Chapter

ADVANCED METHODS USED IN MOLECULAR DYNAMICS SIMULATION OF MACROMOLECULES

Hiqmet Kamberaj*
Department of Computer Engineering
Faculty of Engineering
International Balkan University
Skopje, Republic of Macedonia

Abstract

Molecular dynamics (MD) simulations at atomic level have widely been used in studying macromolecular systems, such as protein, DNA and their complexes, mainly because the classical statistical mechanic's laws can explain different phenomena occurring at specified experimental conditions. In this study, we will present the most advanced methods used in the MD simulation of macromolecular systems. Furthermore, a discussion of applications of these methods and perspective on developing new approaches will be introduced. This study aims to review the methods that are developed to enhance the conformation sampling of molecular simulations, in particular, for observing rare events in complex molecular systems. In the summary, we also present a discussion and perspective

*Corresponding Author Email: h.kamberaj@gmail.com

on the methods described in this chapter and propose the new possible improvement of these approaches, which could result in further enhancement of conformation sampling.

Keywords: Molecular dynamics simulation, enhanced sampling, rare events, conformation transitions.

1. Introduction

MD approach at atomic resolution is often used to study complex biomolecular systems [126], mainly because the classical statistical mechanic's laws can explain different phenomena occurring at specified experimental conditions [194]. In particular, MD is used to study the internal fluctuations [1, 126], protein folding dynamics [154], transition path sampling [142], protein-DNA, protein-protein and protein-ligand complexes, and free energy calculations [165]. However, standard MD simulation has limited time and size scale, which makes it difficult to study typical phenomena of macromolecular systems, such as slow conformation motions [32, 138]. Therefore, it has been argued elsewhere [60, 127] that these limitations may be avoided by employing new statistical and computational approaches to be studied efficiently.

There have been different efforts in developing new approaches for enhancement of conformation sampling of simulations using MD technique, as discussed elsewhere [194]. These MD approaches are used in many applications for lowering conformation transition barriers by increasing the rate of rare events occurrence by introducing a bias that can be rigorously removed a posterior, or even without the bias term. Different approached have been used to introduce bias during MD simulation, such as by changing the shape of potential energy surface until a (quasi) flat landscape is obtained [11, 73, 93, 110], using soft-core potential interactions [93], conformational flooding [70], (geometrical) constraints [199], or using Tsallis dynamics [7]. Other approaches include parallel tempering, such as replica-exchange [22,46,53,119,197], multicanonical algorithms [206], and swarm-like dynamics [84, 85, 91, 107]. For gaining an increase in both time a size scale of the systems, the coarse-grained models have also shown a great interest, for instance, by decreasing the number of interacting particles [94, 96, 120, 137, 166, 167, 181–183, 211], or reducing

the dimensional space to only essential degrees of freedom [102, 114, 170]. Recently [45, 56], using computer engineering, longer MD simulation runs have been reported scaling from hundreds of microseconds to milliseconds timescale. Besides, development of multiple time step integration numerical schemes, such as reference system propagator algorithm (RESPA), have provided other approaches for extending the time scales of MD simulations [121, 125, 188].

In this chapter, we will describe in details some of these methods which are most often used to improve the sampling of configuration space in the MD simulations. We aim to critically review these methods and provide a discussion and perspective of the approaches introduced here. Also, we will further discuss possible improvements of some these methods, which could yield an increase of sampling efficiency of MD simulation in studying more complex phenomena of macromolecular systems.

2. Multiple time step integrator

MD simulations of complex molecular systems, such as biomolecules characterized by multiple time scales, show some disadvantage due to the small time steps used to ensure the stability of numerical integration of the fast motions. Hence, too many time steps are needed for observation of slow conformation transitions, which practically requires a large number of force computations. For these reasons, the Reference System Propagator Algorithm (RESPA) method is introduced to reduce computational efforts for simulations of such system [185–187, 189]. The time-reversible forms of the RESPA methods have also been developed, named r-RESPA, which have shown to be very stable concerning the order and stability of numerical integrators [121]. The r-RESPA, which will be discussed below in more details, uses Trotter factorization of the classical Liouville propagation operator [34, 80, 178].

Following the discussion in literature [121] (see also Ref. [128]), for a system with f degrees of freedom the Liouville operator, L, is defined as

$$iL = \{\cdots, H\} = \sum_{j=1}^{f} \left[\dot{x}_j \frac{\partial}{\partial x_j} + \dot{p}_j \frac{\partial}{\partial p_j} \right] \tag{1}$$

where Cartesian coordinates are used with $(x_j, p_j) \equiv \Gamma$ the position and conjugate momenta of the system, $\dot{p}_j$ gives the force along the jth direction, and

$\{\cdots\}$ represents the Poisson bracket of the system. L is a linear Hermitian operator of square integrable function on the phase space of Γ. The time propagation operator as a function of L is defined by

$$U(t) = \exp(iLt)$$

which is a unitary: $U(-t) = U^{-1}(t)$. The position and conjugate momenta state point of the system at a given time t is defined as $\Gamma(t) = U(t)\Gamma(0)$, which allows determining one time step propagation as the following:

$$\Gamma(\Delta t) = \exp(iL\Delta t)\Gamma(0)$$

where $\Delta t = t/P$ is the size of a time step. Here, t is the total evolution time and P are the number of integration points.

By splitting the Liouville operator into n different terms, like the following:

$$iL = \sum_{k=1}^{n} iL_k$$

and use the Trotter factorization scheme [81], then the propagator becomes

$$U(t) = \left\{ \left[\sum_{k=1}^{n-1} U_k(\Delta t/2) \right] U_n(\Delta t) \times \left[\sum_{k=1}^{n-1} U_{n-k}(\Delta t/2) \right] \right\}^{P} + O(t^3/P^2) \tag{2}$$

where $U_k(h) = \exp(iL_k h)$. Denoting

$$G(\Delta t) = \left[\sum_{k=1}^{n-1} U_k(\Delta t/2) \right] U_n(\Delta t) \times \left[\sum_{k=1}^{n-1} U_{n-k}(\Delta t/2) \right]$$

As shown in Ref. [121], $G(\Delta t)G(-\Delta t) = 1$, therefore, $G(\Delta t)$ generates time-reversible dynamics.

The multiple time step integrator is based on splitting the system into the fast and slow degrees of freedom. Equivalently, decomposing the forces entering

into the equations of motion into long-range forces, $F_l(\mathbf{r})$ and short-range forces $F_s(\mathbf{r})$ [121]:

$$F(\mathbf{r}) = F_s(\mathbf{r}) + F_l(\mathbf{r})$$

The short-range forces in the system are related to the slow degrees of freedom, and thus, they determine the multiple time step of the integrator δt. On the other hand, the long-range forces are related to the fast degrees of freedom, and thus, they determine the most extended time step of the integrator Δt. The relationship is established as

$$\delta t = \frac{\Delta t}{N_{MTS}} \tag{3}$$

where N_{MTS} is the number of multiple steps. Here, the short-range forces are calculated every time step δt, and long-range forces are calculated after every N_{MTS} time steps (i.e., every time step Δt). Hence, the degrees of freedom are advanced using Δt as a time step. In the r-RESPA implementation, this procedure decreases the number of calls for forces calculations, which reduces, in turn, the overall computational time.

The basic idea of r-RESPA implementation, as discussed elsewhere [121, 125, 188], is on determining a reference system force $F_s(\mathbf{r})$ for short range interactions. Then, Eq. (1) can be written in the following form:

$$\begin{aligned} iL &= \sum_{j=1}^{f} \left(\dot{x}_j \frac{\partial}{\partial x_j} + F_s(x_j) \frac{\partial}{\partial p_j} + F_l(x_j) \frac{\partial}{\partial p_j} \right) \\ &= iL_s + \sum_{j=1}^{f} F_l(x_j) \frac{\partial}{\partial p_j} \end{aligned} \tag{4}$$

and the propagator operator is factorized as

$$\begin{aligned} G(\Delta t) = &\prod_{j=1}^{f} \exp\left(\frac{\Delta t}{2} F_l(x_j) \frac{\partial}{\partial p_j} \right) \\ &\times \exp\left(iL_s \Delta t \right) \\ &\times \prod_{j=1}^{f} \exp\left(\frac{\Delta t}{2} F_l(x_j) \frac{\partial}{\partial p_j} \right) \end{aligned} \tag{5}$$

where the operator $\exp(iL_s\Delta t)$ propagates the state vector using the short range forces with a shorter time step δt (see Eq. 3). Here, this operator is factorized using the Trotter formula [121]:

$$\exp(iL_s\Delta t) = \left[\prod_{j=1}^{f}\exp\left(\frac{\delta t}{2}F_s(x_j)\frac{\partial}{\partial p_j}\right)\right. \times \prod_{j=1}^{f}\exp\left(\delta t F_s(x_j)\dot{x}_j\frac{\partial}{\partial x_j}\right) \left.\times \prod_{j=1}^{f}\exp\left(\frac{\delta t}{2}F_s(x_j)\frac{\partial}{\partial p_j}\right)\right]^{N_{MTS}} \tag{6}$$

Here, N_{MTS} is usually chosen *a priory* to guarantee the stability of numerical integrator [121]. Usually, when the operator $G(\Delta t)$ is applied to an initial state $(\mathbf{r}(0),\mathbf{p}(0))$, it gives a solution for both position and velocity similar to Verlet numerical integrator [121].

Following the discussions in Refs. [121,129,198], for a Lennard-Jones type of fluid, exists only the translational relaxation time characteristic. In that case, the integration time step can easily be chosen. On the other hand, for biomolecules, indeed there exists more than one time-scale. For example, in addition to the translational and rotational relaxation times, there exists the time characterizing intra-molecular motion, such as bond stretching, angle bending, and dihedral angle motion. Furthermore, the inter-molecular motion, including van der Waals and electrostatic interactions, is of the typical timescale of one or more orders in magnitude larger than intra-molecular motion. In such cases, the system is characterized by stiff nonlinear differential equations, which require the use of a small enough time step to observe fast motion, if treated using one time-scale.

Other systems that are characterized by more than one timescales are those consisting of high-frequency oscillators interacting with a bath of slow motion [189], and the systems consisting of large mass particles (slow degrees of freedom) interacting with lighter ones (fast particles) [187].

The method is also used to treat systems coupled to a Nosé heat bath [133, 169, 200] used to keep temperature and/or pressure fixed during MD simulations. Here, heat bath includes extra fast degrees of freedom into the system, treated using multiple time stepping algorithms [121], typically, two or

more time steps. The method has been used by many molecular simulation software codes in performing simulations of complex systems, for example, in CHARMM program [27].

However, the approach is limited by the so-called resonance phenomena, which restricts the use of time steps higher than $\Delta t < 8$ fs by r-RESPA in MD simulations of biomolecular systems [24, 118, 161]. It must be noted that not just time-reversible integrators, but also multiple time step symplectic integrators [164] show numerical instability limiting the use of large time steps [201]. According to Ref. [161], the resonance phenomena is the result of using the perturbation techniques to derive the numerical integrators. To overcome these problems, numeric methods have been introduced to increase the time steps in molecular dynamics simulations. For example, non-symplectic Langevin Molly (LM) integration method [97] and the so-called LN integrator, which combines the force separation approach with Langevin dynamics [18, 19]. These methods allow using more substantial time steps in MD simulations using stochastic approaches to increase the numerical stability of integration. A stable version of r-RESPA integrator has also been introduced, named the Targeted Mollified Impulse method [117], which includes the Langevin dynamics to improve the accuracy of multiple time stepping integrator.

In Ref. [125], authors discuss a reversible, resonance-free integrator which allows for using time steps of the order up to 100 fs or even larger depending on the time length correlations studied. This integrator uses non-Hamiltonian dynamics, which are shown to sample a canonical distribution of physical configuration space [125]

$$(q_1, q_2, \cdots, q_{3N}) \equiv ((x_1, y_1, z_1), \cdots, (x_N, y_N, z_N))$$

Here, we have written the equations of motion governing dynamics by modifying those given in Ref. [125] as the following:

$$\begin{aligned} \dot{q}_i &= \frac{p_i}{m_i}, \\ \dot{p}_i &= F_i - \lambda_i p_i - P_1^{(s)} p_i \\ \dot{\eta}^{(i)} &= -\sum_{j=1}^{L} \left[\frac{Q_1^{(i)} (\xi_{1,j}^{(i)})^2}{k_B T} \xi_{2,j}^{(i)} - \sum_{k=2}^{M} \xi_{k,j}^{(i)} \right] \end{aligned} \tag{7}$$

$$\dot{\xi}^{(i)}_{1,j} = -\xi^{(i)}_{1,j}\xi^{(i)}_{2,j} - \lambda^{(i)}_b \xi^{(i)}_{1,j} - \lambda_i \xi^{(i)}_{1,j}$$
$$j = 1, \cdots, L$$
$$\dot{\xi}^{(i)}_{k,j} = \frac{G^{(i)}_{k,j}}{Q^{(i)}_k} - \xi^{(i)}_{k+1,j}\xi^{(i)}_{k,j}$$
$$j = 1, \cdots, L;\ k = 2, \cdots, M-1$$
$$\dot{\xi}^{(i)}_{M,j} = \frac{G^{(i)}_{M,j}}{Q^{(i)}_{M,j}}, \quad j = 1, \cdots, L$$
$$\dot{s}_{i,k} = P^{(s)}_{i,k}, \quad k = 1, 2, \cdots, M$$
$$\dot{P}^{(s)}_{i,k} = \frac{\Gamma_{i,k}}{W_{i,k}} - P^{(s)}_{i,k+1}P^{(s)}_{i,k}, \quad k = 1, 2, \cdots, M-1$$
$$\dot{P}^{(s)}_{i,M} = \frac{\Gamma_{i,M}}{W_{i,M}},$$

for $i = 1, 2, \cdots f$ ($f = 3N$), where

$$\lambda_i = \frac{1}{2K(p,\xi)}\left(\frac{p_i F_i}{m_i}\right) \tag{8}$$

$$\lambda^{(i)}_b = \frac{1}{2K(p,\xi)}\left(-\frac{L-1}{L}\sum_{j=1}^{L} Q^{(i)}_1 \xi^{(i)}_{2,j}(\xi^{(i)}_{1,j})^2\right) \tag{9}$$

$$2K(p,\xi) = \frac{p_i^2}{m_i} + \frac{L-1}{L}\sum_{j=1}^{L} Q^{(i)}_1 (\xi^{(i)}_{1,j})^2 \tag{10}$$

which ensures that maximum total kinetic energy accumulated in each degree of freedom is Lk_BT. In Eq. 7, M is the Nosé-Hoover chain length of thermostats, $\xi^{(i)}_{k,j}$ ($k = 1, \cdots, M$ and $j = 1, \cdots, L$) are the thermostat velocities associated with Lagrangian multiplier along the i degrees of freedom and $\eta^{(i)}$ is the corresponding thermostat coordinate, which is used to control the accumulated kinetic energy fluctuations. F_i is Newton's force on the i degrees of freedom. The thermostat forces $G^{(i)}_{k,j}$ are defined as

$$G^{(i)}_{k,j} = Q^{(i)}_{k-1}(\xi^{(i)}_{k-1,j})^2 - k_BT, \quad (k = 2, 3, \cdots, M) \tag{11}$$

for $(j = 1, \cdots, L)$, were L is an adjustable parameter. Here, k_B is the Boltzmann constant and $Q_k^{(i)}$ determine fictitious thermostat masses optimized in Refs. [64, 168]:

$$Q_k^{(i)} = k_B T \tau, \quad k = 1, 2, \cdots, M \tag{12}$$

where τ is a time scale associated with the thermostat. λ_i and $\lambda_b^{(i)}$ are the Lagrangian multipliers which are determined such that equations of motion have to satisfy the following constraint:

$$2K(p, \xi) = L k_B T$$

In Eqs. 7, $s_{i,k}$ and $P_{i,k}^{(s)}$ are, respectively, the thermostat coordinates and their associated velocities ($k = 1, \cdots, M$) at temperature T for ith degrees of freedom of the real system. Thermostat forces $\Gamma_{i,k}$ are defined as

$$\begin{aligned} \Gamma_{i,1} &= \frac{p_i^2}{m_i} - k_B T \\ \Gamma_{i,k} &= W_{i,k-1} \left(P_{i,k-1}^{(s)} \right)^2 - k_B T, \quad k = 2, \cdots, M \end{aligned} \tag{13}$$

$W_{i,k}$ are thermostat masses determined by Eq. 12, as in Refs. [64, 168].

Eqs. 7 can be numerically solved using the Liouville operator formalism and Trotter factorization schemes as suggested elsewhere [125] (and the references therein). The classical Liouville operator can be expressed as:

$$\begin{aligned} iL = \sum_{d=1}^{f} \Bigg[& \dot{v}_d \frac{\partial}{\partial v_d} + \dot{q}_d \frac{\partial}{\partial q_d} \\ & + \sum_{k=1}^{M} \sum_{j=1}^{L} \dot{\xi}_{k,j}^{(d)} \frac{\partial}{\partial \xi_{k,j}^{(d)}} + \dot{\eta}^{(d)} \frac{\partial}{\partial \eta^{(d)}} \\ & + \sum_{k=1}^{M} \dot{P}_{d,k}^{(s)} \frac{\partial}{\partial P_{d,k}^{(s)}} + \dot{s}_{d,k} \frac{\partial}{\partial s_{d,k}} \Bigg] \end{aligned} \tag{14}$$

which can then be decomposed for every degree of freedom d as

$$iL = \sum_{d=1}^{f} iL^{(d)} = \sum_{d=1}^{f} \left[iL_1^{(d)} + \sum_{n=1}^{N_d} iL_{2,n}^{(d)} + iL_{NHC}^{(d)} \right]$$

where

$$
\begin{aligned}
iL_1^{(d)} &= v_d \frac{\partial}{\partial q_d} + \dot{\eta}^{(d)} \frac{\partial}{\partial \eta^{(d)}} + \sum_{k=1}^{M} \dot{s}_{d,k} \frac{\partial}{\partial s_{d,k}} \\
iL_{2,n}^{(d)} &= \left(\frac{F_d^{(n)}}{m_d^{(n)}} - \lambda_d^{(n)} v_d^{(n)} - P_{d,1}^{(s)} v_d^{(n)} \right) \frac{\partial}{\partial v_d^{(n)}} \\
&\quad + \frac{\Gamma_{d,1}^{(n)}}{W_{d,1}} \frac{\partial}{\partial P_{d,1}^{(s)}} - \sum_{j=1}^{L} \lambda_d^{(n)} \xi_{1,j}^{(d)} \frac{\partial}{\partial \xi_{1,j}^{(d)}} \quad n = 1, 2, \cdots, N_d \\
iL_{NHC}^{(d)} &= \sum_{k=2}^{M} \sum_{j=1}^{L} \frac{G_{k,j}^{(d)}}{Q_{k,j}^{(d)}} \frac{\partial}{\partial \xi_{k,j}^{(d)}} - \sum_{k=1}^{M-1} \sum_{j=1}^{L} \xi_{k,j}^{(d)} \xi_{k+1,j}^{(d)} \frac{\partial}{\partial \xi_{k,j}^{(d)}} \\
&\quad - \sum_{j=1}^{L} \lambda_b^{(d)} \xi_{1,j}^{(d)} \frac{\partial}{\partial \xi_{1,j}^{(d)}} \\
&\quad + \sum_{k=2}^{M} \frac{\Gamma_{d,k}}{W_{d,k}} \frac{\partial}{\partial P_{d,k}^{(s)}} - \sum_{k=1}^{M-1} P_{d,k}^{(s)} P_{d,k+1}^{(s)} \frac{\partial}{\partial P_k^{(s)}}
\end{aligned}
\tag{15}
$$

where N_d is the number of parts that the force on every degree of freedom can be split, that is

$$
F_d = \sum_{n=1}^{N_d} F_d^{(n)}
$$

for each degree of freedom d ($d = 1, 2, \cdots, f$). Here, v_d is the velocity of the d-the degree of freedom, $v_d \equiv \dot{q}_d = p_d / m_d$. Note that it is assumed that the force's strength is decreasing with n. Introducing the multiple time step parameters [125]:

$$
\begin{aligned}
\delta t &= \frac{\Delta t}{N_{MTS}}, \quad N_{MTS} = \prod_{n=1}^{N_d} s_n \\
s_{N_d} &= 1, \quad w_n = \prod_{k=1}^{n-1} s_k, \quad w_1 = 1
\end{aligned}
\tag{16}
$$

Using the Trotter factorization scheme for classical Liouville operator, as suggested in Ref. [125], then the approximation of true evolution can be written

as:

$$\Gamma(\Delta t) \approx \left\{ e^{i\tilde{L}_{N_d}^{(t)}\delta t} \cdots \left(e^{i\tilde{L}_2^{(t)}\delta t} \left[e^{i\tilde{L}_1\delta t} \right]^{s_1-2} \times e^{i\tilde{L}_2\delta t} \right)^{s_2-2} \cdots e^{i\tilde{L}_{N_d}\delta t} \right\} \Gamma(0) \tag{17}$$

where

$$\begin{aligned} e^{i\tilde{L}_k\delta t} &= e^{iL_{NHC}^{(d)}\frac{\delta t}{2}} \\ &\times e^{iL_{2,1}^{(d)}\frac{\delta t}{2}} e^{iL_1^{(d)}\delta t} e^{\sum_{n=1}^{k} iL_{2,n}^{(d)} w_n \frac{\delta t}{2}} \\ &\times e^{iL_{NHC}^{(d)}\frac{\delta t}{2}} \end{aligned} \tag{18}$$

for each degree of freedom d ($d = 1, 2, \cdots, f$). Thus, the weak or long-range forces correspond to large values of n, and hence are calculated less often, but they are weighted with larger w_n to equalize they time step with that of short-range forces, where

$$s_n = \frac{w_{n+1}}{w_n}$$

gives the ratio of strengths between $(n+1)$ and n forces. It can be seen that the number of n force evaluations is N_{MTS}/w_n. The error in one time step is $O(\Delta t^3)$, and for the entire trajectory of length t, it is $O(t\Delta t^2)$ [125]. Analytical solutions can be obtained for each of the exponential factorized parts of the classical Liouville operator using the following relations:

$$\begin{aligned} \exp\left(a\frac{\partial}{\partial x}\right) f(x) &= f(x+a), \\ \exp\left(ax\frac{\partial}{\partial x}\right) f(x) &= f\left(e^a x\right) \end{aligned} \tag{19}$$

where a is a constant. Furthermore, Nosé-Hoover part iL_{NHC} of the classical operator can also be decomposed using Trotter factorization schemes as suggested in Ref. [63].

The method, using an unmodified version of Eq. 7 as in Ref. [125], has been implemented in the PINY-MD software [122] and it is applied for different

test systems, including a protein studied in *vacuo* using CHARMM22 force field [100]. The results published in Ref. [125] have shown that large time steps of $\Delta = 100\ fs$ provided perfect agreement with other methods using much smaller time steps. Efforts should be made to also include the solvent as a part of the system and check the efficiency of the method in the simulation of large macromolecular systems in the solvent. Future work should also focus on the comparison of the efficiency of sampling conformation equilibrium space of such complex systems using other methods discussed below or combining this method with other enhanced sampling techniques.

3. Generalized ensemble methods

It has been suggested [76] that *generalized-ensemble* can be used for a better sampling of configurations characterized by lower energies in computer simulations. This class of methods includes approaches, such as multicanonical sampling [21, 22], the broad histogram method [40, 41], Wang-Landau algorithm [196], Tsallis weights methods [184], and parallel tempering or replica exchange method [62, 106, 140]. These methods are often used to study the dynamics of biomolecular systems [77].

All of the above mentioned generalized-ensemble approaches have the same starting point, that is, the replacement of canonical Boltzmann-like weights at temperature T

$$\exp(-\beta\Delta E)$$

with non-Boltzmann weights, which allows the system escaping from the local minimum states. Here, ΔE represents the energy barrier height and β is the inverse temperature of the simulation, $\beta = 1/k_B T$.

In the canonical ensemble (characterized by fixed N, V, and T), each state point, $(\mathbf{r}, \mathbf{p})$, in the phase space is associated with a Boltzmann weight, which is defined in terms of the Hamiltonian function $H(\mathbf{r}, \mathbf{p})$:

$$W_B(\mathbf{r}, \mathbf{p}, \beta) = \exp(-\beta H(\mathbf{r}, \mathbf{p})) \tag{20}$$

Since momentum $\mathbf{p}$ and coordinates $\mathbf{r}$ are independent, we can integrate according to the momentum space Eq. 20, and re-write the Boltzmann factor as a

function of the instantaneous value, E, of the potential energy function $U(\mathbf{r})$:

$$W_B(E,\beta) = \exp(-\beta E) \tag{21}$$

The probability distribution function of a canonical ensemble is proportional to the product of $W_B(E,\beta)$ and the density of states $\Omega(E)$:

$$P(E,\beta) \propto \Omega(E) W_B(E,\beta)$$

Here, $\Omega(E)$ is a monotonically increasing function of the energy E. Since $W_B(E,\beta)$ is a monotonically decreasing function of E, then $P(E,\beta)$ has a Gaussian shape distribution with a maximum around average energy E for a fixed inverse temperature β. In a typical MD simulation, due to sampling problems, accurate calculation of $\Omega(E)$ is not possible, especially, at low temperatures and complex systems, which can be trapped at some local minimum energy state.

Here, we will discuss how these weights are chosen for those methods which are most often used in molecular dynamics simulations.

3.1. Multicanonical sampling method

The main aim of the multicanonical ensemble (the so-called *MUCA*) is to multiply the states with a non-Boltzmann multicanonical factor, $W_{\mathrm{mu}}(E)$, which yields a uniform probability energy distribution, $P_{\mathrm{mu}}(E)$ [21, 22]:

$$P_{\mathrm{mu}}(E) \propto \Omega(E) W_{\mathrm{mu}}(E) \equiv \text{constant} \tag{22}$$

Since probability is uniform (i.e., flat), the multicanonical ensemble achieves free random walks is in the potential energy space. In this way the system is able to escape faster any local energy minimum state, hence enhancing the configuration phase space sampling in an MD simulation. From Eq. 22, we can calculate the non-Boltzmann weight as

$$W_{\mathrm{mu}}(E) \equiv \exp(-\beta E_{\mathrm{mu}}(E,\beta_0)) \propto \frac{1}{\Omega(E)} \tag{23}$$

where $E_{\mathrm{mu}}(E,\beta_0)$ is the multicanonical potential energy function given by

$$E_{\mathrm{mu}}(E,\beta_0) = k_B T_0 \ln \Omega(E) = \frac{1}{k_B \beta_0} S(E) \tag{24}$$

where $S(E) = k_B \ln \Omega(E)$ is the entropy function of the microcanonical ensemble and β_0 is the multicanonical inverse temperature.

The density of states is practically unknown *a priory*, therefore, the non-Boltzmann's weights $W_{\rm mu}(E)$ are determined, in general, using short MD simulation runs [21,22], and this is one of the limitations of *standard* multicanonical ensemble approach, which can be overcome by combining MUCA with other methods as discussed in the following sections.

The implementation of the MUCA in MD simulation is conveniently introduced by modifying equations of motion with new forces, $\tilde{F}_i$, acting on particles as [17,78,131]:

$$\tilde{F}_i = -\frac{\partial E_{\rm mu}(E,\beta_0)}{\partial q_i} \tag{25}$$

$$= \frac{\partial E_{\rm mu}(E,\beta_0)}{\partial E} F_i, \quad i = 1,2,\cdots,f \tag{26}$$

where F_i is the Newton force acting on the i degree of freedom. Eqs. 7 describing the dynamics of a system in the multicanonical ensemble are re-written as

$$\begin{aligned}
\dot{q}_i &= \frac{p_i}{m_i} \\
\dot{p}_i &= \frac{\beta(E)}{\beta_0} F_i - \lambda_i p_i - P_1^{(s)} p_i \\
\dot{\eta}^{(i)} &= -\sum_{j=1}^{L} \left[\frac{Q_1^{(i)} (\xi_{1,j}^{(i)})^2}{k_B T} \xi_{2,j}^{(i)} - \sum_{k=2}^{M} \xi_{k,j}^{(i)} \right] \\
\dot{\xi}_{1,j}^{(i)} &= -\xi_{1,j}^{(i)} \xi_{2,j}^{(i)} - \lambda_b^{(i)} \xi_{1,j}^{(i)} - \lambda_i \xi_{1,j}^{(i)} \\
& \quad j = 1,\cdots,L \\
\dot{\xi}_{k,j}^{(i)} &= \frac{G_{k,j}^{(i)}}{Q_k^{(i)}} - \xi_{k+1,j}^{(i)} \xi_{k,j}^{(i)} \\
& \quad j = 1,\cdots,L;\ k = 2,\cdots,M-1 \\
\dot{\xi}_{M,j}^{(i)} &= \frac{G_{M,j}^{(i)}}{Q_{M,j}^{(i)}} \\
& \quad j = 1,\cdots,L
\end{aligned} \tag{27}$$

$$\dot{s}_{i,k} = P_{i,k}^{(s)}, \quad k = 1, 2, \cdots, M$$
$$\dot{P}_{i,k}^{(s)} = \frac{\Gamma_{i,k}}{W_{i,k}} - P_{i,k+1}^{(s)} P_{i,k}^{(s)}, \quad k = 1, 2, \cdots, M-1$$
$$\dot{P}_{i,M}^{(s)} = \frac{\Gamma_{i,M}}{W_{i,M}},$$

where β defines the simulation inverse temperature, such as

$$\beta(E_0) = \frac{1}{k_B} \left(\frac{\partial S(E)}{\partial E} \right)_{E_0}$$

The multicanonical weighting factor is usually determined by short trial MD simulation runs at high temperature T_0 using a canonical ensemble [20, 136], as described by Eqs. 7. From these trial runs, we can then determine

$$\begin{cases} E_{mu}^{(1)}(E, \beta_0) = E \\ W_{mu}^{(1)}(E, \beta_0) = W_B(E, \beta_0) = \exp(-\beta_0 E) \end{cases}$$

A maximum value of energy E_{max} is determined as an average of potential energy function at temperature T_0:

$$E_{max} = \langle E \rangle_{T_0}$$

Then, for $E \leq E_{max}$, a flat energy distribution is achieved, and for $E > E_{max}$, we obtain the canonical ensemble distribution at T_0. At every MD time step, t, the probability distribution weighting factor is given by:

$$W^{(t)}(E, \beta_0) = \exp\left(-\beta_0 E^{(t)}(E, \beta_0) \right)$$

Then, a histogram $N^{(t)}(E)$ is accumulated for distribution $P_{mu}^{(t)}(E)$ of potential energy. Denoting by $E_{min}^{(t)}$ the minimum energy value obtained until the t time step. For the $(t+1)$ time step, the multicanonical potential energy is obtained as

$$E_{mu}^{(t+1)}(E, \beta_0) = \begin{cases} E, \quad E \geq E_{max} \\ E_{mu}^{(t)}(E, \beta_0) + \dfrac{1}{\beta_0} \ln\left(N^{(t)}(E) \right) - c^{(t)}, \\ \quad E_{min}^{(t)} \leq E < E_{max} \\ \dfrac{\beta^{(t+1)}(E_{min}^{(t)})}{\beta_0} \left(E - E_{min}^{(t)} \right) \\ + E_{mu}^{(t+1)}(E_{min}^{(t)}, \beta_0), \quad E < E_{min}^{(t)} \end{cases} \tag{28}$$

where $c^{(t)}$ are used to ensure the continuity of energy function at $E = E_{max}$, determined as

$$c^{(t)} = \frac{1}{\beta_0} \ln \left(N^{(t)}(E_{max}) \right)$$

The MD simulation continues until a reasonably flat potential energy function is obtained, which is determined by comparing the values of energy for all $E < E_{max}$ and requiring to be of the same order of magnitude. After this convergence is reached, $E_{min}^{(t)}$ should be equal to the global minimum potential energy function value. Note that during MD simulation, a polynomial or sometimes a cubic spline function [210] is used to fit the histograms each MD simulation time step [131].

Long MD simulation in a multicanonical ensemble is performed, after the optimal weighting factor is obtained. Then, the ensemble average of any physical quantity, $\mathcal{A}$, is determined using the Weighted Histogram Analysis Method (WHAM) [59], which is described in details, for the general case, in Section 3.10..

3.2. The Wang-Landau multicanonical method

In the Wang-Landau method (WLM), a random walk in the energy space with probability proportional to the density of states $\Omega(E)$:

$$P(E) \propto \frac{1}{\Omega(E)} \tag{29}$$

generates a flat energy distribution [196]. To achieve this the estimated density of states is modified systematically until a flat distribution is produced in the energy space. In this procedure, simultaneously, the density of states converges to the true value, by controlling a so-called modification factor at each iteration step. At the end of the simulation, this modification factor becomes very close to one, representing a random walk with the true density of states [196].

Initially, the density of states, $\Omega(E)$, is unknown, therefore, it is set to one:

$$\Omega(E) = 1, \quad \forall E$$

Then, a sampling of energy space is performed with a probability given by Eq. 29. In general, if E_1 and E_2 are two energy states, then the transition probability from state E_1 to E_2 is

$$P(E_1 \rightarrow E_2) = \min\left[\frac{\Omega(E_1)}{\Omega(E_2)}, 1\right] \tag{30}$$

Every time an energy state E is visited, we multiply existing density of states by the factor $\gamma > 1$ [196]

$$\Omega(E) \rightarrow \Omega(E)\gamma \tag{31}$$

or in algorithmic scale

$$\ln(\Omega(E)) \rightarrow \ln(\Omega(E)) + \ln(\gamma) \tag{32}$$

If the move is rejected, then E remains unchanged, and we modify the current $\Omega(E)$ with the same factor γ. In the first publication [196], the suggested initial value of factor γ is

$$\gamma = \gamma_0 = e^1 = 2.71828$$

which allows faster convergence of $\Omega(E)$ to the true density of states even for a very large system. On the other hand, as discussed in Ref. [196], if γ_0 is too small, then the convergence is extremely slow. However, values of γ_0 being too large will produce high statistical errors.

During the simulation a histogram $H(E)$ is accumulated, representing counts for every visited energy bin, E, by the system. After histogram becomes flat in the sampled energy range, we say $\Omega(E)$ has converged to the true value with an accuracy proportional to the factor $\ln(\gamma)$. Then, the modification factor is decreased according to [196]

$$\gamma_1 = \sqrt{\gamma_0}$$

At this moment, the histogram is reset, and the random walk sampling restarts. Now, the density of states is multiplied by a smaller value of factor γ_1 at each step. The algorithm continues in this way, and each time that the histogram becomes flat, the modification factor is decreased as [196]

$$\gamma_{i+1} = \sqrt{\gamma_i}$$

The algorithm stops when

$$\gamma_{\text{final}} = \exp\left(10^{-8}\right) \approx 1.00000001$$

It can be seen that the accuracy is controlled by estimating the density of states and the length of a simulation by a factor γ. In addition to γ_{final}, the accuracy of estimating $\Omega(E)$ depends on the complexity and size of the system, criterion of the flat histogram, and algorithm's implementation [196].

In MD simulations WLM is implemented by modifying equations of motion for the multicanonical method, as in Refs. [78,131]. Here, we purpose to modify equations of motion, given by Eqs. 27, as the following

$$\begin{aligned}
\dot{q}_i &= \frac{p_i}{m_i} \qquad (33)\\
\dot{p}_i &= \frac{\beta}{\beta_0} F_i - \lambda_i p_i - P_1^{(s)} p_i \\
\dot{\eta}^{(i)} &= -\sum_{j=1}^{L} \left[\frac{Q_1^{(i)} (\xi_{1,j}^{(i)})^2}{k_B T} \xi_{2,j}^{(i)} - \sum_{k=2}^{M} \xi_{k,j}^{(i)} \right] \\
\dot{\xi}_{1,j}^{(i)} &= -\xi_{1,j}^{(i)} \xi_{2,j}^{(i)} - \lambda_b^{(i)} \xi_{1,j}^{(i)} - \lambda_i \xi_{1,j}^{(i)} \\
& \quad j = 1, \cdots, L \qquad (34)\\
\dot{\xi}_{k,j}^{(i)} &= \frac{G_{k,j}^{(i)}}{Q_k^{(i)}} - \xi_{k+1,j}^{(i)} \xi_{k,j}^{(i)} \\
& \quad j = 1, \cdots, L;\ k = 2, \cdots, M-1 \\
\dot{\xi}_{M,j}^{(i)} &= \frac{G_{M,j}^{(i)}}{Q_{M,j}^{(i)}} \\
& \quad j = 1, \cdots, L \\
\dot{s}_{i,k} &= P_{i,k}^{(s)}, \quad k = 1, 2, \cdots, M \\
\dot{P}_{i,k}^{(s)} &= \frac{\Gamma_{i,k}}{W_{i,k}} - P_{i,k+1}^{(s)} P_{i,k}^{(s)}, \quad k = 1, 2, \cdots, M-1 \\
\dot{P}_{i,M}^{(s)} &= \frac{\Gamma_{i,M}}{W_{i,M}},
\end{aligned}$$

where

$$\beta_0 = \frac{d \ln \Omega(E)}{dE} \tag{35}$$

with β and β_0 being, respectively, the inverse simulation and multicanonical temperatures. Note that the Wang-Landau application in MD simulation consists in calculation of density of states $\Omega(E)$ from Eq. 31 or Eq. 32, then using Eq. 35 to run MD simulation [179]. In order to calculate accurately $\Omega(E)$ and hence β_0, a histogram bin of energy distribution is estimated, and the bin width will determine the accuracy of calculation of $\Omega(E)$ and β_0 since it defines the ruggedness of energy distribution [179]. To smooth the ruggedness, the energy distribution is approximated by a Gaussian distribution, and then, WHAM can be used to estimate $\Omega(E)$ and multicanonical inverse temperature.

The method has found application to Ising spin lattice systems [113, 196]. It has been used to study the conformation transitions of proteins using confined lattice models [139], protein folding [203], and optimizing temperature distribution in replica exchange method [87].

3.3. Tsallis statistics molecular dynamics method

In the Tsallis statistics molecular dynamics (TSMD) approach [184], the principle of maximum generalized entropy is employed to obtain the generalized statistical mechanic's formalism. The probability weights can be determined as [184]

$$W_T(E,\beta) = [1 + (q-1)\beta(E - E_0)]^{-\frac{q}{q-1}}$$

q is an adjustable parameter taking real values and E_0 is the system's ground energy. Note that $W_T(E,\beta) > 0$. Besides, for $q \to 1$, the Boltzmann's weight can be obtained, and for $q > 1$, probability distribution has longer tails. The long tails of the Tsallis distribution have inspired construction of generalized distributions which will enhance the excursion towards regions with higher energy by decreasing the magnitude of the force close to these regions. This increases the rate of barrier crossing and hence allows the system escaping the local minimum energy states [3, 7, 86, 98, 177].

The aim of Tsallis statistical ensemble is to weight each state by a weighting factor, $W_{\mathrm{T}}(E,\beta)$ [177]:

$$P_{\mathrm{T}}(E,\beta) \propto \Omega(E) W_{\mathrm{T}}(E,\beta) \tag{36}$$

The implementation of Tsallis statistics in MD simulations is obtained by defining the Tsallis weights as the following [177]

$$W_T(E,\beta) = \exp\left(-\beta U_{\mathrm{eff}}\right)$$

where U_{eff} is an effective potential defined as

$$U_{\mathrm{eff}}(E,\beta) = \frac{q}{\beta(q-1)} \ln\left(1+\beta(q-1)(E-E_0)\right) \tag{37}$$

In the new generalized ensemble, MD simulations use the new potential function U_{eff}, which replaces the old one E. The new forces that drive Newton's equations of motion are written as [177]

$$\begin{aligned}
\tilde{\mathbf{F}}_i &= -\frac{\partial U_{\mathrm{eff}}(E,\beta)}{\partial \mathbf{q}_i} \\
&= \frac{\partial U_{\mathrm{eff}}(E,\beta)}{\partial E}\mathbf{F}_i \\
&= \frac{1}{1+\beta(q-1)(E-E_0)}\mathbf{F}_i
\end{aligned}$$

$\mathbf{F}_i$ is the Newton force on particle i ($i = 1,2,\cdots,N$). Then, the equations of motion describing a generalized canonical ensemble according to Tsallis statistics can be given as the following:

$$\begin{aligned}
\dot{q}_i &= \frac{p_i}{m_i} \\
\dot{p}_i &= \frac{1}{1+\beta(q-1)(E-E_0)}F_i - \lambda_i p_i - P_1^{(s)} p_i \\
\dot{\eta}^{(i)} &= -\sum_{j=1}^{L}\left[\frac{Q_1^{(i)}(\xi_{1,j}^{(i)})^2}{k_B T}\xi_{2,j}^{(i)} - \sum_{k=2}^{M}\xi_{k,j}^{(i)}\right] \\
\dot{\xi}_{1,j}^{(i)} &= -\xi_{1,j}^{(i)}\xi_{2,j}^{(i)} - \lambda_b^{(i)}\xi_{1,j}^{(i)} - \lambda_i \xi_{1,j}^{(i)}
\end{aligned} \tag{38}$$

$$
\begin{aligned}
& j = 1, \cdots, L \\
\dot{\xi}_{k,j}^{(i)} &= \frac{G_{k,j}^{(i)}}{Q_k^{(i)}} - \xi_{k+1,j}^{(i)} \xi_{k,j}^{(i)} \\
& j = 1, \cdots, L;\ k = 2, \cdots, M-1 \\
\dot{\xi}_{M,j}^{(i)} &= \frac{G_{M,j}^{(i)}}{Q_{M,j}^{(i)}} \\
& j = 1, \cdots, L \\
\dot{s}_{i,k} &= P_{i,k}^{(s)}, \quad k = 1, 2, \cdots, M \\
\dot{P}_{i,k}^{(s)} &= \frac{\Gamma_{i,k}}{W_{i,k}} - P_{i,k+1}^{(s)} P_{i,k}^{(s)}, \quad k = 1, 2, \cdots, M-1 \\
\dot{P}_{i,M}^{(s)} &= \frac{\Gamma_{i,M}}{W_{i,M}},
\end{aligned}
$$

TSMD has successfully been employed to different molecular systems, such as simulation of atomic clusters [6, 7], protein folding [95, 162, 191, 207], and molecular docking [208]. The approach has also been implemented with replica exchange method by replacing Boltzmann's weights with Tsallis weighting factors for each replica [98, 163, 190].

3.4. Swarm particle-like molecular dynamics method

As we mentioned above, Eqs. 7 can be used to describe the Nosé-Hoover dynamics [89, 168] of a system of N atoms coupled to a chain of thermostats [64]. Recently [84] a new approach was introduced based on the swarm particle social intelligence, which is tested to improve the conformational sampling [84, 85]. In this approach, in addition to the Newtonian forces, a random force is exerted on each particle [84]. This is similar to Langevin dynamics [160]. In particular, the MD equations of motion given by Eqs. 7 can be modified following Ref. [84] as:

$$
\begin{aligned}
\dot{q}_i &= \frac{p_i}{m_i} \\
\dot{q}_i^{\mathrm{Lbest}} &= \frac{p_i^{\mathrm{Lbest}}}{m_i} \delta \left(U(\mathbf{q}) < U\left(\mathbf{q}^{\mathrm{Lbest}}\right) \right),
\end{aligned}
\tag{39}
$$

$$
\begin{aligned}
\dot{q}_i^{\text{Gbest}} &= \frac{p_i^{\text{Gbest}}}{m_i}\delta\left(U(\mathbf{q}) < U\left(\mathbf{q}^{\text{Gbest}}\right)\right), \\
\dot{p}_i &= F_i - \lambda_i p_i - P_1^{(s)} p_i \\
&\quad + \sum_{j=1}^{m} P_{ij}\left(\gamma_1 u_1 (c_j^{\text{Lbest}} - c_j) + \gamma_2 u_2 (c_j^{\text{Gbest}} - c_j)\right) \\
\dot{p}_i^{\text{Lbest}} &= -\gamma_1 u_1 (q_i^{\text{Lbest}} - q_i) \\
\dot{p}_i^{\text{Gbest}} &= -\gamma_2 u_2 (q_i^{\text{Gbest}} - q_i) \\
\dot{\eta}^{(i)} &= -\sum_{j=1}^{L}\left[\frac{Q_1^{(i)}(\xi_{1,j}^{(i)})^2}{k_B T}\xi_{2,j}^{(i)} - \sum_{k=2}^{M}\xi_{k,j}^{(i)}\right] \\
\dot{\xi}_{1,j}^{(i)} &= -\xi_{1,j}^{(i)}\xi_{2,j}^{(i)} - \lambda_b^{(i)}\xi_{1,j}^{(i)} - \lambda_i \xi_{1,j}^{(i)} \\
&\quad j = 1, \cdots, L \\
\dot{\xi}_{k,j}^{(i)} &= \frac{G_{k,j}^{(i)}}{Q_k^{(i)}} - \xi_{k+1,j}^{(i)}\xi_{k,j}^{(i)} \\
&\quad j = 1, \cdots, L;\ k = 2, \cdots, M-1 \\
\dot{\xi}_{M,j}^{(i)} &= \frac{G_{M,j}^{(i)}}{Q_{M,j}^{(i)}} \\
&\quad j = 1, \cdots, L \\
\dot{s}_{i,k} &= P_{i,k}^{(s)}, \quad k = 1, 2, \cdots, M \\
\dot{P}_{i,k}^{(s)} &= \frac{\Gamma_{i,k}}{W_{i,k}} - P_{i,k+1}^{(s)} P_{i,k}^{(s)}, \quad k = 1, 2, \cdots, M-1 \\
\dot{P}_{i,M}^{(s)} &= \frac{\Gamma_{i,M}}{W_{i,M}},
\end{aligned}
\tag{40}
$$

where the vector $\mathbf{c} = (c_1, c_2, \cdots, c_m)^T$ characterizes the essential degrees of freedom in the system. Projection operator, $\mathbf{P}$, transforms the real coordinates $\mathbf{q}$ to the so-called collective coordinates $\mathbf{c}$ according to:

$$
c_j = \sum_{i=1}^{f} P_{ij} q_i
$$

In Eq. 39, $\{c_j^{\text{Lbest}}\}_{j=1}^m$ and $\{c_j^{\text{Gbest}}\}_{j=1}^m$ are defined as [85]:

$$c_j^{\text{Lbest}} = \sum_{i=1}^{f} P_{ij} q_i^{\text{Lbest}}$$
$$c_j^{\text{Gbest}} = \sum_{i=1}^{f} P_{ij} q_i^{\text{Gbest}}$$

which are updated every time step. Here, $\mathbf{q}^{\text{Lbest}}$ is configuration vector with the lowest value of the potential energy of the system and $\mathbf{q}^{\text{Gbest}}$ is configuration vector of the final state of the system.

In Eq. 39 u_i $(i = 1,2)$ denotes a uniformly distributed random number in $(0,1)$, and γ_1 and γ_2 are adjustable parameters.

In Eq. 39, the δ function is given as:

$$\delta\left(U(\mathbf{q}) < U\left(\mathbf{q}^{\text{Lbest}}\right)\right) = \begin{cases} 1, & \text{if } U(\mathbf{q}) < U\left(\mathbf{q}^{\text{Lbest}}\right) \\ 0, & \text{otherwise} \end{cases}$$

and

$$\delta\left(U(\mathbf{q}) < U\left(\mathbf{q}^{\text{Gbest}}\right)\right) = \begin{cases} 1, & \text{if } U(\mathbf{q}) < U\left(\mathbf{q}^{\text{Gbest}}\right) \\ 0, & \text{otherwise} \end{cases}$$

The augmented dynamical system, which is given by Eq. (39) , sample an equilibrium canonical distribution with conserved total energy given by:

$$\begin{aligned} E_{\text{ext}} = & \sum_{i=1}^{f} \frac{p_i^2}{2m_i} \\ & + \left(\sum_{i=1}^{f} \frac{\left(p_i^{\text{Lbest}}\right)^2}{2m_i} \right) \delta\left(U(\mathbf{q}) < U\left(\mathbf{q}^{\text{Lbest}}\right)\right) \\ & + \underbrace{\left(\sum_{i=1}^{f} \frac{\left(p_i^{\text{Gbest}}\right)^2}{2m_i} \right) \delta\left(U(\mathbf{q}) < U\left(\mathbf{q}^{\text{Gbest}}\right)\right)}_{E_{\text{tot,kin}}} \\ & + \underbrace{U(\mathbf{q}) + \frac{1}{2} \sum_{j=1}^{f} \left[u_1 \gamma_1 \left(q_j^{\text{Lbest}} - q_j\right)^2 + u_2 \gamma_2 \left(q_j^{\text{Gbest}} - q_j\right)^2 \right]}_{U_{\text{bias}}} \end{aligned} \tag{41}$$

$$+\underbrace{\sum_{i=1}^{f}\sum_{k=1}^{M}\left(\frac{Q_{i,k}\xi_{i,k}^2}{2}+k_BTs_{i,k}\right)}_{E_{\text{thermo}}}$$

where f is the total number of degrees of freedom of the system ($f = 3N$). These equations represent an extended phase space of the augmented dynamical system with real variables:

$$\left((q_i,p_i),\left(q_i^{\text{Lbest}},p_i^{\text{Lbest}}\right),\left(q_i^{\text{Gbest}},p_i^{\text{Gbest}}\right)\right),\quad i=1,2,\cdots,f$$

and thermostats variables:

$$(s_{i,k},\xi_{i,k}),\quad i=1,2,\cdots,f;\ k=1,2,\cdots,M$$

In Eq. 41, $E_{\text{tot,kin}}$ is the total kinetic energy of augmented system, U_{bias} is the total potential energy including bias term, and E_{thermo} is the thermostat energy. WHAM is used to recover the equilibrium canonical distribution of the real system [84, 85]. The augmented dynamical system is shown to sample metastable [85] and rare transition events [92], and to enhance the conformation sampling [7, 98]. In Eq. (39), the first bias term steers the system towards the state with the lowest energy, which has been visited at any instant time t and hence enhancing the local basin sampling. Besides, the second bias term indicates the "information" about configuration with the lowest energy ever visited, and hence enhancing the barrier crossing rate.

3.5. Replica exchange method

Another class of methods that use the generalized distributions for sampling the conformation phase space is also the so-called temperature *Replica Exchange Method* (REM) [46, 51, 132, 177, 197]. REM is often used to solve the problems of quasi-ergodicity in simulations of (bio)molecular systems. In REM, replicas representing the system are simulated independently at different temperatures [197]. In particular, consider a system of N atoms each with a mass m_i, position vector $\mathbf{r}_i = (x_i, y_i, z_i)$, and conjugated momentum $\mathbf{p}_i = (p_{xi}, p_{yi}, p_{zi})$. In standard REM the generalized ensemble corresponds to $\mathcal{L}$ independent replications of the original system coupled to $\mathcal{L}$ thermostats at different temperatures. Using Nosé-Hoover dynamics [89, 168], each replica is in equilibrium with a

chain of thermostats [64] and the equations of motion are given here as the following for each replica α:

$$
\begin{aligned}
\dot{q}_{i,\alpha} &= \frac{p_{i,\alpha}}{m_i} \\
\dot{p}_{i,\alpha} &= F_{i,\alpha} - \lambda_{i,\alpha} p_{i,\alpha} - P_1^{(\alpha,s)} p_{i,\alpha} \\
\dot{\eta}^{(i,\alpha)} &= -\sum_{j=1}^{L} \left[\frac{Q_1^{(i,\alpha)} (\xi_{1,j}^{(i,\alpha)})^2}{k_B T_\alpha} \xi_{2,j}^{(i,\alpha)} - \sum_{k=2}^{M} \xi_{k,j}^{(i,\alpha)} \right] \\
\dot{\xi}_{1,j}^{(i,\alpha)} &= -\xi_{1,j}^{(i,\alpha)} \xi_{2,j}^{(i,\alpha)} - \lambda_b^{(i,\alpha)} \xi_{1,j}^{(i,\alpha)} - \lambda_{i,\alpha} \xi_{1,j}^{(i,\alpha)} \\
&\qquad j = 1, \cdots, L \\
\dot{\xi}_{k,j}^{(i,\alpha)} &= \frac{G_{k,j}^{(i,\alpha)}}{Q_k^{(i,\alpha)}} - \xi_{k+1,j}^{(i,\alpha)} \xi_{k,j}^{(i,\alpha)} \\
&\qquad j = 1, \cdots, L;\ k = 2, \cdots, M-1 \\
\dot{\xi}_{M,j}^{(i,\alpha)} &= \frac{G_{M,j}^{(i,\alpha)}}{Q_{M,j}^{(i,\alpha)}} \\
&\qquad j = 1, \cdots, L \\
\dot{s}_{i,k}^{(\alpha)} &= P_{i,k}^{(\alpha,s)}, \quad k = 1, 2, \cdots, M \\
\dot{P}_{i,k}^{(\alpha,s)} &= \frac{\Gamma_{i,k}^{(\alpha)}}{W_{i,k}^{(\alpha)}} - P_{i,k+1}^{(\alpha,s)} P_{i,k}^{(\alpha,s)}, \quad k = 1, 2, \cdots, M-1 \\
\dot{P}_{i,M}^{(\alpha,s)} &= \frac{\Gamma_{i,M}^{(\alpha)}}{W_{i,M}^{(\alpha)}}
\end{aligned}
\tag{42}
$$

Two neighboring thermostats (*e.g.*, i and j) swap at regular interval of times their configurations (replicas) with probability, P_{acc}, which preserves the detailed balance [177, 197]:

$$
P_{\text{acc}} = \min\left\{1, \exp\left(-(\beta_j - \beta_i)(E_i - E_j)\right)\right\} \tag{43}
$$

where E_i and E_j correspond to the total energies of replicas i and j, respectively. In REM, high-temperature replicas are able to cross more often energy barrier.

On the other hand, low-temperature replicas sample more often potential energy basins. It has been suggested [82] that the number of replicas scales as the square root of system's degrees of freedom. Note that increasing the number of replicas requires longer simulation runtime, which is necessary to optimize the rate of round trips between the two extreme temperatures.

Omitting the solvent degrees of freedom through the use of implicit or hybrid explicit/implicit solvent models [2, 10, 37, 156, 157] have increased the efficiency of REM. Another approach includes the use of separate heat baths for the solute and solvent [204]. It has been argued [2, 148, 156, 157] that reductions in system size may not accurately describe the structure and dynamics of the system.

Other approaches, similar to REM, are also proposed. For instance, the *Hamiltonian Replica Exchange Method* (HREM) [82]. An HREM with biasing the backbone dihedral potentials yielded a reduction in the number of replicas [108]. In resolution HREM approach, which uses implicit solvent models only [47, 149], in addition to different temperature couplings, the replicas exchange a subset of configuration coordinates from a coarse-grained model [47].

Use of temperature scaling for the solvent-solvent and solvent-protein interactions in REM has also shown to reduce the number of replicas [148], which has further been improved by using the Tsallis biasing potential [86].

Efforts have also been made to optimize the distribution of temperatures among the replicas as in Refs. [4, 23, 33, 35, 49, 68, 83, 116, 130, 134, 150, 158, 175–177]. Some of these methods [49, 83, 130, 175, 176] have particularly been important in studying the protein folding/unfolding transitions, which represent a difficult case study in standard REM because of the low rate of accepted swaps between replicas across the transition temperature [90, 176].

To further increase the efficiency of REM, other approaches to REM have also been proposed [87], which aims to obtain a flat generalized probability distribution function in temperature space using the Wang-Landau algorithm [57, 58]. The method addresses two problems of REM: it increases the probability of swapping, and it decreases the bottleneck for exchange at the transition temperature.

Note that a WHAM is used for analyzing the data from all replicas (see Section 3.10.).

3.6. Swarm particle-like replica exchange method

More recently [84], a combination of replica exchange method with Swarm Particle-like Molecular Dynamics (SPMD) is introduced. SPMD showed to improve conformation sampling when applied to Lennard-Jones atomic cluster systems [84] and protein folding problems [85] when combined with replica exchange approach. Here, the equations of motion given in Ref. [85] are modified as the following:

$$\dot{q}_{i,\alpha} = \frac{p_{i,\alpha}}{m_i} \tag{44}$$

$$\dot{q}_{i,\alpha}^{\text{Lbest}} = \frac{p_{i,\alpha}^{\text{Lbest}}}{m_i}\delta\left(U(\mathbf{q}_\alpha) < U\left(\mathbf{q}_\alpha^{\text{Lbest}}\right)\right),$$

$$\dot{q}_{i,\alpha}^{\text{Gbest}} = \frac{p_{i,\alpha}^{\text{Gbest}}}{m_i}\delta\left(U(\mathbf{q}_\alpha) < U\left(\mathbf{q}_\alpha^{\text{Gbest}}\right)\right),$$

$$\begin{aligned}\dot{p}_{i,\alpha} &= F_{i,\alpha} - \lambda_{i,\alpha}p_{i,\alpha} - P_1^{(\alpha,s)}p_{i,\alpha} \\ &\quad + \sum_{j=1}^{m} P_{ij}\Big(\gamma_1 u_1(c_j^{\alpha,\text{Lbest}} - c_j^{(\alpha)}) \\ &\quad + \gamma_2 u_2(c_j^{\text{Gbest}} - c_j^{(\alpha)})\Big)\end{aligned} \tag{45}$$

$$\dot{p}_{i,\alpha}^{\text{Lbest}} = -\gamma_1 u_1(q_{i,\alpha}^{\text{Lbest}} - q_{i,\alpha})$$

$$\dot{p}_{i,\alpha}^{\text{Gbest}} = -\gamma_2 u_2(q_{i,\alpha}^{\text{Gbest}} - q_{i,\alpha})$$

$$\dot{\eta}^{(i,\alpha)} = -\sum_{j=1}^{L}\left[\frac{Q_1^{(i,\alpha)}(\xi_{1,j}^{(i,\alpha)})^2}{k_B T_\alpha}\xi_{2,j}^{(i,\alpha)} - \sum_{k=2}^{M}\xi_{k,j}^{(i,\alpha)}\right]$$

$$\dot{\xi}_{1,j}^{(i,\alpha)} = -\xi_{1,j}^{(i,\alpha)}\xi_{2,j}^{(i,\alpha)} - \lambda_b^{(i,\alpha)}\xi_{1,j}^{(i,\alpha)} - \lambda_{i,\alpha}\xi_{1,j}^{(i,\alpha)}$$
$$j = 1,\cdots,L$$

$$\dot{\xi}_{k,j}^{(i,\alpha)} = \frac{G_{k,j}^{(i,\alpha)}}{Q_k^{(i,\alpha)}} - \xi_{k+1,j}^{(i,\alpha)}\xi_{k,j}^{(i,\alpha)}$$
$$j = 1,\cdots,L;\ k = 2,\cdots,M-1$$

$$\dot{\xi}_{M,j}^{(i,\alpha)} = \frac{G_{M,j}^{(i,\alpha)}}{Q_{M,j}^{(i,\alpha)}}$$

$$
\begin{aligned}
& j = 1, \cdots, L \\
\dot{s}_{i,k}^{(\alpha)} &= P_{i,k}^{(\alpha,s)}, \quad k = 1, 2, \cdots, M \\
\dot{P}_{i,k}^{(\alpha,s)} &= \frac{\Gamma_{i,k}^{(\alpha)}}{W_{i,k}^{(\alpha)}} - P_{i,k+1}^{(\alpha,s)} P_{i,k}^{(\alpha,s)}, \quad k = 1, 2, \cdots, M-1 \\
\dot{P}_{i,M}^{(\alpha,s)} &= \frac{\Gamma_{i,M}^{(\alpha)}}{W_{i,M}^{(\alpha)}},
\end{aligned}
$$

where all the variables have the same meaning as in Eq. 39 for the replica α and $\{c_j^{\text{Gbest}}\}_{j=1}^{m}$ is related to the *global best* coordinates $\mathbf{q}^{\text{Gbest}}$ corresponding to configuration with the lowest energy among all replicas through the projection operator $\mathbf{P}$:

$$
c_j^{\text{Gbest}} = \sum_{i=1}^{f} P_{ij} q_i^{\text{Gbest}}
$$

It has been shown elsewhere [85] that the Eqs. 44 preserve the detailed balance condition. Following Ref. [85], assuming that a Markovian chain of states is formed, the probability of obtaining a trajectory in the configuration space of the replica k can be written as:

$$
P_k(\mathbf{X}_T^k) = \exp\left(-\beta_k E(\mathbf{x}_{k,0})\right) \prod_{t=0}^{T-1} \pi(\mathbf{x}_{k,t} \rightarrow \mathbf{x}_{k,t+1}) \tag{46}
$$

with β_k being the inverse temperature of the thermostat k. In Eq. 46 $E(\mathbf{x}_{k,t})$ is the total energy obtained for the configuration $\mathbf{x}_{k,t}$. Here, $\mathbf{X}_T^k$ represent T replicas of the system:

$$
\mathbf{X}_T^k = \{\mathbf{x}_{k,0} \rightarrow \mathbf{x}_{k,1} \rightarrow \cdots \rightarrow \mathbf{x}_{k,T-1}\}
$$

The initial configurations of each replica are obtained from a canonically distributed with an initial unbiased energy of the system for replica k $E(\mathbf{x}_{k,0})$:

$$
\rho_{init}(\mathbf{x}_{k,0}) = \exp\left(-\beta_k E(\mathbf{x}_{k,0})\right)
$$

In Eq. 46, $\pi(\mathbf{x}_{k,t} \rightarrow \mathbf{x}_{k,t+1})$ is the propagation probability at each time step, which depend on the details of deterministic or stochastic dynamics. In general,

the Markovian transition probability $\pi(\mathbf{x}_{k,t} \rightarrow \mathbf{x}_{k,t+1})$ can have any distribution that conserves the Boltzmann distribution. Here, $\pi(\mathbf{x}_{k,t} \rightarrow \mathbf{x}_{k,t+1})$ represents the action characterized by augmented system given in Eq. (44), which produces a Boltzmann distribution in the extended phase space of variables. In the general case of the Newtonian dynamics, we can write:

$$p(\mathbf{x}_t \rightarrow \mathbf{x}_{t+1}) = \delta\left(\mathbf{x}_{t+1} - \Phi_{\Delta t}(\mathbf{x}_t)\right)$$

where δ is the delta function and $\Phi_{\Delta t}(\mathbf{x}_t)$ is the discrete flow map of one time step Δt propagation operator. In this case, a trajectory can be generated using an initial state sampled from some canonical distribution and then propagating in time using usual Hamiltonian dynamics. Note that for Hamiltonian dynamics is easy to find a time-reversible discrete flow map. On the other hand, when dynamics are governed by Eq. (44), the structure is not symplectic, but still, it is time reversible.

It is important to note that WHAM can be used for analyzing the data from all replicas in the case of REM:SPMD simulation as presented in Refs. [84, 85] (see Section 3.10.).

3.7. Replica exchange multicanonical method

To overcome the problems of standard REM (e.g., the large number of replicas and high computational demands) and of MUCA (e.g., difficulties on determining the weighting factor), a new method has been proposed called *replica exchange multicanonical* (REMUCA) [210]. In REMUCA, the first step is to perform a short replica exchange method simulation (e.g., of $\mathcal{L}$ replicas) to calculate the multicanonical weighting factor. Then, a standard multicanonical simulation run is performed using this weighting factor. The multiple-histogram re-weighting technique [52, 109] can be used to calculate the energy density of states as described in the next section (see 3.10.).

After we obtain the energy density of states, the multicanonical weighting factor is obtained from Eq. 23 and Eq. 24. Note that the multicanonical energy $E_{mu}(E, \beta_0)$ obtained in this way is determined in the range between $\langle E \rangle_{T_1}$ and $\langle E \rangle_{T_\mathcal{L}}$, where T_1 and $T_\mathcal{L}$ denote the lowest and highest temperature, respectively, used in the replica exchange simulation. While outside this interval the potential

energy of multicanonical simulation is determined through extrapolation:

$$E_{mu}^{(0)}(E) = \begin{cases} \left(\dfrac{\partial E_{mu}(E,\beta_0)}{\partial E}\right)_{E_1}(E-E_1) \\ \quad +E_{mu}(E_1,\beta_0), \quad E < E_1 \\ E_{mu}(E,\beta_0), \quad E_1 \le E \le E_{\mathcal{L}} \\ \left(\dfrac{\partial E_{mu}(E,\beta_0)}{\partial E}\right)_{E_{\mathcal{L}}}(E-E_{\mathcal{L}}) \\ \quad +E_{mu}(E_{\mathcal{L}},\beta_0), \quad E > E_{\mathcal{L}} \end{cases} \tag{47}$$

In multicanonical MD simulations, Eq. 25 (or Eq. 27) is used to govern Newton's dynamics where $E_{mu}(E,\beta_0)$ is replaced by $E_{mu}^{(0)}(E)$. Then, after the simulation run, the trajectories are analyzed using WHAM for a single run, as described in the next section.

Eq. 47 can also be written as

$$E_{mu}^{(0)}(E) = \begin{cases} \dfrac{T_0}{T_1}(E-E_1)+T_0S(E_1) = \dfrac{T_0}{T_1}E + \text{constant}, \\ \quad E < E_1 = \langle E\rangle_{T_1} \\ T_0S(E), \quad E_1 \le E \le E_{\mathcal{L}} \\ \dfrac{T_0}{T_{\mathcal{L}}}(E-E_{\mathcal{L}})+T_0S(E_{\mathcal{L}}) = \dfrac{T_0}{T_{\mathcal{L}}}E + \text{constant}, \\ \quad E > E_{\mathcal{L}} = \langle E\rangle_{T_{\mathcal{L}}} \end{cases} \tag{48}$$

and the dynamical equations of motion are defined here as:

$$\dot{q}_{i,\alpha} = \frac{p_{i,\alpha}}{m_i} \tag{49}$$

$$\dot{p}_{i,\alpha} = -\lambda_{i,\alpha}p_{i,\alpha} - P_1^{(\alpha,s)}p_{i,\alpha} + \begin{cases} \dfrac{\beta_1}{\beta_0}F_{i,\alpha}, & E < E_1 = \langle E\rangle_{T_1} \\ \dfrac{\beta_\alpha(E)}{\beta_0}F_{i,\alpha}, & E_1 \le E \le E_{\mathcal{L}} \\ \dfrac{\beta_{\mathcal{L}}}{\beta_0}F_{i,\alpha}, & E > E_{\mathcal{L}} = \langle E\rangle_{T_{\mathcal{L}}} \end{cases}$$

$$\dot{\eta}^{(i,\alpha)} = -\sum_{j=1}^{L}\left[\frac{Q_1^{(i,\alpha)}(\xi_{1,j}^{(i,\alpha)})^2}{k_BT_\alpha}\xi_{2,j}^{(i,\alpha)} - \sum_{k=2}^{M}\xi_{k,j}^{(i,\alpha)}\right]$$

$$\dot{\xi}_{1,j}^{(i,\alpha)} = -\xi_{1,j}^{(i,\alpha)}\xi_{2,j}^{(i,\alpha)} - \lambda_b^{(i,\alpha)}\xi_{1,j}^{(i,\alpha)} - \lambda_{i,\alpha}\xi_{1,j}^{(i,\alpha)}$$

$$j = 1, \cdots, L$$

$$\dot{\xi}^{(i,\alpha)}_{k,j} = \frac{G^{(i,\alpha)}_{k,j}}{Q^{(i,\alpha)}_{k}} - \xi^{(i,\alpha)}_{k+1,j}\xi^{(i,\alpha)}_{k,j}$$

$$j = 1, \cdots, L;\ k = 2, \cdots, M-1$$

$$\dot{\xi}^{(i,\alpha)}_{M,j} = \frac{G^{(i,\alpha)}_{M,j}}{Q^{(i,\alpha)}_{M,j}}$$

$$j = 1, \cdots, L$$

$$\dot{s}^{(\alpha)}_{i,k} = P^{(\alpha,s)}_{i,k}, \quad k = 1, 2, \cdots, M$$

$$\dot{P}^{(\alpha,s)}_{i,k} = \frac{\Gamma^{(\alpha)}_{i,k}}{W^{(\alpha)}_{i,k}} - P^{(\alpha,s)}_{i,k+1}P^{(\alpha,s)}_{i,k}, \quad k = 1, 2, \cdots, M-1$$

$$\dot{P}^{(\alpha,s)}_{i,M} = \frac{\Gamma^{(\alpha)}_{i,M}}{W^{(\alpha)}_{i,M}}$$

From Eq. 23, the Boltzmann's factor depends on temperature T and energy E, and hence, scaling potential energy (and so the force) by a constant κ is similar to scaling the temperature by $1/\kappa$ [78, 155]. Therefore, $E^{(0)}_{mu}$ given by Eq. 47 (or Eq. 48) generates a canonical ensemble distribution at $T = T_1$, multicanonical ensemble distribution for $E_1 \leq E \leq E_{\mathcal{L}}$, and canonical ensemble distribution simulation at $T = T_{\mathcal{L}}$ for $E > E_{\mathcal{L}}$.

3.8. Multicanonical replica exchange method

The replica exchange method in multicanonical simulation, REMUCA, can also be introduced as a multicanonical replica exchange method (MUCAREM) [210]. In MUCAREM, the final MD simulation run is a replica exchange with fewer replicas, say $\mathcal{L}$, in contrast to REMUCA, where the final run is a standard multicanonical MD simulation. Since the degree of energy probability distribution overlapping in a multicanonical simulation is higher compared to canonical one, a fewer number of replicas are needed for the final simulation run to guarantee an optimal distribution of replicas among thermostats.

In MUCAREM, similar to REMUCA, short replica exchange MD simula-

tions are performed with L replicas and L thermostats, covering a temperature range from T_1 to T_L. During this short simulation run, we can estimate the energy density of states $\Omega(E)$ for all range of energy using WHAM techniques. After we define the density of states $\Omega(E)$, we can chose $\mathcal{L}$ pairs of thermostats with temperatures $(T_L^{(m)}, T_H^{(m)})$, for $m = 1, 2, \cdots, \mathcal{L}$, where $T_L^{(m)} < T_H^{(m)}$. In practice, the temperatures are arranged such that ensure sufficient overlapping between neighboring pairs. Here, we have $T_L^{(1)} = T_1$ and $T_H^{(\mathcal{L})} = T_{\mathcal{L}}$, and

$$\begin{cases} E_L^{(m)} = \langle E \rangle_{T_L^{(m)}}, \\ E_H^{(m)} = \langle E \rangle_{T_H^{(m)}}, \quad m = 1, 2, \cdots, \mathcal{L} \end{cases} \tag{50}$$

Then, we chose $\mathcal{L}$ thermostats at temperatures $T_1, T_2, \cdots, T_{\mathcal{L}}$ and assign to each the multicanonical potential as

$$E_{mu}^{(m)}(E) = \begin{cases} \left(\dfrac{\partial E_{mu}(E,T_m)}{\partial E}\right)_{E_L^{(m)}} (E - E^{(m)}) \\ \quad + E_{mu}(E_L^{(m)}, T_m), \quad E < E_L^{(m)} \\ E_{mu}(E, T_m), \quad E_L^{(m)} \leq E \leq E_H^{(m)} \\ \left(\dfrac{\partial E_{mu}(E,T_m)}{\partial E}\right)_{E_H^{(m)}} (E - E^{(m)}) \\ \quad + E_{mu}(E_H^{(m)}, T_m), \quad E > E_H^{(m)} \end{cases} \tag{51}$$

where the multicanonical potential energy, $E_{mu}(E,T)$, is obtained for the entire interval of energy. Also, this choice of $E_{mu}^{(m)}(E)$ generates a canonical distribution at $T = T_L^{(m)}$ for $E < E_L^{(m)}$, a multicanonical distribution for $E_L^{(m)} \leq E \leq E_H^{(m)}$, and a canonical distribution simulation run at $T = T_H^{(m)}$ for $E > E_H^{(m)}$.

In final step of MUCAREM, the production run is defined as a replica exchange simulation with $\mathcal{L}$ different thermostats at temperatures $T_1, T_2, \cdots, T_{\mathcal{L}}$ and multicanonical potential energies, $E_{mu}^{(1)}(E), E_{mu}^{(2)}(E), \cdots, E_{mu}^{(\mathcal{L})}(E)$. The transition probability of swapping two replicas of neighboring temperatures is given by

$$w(x_m^{(i)} \mid x_{m+1}^{(j)}) = \begin{cases} 1, & \Delta \leq 0 \\ \exp(-\Delta), & \Delta > 0 \end{cases} \tag{52}$$

where

$$\Delta = \beta_{m+1}\left[E_{mu}^{(m+1)}(E(q^{(i)})) - E_{mu}^{(m+1)}(E(q^{(j)}))\right] - \beta_m\left[E_{mu}^{(m)}(E(q^{(i)})) - E_{mu}^{(m)}(E(q^{(j)}))\right] \tag{53}$$

Here, the multicanonical potential energies, $E^{(m)}(E(q^{(j)}))$ and $E^{(m+1)}(E(q^{(i)}))$, have to be calculated since $E^{(m)}(E)$ has different values for m [209].

Using the multiple-histogram reweighting method, the canonical distribution can be obtained [52, 109], as presented in Section 3.10..

3.9. Tsallis replica exchange methods

Tsallis's weight factor for a configuration $\mathbf{q}$ at inverse temperature β_ℓ has the following general form [98]:

$$W_\ell(\mathbf{q}) = \left[1 - \beta_\ell(1-q_\ell)(U(\mathbf{q}) - U_\ell^{(0)})\right]^{\frac{q_\ell}{1-q_\ell}} \tag{54}$$

where $U_\ell^{(0)}$ is a reference minimum value of the potential energy $U(\mathbf{q})$ at Tsallis entropy parameter q_ℓ of replica ℓ associated with thermostat held at inverse temperature β_ℓ. Note that in the limit when $q_\ell \to 1$, the Boltzmann's weight of standard MD simulation can be obtained:

$$W_B(\mathbf{q}) \sim \exp\left(-\beta_\ell(U - U^{(0)})\right)$$

There may exist two approaches of applying Tsallis-like replica exchange method. In the first implementation (named q-REM) L replicas are initially setup running at different values of q_ℓ for $\ell = 1, 2, \cdots, L$ and equal temperature T as proposed [163]. While in a second implementation, the Tsallis-like dynamics is incorporated with REM in the form of the generalized ensemble (TSREM) [190]. In TSREM implementation, L replicas are associated with L thermostats at different inverse temperatures β_ℓ and Tsallis entropy parameter q_ℓ ($\ell = 1, 2, \cdots, L$). Usually, in both implementations the reference replica samples the phase space using MD simulation with original potential energy function, that is, it has $q_\ell = 1$ and $\beta_\ell = \beta_0$, where β_0 is the required inverse

temperature. On the other hand, the other replicas sample using biased potential energy function associated with effective potential energy that is given by Eq. 37 (i.e., $q_\ell > 1$ for all replicas $\ell > 1$).

In Tsallis-like REM a swapping between two neighboring replicas 1 and 2 has an acceptance probability given by

$$P_{acc}(1 \leftrightarrow 2) = \min\left[1, \exp\left(-\Delta_{12}\right)\right] \tag{55}$$

where

$$\begin{aligned}\Delta_{12} = \beta_2 \left(U_{\text{eff},2}(\mathbf{q}_1; q_2, T_2) - U_{\text{eff},2}(\mathbf{q}_2; q_2, T_2)\right) \\ + \beta_1 \left(U_{\text{eff},1}(\mathbf{q}_2; q_1, T_1) - U_{\text{eff},1}(\mathbf{q}_1; q_1, T_1)\right)\end{aligned} \tag{56}$$

where $U_{\text{eff},1}$ and $U_{\text{eff},2}$ are the effective Tsallis potential energy of replica 1 and 2, respectively.

Close to the barrier regions, the magnitude of the force, because of lowering the barriers, is reduced for q larger than one [7,147]. Therefore, in these regions resistance on the particles decreases and the barrier crossing rates increases. However, the largest value of q has to be carefully determined after some preliminary test runs depending on the system. It is empirically suggested [86,192] an upper value as $q = 1 + 1/f$, where f is the number of degrees of freedom.

A more general approach for optimization of q values has been suggested in Ref. [193]. Based on this approach, expression of Δ_{12} in Eq. 56 is written as

$$\Delta_{12} = \frac{1}{k_B} \int_{\mathbf{q}_2}^{\mathbf{q}_1} \left[\frac{1}{\mathcal{T}_2(z)} - \frac{1}{\mathcal{T}_1(z)}\right] dz \tag{57}$$

where $\mathcal{T}$ is the effective Tsallis temperature given as

$$k_B \mathcal{T}(\mathbf{q}; q, T) = \left(\frac{\partial\left(\beta U_{\text{eff}}\right)}{\partial U}\right)_U^{-1}$$

The performance of either q-REM or TSREM will directly depend on the rate of accepted replica attempted swaps, defined by the average P_{acc} of acceptance probability of each swapping attempt:

$$P_{acc}(1 \leftrightarrow 2) = \int d\mathbf{q}_1 d\mathbf{q}_2 P_1(\mathbf{q}_1) P_2(\mathbf{q}_2) \min\left[\exp\left(-\Delta_{12}\right)\right]$$

where $P_i(\mathbf{q}_i)$ is the probability that replica i has a configuration $\mathbf{q}_i$ or energy $U(\mathbf{q}_i)$, which can be written

$$P_i(\mathbf{q}_i) \equiv P_i(U) = \Omega(U) W_{T,i}(U) \tag{58}$$

where $\Omega(U)$ represents the energy density of states and $W_{T,i}(U)$ Tsallis weighting factor of replica i. Eq. 58 can further be simplified as [193]:

$$\begin{aligned} P_{acc}(1 \leftrightarrow 2) = & \int dU dU' \theta(\Delta_{12}) P_1(U) P_2(U') \\ & + \int dU dU' \theta(-\Delta_{12}) P_1(U') P_2(U) \end{aligned} \tag{59}$$

where $\theta(\Delta)$ denotes the Heaviside step function:

$$\theta(x) = \begin{cases} 1, & x > 0 \\ 0, & \text{otherwise} \end{cases}$$

and $\Delta_i(U,U') = -\Delta_i(U',U)$, from which we can obtain [193]:

$$P_{acc}(1 \leftrightarrow 2) = 2 \int dU dU' P_1(U) P_2(U') \theta(\Delta_{12}) \tag{60}$$

The optimal $P_{acc}(1 \leftrightarrow 2)$ is obtained by maximizing the overlapping integrals between two neighboring Tsallis probability distribution functions. For that, we can approximate the function of $P_i(U)$ by local expansion around the stationary point U_0 as [193]:

$$\ln P_i(U) = \ln P_i(U_0) - \frac{1}{2\sigma_q} (U - U_0)^2 + \cdots \tag{61}$$

where U_0 is the energy where both Tsallis effective temperature and statistical temperature, $T_S(U)$ are equal:

$$T_S(U_0) = \mathcal{T}(U_0)$$

where

$$T_S(U_0) = \left(\frac{\partial S}{\partial U} \right)^{-1}_{U_0}$$

where $S(U)$ is the microcanonical entropy:

$$S(U) = k_B \ln \Omega(U)$$

In Eq. (61) σ_q denotes the width of Tsallis probability density function at Gaussian approximation given by [193]:

$$\sigma_q(U_0) = \left[\frac{T_S'(U)}{T_S^2(U)} - \frac{\mathcal{T}_i'(U)}{\mathcal{T}_i^2(U)} \right]^{-1}_{U_0} \tag{62}$$

where

$$T_S'(U) = \frac{\partial T_S}{\partial U},\ \mathcal{T}_i'(U) = \frac{\partial \mathcal{T}_i}{\partial U}.$$

If $T_S(U)$ is assumed to be linear function of U around U_0, then $T_S'(U)$ is a constant. Furthermore, the equivalence between the microcanonical and canonical ensembles indicates that this constant is $1/C_V(T_0)$, thus

$$T_S'(U_0) = 1/C_V(T_0)$$

Then, we obtain [193]

$$\sigma_q(U_0) = \frac{\sigma_0}{1-\kappa} \tag{63}$$

where

$$\kappa = (q_i - 1)C_V(T_0),\ \sigma_0 = T_0^2 C_V(T_0)$$

where $T_0 = T_S(U_0)$ and σ_0 is the Gaussian width of the canonical probability density function at temperature T_0.

It can be seen from Eq. 63 that in the limit of $q_i \to 1$, which is the limit of Boltzmann distribution, $\sigma_q \to \sigma_0$, corresponding to a Gaussian distribution. Moreover, if $1 < q_i < 1 + q_c$, where

$$q_c = T_S'(U_0) = \frac{1}{C_V(T_0)}$$

then $\sigma_q > \sigma_0$ (Eq. 63), and hence Tsallis probability density function, $P_i(U)$, has a broader distribution than canonical function at T_0. Whereas, for $q_i < 1$, $P_i(U)$ becomes narrower compare to canonical probability density function.

However, in both cases, Tsallis distribution has its maximum at stable point U_0 as canonical distribution at T_0. For $q_i = 1+q_c$, the Tsallis effective temperature, $\mathcal{T}(U)$, it is tangential to canonical temperature $T_S(U)$ function at U_0, and $P_i(U)$ is locally flat around U_0, indicating only marginal stability. Thus, the choice $q_i = 1+q_c$ generates the most delocalized Tsallis distribution for standard Tsallis MD simulation run. For $q_i > q_c$, the local minimum of $P_i(U)$, namely U_0, is an unstable crossing point [193].

The WHAM is used to estimate the averages of unbiased system quantities at required temperature T_0 [86]. The configuration probability density for each replica k ($k = 1,2,\cdots,K$) at inverse temperature β_ℓ ($\ell = 1,2,\cdots,L$) is written as

$$P(U;q_k,\beta_\ell) = \frac{1}{Z_{k\ell}^{q_k}} \times \Omega(U)\left[1+\beta_\ell(q_k-1)(U-U^{(0)})\right]^{\frac{q_k}{1-q_k}} \tag{64}$$

where U is unbiased potential energy, $\Omega(U)$ is density of states, $Z_{k\ell}$ is configuration partition function:

$$Z_{k\ell} = \int dU\Omega(U)\left[1+\beta_\ell(q_k-1)(U-U^{(0)})\right]^{q_l/(1-q_k)}$$

The canonical distribution at the required inverse temperature β_0 is given by

$$P(U;\beta_0) = f_{k\ell}P(U;q_k,\beta_\ell)\exp\left(\beta_\ell U_k^{bias} - \beta_0 U\right) \tag{65}$$

where U_k^{bias} is the bias potential energy function of replica k and

$$f_{k\ell} = \frac{Z_{k\ell}^{q_k}}{Z_0}$$

or

$$P(U;q_k,\beta_\ell) = f_{k\ell}^{-1}P(U;\beta_0)\exp\left(-\left(\beta_\ell U_k^{bias} - \beta_0 U\right)\right) \tag{66}$$

Further details of the method are presented in Section 3.10..

3.10. The weighted histogram analysis method

The WHAM is often used to analyze the data from replica exchange molecular dynamics simulation. This is considered an efficient technique of data processing since it combines all the data from replicas. In WHAM, it assumed that K copies of the same system (namely the replicas) are in equilibrium with L thermostats at inverse temperature β_ℓ ($\ell = 1,2,\cdots,L$). In addition, to each replica unbiased potential energy, $U_\ell(\mathbf{q}_k)$ ($\ell = 1,2,\cdots,L$; $k = 1,2,\cdots,K$) a biasing potential energy term is added $\Delta U_\ell(\mathbf{q}_k)$. Then, a histogram of M bins is created for the unbiased potential energy combining all of the replicas, with U_m ($m = 1,2,\cdots,M$) being the energy at the center of the bin. Thus, for each replica k and histogram unbiased potential energy bin m we count the number of independent snapshots, namely H_{km}. The probability of observing the system at energy bin m and thermostat ℓ is defined as [59]:

$$P_{\ell m} = Z_\ell^{-1} C_{\ell m} \Omega_m e^{-\beta_0 U_m} \tag{67}$$

where $\Omega_m = \Omega(U_m)$ is the density of states at the energy bin m and the constant $C_{\ell m}$ determines both the effect of temperature and biasing potential in probability distribution as:

$$C_{\ell m} = \exp\left(-\left(\beta_\ell - \beta_0\right) U_m\right) \times \exp\left(-\beta_\ell \Delta U_\ell\right) \tag{68}$$

In Eq. 67, $\Omega_m e^{-\beta_0 U_m}$ gives the unbiased probability of the bin m at the target temperature and Z_ℓ is the partition function at β_ℓ. Note that $\sum_{m=1}^{M} P_{\ell m} = 1$ must be satisfied. Combining Eq. 67 and Eq. 68 we obtain:

$$P_{\ell m} = \Omega_m e^{-\beta_\ell (U_m^{\text{bias}} - F_\ell)} \tag{69}$$

where U_m^{bias} gives the value of biased potential energy at the center of bin m and F_ℓ is the Helmholtz free energy, which has to be estimated, given as

$$F_\ell = -(1/\beta_\ell) \ln Z_\ell$$

Let $n_{k\ell}$ be the number of saved snapshots from replica k visiting thermostat ℓ, then the accumulated probability density for energy bin m can be determined as:

$$P_m = \Omega_m \sum_{k=1}^{K} \sum_{\ell=1}^{L} \frac{n_{k\ell}}{N_k} e^{-\beta_\ell (U_m^{\text{bias}} - F_\ell)} \tag{70}$$

where N_k is the total number of saved snapshots from the replica k. P_m can also be approximated as [31]:

$$P_m \approx \sum_{k=1}^{K} \frac{H_{km}}{N_k} \tag{71}$$

Using the last two equations, we obtain

$$\begin{aligned} \Omega_m &= \frac{\sum_{k=1}^{K} H_{km}}{\sum_{k=1}^{K}\sum_{\ell=1}^{L} n_{k\ell} e^{-\beta_\ell (U_m^{\text{bias}} - F_\ell)}} \\ F_\ell &= -\frac{1}{\beta_\ell} \ln \sum_{m=1}^{M} \Omega_m e^{-\beta_\ell U_m^{\text{bias}}} \end{aligned} \tag{72}$$

To take into account any possible correlations between configurations saved from simulations, the histogram bin statistical inefficiency for each energy bin m from replica k, g_{km}, can be introduced [31], which determines the effective number of snapshots from replica k with unbiased potential energy falling in bin m, H_{km}^{eff}, and the effective number of snapshots from replica k in equilibrium with thermostat ℓ, $n_{k\ell}^{\text{eff}}$:

$$H_{km}^{\text{eff}} = \frac{H_{km}}{g_{km}}; \quad n_{k\ell}^{\text{eff}} = \frac{n_{k\ell}}{g_{km}}$$

Then, the estimated value of the density of states $\hat{\Omega}_m$ is given as:

$$\begin{aligned} \hat{\Omega}_m &= \frac{\sum_{k=1}^{K} H_{km}^{\text{eff}}}{\sum_{k=1}^{K}\sum_{\ell=1}^{L} n_{k\ell}^{\text{eff}} e^{-\beta_\ell (U_m^{\text{bias}} - F_\ell)}} \\ F_\ell &= -\frac{1}{\beta_\ell} \ln \sum_{m=1}^{M} \hat{\Omega}_m e^{-\beta_\ell U_m^{\text{bias}}} \end{aligned} \tag{73}$$

From Eq. 73, $\hat{\Omega}_m$ depends on F_ℓ, and F_ℓ also depends on $\hat{\Omega}_m$. Therefore, F_ℓ and $\hat{\Omega}_m$ are usually determined iteratively from Eqs. 73, starting from some

arbitrary choice $F_\ell = 0$ ($\ell = 1,2,\cdots,L$) and continuing until a convergence is reached. The statistical error $\sigma^2_{\hat{\Omega}_m}$ of $\hat{\Omega}_m$ is given by [31]:

$$\sigma^2_{\hat{\Omega}_m} = \frac{\hat{\Omega}_m}{\sum_{k=1}^{K}\sum_{\ell=1}^{L} n_{k\ell}^{\text{eff}} e^{-\beta_\ell(U_m^{\text{bias}}-F_\ell)}} \tag{74}$$

The estimated average value of any physical quantity A of the system at the target inverse temperature β_0 is computed by summing the weighted values from all configurations:

$$\hat{A}(\beta_0) = \frac{\sum_{k=1}^{K}\sum_{n=1}^{N_k} W_{kn}(\beta_0)A_{kn}}{\sum_{k=1}^{K}\sum_{n=1}^{N_k} W_{kn}(\beta_0)} \tag{75}$$

In Eq. (75) $W_{kn}(\beta_0)$ are the weights given by

$$W_{kn}(\beta_0) = \sum_{m=1}^{M} \frac{\hat{\Omega}_m}{H_{km}} e^{-\beta_0 U_m}$$

The chain rule of error propagation is used to obtain the statistical error of $\hat{A}(\beta_0)$ [31]:

$$\sigma^2_{\hat{A}} = \left(\frac{\langle X\rangle}{\langle Y\rangle}\right)^2 \left(\frac{\sigma_X^2}{(\langle X\rangle)^2} + \frac{\sigma_Y^2}{(\langle Y\rangle)^2} - 2\frac{\sigma_{XY}^2}{\langle X\rangle\langle Y\rangle}\right) \tag{76}$$

where

$$\langle X\rangle = \frac{1}{N_k}\sum_{n=1}^{N_k} W_n(\beta_0)A_n \tag{77}$$

$$\langle Y\rangle = \frac{1}{N_k}\sum_{n=1}^{N_k} W_n(\beta_0) \tag{78}$$

$$\sigma_X^2 = \frac{g_X}{N_k(N_k-1)}\sum_{n=1}^{N_k} \left(W_n(\beta_0)A_n - \langle X\rangle\right)^2 \tag{79}$$

$$\sigma_Y^2 = \frac{g_Y}{N_k(N_k-1)}\sum_{n=1}^{N_k} \left(W_n(\beta_0) - \langle Y\rangle\right)^2 \tag{80}$$

$$\sigma_{XY}^2 = \frac{g_{XY}}{N_k(N_k-1)} \sum_{n=1}^{N_k} (W_n(\beta_0)A_n - \langle X \rangle) \times (W_n(\beta_0) - \langle Y \rangle) \tag{81}$$

Here, $g_{X(Y,XY)}$ are the statistical inefficiencies determined from (auto)correlation functions of replica exchange simulations.

If $\Delta U_\ell = 0$ ($\ell = 1, 2, \cdots, L$), the standard WHAM of replica exchange simulations is obtained, discussed already in the literature. [31]

4. Metadynamics method

Metadynamics method has been developed by Alessandra Laio & Michele Parrinello [111]. The method consists on finding a limited number of essential collective coordinates, c_i ($i = 1, 2, \cdots, m$), upon which the free energy depends on $F(\mathbf{c})$. At any moment of time t, the free energy surface is explored based on the dynamical equations of motion determined by the forces acting on the system:

$$f_i^t = -\frac{\partial F}{\partial c_i^t}$$

In metadynamics method, a bias potential function $U_{\text{bias}}(\mathbf{c})$ is constructed that is added to the Hamiltonian function of system. This bias potential is written as a sum of Gaussian distributions, which are added at any time t of the trajectory in subspace expanded by collective coordinate as generated during the MD simulation. The main effect of the bias potential is to not allow the system visiting configurations that are already explored. The mathematical form of the bias potential is [13, 111]

$$U_{\text{bias}}(\mathbf{c}, t) = \int_0^t dt' \omega \exp\left(-\sum_{i=1}^{m} \frac{(c_i(t) - c_i(t'))^2}{2\sigma_i^2} \right)$$

Here, ω gives the rate of energy change on time:

$$\omega = \frac{W}{\tau}$$

where W and τ are the Gaussian height and the time interval of deposition, respectively [13]. Here, σ_i is the Gaussian distribution width of collective coordinate i. W, τ and σ are adjustable parameters to optimize the algorithm [111].

The main benefit of using metadynamics method is being able to escape from the local minimum free energy metastable states, and hence it increases the rate of sampling rare events. Besides, the metadynamics method allows sampling of new reaction pathways after the system escapes local minimum states [13]. In the metadynamics method, there is no need for *a priory* knowledge of the exact topology of the free energy landscape.

After a certain long time, metadynamics technique will eventually give a bias potential U_{bias} [112]:

$$U_{\text{bias}}(\mathbf{c}, t \to \infty) = -F(\mathbf{c}) + C$$

where C is an integration constant, $F(\mathbf{c})$ is the underlying free energy of system, defined as

$$F(\mathbf{c}) = -\frac{1}{\beta} \ln \left(\int d\mathbf{r} \delta(\mathbf{c} - \mathbf{c}(\mathbf{r})) e^{-\beta U(\mathbf{r})} \right)$$

where $\mathbf{r}$ is the vector of coordinates, $\beta = 1/k_B T$ and $U(\mathbf{r})$ is the potential energy function. This formula has been tested for simplified models [112] and for other complex systems [61, 111]. A formal proof of this expression is shown in Ref. [29].

The free energy surface can be obtained up to an uncertainty, which is inversely proportional to the inverse temperature β and intrinsic diffusion coefficient D of the system in collective coordinates subspace [61, 112]:

$$\varepsilon \propto \left(\frac{\omega}{D\beta} \right)^{1/2}$$

In practical applications of metadynamics method, ε is estimated by comparing different independent simulation runs [8, 15, 151] or using block averaging [141].

The main advantage of using metadynamics method include parallelization, which is an intrinsic property of metadynamics. For instance, one can run multiple interacting copies of metadynamics simulations for reconstruction of a free energy surface, where every simulation contributes to the time-dependent potential [153]. The implementation of method yields an algorithm that scales

very well linearly with the number of processors, independent on the type of processor.

However, there are two disadvantages of the metadynamics simulations. The first, in single metadynamics simulation the convergence of bias potential U_{bias} is not reached to a constant value, but it oscillates about a constant value, making the criteria for stopping the simulation too tricky in practice [13]. The second disadvantage is related to the identification of collective coordinates for describing complex topological free energy is very difficult.

The well-tempered metadynamics method [14] provides a solution to the first problem of standard metadynamics. In this method, the rate of bias accumulation decreases over the course of the simulation, which is made possible using this expression for the bias potential:

$$U'_{\text{bias}}(\mathbf{c},t) = k_B \Delta T \ln\left(1 + \frac{\omega N(\mathbf{c},t)}{k_B \Delta T}\right)$$

where $N(\mathbf{c},t)$ is a histogram accumulated during the simulation for collective variables $\mathbf{c}$ and ΔT a free parameter with dimensions of temperature. The derivative $U'_{\text{bias}}(\mathbf{c},t)$ with respect to time t is

$$\dot{U}'_{\text{bias}}(\mathbf{c},t) = \frac{\omega \delta_{\mathbf{c},\mathbf{c}(t)}}{1 + \dfrac{\omega N(\mathbf{c},t)}{k_B \Delta T}} = \omega \exp\left(-\frac{U(\mathbf{c},t)}{k_B \Delta T}\right) \delta_{\mathbf{c},\mathbf{c}(t)}$$

The new approach can easily be related to standard metadynamics by taking $\delta_{\mathbf{c},\mathbf{c}(t)}$ to be a Gaussian function. This is practically implemented by defining the height of a Gaussian W as

$$W = \omega \tau \exp\left(-\frac{U_{\text{bias}}(\mathbf{c},t)}{k_B \Delta T}\right)$$

There are two main characteristics of the well-tempered metadynamics compare to standard metadynamics [13]. The first, rate of bias accumulation decreases with simulation time as $1/t$ and the deviations from the equilibrium dynamics are small. Secondly, a convergence of the bias potential is reached up to a constant value, C, though a complete compensation of the free energy surface is not obtained:

$$U_{\text{bias}}(\mathbf{c}, t \to \infty) = -\frac{\Delta T}{T + \Delta T} F(\mathbf{c}) + C$$

In long run simulation, the collective coordinates probability distribution becomes:

$$P(\mathbf{c}) \propto \exp\left(-\frac{F(\mathbf{c})}{k_B(T+\Delta T)}\right)$$

where for $\Delta T \to 0$, standard MD simulations are recovered, and for $\Delta T \to \infty$ we have standard metadynamics. In contrast to standard metadynamics, in well-tempered metadynamics, the bias potential converges to a finite value during one run. For all other values of ΔT the extent of free energy surface exploration is determined by adjusting ΔT [13].

5. Umbrella sampling methods

Umbrella sampling method is developed in Refs. [65,180]. This technique adds a bias term to the potential energy function applied to the system, which ensures efficient sampling along the path of a reaction coordinate. The bias term can be added in a single simulation run or in multiple copies of simulations (often called *windows*) with overlapping distributions. The umbrella sampling method aims to connect regions of phase space that are separated by high energy barriers, which is a reason for naming it the *umbrella sampling*.

The bias potential as a function of the reaction coordinate, let say q, has the following form:

$$U_{\text{bias}}(q) = -F(q)$$

where $F(q)$ is not usually known *a priory*. The umbrella sampling aims to calculate the $F(q)$ by employing two main types of bias potentials, such as harmonic biases consisting of a set of windows centered at different points along q and adaptive bias modeled to match $-F(q)$ in only one window enveloping the whole range of q.

5.1. Harmonic bias potentials

In this approach, the entire range of values of q is divided into a subset of small size windows, N_w. Then, a bias potential function is employed in every window

allowing the system to fluctuate around a reference point q_i^{ref} centered at the window i of the form [103]:

$$U_{\text{i,bias}}(q) = \frac{k_i}{2}\left(q - q_i^{\text{ref}}\right)^2 \tag{82}$$

Free energy curves calculated in this way are combined using WHAM technique [52, 109]. From Eq. (82), the bias potential U_{bias} is characterized by k_i, which can also be adjusted depending on the window, number of windows, N_w, and q_i^{ref}.

In general, q_i^{ref} are chosen uniformly distributed in all range of q. In practice [55], there is a compromise between the statistical errors and computational time required. For instance, increasing the number of windows results in a smaller statistical error, but longer computational time is needed. However, the advantage is that MD simulations in each window are completely independent, and hence they can run in parallel by producing multiple MD simulation copies, which takes into account the advantage of the parallelism of computer architecture available. This approach has already been used in atomistic simulations of protein folding [135].

Combination of the umbrella sampling with replica exchange have also been suggested [12] to improve the conformational sampling. Strength, k_i, of the bias potential, has to be chosen before simulation runs, such that the bias potential allows steering the system across potential energy barriers. On the other hand, too large k_i will cause very narrow probability distributions. Often, if the probability distributions have too large gaps between the windows, then additional windows could also be inserted. Overlapping between distributions at each window has to be sufficiently large for WHAM and it could also be advantageous in umbrella integration [104]. Large values of k_i can also lead to instability of numerical integration of the equations of motion unless small time steps are used. In addition, for too large values of k_i only the configurations with high energies will be sampled [174].

The statistical error can also be derived analytically as a function of k_i [104]. Location of the next sampled window (q_{i+1}^{ref}) can be chosen as [202]

$$q_{i+1}^{\text{ref}} = q_i^{\text{ref}} + w_i$$

where w_i is the window width. Alternatively, the experimental data can also be used to determine optimized values of bias parameters [124].

5.2. Adaptive bias umbrella sampling

On the other hand, this method aims to screen the entire interval of interest for the reaction coordinate q in a single simulation run [16, 17, 88, 110, 123], by choosing a bias potential of the following form:

$$U_{\text{bias}}(q) = -F(q)$$

Adding this bias term to the potential yields an exactly flat energy surface, and hence the resulting probability distribution is uniform along q. Usually, the simulations start with an initial guess for $U_{\text{bias}}(q)$, because $F(q)$ is not known a priory. Then, iteratively, $U_{\text{bias}}(q)$ is improved until a complete uniform distribution is obtained in q space.

6. Transition path sampling methods

In typical chemical reactions in solutions, the difficulty of computer simulations is in understanding the rare events occurring in complex systems when moving from one basin of attraction to another on a multimodal potential energy landscape. In particular, determining the transition state of these processes as a function of order parameters, which have also to be defined, is notoriously difficult problem [43], which will allow sampling using the standard MD simulations starting from this initial state [5,30,75,105]. The disadvantage of this approach is that the transition state is not always known *a priory*. Furthermore, because of the high dimensionality of the problem phase space, the energy landscape has a complex topology with many transition states, and hence the reaction coordinate may not be represented accurately by the order parameter [42]. On the other hand, transition path sampling [26,42], as an alternative method, does not require previous knowledge of transition states, but it relies on the calculation of isomorphic reversible work from reactive flux correlation functions.

In this approach, $L+1$ copies of the trajectories in phase space characterize a path in space-time:

$$\mathbf{X}_L = \{\mathbf{x}_0 \to \mathbf{x}_1 \to \cdots \to \mathbf{x}_L\}$$

Here, $\mathbf{x}_t$ $(t = 0, 1, \cdots, L)$ represent points in a 2D-dimensional phase space. A relationship between the sequence time t and physical simulation run time ex-

ists depending on the transition path [42], which is represented by $2D(L+1)$ coordinates.

If we assume visited states form a Markovian chain, then the probability of simulating the trajectories is given by

$$\mathcal{P}(\mathbf{X}_L)\exp\left(-\beta E(\mathbf{x}_0)\right)\prod_{t=0}^{L-1} p(\mathbf{x}_t \to \mathbf{x}_{t+1}) \tag{83}$$

where β is the inverse simulation temperature and $E(\mathbf{x}_t)$ is the total energy at configuration $\mathbf{x}_t$. Here, the initial configuration (at $t = 0$) is generated from a canonical distribution [42]:

$$\rho_{init}(\mathbf{x}_0) = \exp\left(-\beta E(\mathbf{x}_0)\right)$$

In Eq. (83), $p(\mathbf{x}_t \to \mathbf{x}_{t+1})$ gives the transition probability for each time step, which is determined based on the natural dynamics. Usually, any Markovian transition probability $p(\mathbf{x}_t \to \mathbf{x}_{t+1})$ is such that it should obey to the Boltzmann distribution and is normalized [42]. Typically, two approaches are proposed as a choice for $p(\mathbf{x}_t \to \mathbf{x}_{t+1})$: the Markovian action and Langevin action [42]. If the natural dynamics are governed by the Newton's equations of motion for the Hamiltonian systems, then

$$p(\mathbf{x}_t \to \mathbf{x}_{t+1}) = \delta\left(\mathbf{x}_{t+1} - \Phi_{\Delta t}(\mathbf{x}_t)\right)$$

where δ is the delta function and $\Phi_{\Delta t}(\mathbf{x}_t)$ is the discrete flow map of one-time step propagation.

A transition path sampling can be performed by applying the constraints $h_A(\mathbf{x}_0)$ and $h_B(\mathbf{x}_L)$ at the endpoints in path probability as the following:

$$\begin{aligned} \mathcal{P}_{AB}(\mathbf{X}_L) &\equiv Z_{AB}^{-1}(L)\, h_A(\mathbf{x}_0)\mathcal{P}(\mathbf{X}_L)h_B(\mathbf{x}_L) \\ Z_{AB}(L) &= \int d^L\mathbf{x}\, h_A(\mathbf{x}_0)\mathcal{P}(\mathbf{X};L)h_B(\mathbf{x}_L) \end{aligned} \tag{84}$$

where $h_A(\mathbf{x}_0)$ and $h_B(\mathbf{x}_L)$ are indicator functions defined as

$$h_{A,B}(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \in A,B \\ 0 & \text{if } \mathbf{x} \notin A,B \end{cases}$$

Here, $h_A(\mathbf{x}_0)$ constraints the trajectory path to start in the region A (i.e., reactant) and $h_B(\mathbf{x}_L)$ constraints the trajectory path to stop in the region B (i.e., product).

Typically, L steps are used to take the system from states A to the state B, with action defined by Eq. (84).

This approach has been used to probe the dynamics of folding pathways for the C-terminal β-hairpin of protein G-B1 using MD simulation at room temperature of protein in explicit solvent [25].

It can be suggested for time propagation of the system to be governed by the swarm particle-like dynamics given in the following form ($i = 1, 2, \cdots, f$):

$$
\begin{aligned}
\dot{x}_i &= \frac{p_i}{m_i} \qquad (85)\\
\dot{p}_i &= F_i - \lambda_i p_i - P_1^{(s)} p_i \\
&\quad + \gamma_1 u_1 (x_i^{\text{Lbest}} - x_i) + \gamma_2 u_2 (x_i^{\text{Gbest}} - x_i) \\
\dot{\eta}^{(i)} &= -\sum_{j=1}^{L} \left[\frac{Q_1^{(i)} (\xi_{1,j}^{(i)})^2}{k_B T} \xi_{2,j}^{(i)} - \sum_{k=2}^{M} \xi_{k,j}^{(i)} \right] \\
\dot{\xi}_{1,j}^{(i)} &= -\xi_{1,j}^{(i)} \xi_{2,j}^{(i)} - \lambda_b^{(i)} \xi_{1,j}^{(i)} - \lambda_i \xi_{1,j}^{(i)} \\
&\quad j = 1, \cdots, L \\
\dot{\xi}_{k,j}^{(i)} &= \frac{G_{k,j}^{(i)}}{Q_k^{(i)}} - \xi_{k+1,j}^{(i)} \xi_{k,j}^{(i)} \\
&\quad j = 1, \cdots, L;\ k = 2, \cdots, M-1 \\
\dot{\xi}_{M,j}^{(i)} &= \frac{G_{M,j}^{(i)}}{Q_{M,j}^{(i)}} \\
&\quad j = 1, \cdots, L \\
\dot{s}_{i,k} &= P_{i,k}^{(s)}, \quad k = 1, 2, \cdots, M \\
\dot{P}_{i,k}^{(s)} &= \frac{\Gamma_{i,k}}{W_{i,k}} - P_{i,k+1}^{(s)} P_{i,k}^{(s)}, \quad k = 1, 2, \cdots, M-1 \\
\dot{P}_{i,M}^{(s)} &= \frac{\Gamma_{i,M}}{W_{i,M}},
\end{aligned}
$$

where $\mathbf{x}^{\text{Gbest}}$ represents the coordinates of final state B (product), i.e., $\mathbf{x}_L$. Then, we can generate different trajectories starting in the region A (reactant) and bi-

ased towards the end in final region B. As advantage this method does not postulate *a priory* knowledge of the transition state.

7. Accelerated molecular dynamics method

Accelerated molecular dynamics approach is proposed as a robust method to bias the potential energy function to efficiently enhance the barrier crossing during the simulations [72], based on previously introduced methods [70,152,195]. The method has been described in details elsewhere [71,72]. In this approach, a reference boost potential energy U_0 term is defined with a value slightly lower in magnitude than the local potential energy minimum [72], then each step of simulations the potential energy $U(\mathbf{r})$ is modified by a continuous non-negative bias potential $\Delta U(\mathbf{r})$ as [71,72]

$$U_{\text{bias}}(\mathbf{r}) = U(\mathbf{r}) + \Delta U(\mathbf{r}) \tag{86}$$

where the bias term is given as

$$\Delta U(\mathbf{r}) = \begin{cases} \dfrac{(U_0 - U(\mathbf{r}))^2}{(U_0 - U(\mathbf{r})) + \alpha} & \text{If } U(\mathbf{r}) < U_0 \\ 0 & \text{Otherwise} \end{cases}$$

where α is used to adjust the depth of potential energy minimum and modulate local smoothness of the energy basins of U_{bias}. In this approach, the bias term $\Delta U(\mathbf{r})$ raises the potential surfaces near the minimum states and leaves unaffected surface points near the barriers.

Another form of the bias term $\Delta U(\mathbf{r})$ has also been proposed [39], such as

$$U_{\text{bias}}(\mathbf{r}) = U(\mathbf{r}) - \Delta U(\mathbf{r}) \tag{87}$$

where the bias term is given as

$$\Delta U(\mathbf{r}) = \begin{cases} \dfrac{(U(\mathbf{r}) - U_0)^2}{(U(\mathbf{r}) - U_0) + \alpha} & \text{If } U(\mathbf{r}) \geq U_0 \\ 0 & \text{Otherwise} \end{cases}$$

In this new MD simulation approach, transitions are accelerated by lowering the barriers instead. With increasing α the modified landscape becomes rougher,

and it moves closer to the original potential. It is interesting to note that taking into account the relationship between average potential energy surface roughness and the diffusivity, the method allows acquiring approximately the kinetics of original potential energy landscape [44, 74, 205].

The method has successfully been used to study the sampling of slow diffusive conformation transitions of torsion angles for biomolecules in timescales longer than milliseconds [71]. The approach is also tested in the *ab initio* molecular dynamics simulations [28]. The method is efficiently used to increase accuracy and convergence of free energy computations in condensed-phase systems when combined with thermodynamic integration simulations [39]. The approach is also used as replica exchange by varying the degree of acceleration among the replicas for gas-phase model systems [50]. The approach is used to study the waters contribution to the energetic roughness from peptide dynamics [99]. Besides, the method has also shown to retrieve the kinetic rate constant when applied in simulations of the helix to beta strand transition of alanine dipeptide in explicit solvent [38]. Recently [146], the approach has been used to access conformation changes in time scales of milliseconds for bovine pancreatic trypsin inhibitor protein emphasizing one of the method's advantage for not needing prior knowledge of free energy landscape or reaction coordinate.

8. Conformational flooding method

Helmut Grubmüller [70] introduced a new approach examining conformation transitions in complex macromolecular systems at the atomistic level. In this method, first, the so-called conformation space [9] is defined for the system characterized by the Hamiltonian H as a restricted region of the configuration phase space, in which system spends a long time. Typically, this determines the time needed for the system to sample enough phase space for the correct determination of the statistical averages [70].

In this confined space, the free energy landscape is determined in order to find an effective Hamiltonian H^k_{eff}, where k indicates one of the subspaces in configuration space. Here, it is assumed that this configuration space is made up by regions of low free energy F, which are separated from each other by high energy barriers of order ΔF. The free energy landscape of the subspace is expressed in terms of the so-called collective coordinates [1, 54, 66, 69], charac-

terized by the vector

$$\mathbf{c} = (c_1, c_2, \cdots, c_m)^T$$

where m is the number of essential degrees of freedom of the configuration subspace. This describes a coarse-graining of the configuration space, leaving out $3N-m$ degrees of freedom, with N being the total number of atoms. According to [70], in the subspace describe by the collective coordinates, the conformation space density $\tilde{\rho}(\mathbf{c})$ is defined as

$$\tilde{\rho}(\mathbf{c}) = \int d\mathbf{x}' \rho(\mathbf{x}') \delta\left(\mathbf{c} - \mathbf{c}(\mathbf{x}')\right) \tag{88}$$

where $\mathbf{x}$ is a $3N$-dimensional Cartesian vector of the N particle positions, and $\rho(\mathbf{x})$ is the configuration space density. Hence, the free energy landscape can be evaluated as [70]:

$$F(\mathbf{c} = -\frac{1}{\beta} \ln \tilde{\rho}(\mathbf{c})$$

Calculation of the $\tilde{\rho}(\mathbf{c})$ requires knowledge of $\rho(\mathbf{x})$, which is an integration on a $3N$ configuration space difficult to be evaluated in practice, since the system has to be ergodic [70]. However, in time scales covered by MD simulations (typically of order a few hundred nanoseconds), systems (e.g. biomolecular systems) are non-ergodic at all time scales (see for example [32] or [70] and the references therein).

In [70], the configuration density of the subspace k, $\rho_k(\mathbf{x})$ is approximated as

$$\rho_k(\mathbf{x}) = Z^{-1} \exp\left(-\frac{1}{2}(\mathbf{x}-\bar{\mathbf{x}})^T \mathbf{C}^{-1} (\mathbf{x}-\bar{\mathbf{x}})\right) \tag{89}$$

In Eq. (89), the partition function, Z, is given by

$$Z = \int d\mathbf{x} \exp\left(-\frac{1}{2}(\mathbf{x}-\bar{\mathbf{x}})^T \mathbf{C}^{-1} (\mathbf{x}-\bar{\mathbf{x}})\right)$$

Here $\mathbf{C}$ is the covariance matrix, which is a $\Re^{3N\times 3N}$ symmetric and positive matrix, defined as [70] (and the references therein):

$$\mathbf{C} = \langle (\mathbf{x}-\bar{\mathbf{x}})(\mathbf{x}-\bar{\mathbf{x}})^T \rangle_k$$

where

$$\bar{\mathbf{x}} = \langle \mathbf{x} \rangle_k$$

Here, $\langle \cdots \rangle_k$ denotes ensemble average in the configuration subspace k. The matrix $\mathbf{C}$ is calculated based on MD trajectories, which then is diagonalized:

$$\mathbf{C} = \mathbf{E}^T \Lambda^{-1} \mathbf{E} \tag{90}$$

where $\mathbf{E}$ and Λ are the matrix of eigenvectors and the diagonal matrix of eigenvalues, respectively. Projection of the Cartesian coordinates fluctuations along the space spanned by eigenvectors $\mathbf{E}$ is

$$\mathbf{q} = \mathbf{E}^T (\mathbf{x} - \bar{\mathbf{x}})$$

Equation 89 can be simplified as

$$\rho_k(\mathbf{x}) = Z^{-1} \exp\left(-\frac{1}{2} \mathbf{q}^T \Lambda \mathbf{q} \right) \tag{91}$$

Coarse-graining of the configuration subspace k allows the definition of m essential collective coordinates [1]

$$\mathbf{c} = (q_1, q_2, \cdots, q_m)^T$$

with the largest eigenvalues, which characterize the low-frequency fluctuation modes, and the remaining $3N - m$ eigenvalues, which characterize the high-frequency fluctuation modes, are assumed not to influence the conformation transitions [70]. The justification of this choice is based on the fact that the first m eigenmodes are anharmonic and with high amplitude, whereas the other $3N - m$ eigenmodes are essentially harmonic and with small amplitudes, and hence only m coordinates will dominate the collective motions in biomolecular systems (see for example discussion in Ref. [70] and the references therein.)

Thus, the conformation subspace density is defined as

$$\tilde{\rho}_k(\mathbf{c}) = \tilde{Z}_c^{-1} \exp\left(-\frac{1}{2} \mathbf{c}^T \Lambda_c \mathbf{c} \right)$$

where Λ_c is a reduced matrix of m diagonal elements and $\tilde{Z}_c$ is the corresponding subspace partition function. Then, the effective Hamiltonian becomes,

$$H_{\text{eff}}^k(\mathbf{c}) = \frac{1}{2\beta} \mathbf{c}^T \Lambda_c \mathbf{c} \tag{92}$$

This coarse-grained model for the configuration subspace H_{eff}^{k} is fundamental in designing the so-called flooding potential $V_{\text{fl}}(\mathbf{c})$ chosen as a multivariate Gaussian in order to fulfill the criteria discussed in Ref. [70]:

$$V_{\text{fl}}(\mathbf{c}) = E_{\text{fl}} \exp\left(-\frac{1}{2}\mathbf{c}^{T}\Lambda_{\text{fl}}\mathbf{c}\right) \tag{93}$$

where E_{fl} characterizes the strength of $V_{\text{fl}}(\mathbf{c})$ and Λ_{fl} determines the shape of $V_{\text{fl}}(\mathbf{c})$ in conformation space, which is chosen to be

$$\Lambda_{\text{fl}} = \Lambda_{c}/\gamma^{2}$$

where γ is a proportionality constant specified as a function of V_{fl} [70]:

$$\gamma = (\beta E_{\text{fl}})^{1/2}$$

The method is applied to probe conformation transitions in argon clusters and simplified protein model [70]. Other examples used to demonstrate the application of flooding to accelerate conformation transitions and chemical reactions are also examined [115].

9. Discussion and perspectives

Both industry and academic research are often using the molecular dynamics technique and its variants to a wide range of problems and systems, from inorganic and organic fluids to macromolecular. Yet, there are several issues identified in applying molecular dynamics simulations to these systems as we probed in this work, such as time and size scale, and rare events.

The recent advances in parallel supercomputing have made possible to brace larger spatial scales, but increasing simulation timescales remains still a challenge. For instance, simulations spanning up to microseconds in the lifetime of a macromolecule need to cover billions of numerical integration time steps, which is a challenge from the computation point of view. This is because in simulations of biomolecules, at each time step, only a relatively small amount of computation can be run on parallel among a large number of processors. Hence, indeed, billions of simulation time steps can only be executed in a considerable amount of time. In nowadays, for molecular dynamics method development scientists, in particular, a significant challenge is to effectively employ

the computers of near future to perform simulations of systems with millions of atoms [79, 143, 171–173].

Another approach is exploitation of hardware parallel supercomputers for MD simulations with processors that can execute traditional MD codes orders of magnitude faster, such as Anton supercomputer [159].

In long MD simulation runs up to milliseconds timescale using fully atomistic physical models, force field accuracy will determine the overall accuracy achieved. Very recent studies [145] (and the references therein) have shown that prediction of native-structures and folding rated can be more robust concerning error compare to the potential energy function. Moreover, the numerical integrator used in MD simulations should guarantee energy conservation and stable trajectories in long time scale simulations [67, 101]. Large computer storage is also needed to save, analyze, and better computer graphics to handle a large amount of data produced.

Coarse-grained models of macromolecular systems have probed problems of biologically relevant size and timescale during simulations when combined with computer power [182]. However, the coarse-grained models of proteins remain demanding, because of the challenges in building useful energy potential functions representing all the physics of interactions [102, 114, 170]. The most tested coarse-grained model is the bead-spring model of polymers. Solvent effects are added using Brownian dynamics [48] or Stokesian dynamics [144].

In this study, we attempted to give a big picture of the methods used to enhance sampling in molecular dynamics simulations and thus being able to simulate rare events for complex molecular systems. Also, we presented the advantages and limitations of each method. Our final message from this study is that probing relevant time and size scales of (bio)physical and chemical phenomena in macromolecular systems may need new statistical models of data processing and computational theoretical models to allow studying them efficiently [60, 127].

Acknowledgments

The author thanks International Balkan University for the support.

CONFLICT OF INTEREST

The author declares that there is no conflict of interest regarding the publication of this chapter.

References

[1] A. Amadei, A. B. M. Linssen, and H. J. C. Berendsen, *Essential dynamics of proteins*, Proteins: Struct., Funct., Genet. **17** (1993), 412–425.

[2] A.E. Garcia and J.N. Onuchic, *Folding a protein in a computer: An atomic description of the folding/unfolding of protein A*, Proc. Natl. Acad. Sci. USA **100** (2003), 13898–13903.

[3] A. Karolak and A. van der Vaart, *Importance of local interactions for the stability of inhibitory helix 1 of Ets-1 in the apo state*, Biophys. Chem. **165-166** (2012), no. 3, 74–8.

[4] A. Kone and D.A. Kofke, *Selection of temperature intervals for parallel-tempering simulations*, J. Chem. Phys. **122** (2005), 206101.

[5] J. B. Anderson, *Statistical theories of chemical reactions. Distributions in the transition region*, J. Chem. Phys. **58** (1973), 4684.

[6] I. Andricioaei and J. E. Straub, *Generalized simulated annealing algorithms using Tsallis statistics: Application to conformational optimization of a tetrapeptide*, Phys. Rev. E **53** (1996), R3055–R3058.

[7] _____, *On Monte Carlo and molecular dynamics methods inspired by Tsallis statistics: Methodology, optimization, and application to atomic clusters*, J. Chem. Phys. **107** (1997), no. 21, 9117–9124.

[8] S. Angioletti-Uberti, M. Ceriotti, P. D. Lee, and M. W. Finnis, *Solid-liquid interface free energy through metadynamics simulations*, Phys. Rev. B **81** (2010), 125416.

[9] A. Ansari, J. Berendzen, D. Braunstein, B. R. Cowen, H. Frauenfelder, M. K. Hong, I. E. T. Iben, J. B. Johnson, P. Ormos, T. B. Sauke, R. Scholl,

A. Schulte, P. J. Steinbach, J. Vittitow, and R. D. Young, *Rebinding and relaxation in the myoglobin pocket*, Biophys. Chem. **26** (1987), 337–355.

[10] A. Okur, L. Wickstrom, M. Layten, R. Geney, K. Song, V. Hornak, and C.J. Simmerling, *Improved efficiency of replica exchange simulations through use of hybrid explicit/implicit salvation model*, J. Chem. Theory and Comput. **2** (2006), 420–433.

[11] A. Piela, J. Kostrowicki, and H. A. Scheraga, *On the multiple-minima problem in the conformational analysis of molecules: Deformation of the potential energy hypersurface by the diffusion equation method*, J. Phys. Chem. **93** (1989), 3339–3346.

[12] S. Auer and D. Frenkel, *Numerical prediction of absolute crystallisation rates in hard-sphere colloids*, J. Chem. Phys. **120** (2004), no. 6, 3015–3029.

[13] A. Barducci, M. Bonomi, and M. Parrinello, *Metadynamics*, Advanced Review **1** (2011), 826–843.

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

[15] A. Barducci, R. Chelli, P. Procacci, V. Schettino, F. L. Gervasio, and M. Parrinello, *Metadynamics simulation of prion protein: Beta-structure stability and the early stages of misfolding*, J. Am. Chem. Soc. **128** (2006), 2705–2710.

[16] C. Bartels and M. Karplus, *Multidimensional adaptive umbrella sampling: Applications to main chain and side chain peptide conformations*, J. Comput. Chem. **18** (1997), 1450–1462.

[17] _____, *Probability distributions for complex systems: Adaptive umbrella sampling of the potential energy*, J. Phys. Chem. B **102** (1998), 865–880.

[18] E. Barth and T. Schlick, *Extrapolation versus impulse in multiple-time stepping schemes. II. Linear analysis and applications to Newtonian and Langevin dynamics*, J. Chem. Phys. **109** (1998), no. 5, 1633–1642.

[19] ———, *Overcoming stability limitations in biomolecular dynamics. I. Combining force splitting via extrapolation with Langevin dynamics in LN*, J. Chem. Phys. **109** (1998), no. 5, 1617–1632.

[20] B. A. Berg and T. Celik, *New approach to spin-glass simulations*, Phys. Rev. Lett. **69** (1992), 2292–2295.

[21] B. A. Berg and T. Neuhaus, *Multicanonical algorithms for first order phase transitions*, Phys. Lett. B **267** (1991), no. 2, 249 – 253.

[22] ———, *Multicanonical ensemble: A new approach to simulate first-order phase transitions*, Phys. Rev. Lett. **68** (1992), 9–12.

[23] B.A. Berg, *Markov chain Monte Carlo simulations and their statistical analysis*, World Scientific, Singapore, 2004.

[24] T. C. Bishop, R. D. Skeel, and K. Schulten, *Difficulties with multiple time stepping and fast multipole algorithm in molecular dynamics*, J. Comput. Chem. **18** (1997), no. 14, 1785–1791.

[25] P. G. Bolhuis, *Transition-path sampling of* β*-hairpin folding*, Proc. Natl. Acad. Sci. USA **100** (2003), no. 21, 12129–12134.

[26] P. G. Bolhuis, C. Dellago, P. L. Geissler, and D. Chandler, *Transition-path sampling: Throwing ropes over mountains in the dark*, J. Phys.: Condens. Matter **12** (2000), A147–A152.

[27] B. R. Brooks, C. L. Brooks, A. D. MacKerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus, *CHARMM: The biomolecular simulation program*, J. Comput. Chem. **30** (2009), no. 10, 1545–1614.

[28] D. Bucher, L. C. T. Pierce, J. A. McCammon, and P. R. L. Markwick, *On the use of accelerated molecular dynamics to enhance configurational*

sampling in ab initio simulations, J. Chem. Theory Comput. **7** (2011), 890–897.

[29] G. Bussi, A. Laio, and M. Parrinello, *Equilibrium free energies from non equilibrium metadynamics*, Phys. Rev. Lett. **96** (2006), 090601.

[30] D. Chandler, *Statistical mechanics of isomerization dynamics in liquids and the transition state approximation*, J. Chem. Phys. **68** (1978), no. 6, 2959–2970.

[31] J. D. Chodera, W. C. Swope, J. W. Pitera, C. Seok, and K. A. Dill, *Use of the weighted histogram analysis method for the analysis of simulated and parallel tempering simulations*, J. Chem. Theory Comput. **3** (2007), 26–41.

[32] J. B. Clarage, T. Romo, B. .K. Andrews, B. M. Pettitt, and G. N. Jr. Phillips, *A sampling problem in molecular dynamics simulations of macromolecules*, Proc. Natl. Acad. Sci. USA **92** (1995), 3288–3292.

[33] C. Predescu, M. Predescu, and C. Ciobanu, *The incomplete beta function law for parallel tempering sampling of classical canonical systems*, J. Chem. Phys. **120** (2004), no. 9, 4119–4128.

[34] M. Creutz and A. Goksch, *Higher-order hybrid Monte Carlo algorithms*, Phys. Rev. Lett. **63** (1989), 9–12.

[35] D.A. Kofke, *On the acceptance probability of replica-exchange Monte Carlo trials*, J. Chem. Phys. **117** (2002), no. 15, 6911–6914, With erratum in Ref. [36].

[36] ———, *Erratum: On the acceptance probability of replica-exchange Monte Carlo trials*, J. Chem. Phys. **120** (2004), no. 22, 10852–10852.

[37] D. Bashford and D.A. Case, *Generalized Born model of macromolecular salvation effects*, Annu. Rev. Phys. Chem. **51** (2000), 129–152.

[38] C. A. F. de Oliveira, D. Hamelberg, and J. A. McCammon, *Estimating kinetic rates from accelerated molecular dynamics simulations: Alanine dipeptide in explicit solvent as a case study*, J. Chem. Phys. **127** (2007), 175105.

[39] ———, *Coupling accelerated molecular dynamics methods with thermodynamic integration simulations*, J. Chem. Theory Comput. **4** (2008), 1516–1525.

[40] P. M. C. de Oliveira, *Broad Histogram Simulation: Microcanonical Ising Dynamics*, Int. J. Mod. Phys. C **9** (1998), 497–503.

[41] P. M. C. de Oliveira, T. J. P. Penna, and H. J. Herrmann, *Broad histogram method*, Braz. J. Phys. **26** (1996), 677–683.

[42] C. Dellago, P. G. Bolhuis, F. S. Csajka, and D. Chandler, *Transition path sampling and the calculation of rate constants*, J. Chem. Phys. **108** (1998), no. 5, 1964–1977.

[43] D. J. Wales, *Perspective: Insight into reaction coordinates and dynamics from the potential energy landscape*, J. Chem. Phys. **142** (2015), 130901–12.

[44] U. Doshi and D. Hamelberg, *Extracting realistic kinetics of rare activated processes from accelerated molecular dynamics using Kramer's theory*, J. Chem. Theory Comput. **7** (2011), 575–581.

[45] R. O. Dror, D. H. Arlow, P. Maragakis, T. J. Mildorf, A. C. Pan, H. Xu, and D. W. Borhani, *Activation mechanism of the* β*2-adrenergic receptor*, Proc. Natl. Acad. Sci. USA **108** (2011), 18684–18689.

[46] D. J. Earl and M. W. Deem, *Parallel tempering: Theory, applications, and new perspectives*, Phys. Chem. Chem. Phys. **7** (2005), 3910.

[47] E. Lyman, F. M. Ytreberg, and D. M. Zuckerman, *Resolution exchange simulation*, Phys. Rev. Lett. **96** (2006), 028105.

[48] D. L. Ermak and J. A. McCammon, *Brownian dynamics with hydrodynamic interactions*, J. Chem. Phys. **64** (1978), 1352.

[49] F.A. Escobedo and F.J. Martinez-Veracoechea, *Optimized expanded ensembles for simulations involving molecular insertions and deletions. i. closed systems.*, J. Chem. Phys. **127** (2007), 174103.

[50] M. Fajer, D. Hamelberg, and J. A. McCammon, *Replica-exchange accelerated molecular dynamics (REXAMD) applied to thermodynamic integration*, J. Chem. Theory Comput. **4** (2008), 1565–1569.

[51] M. Falcioni and M. W. Deem, *A biased Monte Carlo scheme for zeolite structure solution*, J. Chem. Phys. **110** (1999), no. 3, 1754.

[52] A. M. Ferrenberg and R. H. Swendsen, *Optimized Monte Carlo data analysis*, Phys. Rev. Lett. **63** (1989), 1195–1198.

[53] D. D. Frantz, D. L. Freeman, and J. D. Doll, *Reducing quasi-ergodic behavior in Monte Carlo simulations by J-walking: Applications to atomic clusters*, J. Chem. Phys. **93** (1990), no. 4, 2769–2784.

[54] H. Frauenfelder, P. J. Steinbach, and R. D. Young, *Conformational relaxation in proteins*, Chem. Scr. **23A** (1989), 145–150.

[55] D. Frenkel and B. Smit, *Understanding molecular simulations: From algorithms to applications*, second ed., Academic Press, San Diego, CA, 2002.

[56] M. S. Friedrichs, P. Eastman, V. Vaidyanathan, M. Houston, S. Legrand, A. L. Beberg, D. L. Ensign, C. M. Bruns, and V. S. Pande, *Accelerating molecular dynamics simulation on graphics processing units*, J. Comput. Chem. **30** (2009), 864–72.

[57] F. Wang and D.P. Landau, *Determining the density of states for classical statistical models: A random walk algorithm to produce a flat histogram*, Phys. Rev. E **64** (2001), 056101–16.

[58] ———, *Efficient, multiple-range random walk algorithm to calculate the density of states*, Phys. Rev. Lett. **86** (2001), no. 10, 2050–2053.

[59] E. Gallicchio, M. Andrec, A. K. Felts, and R. M. Levy, *Temperature weighted histogram analysis method, replica exchange, and transition paths*, J. Phys. Chem. B **109** (2005), 6722–6731.

[60] G. Ciccotti and E. Vanden-Eijnden, *The trees and the forest. Aims and objectives of molecular dynamics simulations*, Eur. Phys. J. Special Topics **224** (2015), 2515–2518.

[61] F. L. Gervasio, A. Laio, and M. Parrinello, *Flexible docking in solution using metadynamics*, J. Am. Chem. Soc. **127** (2005), 2600–2607.

[62] G. J. Geyer, *Practical Markov Chain Monte Carlo*, Stat. Sci. **7** (1992), no. 4, 473–483.

[63] G. J. Martyna, M. E. Tuckerman, D. J. Tobias, and M. L. Klein, *Explicit reversible integrators for extended systems dynamics*, Mol. Phys. **87** (1996), no. 5, 1117–1157.

[64] G. J. Martyna, M. L. Klein, and M. E. Tuckerman, *Nosé-Hoover chains: The canonical ensemble via continuous dynamics*, J. Chem. Phys. **97** (1992), no. 4, 2635–2643.

[65] G.M. Torrie and J. P. Valleau, *Nonphysical sampling distributions in Monte-Carlo free-energy estimation: Umbrella sampling*, J. Comput. Phys. **23** (1977), no. 2, 187–199.

[66] N. Go and H. A. Scheraga, *Analysis of the contribution of internal vibrations to the statistical weights of equilibrium conformations of macromolecules*, J. Chem. Phys. **51** (1969), no. 11, 4751–4767.

[67] S. K. Gray, D. W. Noid, and B. G. Sumpter, *Symplectic integrators for large scale molecular dynamics simulations: A comparison of several explicit methods*, J. Chem. Phys. **101** (1994), no. 5, 4062–4072.

[68] D. Gront and A. Kolinski, *Efficient scheme for optimization of parallel tempering Monte Carlo method*, J. Phys.: Condens. Matter **19** (2007), 036225.

[69] H. Grubmüller, Ph.D. thesis, Technische Universität München, 1994.

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

[71] D. Hamelberg, C. A. F. de Oliveira, and J. A. McCammon, *Sampling of slow diffusive conformational transitions with accelerated molecular dynamics*, J. Chem. Phys. **127** (2007), no. 15, 155102.

[72] D. Hamelberg, J. Mongan, and J. A. McCammon, *Accelerated molecular dynamics: A promising and efficient simulation method for biomolecules*, J. Chem. Phys. **120** (2004), 11919.

[73] D. Hamelberg, J. Mongan, and J.A. McCammon, *Accelerated molecular dynamics: A promising and efficient simulation method for biomolecules*, J. Chem. Phys. **120** (2004), no. 24, 11919–11929.

[74] D. Hamelberg, T. Shen, and J. A. McCammon, *Relating kinetic rates and local energetic roughness by accelerated molecular-dynamics simulations*, J. Chem. Phys. **122** (2005), 241103.

[75] P. Hänggi, P. Talkner, and M. Borkovec, *Reaction-rate theory: Fifty years after Kramers*, Rev. Mod. Phys. **62** (1990), 251.

[76] U. H. E. Hansmann and Y. Okamoto, *Prediction of peptide conformation by multicanonical algorithm: New approach to the multiple-minima problem*, J. Comput. Chem. **14** (1993), 1333–1338.

[77] ———, *In annual reviews in computational physics VI*, (World Scientific, Singapore, 1999.

[78] U. H. E. Hansmann, Y. Okamoto, and F. Eisenmenger, *Molecular dynamics, Langevin and hydrid Monte Carlo simulations in a multicanonical ensemble*, Chem. Phys. Lett. **259** (1996), 321–330.

[79] D. .J. Hardy, J. .E. Stone, K. L. Vandivort, D. Gohara, C. Rodrigues, and K. Schulten, *Fast molecular electrostatics algorithms on GPUs*, GPU Computing Gems (In Wen mei Hwu, ed.), Morgan Kaufmann Publishers, 2011, pp. 43–58.

[80] H.D. Raedt and B.D. Raedt, *Applications of the generalized Trotter formula*, Phys. Rev. A **28** (1983), no. 6, 3575.

[81] H.F. Trotter, *On the product of semi-groups of operators*, Proc. Am. Math. Soc. **10** (1959), 545–551.

[82] H. Fukunishi, O. Watanabe, and S. Takada, *On the Hamiltonian replica exchange method for efficient sampling of biomolecular systems: Aplacation to protein structure prediction*, J. Chem. Phys. **116** (2002), 9058–9067.

[83] H.G. Katzgraber, S. Trebst, D.A. Huse, and M. Troyer, *Feedback-optimized parallel tempering Monte Carlo*, J. Stat. Mech. (2006), P03018.

[84] H. Kamberaj, *Conformational sampling enhancement of replica exchange molecular dynamics simulations using swarm particle intelligence*, J. Chem. Phys. **143** (2015), 124105–8.

[85] _____, *Faster protein folding using enhanced conformational sampling of molecular dynamics simulation*, J. Mol. Graph. Model. **81** (2018), 32–49.

[86] H. Kamberaj and A. van der Vaart, *Multiple scaling replica exchange for the conformational sampling of biomolecules in explicit water*, J. Chem. Phys. **127** (2007), 234102–234109.

[87] _____, *An optimised replica exchange method for molecular dynamics simulations*, J. Chem. Phys. **130** (2009), 074904.

[88] R. W. .W. Hooft, B. P. van Eijck, and J. Kroon, *An adaptive umbrella sampling procedure in conformational analysis using molecular dynamics and its application to glycol*, J. Chem. Phys. **97** (1992), 6690–6694.

[89] W.G. Hoover, *Canonical dynamics: Equilibrium phase-space distributions*, Phys. Rev. A **31** (1985), 1695–1697.

[90] X. Huang, M. Hagen, B. Kim, R.A. Friesner, R. Zhou, and B.J. Berne, *Replica exchange with solute tempering: efficiency in large scale systems*, J. Phys. Chem. B **111** (2007), 5405–5410.

[91] T. Huber and W. F. van Gunsteren, *SWARM-MD: Searching conformational space by cooperative molecular dynamics*, J. Phys. Chem. A **102** (1998), 5937–5943.

[92] G. Hummer and A. Szabo, *Free energy profiles from single-molecule pulling experiments*, Proc. Natl. Acad. Sci. USA **107** (2010), no. 50, 21441–21446.

[93] P. H. Hünenberger and W. F. van Gunsteren, *Computer simulation of biomolecular systems, theoretical and experimental applications*, Kluwer, Dordrecht, The Netherlands, 1997.

[94] I. Bahar and R. L. Jernigan, *Inter-residue potentials in globular proteins and the dominance of highly specific hydrophilic interactions at close separations*, J. Mol. Biol. **266** (1997), 195–214.

[95] I. Fukuda and H. Nakamura, *Tsallis dynamics using the Nosé-Hoover approach*, Phys. Rev. E **65** (2002), 026105.

[96] A. Irbäck, F. Sjunnesson, and S. Wallin, *Three-helix-bundle protein in a Ramachandran model*, Proc. Natl. Acad. Sci. U.S.A. **97** (2000), 13614–13618.

[97] J. A. Izaguirre, D. P. Catarello, J. M. Wozniak, and R. D. Skeel, *Langevin stabilization of molecular dynamics*, J. Chem. Phys. **114** (2001), no. 5, 2090–2098.

[98] J. Kim and J. E. Straub, *Optimal replica exchange method combined with Tsallis weight sampling*, J. Chem. Phys. **130** (2009), 144114–11.

[99] Q. Johnson, U. Doshi, T. Chen, and D. Hamelberg, *Water's contribution to the energetic roughness from peptide dynamics*, J. Chem. Theory Comput. **6** (2010), 2591–2597.

[100] A. D. MacKerell Jr, D. Bashford, M. Bellot, R. L. Dunbrack Jr., J. D. Evanseck, M. J. Field, S. Fisher, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus, *All-atom empirical potential for molecular modelling and dynamics studies of proteins*, J. Phys. Chem. B **102** (1998), no. 18, 3586–3616.

[101] H. Kamberaj, Ph.D. thesis, Manchester Metropolitan University, Manchester, UK, May 2005.

[102] _____, *A theoretical model for the collective motion of proteins by means of principal component analysis*, Cent. Eur. J. Phys. **9** (2011), no. 1, 96–109.

[103] J. Kästner, *Umbrella sampling*, Advanced Review **00** (2011), 1–11.

[104] J. Kästner and W. Thiel, *Analysis of the statistical error in umbrella sampling simulations by umbrella integration*, J. Chem. Phys. **124** (2006), 234106(1–7).

[105] J. C. Keck, *Statistical investigation of dissociation cross-sections for diatoms*, Discuss. Faraday Soc. **33** (1962), 173–182.

[106] K. Hukushima and K. Nemoto, *Exchange Monte Carlo method and application to spin glass simulations*, J. Phys. Soc. Jpn. **65** (1996), 1604–1608.

[107] K. K. Burusco, N. J. Bruce, I. Alibay, and R. A. Bryce, *Free energy calculations using a swarm-enhanced sampling molecular dynamics approach*, Chem. Phys. Chem. **16** (2015), 3233–3241.

[108] K. Srinivasaraghavan and M. Zacharias, *Enhanced sampling of peptide and protein conformations using replica exchange simulations with a peptide backbone biasing potential*, Proteins: Struct., Funct., Bioinf. **66** (2007), 697–706.

[109] S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, and P. A. Kollman, *The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method*, J. Comput. Chem. **13** (1992), no. 8, 1011–1021.

[110] A. Laio and M. Parrinello, *Escaping free-energy minima*, Proc. Natl. Acad. Sci. USA **99** (2002), 12562–12566.

[111] _____, *Escaping free-energy minima*, Proc. Natl. Acad. Sci. USA **99** (2002), no. 20, 12562–12566.

[112] A. Laio, A. F. L. Rodriguez, F. Gervasio, M. Ceccarelli, and M. Parrinello, *Assessing the accuracy of metadynamics*, J. Phys. Chem. B **109** (2005), 6714–6721.

[113] D. P. Landau, S.-H. Tsai, and M. Exler, *A new approach to Monte Carlo simulations in statistical physics: Wang-Landau sampling*, Am. J. Phys. **72** (2004), no. 10, 1294–1302.

[114] O. F. Lange and H. Grubmüller, *Can principal components yield a dimension reduced description of protein dynamics on long time scales?*, J. Phys. Chem. B **110** (2006), 22842–22852.

[115] O. F. Lange, L. V. Schäfer, and H. Grubmüller, *Flooding in GROMACS: accelerated barrier crossing in molecular dynamics*, J. Comput. Chem. **27** (2006), no. 14, 1693–1702.

[116] X. Li, C.P. O'Brien, G. Collier, N.A. Vellore, F. Wang, and R.A. Latour, *An improved replica-exchange sampling method: Temperature intervals with global energy reassignment*, J. Chem. Phys. **127** (2007), 164116.

[117] Q. Ma and J. A. Izaguirre, *Targeted mollified impulse: A multi scale stochastic integrator for long molecular dynamics simulations*, Multiscale Model. Simul. **2** (2003), no. 1, 1–21.

[118] Q. Ma, J. A. Izaguirre, and R. D. Skeel, *Verlet-I/r-RESPA/Impulse is limited by nonlinear instabilities*, SIAM J. Sci. Comput. **24** (2003), 1951.

[119] E. Marinari and G. Parisi, *Simulated tempering: a new Monte Carlo scheme*, Europhys. Lett. **19** (1992), 451.

[120] J. A. McCammon, S. H. Northrup, M. Karplus, and R. M. Levy, *Helix-coil transitions in a simple polypeptide model*, Biopolymers **19** (1980), 2033–2045.

[121] M.E. Tuckerman, B.J. Berne, and G.J. Martyna, *Reversible multiple time scale molecular dynamics*, J. Chem. Phys. **97** (1992), no. 3, 1990.

[122] M. E. Tuckerman, D. A. Yarne, S. O. Samuelson, A. L. Hughes, and G. J. Martyna, *Exploiting multiple levels of parallelism in molecular dynamics based calculations via modern techniques and software*

paradigms on distributed memory computers, Comput. Phys. Commun. **128** (2000), 333.

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

[124] M. Mills and I. Andricioaei, *An experimentally guided umbrella sampling protocol for biomolecules*, J. Chem. Phys. **129** (2008), 114101(1–13).

[125] P. Minary, M. E. Tuckerman, and G. J. Martyna, *Long time molecular dynamics for enhanced conformational sampling in biomolecular systems*, Phys. Rev. Lett. **93** (2004), no. 150201-4.

[126] M. Karplus and J.A. McCammon, *Molecular dynamics simulations of biomolecules*, Nat. Struct. Biol. **9** (2002), no. 9, 646–652, With corrigenda in Nat. Struct. Biol. 9(10),788 (2002).

[127] M. K. Transtrum, B. B. Machta, K. S. Brown, B. C. Daniels, C. R. Myers, and J. P. Sethna, *Perspective: Sloppiness and emergent theories in physics, biology, and beyond*, J. Chem. Phys. **143** (2015), 010901–13.

[128] M.P. Allen and D.J. Tildesley, *Computer simulation of liquids*, Oxford University Press, 1989.

[129] M. Tuckerman and M. Parrinello, *Integrating the Car-Parrinello equations. I. Basic integration techniques*, J. Chem. Phys. **101** (1994), no. 2, 1302.

[130] W. Nadler and U.H.E. Hansmann, *Generalized ensemble and tempering simulations: A unified view*, Phys. Rev. E **75** (2007), 026109.

[131] N. Nakajima, H. Nakamura, and A. Kidera, *Multicanonical ensemble generated by molecular dynamics simulation for enhanced conformational sampling of peptides*, J. Phys. Chem. B **101** (1997), 817–824.

[132] R. M. Neal, *Sampling from multimodal distributions using tempered transitions*, Statistics and Computing **6** (1996), no. 4, 353–366.

[133] S. Nosé, *A molecular dynamics method for simulations in the canonical ensemble*, Mol. Phys. **52** (1984), 255.

[134] N. Rathore, M. Chopra, and J.J. de Pablo, *Optimal allocation of replicas in parallel tempering simulations*, J. Chem. Phys. **122** (2005), 024111.

[135] H. Nymeyer, S. Gnanakaran, and A. E. Garcia, *Atomistic simulations of protein folding. using the replica exchange algorithm*, Methods in Enzymology **383** (2004), 119–149.

[136] Y. Okamoto and U. H. E. Hansmann, *Thermodynamics of helix-coil transitions studied by multicanonical algorithms*, J. Phys. Chem. **99** (1995), 11276–11287.

[137] S. Oldziej, A. Liwo, C. Czaplewski, J. Pillardy, and H. A. Scheraga, *Optimization of the UNRES force field by hierarchical design of the potential-energy landscape. 2. Off-lattice tests of the method with single proteins*, J. Phys. Chem. B **108** (2004), 16934–16949.

[138] R. Palmer, *Broken ergodicity*, Adv. Phys. **32** (1982), 669–735.

[139] B. Pattanasiri, Y. W. Li, D. P. Landau, T. Wüst, and W. Triampo, *Conformation transitions of a confined lattice protein: A Wang-Landau study*, J. Phys.: Conf. Series **402** (2012), 012048.

[140] T. J. P. Penna, *Traveling salesman problem and Tsallis statistics*, Phys. Rev. E **51** (1995), R1.

[141] J. Pfaendtner, D. Brabduardi, M. Parrinello, T. D. Pollard, and G. A. Voth, *Nucleotide-dependent conformational states of actin*, Proc. Natl. Acad. Sci. USA **106** (2009), 12723–12728.

[142] P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler, *Transition path sampling: Throwing ropes over rough mountain passes, in the dark*, Annu. Rev. Phys. Chem. **53** (2002), 291–318.

[143] J. C. Phillips, J. .E. Stone, K. .L. Vandivort, T. .G. Armstrong, J. M. Wozniak, M. Wilde, and K. Schulten, *Petascale Tcl with NAMD, VMD, and Swift/T*, SC'14 workshop on High Performance Technical Computing in Dynamic Languages, IEEE Press, 2014.

[144] T. N. Phung, J. F. Brady, and G. Bossis, *Stokesian dynamics simulation of brownian suspensions*, J. Fluid Mech. **313** (1996), 181–207.

[145] S. Piana, J. L. Klepeis, and D. E. Shaw, *Assessing the accuracy of physical models used in protein-folding simulations: quantitative evidence from long molecular dynamics simulations*, Curr. Opin. Struct. Biol. **24** (2014), 98–105.

[146] L. C. Pierce, R. S. Ferrer, C. .A. F. de Oliveira, J. A. McCammon, and R. .C. Walker, *Routine access to millisecond time scale events with accelerated molecular dynamics*, J. Chem. Theory Comput. **8** (2012), no. 9, 2997–3002.

[147] A. R. Plastino and C. Anteneodo, *A dynamical thermostatting approach to nonextensive canonical ensembles*, Ann. Phys. **255** (1997), no. 2, 250–269.

[148] P. Liu, B. Kim, R.A. Friesner, and B.J. Berne, *Replica exchange with solute tempering: A method for sampling biological systems in explicit water*, Proc. Natl. Acad. Sci. USA **103** (2005), no. 39, 13749–13754.

[149] P. Liu and G.A. Voth, *Smart resolution replica exchange: An efficient algorithm for exploring complex energy landscapes*, J. Chem. Phys. **126** (2007), no. 4, 045106–6.

[150] C. Predescu, M. Predescu, and C.V. Ciobanu, *On the efficiency of exchange in parallel tempering Monte Carlo simulations*, J. Phys. Chem. B **109** (2005), 4189–4196.

[151] D. Provasi and M. Filizola, *Putative active states of a prototype G-protein-coupled receptor from biased molecular dynamics*, Biophys. J. **98** (2010), 2347–2355.

[152] J. A. Rahman and J. C. Tully, *Puddle-skimming: An efficient sampling of multidimensional configuration space*, J. Chem. Phys. **116** (2002), 8750.

[153] P. Raiteri, A. Laio, F. L. Gervasio, C. Micheletti, and M. Parrinello, *Efficient reconstruction of complex free energy landscapes by multiple walkers metadynamics*, J. Phys. Chem. B. **110** (2006), 3533–3539.

[154] J. Rogal and P. G. Bolhuis, *Multiple state transition path sampling*, J. Chem. Phys. **129** (2008), 224107.

[155] R. Yamamoto and W. Kob, *Replica-exchange molecular dynamics simulation for supercooled liquids*, Phys. Rev. E **61** (2000), 5473–5476.

[156] R. Zhou, *Free energy landscape of protein folding in water: explicit vs. implicit solvent*, Proteins: Struct., Funct., Bioinf. **53** (2003), no. 2, 148–161.

[157] R. Zhou and B.J. Berne, *Can a continuum solvent model reproduce the free energy landscape of a* β*-hairpin folding in water?*, Proc. Natl. Acad. Sci. USA **99** (2002), 12777–12782.

[158] D. Sabo, M. Meuwly, D.L. Freeman, and J.D. Doll, *A constant entropy increase model for the selection of parallel tempering ensembles*, J. Chem. Phys. **128** (2008), 174109.

[159] D. P. Scarpazza, D. J. Ierardi, A. K. Lerer, K. M. Mackenzie, A. C. Pan, J. A. Bank, E. Chow, R. O. Dror, J. P. Grossman, D. Killebrew, M. A. Moraes, C. Predescu, J. K. Salmon, and D. E. Shaw, *Extending the generality of molecular dynamics simulations on a special-purpose machine*, Proceedings of the 27th IEEE International Parallel and Distributed processing Symposium (Boston, MA, USA), IEEE Computer Society, 2013, pp. 933–945.

[160] T. Schlick, *Molecular modeling and simulation*, 2nd ed., Springer, New York, NY, 2010.

[161] T. Schlick, M. Mandziuk, R. D. Skeel, and K. Srinivas, *Nonlinear resonance artefacts in molecular dynamics simulations*, J. Comput. Phys. **140** (1998), no. 1, 1–29.

[162] S. Jang, E. Kim, and Y. Pak, *All-atom level direct folding simulation of a* ββα *miniprotein*, J. Chem. Phys. **128** (2008), 105102.

[163] S. Jang, S. Shin, and Y. Pak, *Replica-exchange method using the generalized effective potential*, Phys. Rev. Lett. **91** (2003), 058305.

[164] R. D. Skeel, G. Zhang, and T. Schlick, *A family of symplectic integrators: stability, accuracy, and molecular dynamics applications*, SIAM J. Sci. Comput. **18** (1997), no. 1, 203–222.

[165] S. L. Seyler and O. Beckstein, *Sampling large conformational transitions: adenylate kinase as a testing ground*, Mol. Sim. **40** (2014), no. 10-11, 855–877.

[166] A. V. Smith and C. K. Hall, *Alpha-helix formation: discontinuous molecular dynamics on an intermediate-resolution protein model.*, Proteins **44** (2001), no. 3, 344–60.

[167] ———, *Assembly of a tetrameric alpha-helical bundle: Computer simulations on an intermediate-resolution protein model*, Proteins **44** (2001), 376.

[168] S. Nosé, *A molecular dynamics method for simulation in the canonical ensemble*, Mol. Phys. **52** (1984), 255.

[169] ———, *A unified formulation of the constant temperature molecular dynamics methods*, J. Chem. Phys. **81** (1984), 511.

[170] M. Stepanova, *Dynamics of essential collective motions in proteins: Theory*, Phys. Rev. E **76** (2007), no. 5, 051918.

[171] J. .E. Stone, D. .J. Hardy, B. Isralewitz, and K. Schulten, *GPU algorithms for molecular modelling*, Scientific Computing with Multicore and Accelerators (D. A. Bader and J. Kurzak, eds.), Chapman & Hall, 2011, pp. 351–371.

[172] J. E. Stone, R. McGreevy, B. Isralewitz, and K. Schulten, *GPU-accelerated analysis and visualisation of large structures solved by molecular dynamics flexible-fitting*, Faraday Discuss. **169** (2014), 265–283.

[173] J. E. Stone, K. L. Vandivort, and K. Schulten, *GPU-accelerated analysis and visualisation on petascale supercomputing platforms*, Proceedings of the 8th International Workshop on Ultrascale Visualization (New York, NY, USA), UltraVis'13, 2013, pp. 6:1–6:8.

[174] T. P. Straatsma and J. A. McCammon, *Multiconfiguration thermodynamic integration*, J. Chem. Phys. **95** (1991), 1175–1188.

[175] S. Trebst, D.A. Huse, and M. Troyer, *Optimizing the ensemble for equilibrium in broad-histogram Monte Carlo*, Phys. Rev. E **70** (2004), 046701.

[176] S. Trebst, M. Troyer, and U.H.E. Hansmann, *Optimized parallel tempering simulations of proteins*, J. Chem. Phys. **124** (2006), 174903.

[177] Y. Sugita and Y. Okamoto, *Replica-exchange molecular dynamics method for protein folding*, Chem. Phys. Lett. **314** (1999), 141–151.

[178] M. Takahashi and M. Imada, *Monte Carlo calculation of quantum systems. II. higher order correction*, J. Phys. Soc. Jpn. **53** (1984), 3765–3769.

[179] T. Nagasima, A. R. Kinjo, T. Mitsui, and K. Nishikawa, *Wang-Landau molecular dynamics technique to search for low-energy conformational space of proteins*, Phys. Rev. E **75** (2007), 066706.

[180] G. M. Torrie and J. P. Valleau, *Monte Carlo free energy estimates using non-Boltzmann sampling: Application to the sub-critical Lennard-Jones fluid*, Chem. Phys. Lett.. **28** (1974), 578–581.

[181] V. Tozzini, *Coarse-grained models for proteins*, Curr. Opin. Struc. Biol. **15** (2005), 144.

[182] V. Tozzini and J. McCammon, *A coarse grained model for the dynamics of the early stages of the binding mechanism of hiv-1 protease*, Chem. Phys. Lett. **413** (2005), 123–128.

[183] V. Tozzini, W. Rocchia, and J. A. McCammon, *Mapping all-atom models onto one-bead coarse-grained models: General properties and applications to a minimal polypeptide model*, J. Chem. Theory Comput. **2** (2006), 667–673.

[184] C. Tsallis, *Possible generalization of Boltzmann-Gibbs statistics*, J. Stat. Phys. **52** (1988), 479–487.

[185] M. .E. Tuckerman and B. J. Berne, *Molecular dynamics algorithm for multiple time scales: Systems with long range forces*, J. Chem. Phys. **94** (1991), 6811.

[186] _____, *Molecular dynamics in systems with multiple time scales: Systems with stiff and soft degrees of freedom and with short and long range forces*, J. Chem. Phys. **95** (1991), 8362.

[187] M. .E. Tuckerman, B. J. Berne, and A. Rossi, *Molecular dynamics algorithm for multiple time scales: Systems with disparate masses*, J. Chem. Phys. **94** (1991), 1465.

[188] M. .E. Tuckerman and G. J. Martyna, *Understanding modern molecular dynamics: Techniques and Applications*, J. Phys. Chem. B **104** (2000), 159–178.

[189] M. .E. Tuckerman, G. J. Martyna, and B. J. Berne, *Molecular dynamics algorithm for condensed systems with multiple time scales*, J. Chem. Phys. **93** (1990), 1287.

[190] T. W. Whitfield, L. Bu, and J. E. Straub, *Generalized parallel sampling*, Physica A: Statistical Mechanics and its Applications **305** (2002), 157–171.

[191] U. H. E. Hansmann and Y. Okamoto, *Generalized-ensemble Monte Carlo method for systems with rough energy landscape*, Phys. Rev. E **56** (1997), no. 2, 2228–2233.

[192] _____, *Tackling the protein folding problem by a generalized-ensemble approach with Tsallis statistics*, Braz. J. Phys. **29** (1999), 187.

[193] A. E. van Giessen and J. E. Straub, *Monte Carlo simulations of polyalanine using a reduced model and statistics-based interaction potentials*, J. Chem. Phys. **122** (2005), 024904.

[194] W. van Gunsteren, D. Bakowies, R. Baron, I. Chandrasekhar, M. Christen, X. Daura, P. Gee, D. P. Geerke, A. Glättli, P. H. Hünenberger, M. A. Kastenholz, C. Oostenbrink, M. Schenk, D. Trzesniak, N. F. A. can der

Vegt, and H. B. Yu, *Biomolecular modeling: Goals, problems, perspectives*, Angew. Chem. Int. Ed. **45** (2006), no. 25, 4064–4092.

[195] A. F. Voter, *Hyperdynamics: Accelerated molecular dynamics of infrequent events*, Phys. Rev. Lett. **78** (1997), 3908.

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

[197] J. S. Wang and R. H. Swendsen, *Replica Monte Carlo simulation of spin glasses*, Phys. Rev. Lett. **57** (1986), 2607–2609.

[198] W.B. Street, D.J. Tildesley, and G. Saville, *Multiple timestep methods in molecular dynamics*, Mol. Phys. **35** (1978), 639.

[199] S. Wells, S. Menor, B. Hespenheide, and M. Thorpe, *Constrained geometric simulation of diffusive motion in proteins*, Phys. Biol. **2** (2005), no. 4, S127–S136.

[200] W.G. Hoover, *Canonical dynamics: Equilibrium phase-space distributions*, Phys. Rev. A **31** (1985), 1695.

[201] S. Wolfram, *A new kind of science*, Wolfram Media, Inc., Champaign, IL, 2002.

[202] T. B. Woolf and A. Grossfield, *Interaction of tryptophan analogs with POPC lipid bilayers investigated by molecular dynamics calculations*, Langmuir **18** (2002), 198–210.

[203] T. Wüst and D. P. Landau, *Optimized Wang-Landau sampling of lattice polymers: Ground state search and folding thermodynamics of HP model proteins*, J. Chem. Phys. **137** (2012), no. 6, 064903.

[204] X. Cheng, G. Cui, V. Hornak, and C. Simmerling, *Modified replica exchange simulation methods for local structure refinement*, J. Phys. Chem. B **109** (2005), 8220–8230.

[205] Y. Xin, U. Doshi, and D. Hamelberg, *Examining the limits of time reweighing and kramer's rate theory to obtain correct kinetics from accelerated molecular dynamics*, J. Chem. Phys. **132** (2010), 224101.

[206] Y. Okamoto, *Designing generalized statistical ensembles for numerical simulations of biopolymers*, J. Mol. Graphics Modell. **22** (2004), 425–439.

[207] Y. Pak and S. Wang, *Folding of a 16-residue helical peptide using molecular dynamics simulation with Tsallis effective potential*, J. Chem. Phys. **111** (1999), 4359.

[208] _____, *Application of a molecular dynamics simulation method with a generalized effective potential to the flexible molecular docking problems*, J. Phys. Chem. B **104** (2000), 354–359.

[209] Y. Sugita, A. Kitao, and Y. Okamoto, *Multidimensional replica-exchange method for free-energy calculations*, J. Chem. Phys. **113** (2000), 6042–6051.

[210] Y. Sugita and Y. Okamoto, *Replica-exchange multicanonical algorithm and multicanonical replica-exchange method for simulating systems with rough energy landscape*, Chem. Phys. Lett. **329** (2000), 261–270.

[211] Y. Ueda, H. Taketomi, and N. Go, *Studies on protein folding, unfolding, and fluctuations by computer simulation. II. A. three-dimensional lattice model of lysozyme*, Biopolymers **17** (1978), 1531–1548.