eq} configuration corresponds to structures having $-150^\circ < \phi < -30^\circ$ and $0^\circ < \psi < 180^\circ$. The C_{7ax} configuration corresponds to structures having $30^\circ < \phi < 130^\circ$ and $-180^\circ < \psi < 0^\circ$.



(a)

spaces. A possible simple metric is the RMSD between the two structures after they are optimally aligned using the Kearsley [122] algorithm. Different choices for the metric are also possible, as e.g. the contact map (CMAP) matrix $C(x)$ defined as $f(r_{ij})$ in equation (10) where r_{ij} is the distance between the i th and j th C_α atoms of the protein backbone, n and m set to 6 and 10, respectively and the cutoff distance r_0 is taken to be $r_0 = 8.5 \text{ \AA}$ [123]. The square distance $\|\dots\|^2$ between a generic state x and a point along the path described by the CMAP $S_C(l)$ is measured in this case as

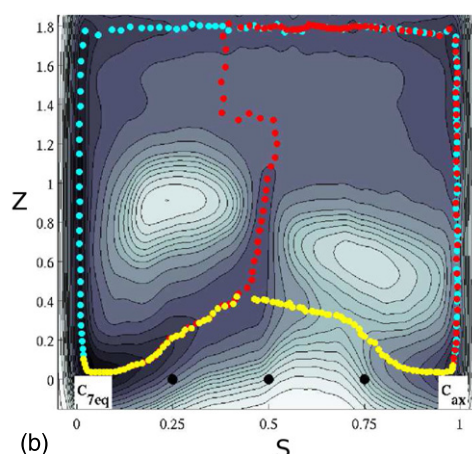
$$\|S_C(x) - S_C(l)\|^2 = \sum_{j>i} (C_{i,j}(x) - C_{i,j}(l))^2, \quad (14)$$

where nearest neighbors are excluded from the sum.

As shown elsewhere [19], the initial guess on the path can be refined at will, eventually finding a rigorous parametrization of the committor. Still, if a totally independent reaction mechanism exists, it will be explored with vanishingly small probability as a transition between the two mechanisms is a ‘rare event’ in paths space. Using $z(x)$ together with metadynamics allows exploring reaction pathways that are further and further from the initial guess, eventually finding a reaction pathway that is completely different [121]. Indeed, independent reaction mechanisms are similar to different free energy minima in path space, and metadynamics can help in escaping local minima. An example taken from [121] of path variables applied to study the reactive coordinates connecting the two minima of alanine dipeptide in vacuum (represented in figure 3) is reported in the following.

In figure 4, panel A the bi-dimensional FES as a function of ϕ and ψ is reported. The yellow line corresponds to the ‘reference path’. In panel B the corresponding FES as a function of S and Z is reported. As the distance from the yellow path increases, two other independent low-free energy paths are found (in red and cyan).

Normal modes. Normal modes and essential coordinates have a long history in enhanced sampling methods, especially for studying biological systems [124]. By analyzing relatively



(b)

Figure 4. Correspondence between minimum free energy paths in Ramachandran plot representation (panel a) and s, z space (panel b). The isoline separation is $1.0 \text{ kcal mol}^{-1}$. The yellow path is the reference path. The cyan and red paths represent alternative low-free energy paths found by PCV.

short MD trajectories or, even more simply, analyzing the topology of the system, one can gather useful information on what are the ‘slow’ or ‘soft’ modes of the system. Then the sampling in the direction of these modes can be enhanced applying various methods [6, 124]. For example, the system can be evolved in at finite temperature in a local minimum of the potential energy surface. The covariance matrix in a suitably chosen space is then evaluated, and its normal modes are used for constructing an appropriate bias potential, that favours the transitions in the direction of the slow modes. The normal modes can be used as CVs in metadynamics. In [125] trajectories of the alanine dipeptide have been analyzed by essential dynamics analysis [124] to trace major collective motions. In this method, first the covariance matrix of the atomic positional deviations is calculated. Diagonalization of the matrix leads to eigenvalues and eigenvectors. The first and second eigenvectors have been used to perform metadynamics leading to a reconstructed FES in good agreement with other computational as well as experimental studies of the model system.

Protein-specific variables. In addition to the variables described so far several other CVs have been developed to explore the free energy surface of peptide (un)folding and aggregation with metadynamics. These are the number of C γ contacts (NC γ), number of C α contacts (NC α or CMAP), number of backbone H-bonds (Nhb), ‘helicity of the backbone’ (Φ_α) and ‘dihedral correlation’ (Φ_{corr}). NC γ , NC α or CMAP and Nhb are defined as coordination numbers (equation (10)), where r_{ij} is the distance between atoms i and j (either the C γ contacts, the C α contacts, the HN, and O backbone atoms or the functional groups of the charged amino acids). $l = 8$, $m = 10$ and $r_0 = 5.0, 6.5$ and 2.0 Å for NC γ , NC α and Nhb, respectively. A slightly different definition of NC α , where the exponents where $l = 6$, $m = 10$ and r_0 was set to 8.5 Å was used to study peptide aggregation in [69].

The helicity of the backbone is defined as [118]

$$\Phi_\alpha = \sum_{i=1}^n \frac{1}{2} [1 + \cos(\phi_i - \phi_0)], \quad (15)$$

where ϕ_i is the backbone dihedral angle of residue i and ϕ_0 is -45° . This variable counts the number of residues that have $\phi_i \sim \phi_0$ and is useful for monitoring the presence of an α -helix.

The dihedral correlation is defined as

$$\Phi_{\text{corr}} = \sum_{i=2}^N \sqrt{[1 + \cos^2(\phi_i - \phi_{i-1})]}, \quad (16)$$

where ϕ_i is the backbone dihedral angle of residue i . In α -helix and β -sheets, successive ϕ dihedrals have approximately the same value. Thus, a large value of Φ_{corr} is related to the presence of secondary structure elements. These CVs have been used in [71, 118] to study the folding of small proteins and in [70] to study the helix-coil transition in DNA.

Collective variables for studying phase transformations. Studying phase transformations by computer simulation is a very important goal and has sometimes been defined the ‘holy grail’ of computational physics. Clearly, in realistic systems it is not possible to reconstruct systematically the energy density of states, as can be done in model systems. Moreover, phase transformations in realistic systems are usually first order transitions, which means that a phase can exist for a finite time also in conditions of metastability. For instance, liquid water can exist also below zero degrees Celsius. Metastability has very severe consequences: phenomena that are ordinary on a human time scale correspond to extremely rare events in computer simulation, so rare that they will simply not take place unless a proper bias is applied. Several techniques have been developed for computing the free energy difference between two ordered (crystalline) phases [120] or a crystalline and a disordered (liquid) phase [126–128] with the aim of predicting the transition temperature or pressure. These techniques are nowadays successfully applied to systems of increasing complexity [129]. Usually these approaches do not explicitly simulate the transformation between the two phases, but rather compute the two free energies separately or rely on phase coexistence [128]. The techniques used to simulate how one phase is converted to the other by nucleation are completely different [130], and usually require biasing one or more order parameters that distinguish the two phases and can drive the appropriate microscopic rearrangements. This approach has also been followed in metadynamics. Usually, appropriate order parameters are biased *together with* the classical variable for phase transitions, the potential energy. For instance, the reader is referred to [46], where the number of five and six membered rings is used to describe the hydrogen bonding pattern of water to study the nucleation of the liquid phase in hexagonal ice. In [49] the nucleation of a Lennard-Jones fluid was studied as a function of the degree of supercooling in a metadynamics simulation where the CVs were the potential energy and the Steinhardt order parameter Q_6 [131]. In [52] ice freezing was studied by using a continuous version of the general Steinhardt parameter and a tetrahedral order parameter ζ . Using these CVs the authors showed homogeneous nucleation and growth of ice at 180 K in the isothermal–isobaric ensemble without the presence of external fields or surfaces.

3. Estimating the error

In order to best allocate the available computational resources to study with metadynamics a given system, it is useful to estimate *a priori* the performance of the method, and choose the parameters to obtain the best possible accuracy in a given simulation time. The accuracy and efficiency of the free energy reconstruction is determined by the Gaussian width δs , the Gaussian height w and the Gaussians deposition time τ_G . Since metadynamics is usually applied for reconstructing free energies in several dimensions, one can in principle fix independently the Gaussian width in each dimension. However, the relative units of different collective variables can always be chosen in such a way that the shape of free energy

well is approximately spherical. For instance, if the region of interest is a single well, this can be done by computing the standard deviations of the CVs in a preliminary finite temperature run. Therefore in the following we will consider only the case of a free energy well with spherical symmetry. The parameter w and τ_G determine the height and the rate at which the Gaussians are placed. In [132] it has been shown that the error on the reconstructed profile is approximately determined by the ratio w/τ_G and not by w and τ_G separately. Indeed, adding a Gaussian of height, say, $0.2 \text{ kcal mol}^{-1}$ every picosecond is approximately equivalent to adding a Gaussian of height $0.1 \text{ kcal mol}^{-1}$ every 0.5 ps , as long as τ_G remains much shorter than the time required to fill the free energy basin. In real-life applications it is not practical to take a very small τ_G as it implies depositing many Gaussians in a short time, meaning that the evaluation of the metadynamics forces a computational bottleneck. This problem, however, is easily solvable. If instead of storing the Gaussian centers, the program stores directly the history-dependent potential on a grid, as proposed in [66], or the small Gaussians are from time to time re-fitted with a few, larger Gaussians, τ_G can be further reduced (see also section 4).

In order to understand how δs and w/τ_G influence the accuracy and construct an explicit expression for the error, consider first the idealized case in which the CVs evolve following an overdamped Langevin dynamics:

$$ds = -\frac{1}{T} D \frac{dF(s)}{ds} dt + \sqrt{2D} dW(t), \quad (17)$$

where $dW(t)$ is a Wiener process and D is the diffusion coefficient. The motion of the walker described by equation (18) is assumed to satisfy reflecting boundary conditions at the boundary of a region Ω . The evolution of this system under the action of metadynamics is modeled adding a history-dependent term to the free energy:

$$ds = -\frac{1}{T} D \frac{d}{ds} \left(F(s) + \int_0^t dt' g(s, s(t')) \right) dt + \sqrt{2D} dW(t), \quad (18)$$

where $g(s, s')$ is a kernel that specifies how fast the metadynamics potential changes. In normal implementation g is a Gaussian of width δs and height w/τ_G (see equation (3)):

$$g(s, s') = \frac{w}{\tau_G} \exp \left(-\frac{(s - s')^2}{2\delta s^2} \right).$$

It should be noted that in real systems the evolution of the CVs is described by a much more complex stochastic differential equation [133, 134] with memory and inertial terms. Still, as will be discussed in the following, the quantitative behavior of metadynamics is reproduced rather precisely by this simple model. This is because, if the CV set is properly chosen, all the relaxation times are smaller than the time required to fill the free energy wells.

Equation (18) describes a non-Markovian process in CV space. In fact, the forces acting on the CVs depend explicitly on their history. Due to this non-Markovian nature it is not clear whether, and in which sense, the system can reach a stationary state under the action of this dynamics. In [43]

a formalism was introduced which allows one to map this history-dependent evolution into a Markovian process in the original variable and in an auxiliary field that keeps track of the visited configurations. Defining

$$\varphi(s, t) = \int_0^t dt' \delta(s - s(t')), \quad (19)$$

equation (18) can in fact be written as

$$d\varphi = \delta(s - s(t)) dt \quad (20)$$

$$ds = -\frac{1}{T} D \left(\frac{d}{ds} \left(F(s) + \int ds' \varphi(s', t) g(s', s(t)) \right) \right) dt + \sqrt{2D} dW(t). \quad (21)$$

These equations are fully Markovian, i.e. the state of the system at time $t + dt$, $(s(t + dt), \varphi(s, t + dt))$, depends only on the state of the system at time t , $(s(t), \varphi(s, t))$. Using this property equation (18) can be rigorously analyzed to obtain, for instance, its long time behavior.

The history-dependent potential at time t is related to $\varphi(s', t)$ by

$$V_G(s, t) = \int ds' \varphi(s', t) g(s, s'). \quad (22)$$

In order to characterize the average properties of a system described by equation (21) it is convenient to consider the probability $P(s, [\varphi], t)$, to observe s and the field realization φ . $P(s, [\varphi], t)$ satisfies a Fokker-Planck equation that can be directly derived from equation (21) using standard techniques [111]. [43] shows that, for large t , $P(s, [\varphi], t)$ converges to a distribution $P_\infty([\varphi])$ that does not depend on s . This distribution is Gaussian in functional space and is given by

$$P_\infty([\varphi]) \propto \exp \left(\frac{D}{2T} \int ds ds' (\varphi(s') - \varphi_0(s')) \partial_s^2 g(s, s') (\varphi(s) - \varphi_0(s)) \right), \quad (23)$$

where $\varphi_0(s')$ is defined in such a way that its convolution with the kernel g gives minus the free energy of the system $F(s)$:

$$\varphi_0(s') : \int ds' \varphi_0(s') g(s, s') = -F(s). \quad (24)$$

Using equation (23) it is straightforward to prove that the average value of $V_G(s, t)$ over several independent metadynamics runs is exactly equal to $-F(s)$. In fact, denoting by $\langle \cdot \rangle_M$ the average over several metadynamics realizations, equation (22) gives

$$\begin{aligned} \langle V_G(s) \rangle_M &= \int ds' g(s, s') \langle \varphi(s') \rangle_M \\ &= \int ds' g(s, s') \int d\varphi P_\infty(\varphi) \varphi \\ &= \int ds' g(s, s') \varphi_0(s') = -F(s). \end{aligned} \quad (25)$$

This property is also illustrated in figure 5 in which the results obtained integrating numerically equation (18) for four

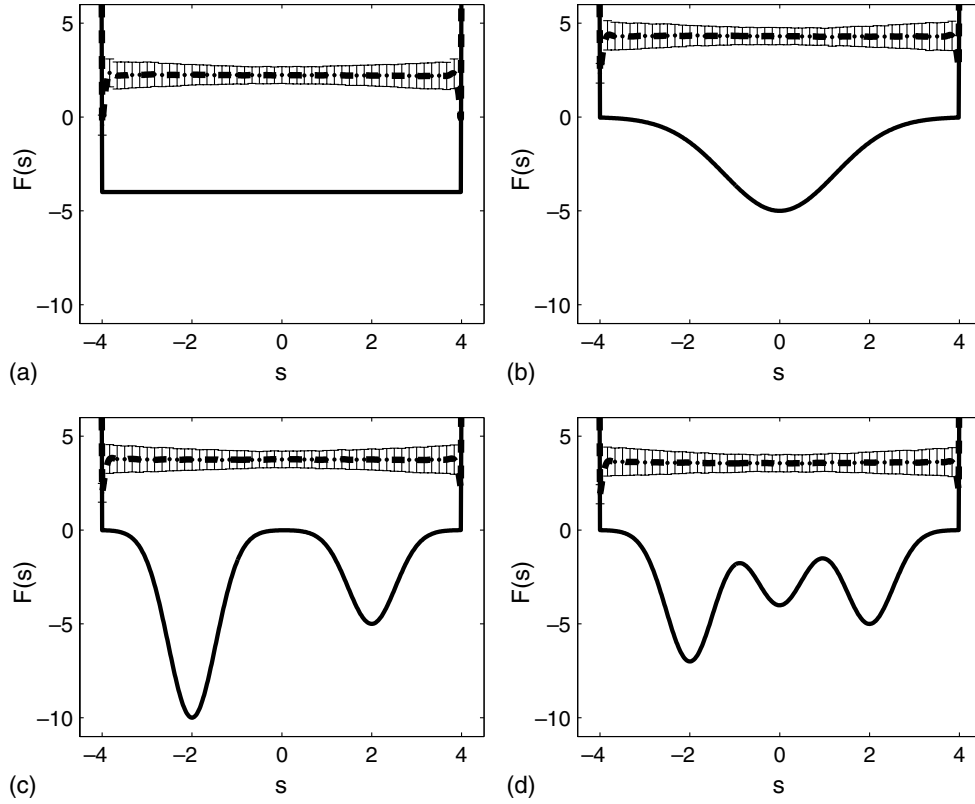


Figure 5. Metadynamics results for four different free energy profiles. (a) $F(s) = -4$; (b) $F(s) = -5 \exp(-(\frac{s}{1.75})^2)$; (c) $F(s) = -5 \exp(-(\frac{s-2}{0.75})^2) - 10 \exp(-(\frac{s+2}{0.75})^2)$; (d) $F(s) = -5 \exp(-(\frac{s}{0.75})^2) - 4 \exp(-(\frac{s}{0.75})^2) - 7 \exp(-(\frac{s+2}{0.75})^2)$. The average $\langle F(s) + V_G(s, t) \rangle$ computed over 1000 independent trajectories is represented as a dashed line, with the error bar given by equation (27). Errors are measured in units of the free energy. The metadynamics parameters are $\delta s = 0.1$ and $w/\tau_G = 4 \times 10^{-4}$. We also have $D = 0.0005$ and $T = 1$ and the CVs satisfy reflecting boundary conditions in a region of length 8. After [132].

different profiles $F(s)$ are shown. The average value of $V_G(s, t) + F(s)$ is represented as a continuous line in all the four profiles, and is constant in all the explored region, as predicted by equation (25).

The metadynamics error in s is given by the expected deviation of $V_G(s, t)$ from $-F(s)$:

$$\varepsilon^2(s) = \langle (V_G(s) + F(s))^2 \rangle_M \quad (26)$$

$$= \langle (V_G(s) - \langle V_G(s) \rangle_M)^2 \rangle_M. \quad (27)$$

This error was estimated numerically for the four profiles in figure 5, repeating several statistically independent metadynamics runs. Remarkably, the error does not depend on $F(s)$, as shown by the error bars, that are indistinguishable in the four cases. Moreover, the error is only marginally larger close to the reflecting walls. Thus, $\varepsilon^2(s)$ can be conveniently characterized by its average over the entire domain

$$\bar{\varepsilon}^2 = \frac{1}{\text{vol}(\Omega)} \int_{\Omega} ds \varepsilon^2(s). \quad (28)$$

Using the explicit expression for the probability to observe a given φ , equation (23), allows computing explicitly $\varepsilon^2(s)$ or $\bar{\varepsilon}^2$ and these turn out to be independent of $F(s)$. The specific value depends only on the metadynamics parameters, on the shape of the domain on which the system is confined, on the diffusion coefficient and on temperature. For example, in a

cubic domain of side S in d dimensions [43] the error is

$$\bar{\varepsilon}^2 = \frac{S^2 w T}{D \tau_G} \left(\frac{\delta s}{S} \right)^d (2\pi)^{d/2} \sum_k \frac{1}{\pi^2 k^2} \times \exp \left(-\frac{k^2 \pi^2}{2} \left(\frac{\delta s}{S} \right)^2 \right), \quad (29)$$

where the sum is performed over all the d dimensional vectors of integers with non-zero norm.

In [132] an alternative expression for the error was deduced by performing extensive numerical simulations of the stochastic differential equation (18). The metadynamics parameters, w/τ_G , $\delta\sigma$, and the system-dependent parameters, T , D and S were systematically varied, and for each choice of the parameters the error was computed by repeating several metadynamics reconstructions. A proper data fit ($\sim 20\%$ of accuracy) was provided by the formula

$$\bar{\varepsilon}_{\text{approx}}^2 = C_d \frac{S^2 w T}{D \tau_G} \frac{\delta s}{S}, \quad (30)$$

where C_d is a constant that depends only on the dimensionality. The two expressions for the error share the same functional dependence on w/τ_G , T , D and S . The dependence on $\delta s/S$ in equation (29) is instead much more complicated. The ratio between the two expressions is approximately a constant as a function of δs only for $d = 1$ and $d = 2$, while significant deviations are observed in higher dimensions.

The dependence of the error on the simulation parameters becomes more transparent if $\bar{\varepsilon}$ is expressed as an explicit function of the total simulation time. Consider in fact a free energy profile $F(s)$ that has to be filled with Gaussians up to a given level F_{\max} , for example the free energy of the lowest saddle point in $F(s)$. The total computational time needed to fill this profile can be estimated as the ratio between the volume that has to be filled and the volume of one Gaussian times τ_G :

$$t_{\text{sim}} \approx \tau_G \frac{F_{\max}}{w} \left(\frac{S}{\delta s} \right)^d. \quad (31)$$

Substituting in equation (29) yields

$$\bar{\varepsilon}^2 \approx \frac{\tau_S}{t_{\text{sim}}} F_{\max} T f_d \left(\frac{\delta s}{S} \right), \quad (32)$$

where $\tau_S \doteq S^2/D$ is the average time required for the CVs to diffuse on a distance S and

$$f_d \left(\frac{\delta s}{S} \right) = (2\pi)^{d/2} \sum_k \frac{1}{k^2 \pi^2} \exp \left(-\frac{k^2 \pi^2}{2} \left(\frac{\delta s}{S} \right)^2 \right)$$

is a function of $\delta s/S$ and of the dimensionality alone. Equation (32) states that the error of a metadynamics reconstruction is inversely proportional to the square root of the total simulation time, measured in units of the diffusion time. The error will be large for slowly diffusing systems, in which the walker takes a long time to explore the CVs space.

The error at fixed simulation time is determined by the pre-factor $f_d(\delta s/S)$, plotted in figure 6 for $d = 2$ and $d = 3$ as a function of $\delta s/S$. For a comparison, the function $f_d^{\text{real}}(\delta s/S) = \bar{\varepsilon}_{\text{fixed}}^2$ is also plotted, where $\bar{\varepsilon}_{\text{fixed}}^2$ is the error measured by equation (28) on several independent metadynamics reconstructions performed in a parabolic well. All runs were performed by changing $\delta s/S$ and choosing w so as to obtain the same filling velocity, i.e. $w \propto (S/\delta s)^d$. Moreover, the parameters of the system are chosen in such a way that $(\tau_S/t_{\text{sim}})F_{\max}T = (S^2 w T / D \tau_G)(\delta s/S)^d = 1$. The function $f_d^{\text{real}}(\delta s/S)$ measured under these conditions is an estimate of the error at fixed filling time in a free energy well. For a fixed total simulation time, the error always decreases as a function of $\delta s/S$ and the higher the dimensionality the more this dependence is significant. Equation (29) is accurate only for small $\delta s/S$ and deviations are observed when δs becomes comparable with the dimension of the well. Indeed, the ‘filling’ behavior of metadynamics depicted in figure 1 can be observed only if δs is smaller than the typical length scale of variation of the free energy. Equation (29) describes the error only when the walker is freely diffusing in a confined region in which reflecting boundary conditions are imposed. For a large δs the error will be large while the walker is filling each well. In conclusion, δs should be chosen as large as possible, but an upper bound is imposed by the features of the free energy profile that has to be filled.

In [132] the performance of metadynamics is estimated for a system consisting of a tetracationic cyclophane (cyclobis(paraquat-p-phenylene)₈⁴⁺) ring and a 1,5-dihydroxy-naphthalene. This system was used to study the unthreading

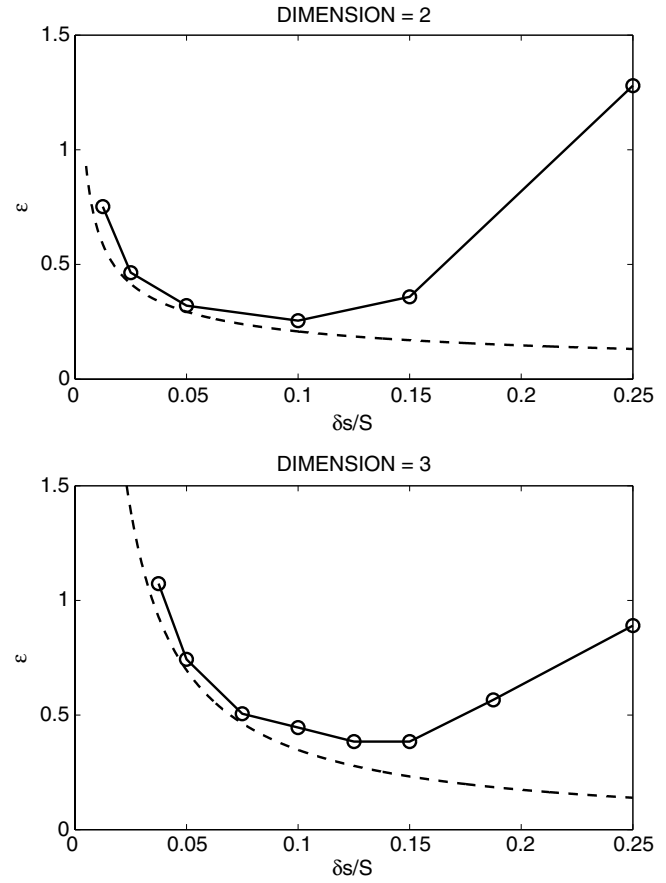


Figure 6. Effect of $\delta s/S$ on the error ε computed with equation (28) for a system in $d = 2$ (top panel) and $d = 3$ (bottom panel), with $S = 8$, $T = 1$ and $D = 0.0005$. The dashed line corresponds to the theoretical error given by equation (29). w/τ_G is chosen so that $(\tau_S/t_{\text{sim}})F_{\max}T = (S^2 w T / D \tau_G)(\delta s/S)^d = 1$.

behavior of [2]-catenane which has recently attracted much interest as fully reversible bistable molecular switch and was used as a basic component in the realization of a molecular elevator [135]. For this system the free energy is computed as a function of two collective variables, the distance between the centroids of the cyclophane and the naphthalene and the coordination number of the naphthalene with the atoms of the acetonitrile. $F(s)$ is first estimated by a two-dimensional umbrella sampling using the WHAM scheme [8] (see [132] for technical details). This free energy has two minima, one corresponding to the threaded state ($s_1 \sim 0$ and $s_2 \sim 6$) and one corresponding to the unthreaded state. The barrier toward the dissociated state is ~ 12 kcal mol⁻¹. On the same system, several metadynamics runs were performed with different choices of w , δs and τ_G and the error was estimated from the deviation of $-V_G(s)$ from the ‘exact’ WHAM free energy. In total, more than fifty different combinations of parameters were tested.

In figure 7 the standard deviation from the WHAM result is compared with the one predicted by equation (30). A correction factor for a finite τ_G derived in [132] is also included. The difference between the predicted and the computed value is always rather small, showing that equation (30) accurately predicts the true metadynamics error.

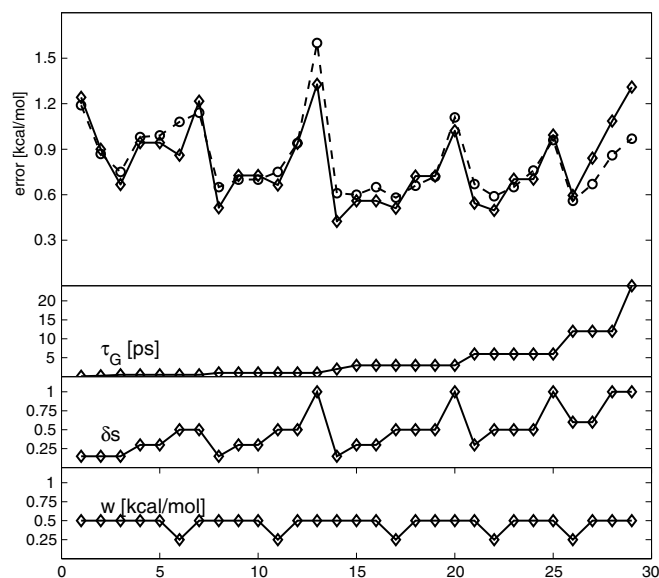


Figure 7. Metadynamics error on the free energy for the pseudo-rotaxane naphthalene complex as a function of the run number. Upper panel: the error computed by equation (28) (dashed line with circles) and the error predicted by equation (30), including the correction factor for a finite τ_G derived in [132] (continuous line with diamonds). Lower panels: the metadynamics parameters w , δs and τ_G as a function of the run number. After [132].

This result requires some comment. Equations (29) and (30) were derived under the assumption that metadynamics can be modeled with a stochastic differential equation of the form equation (18). In this equation the noise is independent of the position in the CV space and there are no inertial effects. These approximations might sound severe, but the example described in this section shows that in real systems the *quantitative* behavior of metadynamics is perfectly reproduced by this equation. This is because the error in metadynamics is due to the relaxation time required for the CV to diffuse through the region in which the free energy is reconstructed, S^2/D . Other characteristic times, that are certainly present in real systems, are averaged out during the metadynamics reconstruction. Once again, it should be underlined that such a behavior depends crucially on the appropriate choice of CVs.

3.1. Reducing the error by averaging several profiles

The expressions for the error derived in this section quantify the typical deviation of a single metadynamics profile $V_G(s, t)$ from $F(s)$. This error is relevant and has to be kept under control if the available computer time allows running the system only for a short time, for example only until it escapes from the initial free energy minimum. In this case the exit free energy barrier is reliable only if $\bar{\epsilon}$ is small. In many relevant cases the metadynamics reconstruction can be continued until the collective variable explores diffusively a large region of CV space, crossing backward and forward several times from the product to the reactant region. A consequence of the derivation presented above, and in particular of equation (23), is that all the profiles $V_G(s, t)$ are equally reliable estimates of the free energy. Thus, as proposed in [112], after the entire FES has

been filled the best possible estimate of the free energy is *not the single profile, but the arithmetic average of all the profiles*. More precisely, if t_F is the time at which the CV starts diffusing in all the relevant region and t_{tot} is the total simulation time, the best estimate of the free energy is

$$F_{\text{meta}}(s, t_{\text{tot}}) = -\frac{1}{t_{\text{tot}} - t_F} \int_{t_F}^{t_{\text{tot}}} V_G(s, t) dt. \quad (33)$$

The standard deviation of this estimate from the free energy decays to zero with a law that is determined by the autocorrelation time of the profiles (see also the results presented in section 4.2). Once again, in slowly diffusing systems, successive profiles are very correlated and even taking the average does not significantly improve the free energy estimate.

In order to take the average over different profiles in equation (33) it is necessary that the dynamics of the CV after t_F is bound in a finite region of CV space. In several cases, this happens automatically, as the free energy outside the ‘relevant’ region is very large. In other cases one has to artificially restrain the system inside the region. This can be done, for example, by adding a suitable harmonic potential that is active when the CVs cross the boundaries of the region.

A different way to converge the free energy profile has been proposed in [136] where the authors propose a ‘well-tempered-metadynamics’ algorithm in which the height of the added Gaussian depends on the underlying bias, decreasing to zero when a given energy threshold is reached. The convergence of this algorithm to the correct free energy profile can be proved rigorously also in the presence of memory effects. An additional advantage of the well-tempered algorithm is that the exploration of the CV space is automatically limited by the control parameters. Thus, it is not necessary to restrain the simulation inside the relevant region.

4. The algorithm in practice

4.1. Implementation

Metadynamics can be easily implemented by using the continuous formulation described in section 1. Here we present and describe a practical example of the Fortran code that can be inserted with minimal effort in most molecular dynamics packages. The call to this subroutine should be inserted in the main MD loop, after the forces are computed and before the positions/velocities are updated. In this example the CV is the distance between atom 1 and atom 2, but it is easy to include other variables and extend the metadynamics to several dimensions. The routine performs three tasks:

- (i) It computes the value of the CV $s = S(x)$.
- (ii) Every τ_G time steps, it stores the value of s in an array that contains the centers of all Gaussians.
- (iii) It computes the derivative of $V_G(S(x), t)$ with respect to x using the chain rule: $(\partial/\partial x)V_G(S(x), t) = (\partial V_G(s, t)/\partial s)(\partial S(x)/\partial x)$. These derivatives are then added to the usual forces on the atoms, biasing the dynamics of the system.

In the example, all the metadynamics parameters are hard-coded, but, of course, their correct value will depend on the units of measure used in the code and on the specific system of interest.

```

subroutine metadynamics(r,f,N,it)
implicit none
integer N ! the number of atoms
integer it ! time step number
real*8 r(3,N) ! positions of the atoms (input)
real*8 f(3,N) ! forces from metadynamics (changed in this subroutine)
!variables for metadynamics
real*8, save :: w=0.1 !height of the Gaussians
real*8, save :: ds=0.1 !width of the Gaussians
integer, save :: tau_G=100 !frequency of Gaussian deposition
integer, save :: NG=0 !number of Gaussians
real*8, save :: s_of_t(1000) !position of the center of the Gaussians
!local variables
real*8 rij(3),s,ds_dr(3,N),gauss,dVg_ds
integer i
! first, compute the value s of the CV and of its derivatives ds_dr
! with respect to the atomic positions. This is the only part that should
! be changed
! to make the code perform metadynamics on another variable
rij(:)=r(:,1)-r(:,2)
s=sqrt(sum(rij(:)**2))
ds_dr(:,1)=rij(:)/s
ds_dr(:,2)=-rij(:)/s
! every tau_G time steps, save the value of s
if(mod(it,tau_G)==0)then
NG=NG+1
if(NG>1000)stop
s_of_t(NG)=s
write(10,'(i6,3f12.4)')it,s,ds,w
endif
! now compute the derivative of the history-dependent potential
! Vg(s,t) with respect to s
dVg_ds=0
do i=1,NG
gauss=w*exp(-(s-s_of_t(i))**2/2/ds**2)
dVg_ds=dVg_ds-gauss*(s-s_of_t(i))/2/ds**2
enddo
! finally, compute the forces on the atoms
f(:,1)=f(:,1)-dVg_ds*ds_dr(:,1)
f(:,2)=f(:,2)-dVg_ds*ds_dr(:,2)
return
end subroutine metadynamics

```

The subroutine produces a file, fort.10, containing the positions of the centers of all Gaussians. This file can be used by an external program to reconstruct the potential $V_G(s)$ on a grid and visualize it, eventually as a function of the number of Gaussians. In this manner it is possible to obtain the free energy profiles as a function of time, as shown in figure 9.

The computational overhead of computing metadynamics forces is usually negligible, unless the number of Gaussians is very large. In fact, at each step it is necessary to compute a number of exponentials that becomes larger and larger as the simulation proceeds. A simple solution to this problem has been proposed in [66], and consists of storing the history-dependent potential on a regular grid. Every time a new Gaussian is added, the potential is updated. At every MD time step, the derivative $\partial V_G(s, t)/\partial s$ is estimated from the

finite difference of V_G evaluated on the two grid points that are closest to s . By this simple trick, the computational overhead of metadynamics does not grow linearly with simulation time but remains constant.

4.2. A practical example: alanine dipeptide in vacuum

In order to discuss how metadynamics works in practice it is interesting to consider a very simple case, alanine dipeptide (ACE-ALA-NME) in vacuum. The system, depicted in figure 3, is simulated with the Amber99SB force field [137] using a modified version of the NAMD MD code [138]. The temperature is maintained at 300 K by a Langevin thermostat. The conformation of alanine dipeptide is satisfactorily described by its two backbone dihedral angles ϕ and ψ . In figure 8, left panel, the free energy is shown as a function of these two angles, estimated with umbrella sampling using the WHAM scheme [8, 139]. Clearly, the free energy depends non-trivially on both variables, but the dependence on ϕ is more important, as the highest barriers correspond to changes in ϕ . In such a case, it is possible to perform metadynamics using only one CV, the angle ϕ . Thus, the history-dependent potential reads

$$V_G(\phi(x), t) = w \sum_{\substack{t' = \tau_G, 2\tau_G, \dots \\ t' < t}} \exp\left(-\frac{(\phi(x) - s(t'))^2}{2\delta s^2}\right), \quad (34)$$

where $s(t) = \phi(x(t))$. A preliminary unbiased run at 300 K is used to choose the Gaussian width δs . As shown in [132] and discussed in section 3 the optimal δs is approximately one half of the fluctuation of the CV in an unconstrained MD run. Indeed, metadynamics cannot reconstruct features of the free energy on a scale that is smaller than δs . This criterion gives, for the ϕ angle, $\delta s = 0.2$ rad. Clearly the fluctuations of the CV may depend on the local free energy minimum in which the system is found. Thus, there is no single best choice for δs and in practice it is convenient to take δs smaller than the smallest observed fluctuation. As shown in [132] and discussed in section 3 the error made on the reconstructed profile depends on the ratio w/τ_G and not on w and τ_G separately. Taking, for instance, $\tau_G = 2$ ps, the Gaussian height w is the only parameter that determines the computational time required for the free energy reconstruction. A useful criterion for choosing w is provided by equation (31), that states the total time required for filling a d -dimensional region of size S and with free energy wells of depth F_{\max} scales approximately as $t_{\text{sim}} \approx \tau_G (F_{\max}/w) (S/\delta s)^d$. If, for instance, one wants to invest approximately 3 ns in the free energy reconstruction and estimates that the free energy wells have a typical depth of ≈ 5 kcal mol⁻¹, an optimal choice would be $w \sim 0.1$ kcal mol⁻¹. The typical deviations of $V_G(s, t)$ from $F(s)$ will be of the order of $\sqrt{wT\delta sS/D\tau_G}$ (see equation (30)). If necessary, one can preliminarily estimate the value of the diffusion coefficient D , evaluate this error and decide whether the accuracy is sufficient or excessive, and eventually change the value of the parameters.

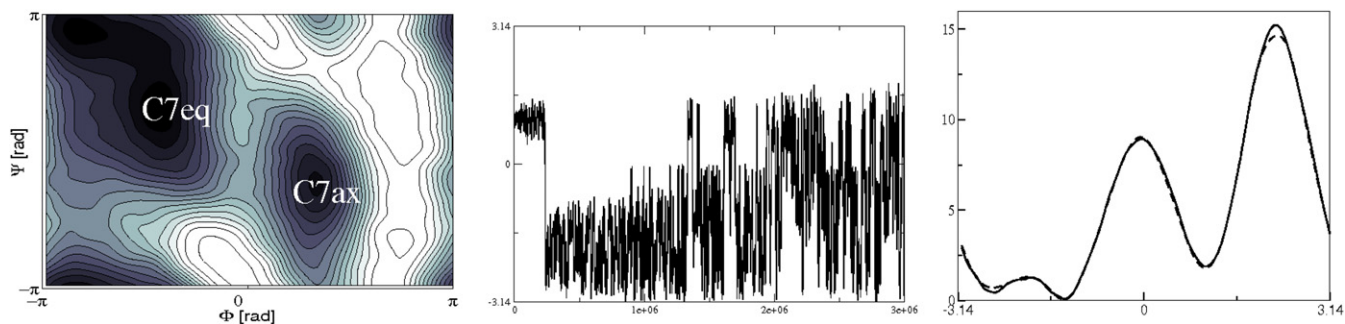


Figure 8. Metadynamics results for the alanine dipeptide. Upper left panel, a 2D FES as a function of ϕ and ψ obtained by umbrella sampling. Central panel: the first 2 ns of metadynamics trajectory on the CV ϕ . Right profile: the dashed curve is the result of averaging the profiles obtained by 1D metadynamics; the continuous curve is the 1D projection of the above 2D FES.

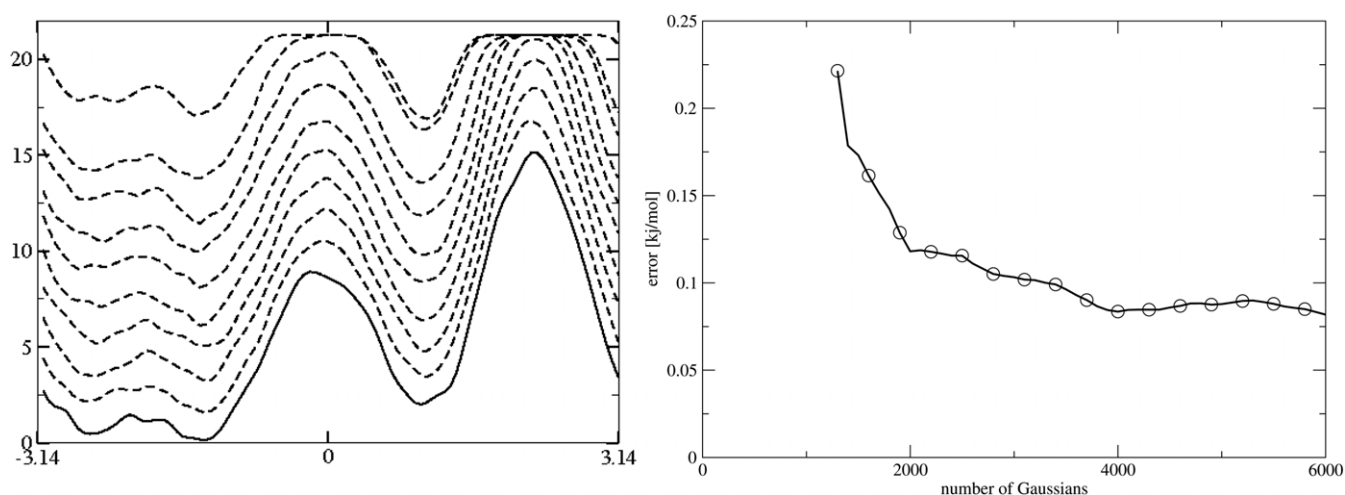


Figure 9. Left panel: reconstructed free energy profile as a function of time. The metadynamics run was 8 ns long; the profiles have been reconstructed every 0.8 ns. The continuous black curve is the history-dependent potential at the end of the simulation. Right panel: standard deviation of the metadynamics free energy with respect to the umbrella sampling result as a function of time calculated for the alanine dipeptide ϕ CV.

The central panel of figure 8 shows the value of ϕ during the first 2 ns of metadynamics performed with these parameters. The initial configuration corresponds to a $C7_{ax}$ conformation. As the history-dependent potential is added, the fluctuations of ϕ increase, until the system ‘fills’ the free energy basin to reach a $C7_{eq}$ conformation, with a behavior similar to that in figure 1. After approximately 0.8 ns, the added potential has mostly filled the two energy basins and the CV starts to show a diffusive behavior, visiting repeatedly the two basins. The reconstructed free energy profile as a function of time is shown in the right panel of figure 9. The metadynamics trajectory was run for a total of 8 ns. Accumulated profiles at 0.8 ns intervals are shown. It can be seen that after 2.4 ns the profile has reached an almost complete convergence, and the history-dependent potential defined by the sum of Gaussians is qualitatively similar to the free energy profile. Deviations are mainly due to the volume of the ‘computational sand’, the Gaussians, that are continuously thrown. Using smaller Gaussians would improve the accuracy of free energy reconstruction, as discussed in section 3.

It is also clear from figure 8 that, as discussed in section 3 all the profiles after 2.4 ns are equally reliable estimates of the free energy. Thus the best possible estimate of the free

energy is

$$F_{\text{meta}}(\phi, t_{\text{tot}}) = -\frac{1}{t_{\text{tot}} - t_F} \int_{t_F}^{t_{\text{tot}}} V_G(\phi, t) dt$$

with $t_F = 2.4$ ns. The averaged profile $F_{\text{meta}}(\phi, t_{\text{tot}})$ is shown in figure 8 right panel, black dashed line. It is almost indistinguishable from the profile calculated by integrating over ψ the ‘exact’ 2D profile obtained by umbrella sampling. Figure 9, right panel, reports the standard deviation of the metadynamics free energy with respect to the umbrella sampling result $F(\phi)$,

$$\varepsilon(t) = \left[\frac{1}{2\pi} \int_0^{2\pi} d\phi [F_{\text{meta}}(\phi, t) - F(\phi)]^2 - \left(\frac{1}{2\pi} \int_0^{2\pi} d\phi (F_{\text{meta}}(\phi, t) - F(\phi)) \right)^2 \right]^{1/2}$$

as a function of t_{tot} . Clearly, taking the average over a longer and longer t_{tot} systematically improves the estimate.

Systematic errors due to memory, inertia, lack of adiabatic separation, etc, are, in this case, negligibly small. This happens even if the choice of the CV is not optimal, as a relevant degree of freedom, the angle ψ , is not included in the bias. Of course

the situation would be totally different if metadynamics was performed using ψ as a collective variable. In this case the free energy reconstruction would be affected by large systematic error and hysteresis, and would probably never converge in a finite simulation time.

In conclusion, a successful metadynamics run requires following the steps that have already been described in the previous sections, and that are summarized here:

- (i) Equilibrate the system by normal molecular dynamics at the temperatures of interest.
- (ii) Choose the collective variables (see section 2), taking into account a few general rules: (i) the performance of metadynamics is optimal with two or three variables. With one variable the method is practically equivalent to thermodynamic integration [2, 3] or WHAM [7–9]. With more than three variables the time required to fill the free energy surface becomes very long. (ii) The CVs should clearly distinguish the metastable states that are already known before starting the simulation: the initial state and, if these are known, the final and the intermediate states.
- (iii) Choose the metadynamics parameters (see section 3), keeping in mind that: (i) the width(s) δs is (are) chosen by monitoring the value of standard deviation the CV(s) in a finite temperature run. (ii) The height w and the time between two successive Gaussians τ_G are not independent parameters, but the accuracy is determined by their ratio according to equation (30).
- (iv) Check the convergence of the results. If one uses metadynamics for reconstructing the free energy in a finite region of CV space, the estimate is reliable only if the observed behavior is similar to that shown in figure 9: the history-dependent potentials evaluated at different times must be approximately similar. When this is the case, the best possible estimate of the free energy is given by the average of all the profiles, according to the procedure described in section 3.1.

5. Extensions

As underlined in the previous sections, metadynamics requires selecting, before starting the simulation, a set of two or three CVs that are used for constructing the history-dependent bias. The method provides useful and accurate results only if this set includes all the relevant (slow) variables. Unfortunately, there is no way to ensure that this condition is satisfied before performing the simulation. Moreover, in several important cases the ‘relevant’ variables are not two or three, but many more. In these cases, the method cannot be applied, as filling a free energy surface in many dimensions would be computationally too expensive. In this section we describe some recent extensions of the methodology that are aimed at alleviating this serious drawback. All the approaches described here are based on running metadynamics simultaneously on several replicas of the system, gathering the information from all the replicas in order to improve the statistics of the free energy estimate.

5.1. Multiple walkers

The *multiple walkers* version of metadynamics [140] was the first approach taking advantage of running metadynamics on multiple replicas of the system simultaneously. It was originally developed with the purpose to speed up free energy calculations in loosely coupled parallel machines. It is based on running multiple interacting simulations, *walkers*, all contributing to construct the same metadynamics bias, that is shared by all walkers. Denoting by $x_i(t)$ the trajectory of the walker i , the history-dependent potential at time t is

$$V_G(S(x), t) = w \sum_{t' < t} \sum_{i=1}^{\text{NR}} \exp \left(- \frac{(S(x) - s_i(t'))^2}{2\delta s^2} \right),$$

where $s_i(t) = S(x_i(t))$. This potential biases the dynamics of all walkers simultaneously. In [140] it was shown that the error on the reconstructed free energy does not depend on the number of walkers, and is still approximately given by equation (30). Since NR walkers contribute simultaneously to the history-dependent potential, any free energy profile is reconstructed in $1/\text{NR}$ of the time. This leads to a linear scaling algorithm even on inexpensive loosely coupled clusters of PCs. Indeed the only information that has to be shared among the walkers is the list of added Gaussians. In a Linux cluster this can be easily achieved by sharing a single file containing the list. The same article [140] reports also that the accuracy and stability of the method can be improved by combining it with a weighted histogram analysis [8, 112].

5.2. Parallel tempering metadynamics

Parallel tempering metadynamics (PTmetaD) [113] is another method based on running several metadynamics in parallel. It is based on a combination of parallel tempering [32] and metadynamics and provides a possible solution to the problem of neglected CVs. Multiple replicas of metadynamics are run at different temperatures. In the spirit of the replica exchange method, an exchange of the coordinates of two replicas at adjacent temperatures is attempted with frequency $1/\tau_x$. The acceptance ratio takes into account the fact that different replicas experience different bias potentials [141]. The acceptance ratio for an exchange involving replicas i and j is thus

$$P = \min \left\{ 1, \exp \left[\left(\frac{1}{T_j} - \frac{1}{T_i} \right) (U(r_j) - U(r_i)) + \frac{1}{T_i} (V_i(s(r_i)) - V_i(s(r_j))) + \frac{1}{T_j} (V_j(s(r_j)) - V_j(s(r_i))) \right] \right\},$$

where U is the ordinary potential, r_i and T_i are the coordinates and the inverse temperature of replica i and V_i is the bias potential of replica i before the exchange. If the move is accepted, the coordinates are exchanged and the momenta are rescaled as

$$r'_i = r_j; \quad p'_i = \sqrt{\frac{T_i}{T_j}} p_j, \quad (35)$$

$$r'_j = r_i; \quad p'_j = \sqrt{\frac{T_j}{T_i}} p_i. \quad (36)$$

In PTmetaD the free energy profile is filled in parallel at all temperatures, and the dynamics of the system rapidly becomes diffusive in CV space. Combining parallel tempering and metadynamics in this way leads to a very efficient approach. Parallel tempering is improved because metadynamics explores high free energy regions within each replica ultimately leading to a more reliable estimate of the height of the relevant free energy barriers. Metadynamics is strengthened as the parallel tempering improves the sampling over degrees of freedom not explicitly included in the CV space. PTmetaD has been used in [32] to reconstruct the FES of the β -hairpin folding and in [69] to study the dynamics of aggregation of the peptides that form the folding nucleus of the HIV-1 protease.

A similar scheme can be used to combine solute tempering [142] or other variants of parallel tempering with metadynamics. In [142] a combination of metadynamics with solute tempering was used to reconstruct the free energy landscape of the α -helix of protein G. The results obtained are in good agreement with many experimental observations.

5.3. Bias exchange

Bias exchange (BE) [118] is another technique based on the combined use of replica exchange and metadynamics. The method is an extension of replica exchange metadynamics (see section 5.2) designed in order to allow the simultaneous reconstruction of the free energy as a function of a large number of variables. One runs in parallel, at the same temperature, a large number NR of molecular dynamics, biasing each replica with a metadynamics potential acting on just one or two CVs. If all the variables were relevant for describing the process and the replicas were run independently all the metadynamics would be affected by hysteresis and systematic errors. BE, at fixed time intervals, attempts to swap the bias potentials between pairs of replicas. Let us denote by S_k , $k = 1, \dots, \text{NR}$ the NR different CV sets on which the metadynamics biases are constructed. For two replicas of coordinates x_i and x_j and bias potentials $V_G(S_k(x_i), t)$ and $V_G(S_l(x_j), t)$, an exchange move consists of swapping $V_G(S_k, t)$ and $V_G(S_l, t)$. The move is accepted with a probability

$$P = \min \left\{ 1, \exp \left[\frac{1}{T} (V_G(S_k(x_i), t) + V_G(S_l(x_j), t) - V_G(S_k(x_j), t) - V_G(S_l(x_i), t)) \right] \right\}.$$

In this way, each trajectory evolves through the high dimensional free energy landscape sequentially biased by low dimensional potentials acting on one or two CVs at each time. Due to the multidimensional nature of the bias, the dynamics is able to explore a complex free energy landscape with great efficiency. A simple example will clarify how the swap moves are defined. Consider a system in which the relevant variables are two, e.g. two dihedral angles ψ and ϕ . Two replicas are then run in parallel, one with a bias on ψ and the other on ϕ . At a given time t a swap is attempted.

If the move is accepted, the simulation proceeds as follows:

	Coordinates	Before the swap		After the swap	
		CV	Bias potential	CV	Bias potential
Replica 1	x_1	ψ	$V_G(\psi(x_1), t)$	ϕ	$V_G(\phi(x_1), t)$
Replica 2	x_2	ϕ	$V_G(\phi(x_2), t)$	ψ	$V_G(\psi(x_2), t)$

When a swap is accepted the bias potential acting on the two replicas changes direction. Simultaneously, the values of the CVs involved in the swap perform a jump from $\psi(x_1)$ to $\psi(x_2)$ and from $\phi(x_1)$ to $\phi(x_2)$ in the example above. These jumps greatly help in decorrelating the dynamics and finally have the effect of improving the accuracy of the free energy estimate. The efficiency of the method is enhanced if the swaps are attempted not too frequently, say every 10 ps in classical MD simulations. Each replica should be given sufficient time to relax under the effect of the added bias, before attempting a swap that could change the bias direction.

In BE the result of the simulation is not a free energy as a function of NR CVs, but NR low dimensional projections of the free energy. The swap moves have the effect of reducing the hysteresis on the metadynamics potentials, that tend to behave as in the example of figure 8 also in difficult cases. Of course, hysteresis can be eliminated only if all the relevant variables are included and are biased by at least one replica. Thus, the approach of [118] does not solve (not even in principle) the major drawback of metadynamics, namely the necessity of ‘inventing’ before the simulation an appropriate set of CVs. Still, the approach allows treating simultaneously much more CVs than normal metadynamics, and this is of great help in setting up a simulation assuming as little as possible on the reaction mechanism.

This methodology was successfully used to reversibly fold a small protein, the tryptophan cage, using an all-atom force field and eight replicas [118]. The variables that were used are general, such as the number of backbone–backbone hydrogen bonds, the fraction of α -helical content, etc (see section 2.1), and did not require in their definition the knowledge of the folded structure. For this small system, within ~ 300 nanoseconds of simulation time, it was possible to identify the folded state and folding intermediates within statistical accuracy starting from a completely extended conformation. The same approach was also applied to somewhat larger systems, the Villin and Advillin C-terminal headpiece (cHP) and their Pro62Ala mutants [71]. The cHP is a 36 aminoacids thermostable domain composed of three tightly packed alpha helices. BE simulations of the Villin cHP predict a native fold in good agreement with the nuclear magnetic resonance data and indicate that the P62A mutation destabilizes the native fold to the point that it is no longer observed. Instead in Advillin cHP this mutation significantly alters the entropic contribution to the folding free energy but the mutant is still folded at room temperature. Finally, BE was applied to study the hybridization of DNA, and predict the free energies of the different helices that the system can form [70].

6. Conclusions and outlook

Since metadynamics first appeared in 2002 the use of the method among the computational community has steadily spread, as shown by an increasing number of citations. The many powerful extensions that have been recently introduced, the continuous ongoing effort to solve the weaknesses of the approach and the availability of an open source plugin for common MD programs [142] are a clear sign of a thriving metadynamics community. In our opinion this is not due to its uniqueness, as there are competing methods based on similar ideas [12, 35, 40, 41], but to the simplicity of implementation, and to the relatively easy manner in which one can control its accuracy and efficiency. By simply changing the dimensionality of the Gaussians entering in the history-dependent potential one can pass continuously from a fast (but coarse) exploration of the configuration space to an accurate (but expensive) evaluation of the free energy. This versatility has been exploited in applications, and the method is very often used not only for computing free energies, but also for predicting new structures, new reaction pathways, etc. For example in crystal structure prediction [54] the method is used only with the aim of efficiently exploring the configuration space and the fact that the sum of Gaussians provides an estimate of the free energy is not used at all.

In other popular algorithms for reconstructing the free energy, such as thermodynamic integration [2, 3] and WHAM [7–9], the free energy is reconstructed following a predefined scheme designed for covering deterministically all the CV space. In metadynamics, instead, the free energy is reconstructed recursively, starting from the bottom of a free energy well by a history-dependent random walk that explores a larger and larger portion of configuration space. This makes the algorithm very efficient as, by construction, the dynamics explores first the low-free energy regions, which are also the most interesting, avoiding spending time in regions that are irrelevant. Another advantage of metadynamics is that during the free energy reconstruction each point in CV space is explored several times during the simulation (see, for instance, the CV trajectory in figure 8). This greatly enhances de-correlation and leads to an improved accuracy in the free energy estimate. This feature is also exploited in the adaptive force bias, a method that shares many strengths with metadynamics [12].

The algorithm has significantly evolved since it was first introduced, and several variants and improvements have been developed. In practical applications it has proved to be more efficient to evolve the collective variables in a continuous fashion (see section 1.1). Moreover, a proof has been given that in conditions of adiabatic separation the method provides an unbiased estimator of the free energy (see section 3). As discussed in sections 1.1 and 1.2, adiabatic separation can be enforced, at least in principle, by using the Lagrangian or the discrete version of the algorithm. However, if the CVs are properly chosen, systematic errors deriving from memory effects on the s dynamics are negligible and the history-dependent potential provides a reliable estimate of the

free energy even if adiabatic separation is not enforced. An example of such a behavior has been presented in section 4.2, another example has been discussed in section 3. Once again, it is important to note that this extremely useful property holds only if the appropriate set of CVs is chosen. If an important variable is forgotten the bias potential evolves in an unpredictable manner that is determined by the transitions in the hidden variable. At the moment a theoretical framework for understanding and quantifying the degradation of the accuracy in the presence of memory effects and/or hidden variables is not available. One has to rely on empirical rules like the ones presented in section 4 ('the metadynamics potential is growing evenly') to verify if the choice of the CVs is appropriate and the method is converging.

The necessity to choose *a priori* a limited number of CVs is possibly the major drawback of the method, as there is no manner to ensure that the choice is appropriate before performing the simulation. The choice of the CVs for a completely new system is guided only by chemical or physical intuition, and in several cases one has to proceed by trial and error. Often a preliminary analysis based on methods that can cope with a fully unknown reaction mechanism, such as the ones listed in point (ii) of the introduction, might be preferable to a 'blind' metadynamics performed with a wrong variable. Moreover, in many important cases, such as proteins, the 'relevant' variables are just too many. In these cases, the method in its original formulation cannot be applied, as filling a free energy surface in more than three–four dimensions would be too computationally expensive. All the most recent extensions of metadynamics have been developed with the aim of alleviating this drawback, and, as discussed in section 5, important success has been achieved. In the future more effort will have to be spent in this direction, in order to make the choice of the CVs less and less critical.

Finally, it is important to note that experience has demonstrated that the best strategy for extracting useful numbers from numerical simulations is using synergically several tools exploiting the advantages of each. In this respect, metadynamics is perfectly complementary to other techniques, such as transition path sampling, that are less efficient in sampling, but allow the detail of a transition to be analyzed without choosing *a priori* a reaction coordinate or a set of CVs.

Availability

Currently metadynamics is implemented in numerous MD and FPMD codes, including (but not limited to): gromacs (grometa), NAMD, Amber, cpmd, cp2k. Albeit there is no 'official' code repository to metadynamics, at this moment the most up-to-date semiempirical MD distribution resembling an 'official' code repository is grometa (<http://lxmi.mi.infn.it/~provasi/grometa/Site/Welcome.html>) where numerous people contribute to add the latest feature to the code. The corresponding up-to-date FPMD code is cp2k (<http://cp2k.berlios.de/>).

Acknowledgments

A number of people contributed to the development of metadynamics, making the methodology useful for simulating realistic systems of increasing complexity. The methods presented in this review are the result of the combined effort of the authors, Michele Parrinello and (in alphabetic order): Alessandro Barducci, Marco Bernasconi, Mauro Boero, Massimiliano Bonomi, Davide Branduardi, Giovanni Bussi, Rosa Bulo, Matteo Ceccarelli, Davide Donadio, Bernd Ensing, Marcella Iannuzzi, Roman Martoňák, Cristian Micheletti, Artem Oganov, Stefano Piana, Paolo Raiteri and Andras Stirling. We gratefully acknowledge Gianluca Lattanzi and Gareth Tribello for several useful suggestions in writing this manuscript and also Giovanni Ciccotti, David Chandler, Eric Vanden-Eijnden and Michael L Klein for a number of illuminating discussions and many precious suggestions. Finally, a particular acknowledgment goes to Jannis Kevrekidis, who first stimulated us to consider dimensional reduction as a tool for efficiently exploring the phase space.

References

- [1] Warshel A 1991 *Computer Modeling of Chemical Reactions in Enzymes and Solutions* (New York: Wiley)
- [2] Carter E A, Ciccotti G, Hynes J T and Kapral R 1989 *Chem. Phys. Lett.* **156** 472–7
- [3] Sprik M and Ciccotti G 1998 *J. Chem. Phys.* **109** 7737–44
- [4] Bash P A, Singh U C, Brown F K, Langridge R and Kollman P A 1987 *Science* **235** 574–6
- [5] Patey G N and Valleau J P 1975 *J. Chem. Phys.* **63** 2334–9
- [6] Grubmüller H 1995 *Phys. Rev. E* **52** 2893–906
- [7] Ferrenberg A and Swendsen R 1988 *Phys. Rev. Lett.* **61** 2635
- [8] Kumar S, Rosenberg J M, Bouzida D, Swendsen R H and Kollman P A 1995 *J. Comput. Chem.* **16** 1339–50
- [9] Roux B 1995 *Comput. Phys. Commun.* **91** 275–82
- [10] Jarzynski C 1997 *Phys. Rev. Lett.* **78** 2690
- [11] Crooks G 1998 *J. Stat. Phys.* **90** 1481–7
- [12] Darve E and Pohorille A 2001 *J. Chem. Phys.* **115** 9169–83
- [13] Rodriguez-Gomez D, Darve E and Pohorille A 2004 *J. Chem. Phys.* **120** 3563–78
- [14] Gullingsrud J, Braun R and Schulten K 1999 *J. Comput. Phys.* **151** 190–211
- [15] Rosso L, Minary P, Zhu Z and Tuckerman M 2002 *J. Chem. Phys.* **116** 4389–402
- [16] Elber R and Karplus M 1987 *Chem. Phys. Lett.* **139** 375–80
- [17] Henkelman G and Jónsson H 2000 *J. Chem. Phys.* **113** 9978–85
- [18] E W, Ren W Q and Vanden-Eijnden E 2005 *J. Phys. Chem. B* **109** 6688–93
- [19] Maragliano L, Fischer A, Vanden-Eijnden E and Ciccotti G 2006 *J. Chem. Phys.* **125** 024106
- [20] Dellago C, Bolhuis P, Csajka F S and Chandler D 1998 *J. Chem. Phys.* **108** 1964–77
- [21] Dellago C, Bolhuis P and Geissler P 2002 *Adv. Chem. Phys.* **123** 1–78
- [22] Peters B and Trout B L 2006 *J. Chem. Phys.* **125** 054108
- [23] van Erp T, Moroni D and Bolhuis P 2003 *J. Chem. Phys.* **118** 7762–74
- [24] Faradjian A and Elber R 2004 *J. Chem. Phys.* **120** 10880–9
- [25] Allen R, Warren P and ten Wolde P 2005 *Phys. Rev. Lett.* **94** 018104
- [26] Juraszek J and Bolhuis P G 2006 *Proc. Natl Acad. Sci. USA* **103** 15859–64
- [27] Fletcher R and Powell M J D 1963 *Comput. J.* **6** 163–8
- [28] Henkelman G and Jónsson H 1999 *J. Comput. Phys.* **111** 7010–22
- [29] Voter A 1997 *Phys. Rev. Lett.* **78** 3908–11
- [30] Miron R and Fichthorn K 2004 *Phys. Rev. Lett.* **93** 128301
- [31] Barkema G and Mousseau N 1996 *Phys. Rev. Lett.* **77** 4358–61
- [32] Merlitz H and Wenzel W 2002 *Chem. Phys. Lett.* **362** 271–7
- [33] Sugita Y and Okamoto Y 1999 *Chem. Phys. Lett.* **314** 141–51
- [34] Nakajima N, Higo J, Kidera A and Nakamura H 1997 *Chem. Phys. Lett.* **278** 297–301
- [35] Wang F and Landau D P 2001 *Phys. Rev. Lett.* **86** 2050
- [36] Trebst S, Troyer M and Hansmann U 2006 *J. Chem. Phys.* **124** 174903
- [37] Laio A and Parrinello M 2002 *Proc. Natl Acad. Sci. USA* **99** 12562–6
- [38] Theodoropoulos C, Qian Y and Kevrekidis I G 2000 *Proc. Natl Acad. Sci. USA* **97** 9840–3
- [39] Kevrekidis I G, Gear C W and Hummer G 2004 *AIChE J.* **50** 1346–55
- [40] Cvijovic D and Klinowski J 1995 *Science* **267** 664–6
- [41] Huber T, Torda A and van Gunsteren W 1994 *J. Comput. Aided Mol. Des.* **8** 695–708
- [42] Marsili S, Barducci A, Chelli R, Procacci P and Schettino V 2006 *J. Phys. Chem. B* **110** 14011–3
- [43] Bussi G, Laio A and Parrinello M 2006 *Phys. Rev. Lett.* **96** 090601
- [44] Zipoli F, Bernasconi M and Martoňák R 2004 *Eur. Phys. J. B* **39** 41–7
- [45] Iannuzzi M and Parrinello M 2004 *Phys. Rev. Lett.* **93** 025901
- [46] Donadio D, Raiteri P and Parrinello M 2005 *J. Phys. Chem. B* **109** 5421–4
- [47] Donadio D and Bernasconi M 2005 *Phys. Rev. B* **71** 073307
- [48] Di Pietro E, Pagliai M, Cardini G and Schettino V 2006 *J. Phys. Chem.* **110** 13539–46
- [49] Trudu F, Donadio D and Parrinello M 2006 *Phys. Rev. Lett.* **97** 105701
- [50] Laino T, Donadio D and Kuo I F W 2007 *Phys. Rev. B* **76** 195210
- [51] Prestipino S and Giaquinta P V 2008 *J. Chem. Phys.* **128** 114707
- [52] Quigley D and Rodger P M 2008 *J. Chem. Phys.* **128** 154518
- [53] Behler J, Martonak R, Donadio D and Parrinello M 2008 *Phys. Rev. Lett.* **100** 185501
- [54] Martoňák R, Laio A and Parrinello M 2003 *Phys. Rev. Lett.* **90** 75503
- [55] Martoňák R, Laio A, Bernasconi M, Ceriani C, Raiteri P and Parrinello M 2005 *Z. Kristallogr.* **220** 489–98
- [56] Ceriani C, Laio A, Fois E, Gamba A, Martonak R and Parrinello M 2004 *Phys. Rev. B* **70** 113403
- [57] Oganov A, Martonak R, Laio A, Raiteri P and Parrinello M 2005 *Nature* **438** 1142–4
- [58] Martonak R, Donadio D, Oganov A R and Parrinello M 2006 *Nat. Mater.* **5** 623–6
- [59] Martonak R, Donadio D, Oganov A R and Parrinello M 2007 *Phys. Rev. B* **76** 014120
- [60] Karamatzanis P G, Raiteri P, Parrinello M, Leslie M and Price S L 2008 *J. Phys. Chem. B* **112** 4298–308
- [61] Ceccarelli M, Danelon C, Laio A and Parrinello M 2004 *Biophys. J.* **87** 58–64
- [62] Gervasio F L, Laio A and Parrinello M 2005 *J. Am. Chem. Soc.* **127** 2600–7
- [63] Branduardi D, Gervasio F L, Cavalli A, Recantini M and Parrinello M 2005 *J. Am. Chem. Soc.* **127** 9147–55
- [64] Barducci A, Chelli R, Procacci P, Schettino V, Gervasio F and Parrinello M 2006 *J. Am. Chem. Soc.* **128** 2705–10
- [65] Fiorin G, Pastore A, Carloni P and Parrinello M 2006 *Biophys. J.* **91** 2768–77

- [66] Babin V, Roland C, Darden T A and Sagui C 2006 *J. Chem. Phys.* **125** 204909
- [67] Biarnes X, Ardevol A, Planas A, Rovira C, Laio A and Parrinello M 2007 *J. Am. Chem. Soc.* **129** 10686–93
- [68] Kamiya K, Boero M, Tateno M, Shiraishi K and Oshiyama A 2007 *J. Phys.: Condens. Matter* **19** 365220
- [69] Bonomi M, Gervasio F L, Tiana G, Provati D, Broglia R A and Parrinello M 2007 *Biophys. J.* **93** 2813–21
- [70] Piana S 2007 *J. Phys. Chem. A* **111** 12349–54
- [71] Piana S, Laio A, Marinelli F, Van Troys M, Bourry D, Ampe C and Martins J C 2008 *J. Mol. Biol.* **375** 460–70
- [72] Piccinini E, Ceccarelli M, Affinito F, Brunetti R and Jacoboni C 2008 *J. Chem. Theory Comput.* **4** 173–83
- [73] Petraglio G, Bartolini M, Branduardi D, Andrisano V, Recanatini M, Gervasio F L, Cavalli A and Parrinello M 2008 *Proteins: Struct. Funct. Bioinf.* **70** 779–85
- [74] Ceccarelli M, Anedda R, Casu M and Ruggerone P 2008 *Proteins: Struct. Funct. Bioinf.* **71** 1231–6
- [75] Iannuzzi M, Laio A and Parrinello M 2003 *Phys. Rev. Lett.* **90** 238302
- [76] Churakov S, Iannuzzi M and Parrinello M 2004 *J. Phys. Chem. B* **108** 11567–74
- [77] Stirling A, Iannuzzi M, Laio A and Parrinello M 2004 *ChemPhysChem* **5** 1558–68
- [78] Ensing B, Laio A, Gervasio F, Parrinello M and Klein M 2004 *J. Am. Chem. Soc.* **126** 9492–3
- [79] Ensing B, Laio A, Parrinello M and Klein M 2005 *J. Phys. Chem. B* **109** 6676–87
- [80] Boero M, Ikeshoji T, Liew C, Terakura K and Parrinello M 2004 *J. Am. Chem. Soc.* **126** 6280–6
- [81] Stirling A, Iannuzzi M, Parrinello M, Molnar F, Bernhart V and Luinstra G 2005 *Organometallics* **24** 2533–7
- [82] Jug K, Nair N and Bredow T 2005 *Phys. Chem. Chem. Phys.* **7** 2616–21
- [83] Ikeda T, Hirata M and Kimura T 2005 *J. Chem. Phys.* **122** 244507
- [84] Ascuitto E and Sagui C 2005 *J. Phys. Chem. A* **109** 7682–7
- [85] Zipoli F, Bernasconi M and Laio A 2005 *ChemPhysChem* **6** 1772–5
- [86] Boero M, Tateno M, Terakura K and Oshiyama A 2005 *J. Chem. Theory Comput.* **1** 925–34
- [87] Pagliai M, Iannuzzi M, Cardini G, Parrinello M and Schettino V 2006 *ChemPhysChem* **7** 141–7
- [88] Lee J, Ascuitto E, Babin V, Sagui C, Darden T and Roland C 2006 *J. Phys. Chem. B* **110** 2325–31
- [89] Ensing B, De Vivo M, Liu Z, Moore P and Klein M 2006 *Acc. Chem. Res.* **39** 73–81
- [90] Ishikawa T, Nagara H, Kusakabe K and Suzuki N 2006 *Phys. Rev. Lett.* **96** 095502
- [91] Di Valentin C, Pacchioni G and Bernasconi M 2006 *J. Phys. Chem. B* **110** 8357–62
- [92] Blumberger J, Ensing B and Klein M 2006 *Angew. Chem. Int. Ed. Engl.* **45** 2893–7
- [93] Nair N N, Schreiner E and Marx D 2006 *J. Am. Chem. Soc.* **128** 13815–26
- [94] Zipoli F and Bernasconi M 2006 *J. Phys. Chem. B* **110** 23403–9
- [95] Cucinotta C S, Ruini A, Catellani A and Stirling A 2006 *ChemPhysChem* **7** 1229–34
- [96] Boero M, Ikeda T, Ito E and Terakura K 2006 *J. Am. Chem. Soc.* **128** 16798–807
- [97] Michel C, Laio A, Mohamed F, Krack M, Parrinello M and Milet A 2007 *Organometallics* **26** 1241–9
- [98] Rodriguez-Fortea A, Iannuzzi M and Parrinello M 2007 *J. Phys. Chem. C* **111** 2251–8
- [99] Kumar P P, Kalinichev A G and Kirkpatrick R J 2007 *J. Chem. Phys.* **126** 204315
- [100] Urakawa A, Iannuzzi M, Hutter J and Baiker A 2007 *Chem.—Eur. J.* **13** 6828–40
- [101] Stanton C L, Kuo I F W, Mundy C J, Laino T and Houk K N 2007 *J. Phys. Chem. B* **111** 12573–81
- [102] Glezakou V A, Dupuis M and Mundy C J 2007 *Phys. Chem. Chem. Phys.* **9** 5752–60
- [103] Ceriotti M and Bernasconi M 2007 *Phys. Rev. B* **76** 245309
- [104] Schreiner E, Nair N N and Marx D 2008 *J. Am. Chem. Soc.* **130** 2768–70
- [105] Michel C and Milet A 2008 *THEOCHEM J. Mol. Struct.* **852** 54–61
- [106] Car R and Parrinello M 1985 *Phys. Rev. Lett.* **45** 2471
- [107] Parrinello M and Rahman A 1980 *Phys. Rev. Lett.* **45** 1196
- [108] Risken H 1989 *The Fokker–Planck Equation* (Berlin: Springer)
- [109] Andersen H C 1980 *J. Chem. Phys.* **72** 2384–93
- [110] E W and Vanden-Eijnden E 2004 *Metastability, Conformation Dynamics, and Transition Pathways in Complex Systems (Lecture Notes in Computational Science and Engineering)* ed S Attinger and P Koumoutsakos (Berlin: Springer)
- [111] Gardiner C 2004 *Handbook of Stochastic Methods* (Berlin: Springer)
- [112] Micheletti C, Laio A and Parrinello M 2004 *Phys. Rev. Lett.* **92** 170601
- [113] Bussi G, Gervasio F L, Laio A and Parrinello M 2006 *J. Am. Chem. Soc.* **128** 13435–41
- [114] Park M, Laio A, Iannuzzi M and Parrinello M 2006 *J. Am. Chem. Soc.* **128** 11318–9
- [115] Gervasio F, Laio A, Iannuzzi M and Parrinello M 2004 *Chem.—Eur. J.* **10** 4846–52
- [116] Ensing B and Klein M L 2005 *Proc. Natl Acad. Sci. USA* **102** 6755–9
- [117] Wu Y, Schmitt J and Car R 2004 *J. Chem. Phys.* **121** 1193–200
- [118] Piana S and Laio A 2007 *J. Phys. Chem. B* **111** 4553–9
- [119] Allen M and Tildesley D 2001 *Computer Simulation of Liquids* (Oxford: Clarendon)
- [120] Baroni S, de Gironcoli S, Dal Corso A and Giannozzi P 2001 *Rev. Mod. Phys.* **73** 515–62
- [121] Branduardi D, Gervasio F L and Parrinello M 2007 *J. Chem. Phys.* **126** 054103
- [122] Kearsley S K 1989 *Acta Crystallogr. A* **45** 208–10
- [123] Vendruscolo M, Najmanovich R and Domany E 1999 *Phys. Rev. Lett.* **82** 656–9
- [124] Amadei A, Linssen A and Berendsen H 1993 *Proteins: Struct. Funct. Genet.* **17** 412–25
- [125] Spiwok V, Lipovova P and Kralova B 2007 *J. Phys. Chem. B* **111** 3073–6
- [126] Sugino o and Car R 1995 *Phys. Rev. Lett.* **74** 1823–6
- [127] Alfe D, Gillan M J and Price G D 2002 *J. Chem. Phys.* **116** 6170–7
- [128] Tomagnini O, Ercolessi F, Iarlori S, DiTolla F and Tosatti E 1996 *Phys. Rev. Lett.* **76** 1118–21
- [129] Alfe D, Gillan M J and Price G D 1999 *Nature* **401** 462–4
- [130] Auer S and Frenkel D 2001 *Nature* **409** 1020–3
- [131] Steinhardt P J, Nelson D R and Ronchetti M 1983 *Phys. Rev. B* **28** 784–805
- [132] Laio A, Rodriguez-Fortea A, Gervasio F L, Ceccarelli M and Parrinello M 2005 *J. Phys. Chem. B* **109** 6714–21
- [133] Zwanzig R 1961 *Phys. Rev.* **124** 983–92
- [134] Ottinger H 1998 *Phys. Rev. E* **57** 1416
- [135] Badjic J D, Balzani V, Credi A, Silvi S and Stoddart J F 2004 *Science* **303** 1845–9
- [136] Barducci A, Bussi G and Parrinello M 2008 *Phys. Rev. Lett.* **100** 020603
- [137] Hornak V, Abel R, Okur A, Strockbine B, Roitberg A and Simmerling C 2006 *Proteins: Struct. Funct. Bioinf.* **65** 712–25

- [138] Phillips J C, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel R D, Kale L and Schulten K 2005 *J. Comput. Chem.* **26** 1781–802
- [139] Kumar S, Payne P W and Vásquez M 1996 *J. Comput. Chem.* **17** 1269–75
- [140] Raiteri P, Laio A, Gervasio F L, Micheletti C and Parrinello M 2006 *J. Phys. Chem. B* **110** 3533–9
- [141] Yan Q and de Pablo J 1999 *J. Chem. Phys.* **111** 9509–16
- [142] Camilloni C, Provati D, Tiana G and Broglia R A 2008 *Proteins: Struct. Funct. Bioinf.* **71** 1647–54