Eur. Phys. J. B **38**, 651–663 (2004)
DOI: 10.1140/epjb/e2004-00159-0

THE EUROPEAN PHYSICAL JOURNAL B

A path-sampling scheme for computing thermodynamic properties of a many-body system in a generalized ensemble

M. Athènes[a]

Service de Recherches de Métallurgie Physique, Commissariat à l'énergie atomique-Saclay, 91191 Gif-sur-Yvette, France

Received 7 July 2003 / Received in final form 28 January 2004
Published online 8 June 2004 –

Abstract. We propose to compute the thermodynamic properties of many-body systems using a path-sampling Monte Carlo scheme implemented in a generalized path ensemble. Trial paths are generated through an expanded ensemble using a reversible discretization of Langevin's equation of motion. We also show how the systematic errors resulting from the use of a finite time step can rigorously be taken into account in the path-sampling scheme. We find that the degree of convergence of the estimated thermodynamic quantity towards the exact value correlates with the mean acceptance rates of the path-sampling scheme. An application of the path method for simulating glassy systems is finally suggested.

PACS. 82.60.Lf Thermodynamics of solutions – 07.05.Tp Computer modeling and simulation – 64.70.Pf Glass transitions

1 Introduction

Thermodynamic properties of many-body systems can be computed by means of appropriate Monte Carlo processes [1] whose evolution rules must obey two principles, ergodicity and reversibility to insure the numerical convergence of the averages performed over the generated states towards the exact values. The need for saving computing time being a permanent request, the principle of detailed balance is usually satisfied because it insures the direct convergence of generated state distribution towards the equilibrium stationary distribution. In practice, one implements the Metropolis algorithm that consists of rejecting some of the unfavourable generated states (or configurations).

However, computing the thermodynamic potentials requires to implement one of the many methods of generalized ensembles [2,3]. One can for instance couple the many-body system with a known reference system (ideal gas), introduce an adequate "expanded ensemble" describing the the so-introduced hybrid system and then extract the thermodynamic properties by sampling over the expanded ensemble [4–7]. In practice, the various subpartition functions of the expanded ensemble differ from each other by many orders of magnitude. As a result, an appropriate set of bias factors aimed at insuring a subsequent homogeneous sampling over the subensembles included in the expanded ensemble must be introduced. Various iterative procedures [4,8,9] allow to construct the adequate set of bias factors during short preliminary runs. Differences of thermodynamic potentential, such as excess free energies can then be obtained from the biased occupation ratios of subensembles collected during a long run.

The alternative for exploring the phase space of many-body systems consists of generating trajectories by means of an adequate dynamics and sampling over the so-introduced path ensemble [10–14]. The conditions imposed over the dynamics then permit to derive the physical properties of interest from the generated distribution of paths. In practice, it is convenient to impose that the probability flux along a generated trajectory is invariant over a time reversal. For a deterministic path, the flux-reversibility condition would correspond to the invariance of the equations of dynamics with respect to a time reversal. For stochastic paths generated using Langevin-like equations, it corresponds to an extension of detailed balance that accounts for the time symmetry in Markov processes [15,16][1].

The purpose of the present article is to show how to compute thermodynamic properties of a many-body system using a path-sampling scheme for which the trial paths are generated through an expanded ensemble. The challenging issue that motivates this methodological development is to simulate the thermodynamics of glassy systems directly. For example, finding phase coexistence conditions in nuclear glasses is essential for waste management programs. Unfortunately, implementing existing generalized ensemble methods in glassy systems be-

[a] e-mail: mathenes@cea.fr

[1] At variance with detailed balance, the flux-reversibility condition is not compatible with the Metropolis rejection algorithm and does not imply either the diagonalisability of the transition matrix.

comes problematic. Because the microstructure evolves very slowly, the bias factors of the expanded ensemble method should be permitted to vary during the whole production run, which is possible [17] even though non trivial. More importantly, one should also find how to restrain the sampling efficiency in the glass transition region so as to prevent the system from recrystallizing.

The article is organized as follows: the principles of statistical mechanics in expanded ensembles are recalled in Section 2. Section 3 defines a generalized path ensemble and some associated path functions. Section 3.1 describes the way the stochastic paths are constructed by implementing a discretized Langevin's equation. A generalizable approach is indeed proposed in Appendix A for discretizing Langevin-like equations used in path-sampling schemes. In the following section, it is shown how to extract the thermodynamic properties from averages over the constructed path-distribution by means of two distinct methods. The second method is recalled in Appendix B, as it was proposed earlier [18,19] from a view point different from path-sampling. In Section 4, an applicative study of the path ensemble method is presented for a Lennard-Jones fluid whose free energy is calculated. It is finally suggested how to implement the path-sampling methodology in a glassy system.

2 Thermal ensembles

2.1 Canonical ensemble

Let us consider a system of N particles and volume V. We assume that the system is contained in a cubic cell of edge L and define the $6N$ coordinates of a state as

$$\Gamma = (\mathbf{r}, \dot{\mathbf{r}}) = (\mathbf{r}^1, ..., \mathbf{r}^N, \dot{\mathbf{r}}^1, ..., \dot{\mathbf{r}}^N)$$

where the $3N$-vectors $\mathbf{r}$ and $\dot{\mathbf{r}}$ correspond to the particle positions and velocities, respectively. The configurational energy

$$\mathcal{E}(\mathbf{r}) = \sum_{ij} J\left(|\mathbf{r}^i - \mathbf{r}^j|\right) \tag{1}$$

is assumed to be described by a pairwise interaction potential. The summation runs on all particle pairs and $|\mathbf{r}^i - \mathbf{r}^j|$ corresponds to the distance between particles i and j. The Hamiltonian $\mathcal{H}(\mathbf{p}, \mathbf{r}) = \mathbf{p} \cdot \dot{\mathbf{r}} - \mathcal{L}(\dot{\mathbf{r}}, \mathbf{r})$ corresponds to the Legendre transform of the system Lagrangian $\mathcal{L}(\dot{\mathbf{r}}, \mathbf{r}) = \frac{m}{2}|\dot{\mathbf{r}}|^2 - \mathcal{E}(\mathbf{r})$, where the generalized momentum vector is $\mathbf{p} \equiv \partial_{\dot{\mathbf{r}}}\mathcal{L}(\dot{\mathbf{r}}, \mathbf{r})$ and m is the particle mass. Introducing the inverse temperature, β, the canonical partition function is

$$Z(N, V, \beta) = \frac{1}{h^{3N} N!} \int_{-\infty}^{+\infty} d\mathbf{p} \int_0^L d\mathbf{r} \exp\left[-\beta \mathcal{H}(\mathbf{p}, \mathbf{r})\right], \tag{2}$$

where h is Planck's constant. The state density $\mathcal{N}(\mathcal{S}_i) = \exp[-\beta\mathcal{H}(\mathcal{S}_i)]/[h^{3N}N!]$ is called the Boltzmann weight.

2.2 Expanded ensemble

Let $\mathcal{H}^\lambda(\mathbf{p}, \mathbf{r}) = \mathcal{T}(\mathbf{p}) + \lambda\mathcal{E}(\mathbf{r})$ denote a parametrized Hamiltonian where the coupling parameter λ ranges from λ_0 to λ_M and where the kinetic energy $\mathcal{T}(\mathbf{p}) = \frac{1}{2m}|\mathbf{p}|^2$ has not been parametrized to avoid particle momentum rescaling [6]. This allows to define a parametrized isothermal ensemble $\mathcal{Z}_\lambda$ whose partition function is:

$$Z_\lambda = \frac{1}{h^{3N} N!} \int_{-\infty}^{+\infty} d\mathbf{p} \int_0^L d\mathbf{r} \exp\left[-\beta \mathcal{H}^\lambda(\mathbf{p}, \mathbf{r})\right]. \tag{3}$$

Let consider a set $\Lambda = \{\lambda_0, ..., \lambda_n, ...\lambda_M\}$ of $M+1$ increasing values and denote $\mathcal{Z} = \cup_{\lambda\in\Lambda}\mathcal{Z}_\lambda$ the expanded ensemble [4] defined as the union of parametrized isothermal ensembles. The expanded partition function is

$$Z = \sum_{m=0}^{m=M} Z_{\lambda_m} = \sum_{\lambda\in\Lambda}^{\mathcal{S}_k\in\mathcal{Z}_\lambda} \mathcal{N}_\lambda(\mathcal{S}_k) \tag{4}$$

where $\mathcal{N}_\lambda(\mathcal{S}_k) = \exp[-\beta\mathcal{H}^\lambda(\mathcal{S}_k)]/[h^{3N}N!]$ denotes the corresponding state density. The excess free energy can be computed from the following partition function ratio

$$\mathcal{F}_{ex} = -\frac{1}{\beta} \ln \frac{Z_{\lambda_M=1}}{Z_{\lambda_0=0}}, \tag{5}$$

where the "reference" partition function Z_0 which corresponds to a system of N non-interacting particles (ideal gas) possesses an analytical form. Accordingly, the free energy is

$$\mathcal{F} = \mathcal{F}_{ex} - \frac{1}{\beta} \ln \frac{L^{3N}}{\Lambda^{3N} N!} \tag{6}$$

where Λ is de Broglie's wavelength ($\Lambda = \sqrt{h^2\beta/2\pi\, m}$).

In the present study, we will only investigate the computation of the excess free energy. However, any thermodynamic quantity can, in principle, be obtained from a partition function ratio as in equation (5), since it can formally be expressed as a function of the thermodynamic potential and its derivatives. For instance, setting $\lambda_M = 1$ enables to compute the mean internal energy from the following limit

$$\beta\langle\mathcal{E}\rangle = -\lim_{\lambda_0\to 1}\left[(1-\lambda_0)^{-1} \ln \frac{Z_1}{Z_{\lambda_0}}\right]. \tag{7}$$

3 Path ensembles

3.1 Stochastic paths

Let consider the paths $\mathcal{P}_i^j$ consisting of M+1 states $\{(\mathcal{S}_i \in \mathcal{Z}_{\lambda_0=0}), ..., (\mathcal{S}_{k_m} \in \mathcal{Z}_{\lambda_m}), ..., (\mathcal{S}_j \in \mathcal{Z}_{\lambda_M=1})\}$ where λ_m runs through Λ. They are generated througout the expanded ensemble $\mathcal{Z}$ with the stochastic algorithm described below. We note Ω^{01} the "phase space" of such paths and Ω_{ij}^{01} the (infinite) subspace of paths $\mathcal{P}_i^j$ connecting a given $\mathcal{S}_i \in \mathcal{Z}_0$ to a given $\mathcal{S}_j \in \mathcal{Z}_1$.

Path are constructed by means of Langevin's equation of motion [20]. A symplectic Langevin algorithm based on the leapfrog scheme [21] has been derived in Appendix A for a system at constant temperature, volume and particle number and with a conservative potential. In the expanded ensemble, the coupling parameter λ_n is indexed to the step number of the Langevin algorithm. Hence, at a given step, the total force used to update the momenta is the sum of the interparticle force, gradient of the parametrized Hamiltonian $\mathcal{H}^\lambda$ at position $\mathbf{r}_n$,

$$\mathbf{f}_n^\lambda = -\nabla_{\mathbf{r}_n}\mathcal{H}^{\lambda_n}(\mathbf{p}_n,\mathbf{r}_n) = \lambda_n \mathbf{f}_n$$

and the fluctuating Langevin force $\mathbf{l}_n$ generated with probability

$$b_n^\pm = \left[\frac{\alpha\beta\Delta t}{4m\pi\tilde{\gamma}}\right]^{\frac{3N}{2}} \exp\left\{ -\frac{\alpha\beta}{2m}\frac{2}{\tilde{\gamma}\Delta t} \times \left|\mathbf{l}_n\frac{\Delta t}{2} \pm \tilde{\gamma}\frac{\Delta t}{2}\left(\mathbf{p}_{n\mp\frac{1}{2}} \pm \lambda_n\mathbf{f}_n\frac{\Delta t}{2}\right)\right|^2\right\} \quad (8)$$

where $\mathbf{p}_{q+\frac{1}{2}} = m(\mathbf{r}_{q+1}-\mathbf{r}_q)/\Delta t$, $\tilde{\gamma}$ and $\alpha^{-1} = 1-\tilde{\gamma}\Delta t/2$ are respectively the half-step momenta, a discretized friction and a correcting factor (for more details see Appendix A where Eq. (8) corresponds to either Eqs. (34) or (35)). The paths can be generated either from $\mathcal{Z}_0$ to $\mathcal{Z}_1$ with the positive time step Δt or from $\mathcal{Z}_1$ to $\mathcal{Z}_0$ with $-\Delta t$. They consist in $N_{\text{step}} = M-1$ Langevin steps plus two additional half-steps on the momenta at the path extremities. The first and last half-steps are implemented setting the Langevin force to zero so that $\mathbf{p}_{\frac{1}{2}}$ or $\mathbf{p}_{M-\frac{1}{2}}$ can be univocally defined from $\mathcal{S}(i)_0 = \mathcal{S}_i$ or $\mathcal{S}(i)_M = \mathcal{S}_j$, respectively. In this way, the reversibility of a whole Langevin path is preserved.

Substituting

$$\mathbf{p}_{n-\frac{1}{2}} + \mathbf{p}_{n+\frac{1}{2}} = \frac{m}{\Delta t}\left[\mathbf{r}_{n+1}-\mathbf{r}_{n-1}\right] \quad (9)$$

$$-\lambda_n\mathbf{f}_n + \frac{m}{\Delta t^2}(\mathbf{r}_{n+1}+\mathbf{r}_{n-1}-2\mathbf{r}_n) = \mathbf{l}_n \quad (10)$$

into equation (8) permits to regroup the symmetric and asymmetric contributions into the probability $\prod_{n=1}^{M-1} b_n^\pm$ to generate $\mathcal{P}_i^j$ for the positive or negative time directions

$$P_{gen}^\pm(\mathcal{P}_i^j) \propto \exp\left\{-\frac{\beta}{2\omega}\sum_{n=1}^{M-1}\left|\mathbf{l}_n\frac{\Delta t}{2} \pm \omega\frac{\mathbf{r}_{n+1}-\mathbf{r}_{n-1}}{2\Delta t}\right|^2\right\} \propto \exp\left\{-\frac{\beta}{2}\left[\tilde{\mathcal{S}}^{ij} \pm \mathcal{Q}^{i\to j}\right]\right\} \quad (11)$$

where $\omega = \frac{m\tilde{\gamma}\Delta t}{2\alpha^{-1}}$. The symmetric and asymmetric contributions are

$$\mathcal{S}^{ij} = \frac{1}{2}\sum_{n=1}^{M-1}\left\{|\mathbf{l}_n|^2\frac{\Delta t^2}{2\omega} + \left|\frac{\mathbf{r}_{n+1}-\mathbf{r}_{n-1}}{2}\right|^2\frac{2\omega}{\Delta t^2}\right\}$$

$$\mathcal{Q}^{i\to j} = \sum_{n=1}^{M-1}\mathbf{l}_n\cdot\frac{\mathbf{r}_{n+1}-\mathbf{r}_{n-1}}{2} \quad (12)$$

and can be respectively interpreted as a classical mechanical action (defined over a non-classical Lagrangian) and a heat transfered from the thermostat. The two possible probability fluxes can be conveniently expressed as a function of a symmetric path energy $\tilde{\mathcal{U}}^{ij} = \mathcal{H}(j)+\mathcal{H}(i)+\tilde{\mathcal{S}}^{ij}$ and an effective work $\tilde{\mathcal{W}}^{i\to j} = \mathcal{H}(j)-\mathcal{H}(i)-\mathcal{Q}^{i\to j}$, as follows

$$\mathcal{K}_0(\mathcal{P}_i^j) = \mathcal{N}_0(\mathcal{S}_i)P_{gen}^+(\mathcal{P}_i^j) = \kappa\exp\left[-\frac{\beta}{2}\left(\tilde{\mathcal{U}}^{ij}-\tilde{\mathcal{W}}^{i\to j}\right)\right]$$

$$\mathcal{K}_1(\mathcal{P}_i^j) = \mathcal{N}_1(\mathcal{S}_j)P_{gen}^-(\mathcal{P}_i^j) = \kappa\exp\left[-\frac{\beta}{2}\left(\tilde{\mathcal{U}}^{ij}+\tilde{\mathcal{W}}^{i\to j}\right)\right]$$

where $\kappa = \frac{1}{h^{3N}N!}\left[\frac{\alpha\beta\Delta t}{4m\pi\tilde{\gamma}}\right]^{\frac{3N(M-1)}{2}}$.

3.2 Path statistics

A path density and a path distribution, analog of Boltzmann's weight and Gibbs's distribution must now be introduced in order to define the thermodynamic path-ensemble. The density of path $\mathcal{P}_i^j$ is defined to be a probability flux obtained from Langevin's equation [10]. This density can be written as the exponential of an action related to the two introduced variants of Langevin's equation

$$\beta\tilde{\mathcal{U}}\left(\mathcal{P}_i^j\right) = \beta\left[\frac{1}{2}\tilde{\mathcal{U}}^{ij} + \left(\theta-\frac{1}{2}\right)\tilde{\mathcal{W}}^{i\to j}\right].$$

The meaning of the θ parameter will be discussed later. Briefly, this parameter reflects the time asymmetry and is related to the effective work introduced into the present driven system. Note that for conservative systems, the asymmetric effective work was shown to correspond to discretization errors of some orders with respect to the time step [11,12] and was therefore neglected in subsequent studies.

Let $\tilde{Z}_\theta$ designate the normalizing factor for the path analog of the Gibbs distribution. It can be expressed in a compact form by means of the probability fluxes in the positive and negative time direction

$$\tilde{Z}_\theta = \kappa\sum_{\mathcal{P}_i^j\in\Omega^{01}}\exp-\beta\left[\frac{1}{2}\tilde{\mathcal{U}}^{ij} + \left(\theta-\frac{1}{2}\right)\tilde{\mathcal{W}}^{i\to j}\right] \quad (13)$$

$$= \sum_{\mathcal{P}_i^j\in\Omega^{01}}[\mathcal{K}_0(\mathcal{P}_i^j)]^{1-\theta}[\mathcal{K}_1(\mathcal{P}_i^j)]^\theta. \quad (14)$$

The path density comprises a state density contribution $[\mathcal{N}_0(\mathcal{S}_i)]^{1-\theta}[\mathcal{N}_1(\mathcal{S}_j)]^\theta$ ($\mathcal{S}_i\in\mathcal{Z}_0$ and $\mathcal{S}_j\in\mathcal{Z}_1$ belong to $\mathcal{P}_i^j$) that appears delocalized for intermediate θ-values: the effect of the θ parameter is to constrain the distribution of the terminal states of the path. For the two extremal θ-values 0 and 1, the generated states $\mathcal{S}_i\in\mathcal{P}_i^j$ or $\mathcal{S}_j\in\mathcal{P}_i^j$ are distributed according to $\mathcal{Z}_0$ or $\mathcal{Z}_1$, respectively, as a result of the normalization of the generating probabilities. The physical situation for these two values clearly corresponds to transient nonequilibrium thermodynamics. We

temporarily digress from equilibrium to nonequilibrium thermodynamics here because, as will be shown in Section 4, artificially controlling the mean entropy production will enable us to optimize the efficiency of computations aimed at extracting equilibrium thermodynamic properties.

Let $\sigma^{i\to j} = \tilde{\mathcal{W}}^{i\to j} - \mathcal{F}_{ex}$ designate the entropy production that results from increasing the coupling parameter λ from 0 to 1 along path $\mathcal{P}_i^j$. Since the factor $\beta(\theta - 1/2)$ in equation (13) corresponds to the Lagrange multiplier that allows to maximize the path information entropy [22] with a constraint on the mean effective work, standard thermodynamic averages or relations can be defined or derived with respect to θ in this generalized path ensemble. One then defines the mean effective work and the probability of entropy production $\sigma^{0\to 1}$, as $\langle\tilde{\mathcal{W}}^{i\to j}\rangle_\theta = -\beta^{-1}\partial_\theta \log \tilde{Z}_\theta$ and $P^\theta_{ent}(\sigma^{0\to 1}) = \langle\delta(\tilde{\mathcal{W}}^{i\to j} - \mathcal{F}_{ex} - \sigma^{0\to 1})\rangle_\theta$, respectively. Then, for transient non-equilibrium thermodynamics, as for steady-state non-equilibrium thermodynamics [22], Carnot's second principle and the fluctuation theorem concerning the distribution of entropy production [23–25] are direct consequences of a Gibbs statistical approach

$$\int_0^1 d\theta \langle\sigma^{i\to j}\rangle_\theta = 0 \tag{15}$$

$$\partial_\theta\langle\sigma^{i\to j}\rangle_\theta = -\beta\langle\left[\sigma^{i\to j} - \langle\sigma^{i\to j}\rangle_\theta\right]^2\rangle_\theta \leq 0 \tag{16}$$

$$\frac{P^0_{ent}(\sigma^{0\to 1})}{P^1_{ent}(\sigma^{0\to 1})} = \exp\beta\sigma^{0\to 1} \tag{17}$$

which stems again from the normalization of the generating probabilities ($\tilde{Z}_1/\tilde{Z}_0 = \exp -\beta\mathcal{F}_{ex}$).

The second principle results from equation (15) and inequality (16) and states that paths generated with respect to the $\tilde{Z}_0$ or $\tilde{Z}_1$ distribution possesses a positive mean entropy production relative to the positive or negative direction of time, respectively. Moreover, the higher is the mean entropy production and the more disperse are expected to be the entropy productions $\sigma^{i\to j}$ with respect to $\tilde{Z}_0$ or $\tilde{Z}_1$ statistics as a result of equality (16). The fluctuation theorem (17) relates the ratio of the probabilities of having a given entropy production $\sigma^{0\to 1}$ in the positive time direction to that of having $\sigma^{1\to 0} = -\sigma^{0\to 1}$ in the negative time direction. It also shows how the probability of violation of the second principle becomes exponentially small as the system size increases.

3.3 Path-sampling scheme

Let us now focus on the Monte Carlo sampling of the $\tilde{Z}_\theta$ path ensemble. Following Dellago et al. [12], trial paths will be generated either from $\mathcal{Z}_0$ to $\mathcal{Z}_1$ with the positive time direction or from $\mathcal{Z}_1$ to $\mathcal{Z}_0$ with the negative time direction. Let $\mathcal{P}_i^{j'}$ denote a trial path generated from $\mathcal{Z}_0$ to $\mathcal{Z}_1$ with the positive time direction and $P^+_{acc}(\mathcal{P}_i^{j'})$ the a posteriori acceptance probability of the Monte Carlo scheme. This probability must obey a detailed balance equation considered with respect to the $\tilde{Z}_\theta$ distribution

$$\left[\mathcal{K}_0(\mathcal{P}_i^j)^{1-\theta}\mathcal{K}_1(\mathcal{P}_i^j)^\theta\right] P^+_{gen}(\mathcal{P}_i^{j'})P^+_{acc}(\mathcal{P}_i^{j'}) = \left[\mathcal{K}_0(\mathcal{P}_i^{j'})^{1-\theta}\mathcal{K}_1(\mathcal{P}_i^{j'})^\theta\right] P^+_{gen}(\mathcal{P}_i^j)P^+_{acc}(\mathcal{P}_i^j) \tag{18}$$

where the Monte Carlo a priori probabilities coincide with the conditional probabilities $P^+_{gen}(\mathcal{P}_i^j)$ or $P^+_{gen}(\mathcal{P}_i^{j'})$. This feature results from the fact that there exists a bijection between the path subspace $\sum_j \Omega^{01}_{ij}$ and the set of random number sequences used to construct a path in the positive time direction: a random number sequence defines the trial trajectory of the forwards chain, and, the current trajectory must also univocally define the random number sequence that should be used to compute $P^+_{gen}(\mathcal{P}_i^j)$ in the backwards chain. From equation (18), one deduces

$$P^+_{acc}(\mathcal{P}_i^{j'}) = \min\left(1, \left[\frac{\mathcal{K}_1(\mathcal{P}_i^{j'})}{\mathcal{K}_0(\mathcal{P}_i^{j'})}\frac{\mathcal{K}_0(\mathcal{P}_i^j)}{\mathcal{K}_1(\mathcal{P}_i^j)}\right]^\theta\right). \tag{19}$$

An acceptance probability for a trial path $\mathcal{P}_{i'}^j$ constructed with the negative time direction from $\mathcal{S}_j \in \mathcal{Z}_1$ is

$$P^-_{acc}(\mathcal{P}_{i'}^j) = \min\left(1, \left[\frac{\mathcal{K}_0(\mathcal{P}_{i'}^j)}{\mathcal{K}_1(\mathcal{P}_{i'}^j)}\frac{\mathcal{K}_1(\mathcal{P}_i^j)}{\mathcal{K}_0(\mathcal{P}_i^j)}\right]^{1-\theta}\right), \tag{20}$$

which is similarly deduced from a detailed balance condition involving the negative time direction

$$\left[\mathcal{K}_0(\mathcal{P}_i^j)^{1-\theta}\mathcal{K}_1(\mathcal{P}_i^j)^\theta\right] P^-_{gen}(\mathcal{P}_{i'}^j)P^-_{acc}(\mathcal{P}_{i'}^j) = \left[\mathcal{K}_0(\mathcal{P}_{i'}^j)^{1-\theta}\mathcal{K}_1(\mathcal{P}_{i'}^j)^\theta\right] P^-_{gen}(\mathcal{P}_i^j)P^-_{acc}(\mathcal{P}_i^j). \tag{21}$$

Ergodicity in the path ensemble is guaranteed by the fact that paths are generated in both directions of time. However, if trial paths are always generated in the positive time direction, the states $\mathcal{S}_i \in \mathcal{P}_i^j$ from which the trial path $\mathcal{P}_i^{j'}$ is initiated must be permitted to explore $\mathcal{Z}_0$ otherwise ergodicity is not guaranteed. Thus one has to perform a few Metropolis Monte Carlo steps in $\mathcal{Z}_0$ in order to generate a state $\mathcal{S}_{i'} \neq \mathcal{S}_i$ from which the trial path $\mathcal{P}_{i'}^{j'}$ will be generated. If the additional configurational Monte Carlo moves that explores $\mathcal{Z}_0$ obey detailed balance, then the following bias

$$\frac{P_{gen}(\mathcal{S}_{i'}\to\mathcal{S}_i)P_{acc}(\mathcal{S}_{i'}\to\mathcal{S}_i)}{P_{gen}(\mathcal{S}_i\to\mathcal{S}_{i'})P_{acc}(\mathcal{S}_i\to\mathcal{S}_{i'})} = \frac{\mathcal{N}_0(\mathcal{S}_i)}{\mathcal{N}_0(\mathcal{S}_{i'})}, \tag{22}$$

has to be incorporated into the acceptance probability of equation (19). This case will be encountered in Section 4.4 for the monodirectional path-sampling scheme.

3.4 Expanded path method

Access to the free energy profile, i.e. to the values of Z_λ with λ ranging from 0 to 1, can be achieved by extending

the thermodynamic perturbation identity [1,26] to path ensembles. The successive paths generated by the sampling scheme are distributed according to the $\tilde{Z}_\theta$ equilibrium distribution, but they can also be represented with respect to a set of path distributions whose normalizing functions $\hat{Z}_\lambda$ coincide with the partition functions Z_λ.

This leads us to introduce a second path distribution whose path density possesses a localized state density $\mathcal{N}_\lambda(\mathcal{S}_k)$ where $\mathcal{S}_k \in \mathcal{Z}_\lambda$ belongs to $\mathcal{P}_i^j$. Let $\Omega_k^{\lambda\nu}$ denote the subspace of truncated paths connecting $\mathcal{S}_k \in \mathcal{Z}_\lambda$ with any state $\mathcal{S}_l \in \mathcal{Z}_\nu$, and note $P^{\lambda\to\nu}$ the probability to generate the truncated path from $\mathcal{Z}_\lambda$ to $\mathcal{Z}_\nu$ with the positive time direction if $\lambda < \nu$, or the negative time direction if $\lambda > \nu$. Then, taking advantage of the normalization of the generating probabilities, one defines the projection function

$$\hat{Z}_\lambda = \sum_{\mathcal{P}_i^k \in \Omega_k^{\lambda 0} \mathcal{P}_k^j \in \Omega_k^{\lambda 1}}^{\mathcal{S}_k \in \mathcal{Z}_\lambda} P_{gen}^{\lambda\to 0}(\mathcal{P}_i^k)\mathcal{N}_\lambda(\mathcal{S}_k)P_{gen}^{\lambda\to 1}(\mathcal{P}_k^j), \quad (23)$$

that coincides with the partition function Z_λ. The values $\hat{Z}_\lambda/\tilde{Z}_\theta \equiv Z_\lambda/\tilde{Z}_\theta$ can be monitored for a range of λ values in $[0,1]$ by projecting the paths on $\hat{Z}_\lambda$ which gives access to the free energy profile.

As an illustration of the method, the excess free energy can be computed from the following identity

$$\exp\left[-\beta\mathcal{F}_{ex}\right] = \frac{\langle\left[\mathcal{K}_1(\mathcal{P}_{\{s\}})/\mathcal{K}_0(\mathcal{P}_{\{s\}})\right]^{1-\theta}\rangle_\theta}{\langle\left[\mathcal{K}_0(\mathcal{P}_{\{s\}})/\mathcal{K}_1(\mathcal{P}_{\{s\}})\right]^{\theta}\rangle_\theta} \quad (24)$$

where the $\mathcal{P}_{\{s\}}$'s are the successive paths of the Markov chain. The θ subscript with the brackets indicates that the paths are distributed according to the $\tilde{Z}_\theta$ equilibrium statistics. The numerator and the denominator are simultaneously computed. The former corresponds to the $\tilde{Z}_1/\tilde{Z}_\theta$ value while the latter to the $\tilde{Z}_0/\tilde{Z}_\theta$ value, recalling that $\tilde{Z}_0 = \hat{Z}_0 = Z_0$, $\tilde{Z}_1 = \hat{Z}_1 = Z_1$ and $\beta\mathcal{F}_{ex} = -\ln\frac{Z_1}{Z_0}$.

An alternative for computing the excess free energy is the residence weight method [18,19], recalled in Appendix B. This method requires to always reverse the time direction after a Metropolis test.

3.5 Algorithmics

Noting $\mathcal{P}_{\{s\}}$ the paths of the generated chain, we define $\tilde{\mathcal{W}}_{\{s\}} = \tilde{\mathcal{W}}^{i\to j} = -\tilde{\mathcal{W}}^{j\to i}$ the corresponding works considered in the positive time direction ($\mathcal{P}_{\{s\}} \equiv \mathcal{P}_i^j$, $\mathcal{S}_i \in \mathcal{Z}_0$ and $\mathcal{S}_j \in \mathcal{Z}_1$). It will prove convenient to introduce the work function $\phi_{x\{s\}} = \exp\left[\beta(\theta - x)\tilde{\mathcal{W}}_{\{s\}}\right]$, $x \in [0,1]$ and to materialize the time direction with ν that will be either 0 or 1 depending on whether time flows positively or negatively. From $\mathcal{P}_{\{s\}}$, the generation of a trial path $\mathcal{P}'_{\{s+1\}}$ and the determination of the ensuing path $\mathcal{P}_{\{s+1\}}$ are carried out with the following pseudo-programming code

(1) one considers the initial momenta $\mathbf{p}_0 = \mathbf{p}(i)$ or $\mathbf{p}_M = \mathbf{p}(j)$ and computes $\mathbf{p}_{\frac{1}{2}} = \mathbf{p}_0 + \mathbf{f}_0^\lambda \Delta t/(2)$ or $\mathbf{p}_{M-\frac{1}{2}} = \mathbf{p}_M - \mathbf{f}_M^\lambda \Delta t/(2)$, depending on the time direction;
(2) one performs $M-1$ times the following steps: starting from $n = 0$ for the positive time direction, or, from $n = M$ for the negative time direction,
  (i) one updates the increment $n := n \pm 1$,
  (ii) one computes the new position $\mathbf{r}_n = \mathbf{r}_{n\mp 1} \pm \mathbf{p}_{n\mp\frac{1}{2}}\frac{\Delta t}{m}$,
  (iii) one generates the Langevin force $\mathbf{l}_n$ using equation (8),
  (iv) one computes the new momenta $\mathbf{p}_{n\pm\frac{1}{2}} = \mathbf{p}_{n\mp\frac{1}{2}} \pm (\mathbf{f}_n^\lambda + \mathbf{l}_n)\Delta t$;
(3) one computes the final momenta $\mathbf{p}(j') = \mathbf{p}_M = \mathbf{p}_{M-\frac{1}{2}} + \mathbf{f}_M^\lambda \Delta t/(2)$ or $\mathbf{p}(i') = \mathbf{p}_0 = \mathbf{p}_{\frac{1}{2}} - \mathbf{f}_M^\lambda \Delta t/(2)$, depending on the time direction;
(4) one defines the trial path $\mathcal{P}'_{\{s+1\}}$ as $\mathcal{P}_i^{j'}$ or $\mathcal{P}_{i'}^j$ depending on the direction of time ν, evaluates the effective work $\tilde{\mathcal{W}}'_{\{s+1\}}$ along $\mathcal{P}'_{\{s+1\}}$ in the positive time direction and computes $\phi'_{\nu\{s+1\}} = \exp\left[\beta(\theta-\nu)\tilde{\mathcal{W}}'_{\{s+1\}}\right]$;
(5) one chooses $\mathcal{P}_{\{s+1\}}$ randomly between $\mathcal{P}'_{\{s+1\}}$ with probability $\min\left(1, \phi_{\nu\{s\}}/\phi'_{\nu\{s+1\}}\right)$ and $\mathcal{P}_{\{s\}}$ with probability $\max\left(0, 1 - \phi_{\nu\{s\}}/\phi'_{\nu\{s+1\}}\right)$;
(6) one increments the path weights

$$\pi_0 := \pi_0 + \phi_{0\{s+1\}}$$
$$\pi_1 := \pi_1 + \phi_{1\{s+1\}};$$

(7) if the time direction is always reversed after the Metropolis test, one increments the corresponding residence weight

$$\tau_\nu := \tau_\nu + \min\left(\phi_{\nu\{s+1\}}, \phi_{\nu\{s\}}\right)$$

if acceptance, or in case of rejection

$$\tau_\nu := \tau_\nu + \phi'_{\nu\{s+1\}} - \phi_{\nu\{s\}}.$$

Estimates of the excess free energy $\beta\mathcal{F}_{ex}$ are given by $\ln\frac{\pi_1}{\pi_0}$ or $\ln\frac{\tau_1}{\tau_0}$. They are obtained by averaging the effective work over numerous transformations carried out in one or both directions of time.

Note that carrying out transformations in the positive direction of time only with $\theta = 0$ corresponds to a discrete version of Jarzinski's fast-growth method [27,28] (and extended to discrete Metropolis paths [29]). As this method considers continuous-time stochastic dynamics, the direct work $\mathcal{W}^{0\to 1} = \sum_{n=0}^{M-1}\left[\mathcal{H}_n^{\lambda_{n+1}} - \mathcal{H}_n^{\lambda_n}\right]$ is derived instead of the present effective work. The difference between the effective and direct works corresponds to the "discretization error" that results from the use of a finite time step into Langevin's equation. Monitoring the direct work will enable to control the discretization errors and subsequently to optimize the path-sampling efficiency. Note also that the Metropolis acceptation procedure is not taken into account in the fast-growth method, and, as a result, this

method does not achieve importance sampling [19], unlike generalized ensemble methods for which it is the key ingredient.

4 Simulations

4.1 Lennard-Jones system

The path-sampling scheme is implemented in a three-dimensional Lennard-Jonesian fluid and the excess free energy is computed at a relatively high density and low temperature. We have used as a benchmark, the case study in reference [30] and thus used the same set of potential and simulation parameters.

The usual Lennard-Jones interaction potential, for a pair of particles separated by a distance r, is given by

$$J_{LJ}(r) = 4\epsilon\left[(\sigma/r)^{12} - (\sigma/r)^{6}\right] \tag{25}$$

where ϵ is the depth of the potential at its minimum, and σ is the van der Waals diameter of the particle. A modified potential was instead used:

$$J(r) = \begin{cases} a - br^2, & 0 \le r \le 0.8\sigma \\ J_{LJ}(r) + c(r - r_c) - d & 0.8\sigma \le r \le r_c \\ 0 & r_c \le r \end{cases}$$

where the cut-off distance, r_c, is taken to be half the length, L, of a side of the computational cell and the constant a, b, c and d are chosen to preserve the continuity of J and its derivative. The size L of the computational cell is $5.3 \times \sigma$, the temperature is $\beta^{-1} = \epsilon$ and the number of particle, 125. The particle mass is 6.63×10^{-26} kg and determines the time scale in seconds.

4.2 Simulation parameters

The computational efficiency will now be investigated as a function of the simulation parameters, $\{\lambda_n, 0 \le n \le M\}$, $\tilde{\gamma}\Delta t$, Δt, $N_{\text{step}} \equiv M - 1$ and θ. However, the first set of parameters has been chosen from phenomenological arguments. As the largest part of the work required for transforming the system from $\mathcal{Z}_0$ to $\mathcal{Z}_1$ consists of removing short distance interactions in the near ideal gas regime, the choice $\lambda_n = (\frac{n}{M})^2$ appeared appropriate. In practice, it was observed to approximately balance the acceptance rates τ^+_{acc} and τ^-_{acc} with respect to the positive and negative directions of time.

Since a convenient diagnosing tool to check the convergence of the estimated free energy was shown [19] to be the degree of overlap between paths generated in the positive versus negative time directions, alternating the direction of time after each Metropolis test was first imposed and the residence weight method was used with the sampling parameter θ set to 1/2 corresponding to the same intermediate value as in the previous study [19].

4.3 Bidirectional scheme

Table 1 yields the results of the free energy computations for various values of the discretized friction, time step and number of steps. The friction strength via the $\tilde{\gamma}\Delta t$ parameter was first set to 10^{-1}. The number of Langevin steps N_{step} required to obtain high enough a mean acceptance rate and thus converged free energy estimations was, in this case, found to be about 10^6 or 10^7 depending on the time step Δt.

When the number of Langevin steps is 10^6 and the time step is $\Delta t = 1$ fs or $\Delta t = 2$ fs, the discretization errors (measured from the difference between the effective and direct works) are negligible. However, the acceptance rate is low. This is attributed to the insufficient heat transfer towards the reservoir during the too fast quenches or annealing. An additional simulation has been carried out by increasing the number of Langevin steps to $N_{\text{step}} = 10^7$. The mean acceptance rates are higher, due to a larger conversion of the internal energy into heat towards the thermostat: the microstructure relaxes more efficiently during these longer transformations (in both physical and CPU times) and the free energy can be well estimated.

When increasing the time step to $\Delta t = 4$ fs with $N_{\text{step}} = 10^6$, we observe that discretization errors become more important than for $\Delta t = 2$ fs and $N_{\text{step}} = 10^6$ of the previous case. However, the standard deviation for the free energy estimation is much narrower and the mean acceptance rates are much higher than for the previous case. This is attributed to the fact that the system is observed to relax more efficiently during the longer transformations (in physical time but at constant CPU time). The case $\Delta t = 4$ fs and $N_{\text{step}} = 10^6$ is almost as efficient as the case $\Delta t = 2$ fs and $N_{\text{step}} = 10^7$, the standard deviation being slightly larger.

The effect of varying the friction strength $\tilde{\gamma}\Delta t$ on the efficiency was also investigated. Simulation results are also reported in Table 1. The results obtained with higher friction show no increase of efficiency. For instance, simulations carried out with $\tilde{\gamma}\Delta t = 1$ (overdamped or diffusion Langevin dynamics) requires 10^6 to 10^7 Langevin steps to obtain high enough acceptance rates and accurate estimations. At variance, the results obtained with a lower friction ($\tilde{\gamma}\Delta t = 10^{-2}$) show an increased efficiency. In this last case, transformations consisting in $N_{\text{step}} = 10^5$ steps were sufficient to yield relatively high acceptance rates and small standard deviations for the estimated excess free energy. When still lowering the friction strength to $\tilde{\gamma}\Delta t = 10^{-3}$, the efficiency was observed to deteriorate (see Tab. 1 for $\Delta t = 8 \times 10^{-2}$). This effect is attributed to the insufficient dissipation and, hence, heat transfer with the thermostat. As a result, the effective work overestimates the reversible work or equivalently entropy production is positive and important.

Figures 1 and 2 illustrate the compromise that must be achieved when choosing the value of the time step at given friction strength, $\tilde{\gamma}\Delta t = 10^{-2}$, and number of Langevin steps, $N_{\text{step}} = 10^5$, between insufficient dissipation for small time steps and too large discretization errors for large time steps. When the discretization errors

Table 1. Excess free energies per particles estimated using the residence weight method as a function of various simulation parameters. The standard deviation of the estimated free energy is computed by grouping the simulation data in N_{block} blocks for σ_0 and $N_{\text{block}}/10$ blocks for σ_1. Each block consists in N_{cycle} cycles where a cycle corresponds to two transformations, one in the positive direction of time and the other in the negative direction.

$\tilde{\gamma}\Delta t$	$\beta\mathcal{F}_{ex}$/at. $\pm\sigma_0 \pm \sigma_1$	Δt (fs)	N_{step}	N_{cycle}	N_{block}	τ^+_{acc} (%)	τ^-_{acc}(%)
10^{-1}	$-1.2\underline{433} \pm \underline{174} \pm \underline{103}$	1	10^6	10	87	06.20	09.89
	$-1.2\underline{437} \pm \underline{171} \pm \underline{112}$	2	10^6	10	61	08.36	07.54
	$-1.2\underline{722} \pm \underline{107} \pm \underline{021}$	2	10^7	1	100	61.00	51.00
	$-1.26\underline{689} \pm \underline{697} \pm \underline{135}$	4	10^6	10	99	34.74	34.24
5×10^{-1}	$-1.2\underline{490} \pm \underline{206} \pm \underline{006}$	4	10^6	10	26	02.69	02.31
10^0	$-1.1\underline{617} \pm \underline{338} \pm \underline{103}$	4	10^6	10	35	00.00	12.60
	$-1.2\underline{418} \pm \underline{188} \pm \underline{028}$	8	10^6	10	37	12.20	16.20
	$-1.2\underline{711} \pm \underline{233} \pm \underline{088}$	4	10^7	1	80	26.30	21.20
	$-1.2\underline{741} \pm \underline{302} \pm \underline{135}$	8	10^7	1	33	12.10	21.20
10^{-2}	$-1.26\underline{256} \pm \underline{966} \pm \underline{439}$	4	10^5	10	132	21.89	27.27
	$-1.26\underline{648} \pm \underline{681} \pm \underline{246}$	6	10^5	10	69	36.23	33.47
	$-1.26\underline{682} \pm \underline{575} \pm \underline{180}$	8	10^5	10	127	38.66	44.57
10^{-3}	$-1.2\underline{624} \pm \underline{113} \pm \underline{012}$	8	10^5	10	26	13.46	17.69

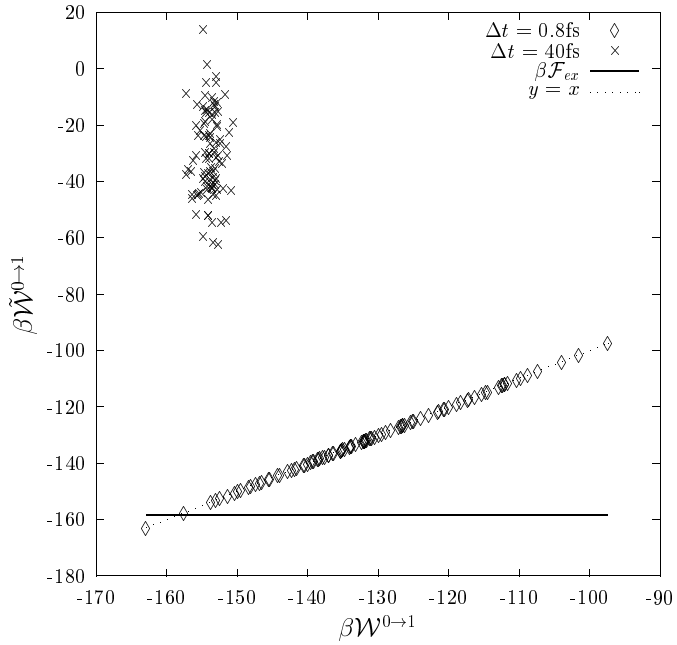

Fig. 1. Normalized effective work $\beta\tilde{\mathcal{W}}^{0\to1}$ after 10^5 Langevin steps as a function of the normalized direct work $\beta\mathcal{W}^{0\to1}$ for a set of 100 quenches.

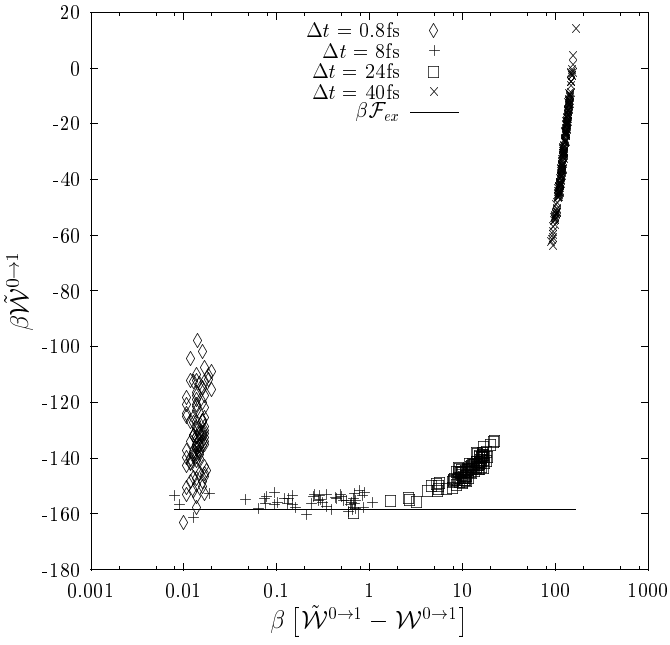

Fig. 2. Normalized effective work $\beta\tilde{\mathcal{W}}^{0\to1}$ after 10^5 Langevin steps as a function of the discretization errors $\beta\left[\tilde{\mathcal{W}}^{0\to1} - \mathcal{W}^{0\to1}\right]$ for a set of 100 quenches.

become too important, the effective works increase dramatically while the standard deviation is much enhanced. Estimated free energies are not reliable in this case.

We now investigate the effect of varying θ on the sampling efficiency. Values ranging from 0 to 1 are possible for θ. However, choosing extremal values would deteriorate the sampling efficiency of the bidirectional scheme which justifies a posteriori the intermediate choice $\theta = 1/2$ made here. To illustrate this point, let us consider that θ is set to 1. Then a path generated in the negative direction of time is first accepted with probability one (as expected from Eq. (20)) but is more likely to be rejected together with the ensuing path generated in the positive direction of time because trial paths will have to be accepted by pairs in the acceptance procedure of equation (19): the probability of insufficient overlapping will be higher yielding smaller mean acceptance rates and a lower computational efficiency. As a consequence, monodirectional schemes should be implemented when the value

Table 2. Excess free energies per particles estimated using the expanded path method as a function of various simulation parameters. Paths have been generated with the positive direction of time only, $\theta = 1$ and $\tilde{\gamma}\Delta t = 10^{-2}$. The standard deviations σ_0 and σ_1 are computed by grouping the corresponding data in $N_{\rm block}$ and $N_{\rm block}/10$ blocks consisting in $2 \times 10^6/N_{\rm steps}$ or $2 \times 10^7/N_{\rm steps}$ trial path generations, respectively.

$N_{\rm step}$	$\beta\mathcal{F}_{ex}$/at. $\pm\sigma_0$ $\pm\sigma_1$	Δt (fs)	$N_{\rm block}$	$\tau^+_{\rm acc}(\%)$
10^5	$-1.26\underline{881} \pm \underline{611} \pm \underline{614}$	0.8	146	00.34
	$-1.2\underline{614} \pm \underline{140} \pm \underline{097}$	2.4	122	02.99
	$-1.2\underline{626} \pm \underline{237} \pm \underline{207}$	8	501	08.08
	$-1.2\underline{661} \pm \underline{193} \pm \underline{151}$	16	200	00.06
	$-1.2\underline{604} \pm \underline{247} \pm \underline{245}$	24	129	0.009
2×10^5	$-1.2\underline{575} \pm \underline{116} \pm \underline{072}$	0.8	168	01.83
	$-1.2\underline{631} \pm \underline{191} \pm \underline{147}$	2.4	170	05.61
	$-1.2\underline{710} \pm \underline{113} \pm \underline{077}$	8	100	22.84
	$-1.2\underline{807} \pm \underline{191} \pm \underline{187}$	16	110	02.25
	$-1.21\underline{153} \pm \underline{768} \pm \underline{338}$	24	232	00.22
5×10^5	$-1.2\underline{653} \pm \underline{132} \pm \underline{115}$	0.8	100	04.00
	$-1.2\underline{594} \pm \underline{159} \pm \underline{132}$	2.4	100	21.25
	$-1.2\underline{661} \pm \underline{149} \pm \underline{126}$	8	100	35.00
	$-1.2\underline{631} \pm \underline{203} \pm \underline{169}$	16	100	03.00
	$-1.1\underline{949} \pm \underline{201} \pm \underline{158}$	24	100	01.25
10^6	$-1.2\underline{694} \pm \underline{132} \pm \underline{116}$	0.8	493	09.72
	$-1.26\underline{881} \pm \underline{973} \pm \underline{562}$	2.4	159	37.18
	$-1.2\underline{670} \pm \underline{112} \pm \underline{087}$	8	237	26.89
	$-1.1\underline{537} \pm \underline{324} \pm \underline{320}$	16	167	04.46
	$-0.4\underline{015} \pm \underline{462} \pm \underline{424}$	24	190	02.62

of the sampling parameter is extremal, the case $\theta = 0$ being analogous.

4.4 Monodirectional scheme

We have only investigated the sampling efficiency for the case where θ is set to 1 and the paths are always generated in the positive direction of time. Table 2 reports free energy estimations obtained using the expanded path method, various time steps and number of steps. The mean acceptance rates τ^+_{acc} are also reported.

In the simulations where $\Delta t = 0.8$ fs with $N_{\rm step}$ equals to 10^5, 2×10^5 and 5×10^5, high standard deviations are obtained for the estimated free energy in spite of small discretization errors. This is because the system has not sufficiently relaxed. In the case where the time step is made too large $\Delta t = 24$ fs, with M equals to 10^5, 2×10^5, 5×10^5 and 10^6, the mean acceptance rates are small resulting from a broad dispersion of the effective works. For intermediate values of the time step and sufficient number of Langevin steps, high acceptance rates and correct free energy estimations can be obtained. Hence, for the monodirectional scheme, the mean acceptance rate also correlates with the degree of convergence of the estimated value.

The observed trend can be rationalized by invoking the nonequilibrium theorems concerning both the mean value and distribution of the entropy production. A broad dispersion of the effective works in the positive direction should result in a a large mean entropy production (Eqs. (15, 16)). Subsequently, almost all paths escape from high probability density regions of the phase space. From the nonequilibrium fluctuation theorem (Eq. (17)), a subsequent path generated in the negative time direction would unlikely return to high probability density regions, but rather continue its excursion in far from equilibrium regions yielding a weak overlap and an important uncertainty on the estimated free energy.

Hence, a narrow dispersion of the effective works measured whether the simulations are carried out with the monodirectional or bidirectional scheme, implies small entropy productions in absolute value, an efficient numerical convergence of the sampling scheme and high a mean acceptance rate. The latter one thus provides a tool for diagnosing the degree of numerical convergence. However, this diagnozing tool may not be a sufficient condition in any cases and one should always analyse standard deviations ultimately so as to make sure that numerical ergodicity in the path space has been reached.

Finally, it is instructive to note that the monodirectional scheme with $\theta = 1$ converges slightly more slowly than the bidirectional scheme with $\theta = 1/2$ as indicated in Table 3. This trend is consistent with the fact that a smaller mean entropy production in absolute value is expected for the intermediate θ-value in [0,1] from equations (15, 16).

Table 3. Excess free energies per particles estimated using the expanded path method (EP) or using the residence weight method (RW). Paths have been generated with either the positive direction of time only, (+), or alternating directions, (A), $\tilde{\gamma}\Delta t = 10^{-2}$, $N_{\rm step} = 10^5$ and $\Delta t = 8$ fs. The standard deviations σ_0 and σ_1 are computed by grouping the collected data in $N_{\rm block}$ and $N_{\rm block}/10$ blocks consisting in 20 or 200 trial path generations, respectively.

Method	$\beta\mathcal{F}_{ex}$/at. $\pm\sigma_0$ $\pm\sigma_1$	$N_{\rm block}$	$\tau_{\rm acc}$ (%)
EP/+	$-1.2\underline{6696} \pm \underline{1543} \pm \underline{0827}$	315	10.60
RW/+	$-1.2\underline{2104} \pm \underline{0871} \pm \underline{0516}$		
EP/A	$-1.2\underline{6941} \pm \underline{1050} \pm \underline{0351}$	336	41.05
RW/A	$-1.2\underline{6852} \pm \underline{0642} \pm \underline{0267}$		

4.5 Towards thermodynamics of metastable systems

The previously described monodirectional scheme may be used as an optimization tool for finding compositions of relative metastability below the glass transition temperature. Metastable states of relatively low energies are constructed by quenching high temperature states, the degree of metastability in the low temperature system being controlled by the quenching rate. Below the glass transition temperature, the structure relaxes so slowly that it is not possible to explore the compositional degrees of freedom directly in the configurational phase space. Composition can only be varied in the high temperature regime above the glass transition temperature before performing a quench.

For this purpose, it should be convenient to consider an expanded grand canonical or expanded Gibbs ensemble $G = \sum_{m=0}^{M} G_{\lambda_m}$ and to generate a path distribution $\{\mathcal{P}_j^i\}_{\theta=1}$ according to the $\tilde{G}_{\theta=1}$ statistics. Then, access to compositions of relative metastability or of coexistence at the low temperature would be guaranteed by the fact that the path terminal states $\mathcal{S}_j \in \mathcal{P}_j^i$ would be directly distributed according to the metastable G_1 distribution, which would result from the fact that θ has been set to 1.

5 Conclusion and perspectives

We have implemented a path-sampling scheme in which the stochastic paths are generated through an expanded ensemble using Langevin's equation of motion. We then showed how to directly compute the thermodynamic properties of a many-body system using either an expanded path method or a residence weight method. The two practical results are that (i) the systematic errors in Langevin's discretized dynamics can be rigorously taken into account in the path-sampling scheme, and that (ii) an efficient numerical convergence of the estimated values towards the exact thermodynamic quantities is insured by small entropy productions (in absolute value) along the generated paths, which requires high a mean acceptance rate for the trial paths in the Metropolis test.

Future studies will focus on the application of the path-sampling methodology to glassy systems where the presence of many local minima prevents from directly exploring the phase space at low temperatures.

A timely correspondence with T. Vlugt about a teaching project allowed me to structure this article. I am also indebted to P.N. Vorontsov-Velyaminov for instructive comments, hospitality and providing me with relevant references. Fruitful discussions with my colleagues G. Adjanor, J.-L. Bocquet, Y. Limoge, L. Martin-Samos and G. Roma are finally acknowledged.

Appendix A: Langevin algorithm

The use of a finite time step in a Langevin dynamics results in ineluctable discretization errors. We show in this appendix how to derive an analytical expression of these errors that can be incorporated into the path-sampling procedure as a bias. Systematic errors will be, in this way, transformed into controllable statistical errors.

The approach consists of deriving the various parameters of the full Langevin equation from the flux-reversibility condition [15] and the form of the stationary distribution which is known a priori. We indeed generalize the way existing Monte Carlo schemes based on the overdamped or diffusion Langevin equation were derived. In these schemes (smart Monte Carlo [31], force-bias Monte Carlo [32], Langevin Monte Carlo [33] and hybrid Monte Carlo Molecular Dynamics [34,35]), the parameters of the deterministic parts were adjusted a posteriori from the detailed balance condition and the discretization errors were also taken into account in the acceptance procedure as a statistical bias.

In Langevin's equation of motion, the generalized force acting upon the particles are the sum of the interparticle forces $\mathbf{f}_n = -\nabla_{\mathbf{r}_n}\mathcal{H}$, the frictional forces and the fluctuating forces, generated from a normal distribution whose exact amplitude will be derived below. The frictional and the fluctuating forces will be combined and described as a single fluctuating force $\mathbf{l}_n$, called the Langevin force. The discretization of Langevin's equation of motion will be based on a leapfrog algorithm. Using the generalized force to increment the momenta at half-step, the leapfrog algorithm with time step Δt can be transposed as follows:

$$\mathbf{r}_n = \mathbf{r}_{n-1} + \mathbf{p}_{n-\frac{1}{2}}\frac{\Delta t}{m} \tag{26}$$

$$\mathbf{p}_{n+\frac{1}{2}} = \mathbf{p}_{n-\frac{1}{2}} + (\mathbf{f}_n + \mathbf{l}_n)\Delta t \tag{27}$$

where $\mathbf{r}_n$ and $\mathbf{p}_{n+\frac{1}{2}}$ are the particle positions and momenta at time $n\Delta t$ and $(n+\frac{1}{2})\Delta t$, respectively. It will be usefull to define integer-step momenta and half-step positions as follows

$$\mathbf{p}_n = [\mathbf{p}_{n-\frac{1}{2}} + \mathbf{p}_{n+\frac{1}{2}}]/2 \tag{28}$$

$$\mathbf{r}_{n+\frac{1}{2}} = [\mathbf{r}_{n+1} + \mathbf{r}_n]/2. \tag{29}$$

Let us first consider the leapfrog algorithm implemented in the microcanonical ensemble. Then, the Langevin force $\mathbf{l}_n$ must be switched off. Neglecting discretization errors, the leapfrog algorithm would clearly be flux-reversible, since there would be equality between the probability fluxes in the positive and negative directions of time: changing the sign of the time step allows to construct the exact reverse trajectory. Imposing now the condition of flux-reversibility to the leapfrog algorithm in the canonical ensemble will allow us to derive the adequate Langevin fluctuating force. The fluctuating force $\mathbf{l}_n$ is generated from two distinct states $\mathcal{S}_{n-\frac{1}{2}} \equiv \left[\mathbf{p}_{n-\frac{1}{2}}, \mathbf{r}_{n-\frac{1}{2}}\right]$ or $\mathcal{S}^-_{n+\frac{1}{2}} \equiv \left[-\mathbf{p}_{n+\frac{1}{2}}, \mathbf{r}_{n+\frac{1}{2}}\right]$ with opposite momenta depending on whether the positive or negative time trajectory is considered. The procedure for generating the Langevin force will thus be adapted so as to make the positive and negative time probability fluxes coincide up to second order with respect to the time step. This requires first to introduce $\mathcal{H}_{n+\frac{1}{2}} = \mathcal{H}(\pm\mathbf{p}_{n+\frac{1}{2}}, \mathbf{r}_{n+\frac{1}{2}})$ the half-step Hamiltonian, b^+_n the conditional probability to flow from $\mathcal{S}_{n-\frac{1}{2}}$ to $\mathcal{S}_{n+\frac{1}{2}}$ and b^-_n the one to flow from $\mathcal{S}^-_{n+\frac{1}{2}}$ to $\mathcal{S}^-_{n-\frac{1}{2}}$. The half-step probability flux ratio between $\mathbf{r}_{n-\frac{1}{2}}$ and $\mathbf{r}_{n+\frac{1}{2}}$ thus corresponds to

$$\frac{\mathcal{K}^{n+\frac{1}{2}}_{n-\frac{1}{2}}}{\mathcal{K}^{n-\frac{1}{2}}_{n+\frac{1}{2}}} = \frac{\exp\left[-\beta\mathcal{H}_{n-\frac{1}{2}}\right] b^+_n}{\exp\left[-\beta\mathcal{H}_{n+\frac{1}{2}}\right] b^-_n}. \tag{30}$$

Let us now describe the generation of the biased fluctuating force used to increment the half-step velocities: a trial fluctuating force is generated $\mathbf{l}_{n'}$ from a normal distribution and is used to construct the corresponding state $\mathcal{S}_{n'}$. So as to insure that the generated state is as representative as possible of the canonical distribution, the generation of the trial state is weighted with the modified state density $C\exp\left[-\alpha\beta H_{n'}\right]$ where a correcting factor α has been introduced here in anticipation of the manipulations to come. In the general case, the weighted distribution can always be constructed using an acceptation-rejection procedure. Here, the position coordinates of the trial state $\mathcal{S}_{n'}$ are independent from the trial fluctuating force $\mathbf{l}_{n'}$ (we have $\mathbf{r}_{n'} \equiv \mathbf{r}_n$ from Eq. (26)), which implies that the density probability distribution can be analytically derived from the product of the two normal distributions:

$$P_1\left[\mathbf{l}_{n'}\right] = \left[\frac{\beta\Delta t}{4m\pi\tilde{\gamma}}\right]^{\frac{3N}{2}} \exp\left[-\frac{\alpha\beta}{2m}\frac{2\alpha^{-1}}{\tilde{\gamma}\Delta t}\left(\mathbf{l}_{n'}\frac{\Delta t}{2}\right)^2\right] \tag{31}$$

$$P_2\left[\mathbf{l}_{n'}\right] = \left[\frac{\alpha\beta\Delta t^2}{8m\pi}\right]^{\frac{3N}{2}} \exp\left[-\frac{\alpha\beta}{2m}|\mathbf{p}_{n'}|^2\right], \tag{32}$$

where the amplitude of the trial fluctuation is controlled by the coefficient, $\tilde{\gamma}$, and $\mathbf{p}_{n'} = \mathbf{p}_{n-\frac{1}{2}} + (\mathbf{f}_n + \mathbf{l}'_n)\Delta t/2$ from equation (27).

Multiplying the two normal probability distributions (Eqs. (31, 32)) yields the conditional probability in the positive direction of time

$$b^+_n = A\exp\left\{-\frac{\alpha\beta}{2m}\left(1+\frac{2}{\alpha\tilde{\gamma}\Delta t}\right)\left|\mathbf{l}_n\frac{\Delta t}{2}\right.\right. \\ \left.\left. + \left(1+\frac{2}{\alpha\tilde{\gamma}\Delta t}\right)^{-1}\left(\mathbf{p}_{n-\frac{1}{2}} + \mathbf{f}_n\frac{\Delta t}{2}\right)\right|^2\right\}, \tag{33}$$

where A is an adequate normalizing constant independent of the state. The standard deviations of the Langevin force can be deduced from the argument of the normal distribution in equation (33)

$$\sigma(\mathbf{l}_n\sqrt{\Delta t}) = \sqrt{\frac{2m\hat{\gamma}}{\alpha\beta}},$$

where $\hat{\gamma} = \frac{2}{\Delta t}\left[1+\frac{2}{\alpha\tilde{\gamma}\Delta t}\right]^{-1}$ is the effective friction. The first moment of the Langevin force corresponds to a frictional force $-\tilde{\tilde{\gamma}}\mathbf{p}$ opposed to the particle velocity. The correcting factor α that determines the two first moments of the Langevin force will be chosen so that the probability flux ratio of equation (30) is as close as possible to 1, which requires to derive this quantity. The half-step conditional probability in the positive direction of time is

$$b^+_n = \left[\frac{\alpha\beta\Delta t}{4m\pi\hat{\gamma}}\right]^{\frac{3N}{2}} \exp\left\{-\frac{\alpha\beta}{2m}\frac{2}{\hat{\gamma}\Delta t} \right. \\ \left. \times\left|\mathbf{l}_n\frac{\Delta t}{2} + \hat{\gamma}\frac{\Delta t}{2}\left(\mathbf{p}_{n-\frac{1}{2}} + \mathbf{f}_n\frac{\Delta t}{2}\right)\right|^2\right\}. \tag{34}$$

The corresponding time-reverse conditional probability can be obtained by similarly expressing the momenta $\mathbf{p}_n$ in equation (32) as a function of $\mathbf{l}_n$ and $\mathbf{p}_{n+\frac{1}{2}}$ by means of equation (28). This is equivalent to substitute $\mathbf{p}_{n-\frac{1}{2}}$ for $-\mathbf{p}_{n+\frac{1}{2}}$ in equation (34). Both ways leads to

$$b^-_n = \left[\frac{\alpha\beta\Delta t}{4m\pi\hat{\gamma}}\right]^{\frac{3N}{2}} \exp\left\{-\frac{\alpha\beta}{2m}\frac{2}{\hat{\gamma}\Delta t} \right. \\ \left. \times\left|\mathbf{l}_n\frac{\Delta t}{2} - \hat{\gamma}\frac{\Delta t}{2}\left(\mathbf{p}_{n+\frac{1}{2}} - \mathbf{f}_n\frac{\Delta t}{2}\right)\right|^2\right\}. \tag{35}$$

The Langevin force in equation (27)

$$\mathbf{l}_n\Delta t = \mathbf{p}_{n+\frac{1}{2}} - \mathbf{p}_{n-\frac{1}{2}} - \mathbf{f}_n\Delta t$$

can be substituted into the probability ratio:

$$b^-_n/b^+_n = \exp\left\{\frac{\alpha\beta}{2m}(1-\hat{\gamma}\Delta t/2)\left[\mathbf{p}_{n+\frac{1}{2}} + \mathbf{p}_{n-\frac{1}{2}}\right]\cdot \right. \\ \left. \times\left[\mathbf{p}_{n+\frac{1}{2}} - \mathbf{p}_{n-\frac{1}{2}} - \mathbf{f}_n\Delta t\right]\right\} \\ = \exp\left\{\frac{\alpha\beta}{(1-\hat{\gamma}\Delta t/2)^{-1}}\mathcal{Q}^+_n\right\} \tag{36}$$

where the effective heat transfered from the thermostat

$$\mathcal{Q}_n^+ = \frac{1}{2m}\left(\mathbf{p}_{n+\frac{1}{2}}^2 - \mathbf{p}_{n-\frac{1}{2}}^2\right) - \frac{\Delta t}{m}\mathbf{p}_n \cdot \mathbf{f}_n$$

has been introduced. Thales's identity

$$\mathbf{p}_n = \frac{m}{2\Delta t}\left[\mathbf{r}_{n+1} - \mathbf{r}_{n-1}\right] = \frac{m}{\Delta t}\left[\mathbf{r}_{n+\frac{1}{2}} - \mathbf{r}_{n-\frac{1}{2}}\right]$$

permits to rewrite the effective heat as

$$\mathcal{Q}_n^+ = \frac{1}{2m}\left(\mathbf{p}_{n+\frac{1}{2}}^2 - \mathbf{p}_{n-\frac{1}{2}}^2\right) - \left[\mathbf{r}_{n+\frac{1}{2}} - \mathbf{r}_{n-\frac{1}{2}}\right] \cdot \mathbf{f}_n.$$

Performing a Taylor expansion of the interatomic potential $\mathcal{E}(\mathbf{r}_n + \mathbf{h}_n)$ around $\mathbf{r}_n$ and substituting $\mathbf{r}_{n\pm\frac{1}{2}} - \mathbf{r}_n$ for $\mathbf{h}_n$ yields

$$\mathcal{E}_{n+\frac{1}{2}} - \mathcal{E}_{n-\frac{1}{2}} = - \left[\mathbf{r}_{n+\frac{1}{2}} - \mathbf{r}_{n-\frac{1}{2}}\right] \cdot \mathbf{f}_n + O\left(\left|\mathbf{r}_{n+\frac{1}{2}} - \mathbf{r}_{n-\frac{1}{2}}\right|^2\right).$$

Assuming now that the forces and velocities are bounded implies that there exists $B > O$ such that $|\mathbf{r}_{n+\frac{1}{2}} - \mathbf{r}_{n-\frac{1}{2}}| < B\Delta t$ for all n which yields $\mathcal{H}_{n+\frac{1}{2}} - \mathcal{H}_{n-\frac{1}{2}} = \mathcal{Q}_n^+ + O(|\Delta t|^2)$. This last identity enables us to choose the adequate correcting factor $\alpha^{-1} = 1 - \hat{\gamma}\Delta t/2$ minimizing the drift of the flux ratio in equation (30). Since $\hat{\gamma}$ depends on both $\tilde{\gamma}$ and α^{-1}, the previous condition yields a second order polynomial $\alpha^{-2} + \alpha^{-1}(\tilde{\gamma}\Delta t/2 - 1) = 0$ with respect to α^{-1}. The unphysical root $\alpha^{-1} = 0$ is to be discarded which implies $\alpha^{-1} = 1 - \tilde{\gamma}\Delta t/2$ and subsequently $\hat{\gamma} = \tilde{\gamma}$. With this choice, the case $\tilde{\gamma}\Delta t = 1$ corresponds to a discretization for the diffusion Langevin equation, since the friction cancels the inertial term. The cases $1 < \tilde{\gamma}\Delta t < 2$ leads to an unphysical dynamics while the cases $\tilde{\gamma}\Delta t \geq 2$ are not possible.

Note that in the previous expansion, the second order term does not cancel because the vector position $\mathbf{r}_n$ is not necessarily the middle of $\mathbf{r}_{n-\frac{1}{2}}$ and $\mathbf{r}_{n+\frac{1}{2}}$. Nevertheless, expanding the potential to second order around $\mathbf{r}_{n+\frac{1}{2}}$,

$$\mathcal{E}_{n+1} - \mathcal{E}_n = (\mathbf{r}_{n+1} - \mathbf{r}_n) \cdot \mathbf{f}_{n+\frac{1}{2}} + O(|\mathbf{r}_{n+1} - \mathbf{r}_n|^3) \quad (37)$$

and then the force $\mathbf{f}_{n+\frac{1}{2}}$ to first order around both $\mathbf{r}_n$ and $\mathbf{r}_{n+1}$

$$\mathcal{E}_{n+1} - \mathcal{E}_n = (\mathbf{r}_{n+1} - \mathbf{r}_n) \cdot (\mathbf{f}_{n+1} + \mathbf{f}_n)/2 + O(|\mathbf{r}_{n+1} - \mathbf{r}_n|^3)$$

allows to cancel the second order term, since $\mathbf{r}_{n+\frac{1}{2}}$ is the middle of $\mathbf{r}_n$ and $\mathbf{r}_{n+1}$. When many Langevin steps are effected, then adequately rearranring the total discretization error implies that an additional order of precision can be obtained. The algorithm is indeed valid to second order.

Note that a more formal way to derive dynamical algorithms consists of decomposing the Liouville operator describing the dynamics [1]. In the present context, our approach can be seen as a second-order Trotter factorisation of Hamilton's equation of motion with a stochastic process on the momenta defined as $\dot{\mathbf{p}} = -\gamma\mathbf{p} + \mathbf{b}$ where γ and $\mathbf{b}$ are a continuous-time friction and a centred normal noise of standard deviation $\sqrt{\frac{2m\gamma}{\beta}}$. Then using a second order Trotter factorization for Hamilton's equation (Verlet's algorithm) and integrating the stochastic process over the time step yields

$$\mathbf{p}_{n+\frac{1}{2}} = \left(\mathbf{p}_{n-\frac{1}{2}} + \mathbf{f}_n\frac{\Delta t}{2}\right)x + \mathbf{b}_n + \mathbf{f}_n\frac{\Delta t}{2} \quad (38)$$

where the discrete centred normal noise $\mathbf{b}_n$ has standard deviation $\sqrt{(1-x^2)\frac{m}{\beta}}$ and $x = \exp -\gamma\Delta t \equiv 1 - \tilde{\gamma}\Delta t$. This connects our discretized friction to the one that would appear in the continuous-time Langevin equation.

Adjusting Langevin's parameters from the flux-reversibility condition as we did is a more intuitive approach that may be fruitfull in complicated dynamics for which the corresponding Liouville operator has not been decomposed into efficient symplectic algorithms yet (for a review see Ref. [36]).

Appendix B: Residence weight method

In the residence weight method, trial paths are alternatively generated in the positive and negative direction of time. Because of this alternation constraint, the successive path of the generated chain can be noted $\mathcal{P}_{2i}^{2i-1}$, $\mathcal{P}_{2i}^{2i+1}$, $\mathcal{P}_{2i+2}^{2i+1}$... with for the corresponding extremal states $\mathcal{S}_{2i} \in \mathcal{Z}_0$ and $\mathcal{S}_{2i+1} \in \mathcal{Z}_1$... The path density of $\mathcal{P}_{2i}^{2i-1}$ will be $\alpha\mathcal{N}_0(\mathcal{S}_{2i})$ and the one of $\mathcal{P}_{2i}^{2i+1}$ will be $\alpha\mathcal{N}_1(\mathcal{S}_{2i+1})$, α being an adequate normalizing factor. The paths are thus alternately considered with respect to the Z_0 or Z_1 subpartition functions that play the role of subprojection functions.

The residence weight algorithm decomposes into three stages: (i) one generates a new trial path $\mathcal{P}_{2i}^{2i-1'}$ with probability $P_{gen}^+(\mathcal{P}_{2i}^{2i-1'})$, initiated from $\mathcal{S}_{2i} \in \mathcal{Z}_0$ of $\mathcal{P}_{2i}^{2i-1}$ projected on Z_0, (ii) one decides whether the trial path is accepted $\mathcal{P}_{2i}^{2i+1} \equiv \mathcal{P}_{2i}^{2i-1'}$ or rejected $\mathcal{P}_{2i}^{2i+1} \equiv \mathcal{P}_{2i}^{2i-1}$. The a posteriori probability associated to this second stage is $P_{acc}^+(\mathcal{P}_{2i}^{2i+1}) = P_{acc}^+(\mathcal{P}_{2i}^{2i-1'})$ or $P_{acc}^+(\mathcal{P}_{2i}^{2i+1}) = 1 - P_{acc}^+(\mathcal{P}_{2i}^{2i-1'})$, depending on whether the path has been accepted or rejected. (iii) One forces a path-reversal (the subprojection function is changed to its complementary).

We will impose that the three-stage algorithm obeys a weighted detailed balance equation written between the two subpartition functions

$$\frac{\alpha}{\tau_0(2i)}\mathcal{N}_0\left(\mathcal{S}_{2i}\right)P_{gen}^+(\mathcal{P}_{2i}^{2i+1})P_{acc}^+(\mathcal{P}_{2i}^{2i+1}) = \frac{\alpha}{\tau_1(2i+1)}\mathcal{N}_1\left(\mathcal{S}_{2i+1}\right)P_{gen}^-\left(\mathcal{P}_{2i-2}^{2i+1}\right)P_{acc}^-\left(\mathcal{P}_{2i-2}^{2i+1}\right) \quad (39)$$

where $\tau_0(2i)$ and $\tau_1(2i+1)$ are the weights to take into account in the sampling procedure when averaging a physical quantity. These weights must be invariant with respect

to a chain-reversal and as they depend on the acceptance probabilities, these latter ones must therefore be chosen judiciously. In the positive time direction, the a posteriori probability to have accepted the trial path is

$$P_{acc}^{+}(\mathcal{P}_{2i}^{2i+1}) = \min\left(1, \left[\frac{\mathcal{K}_1(\mathcal{P}_{2i}^{2i-1'})}{\mathcal{K}_0(\mathcal{P}_{2i}^{2i-1'})}\frac{\mathcal{K}_0(\mathcal{P}_{2i}^{2i-1})}{\mathcal{K}_1(\mathcal{P}_{2i}^{2i-1})}\right]^{\theta}\right)$$

if $\mathcal{P}_{2i}^{2i+1} \equiv \mathcal{P}_{2i}^{2i-1'}$ or

$$P_{acc}^{+}(\mathcal{P}_{2i}^{2i+1}) = 1 - \left[\frac{\mathcal{K}_1(\mathcal{P}_{2i}^{2i-1'})}{\mathcal{K}_0(\mathcal{P}_{2i}^{2i-1'})}\frac{\mathcal{K}_0(\mathcal{P}_{2i}^{2i-1})}{\mathcal{K}_1(\mathcal{P}_{2i}^{2i-1})}\right]^{\theta}$$

if $\mathcal{P}_{2i}^{2i+1} \equiv \mathcal{P}_{2i}^{2i-1}$. This choice corresponds to equation (19) with the same notations of Section 3. Let us define the residence weight in $\mathcal{Z}_0$ as

$$\tau_0(2i) = P_{acc}(\mathcal{P}_{2i}^{2i+1})\left[\frac{\mathcal{K}_0(\mathcal{P}_{2i}^{2i+1})}{\mathcal{K}_1(\mathcal{P}_{2i}^{2i+1})}\right]^{\theta}, \qquad (40)$$

which simplifies to

$$\tau_0(2i) = \begin{cases} \min\left(\left[\frac{\mathcal{K}_0(\mathcal{P}_{2i}^{2i+1})}{\mathcal{K}_1(\mathcal{P}_{2i}^{2i+1})}\right]^{\theta}, \left[\frac{\mathcal{K}_0(\mathcal{P}_{2i}^{2i-1})}{\mathcal{K}_1(\mathcal{P}_{2i}^{2i-1})}\right]^{\theta}\right) \\ \left[\frac{\mathcal{K}_0(\mathcal{P}_{2i}^{2i-1'})}{\mathcal{K}_1(\mathcal{P}_{2i}^{2i-1'})}\right]^{\theta} - \left[\frac{\mathcal{K}_0(\mathcal{P}_{2i}^{2i-1})}{\mathcal{K}_1(\mathcal{P}_{2i}^{2i-1})}\right]^{\theta} \end{cases}, \qquad (41)$$

respectively in case of acceptation ($\mathcal{P}_{2i}^{2i+1} \equiv \mathcal{P}_{2i}^{2i-1'}$) or rejection ($\mathcal{P}_{2i}^{2i+1} \equiv \mathcal{P}_{2i}^{2i-1}$). This residence weight is invariant under a chain-reversible owing to the symmetric role played by the indices $2i+1$ and $2i-1$.

Similarly, with trial paths generated in the negative time direction from $\mathcal{Z}_1$ and the path chain considered in the reverse direction, using an acceptation probability corresponding to equation (20) yields the residence weights, in case of acceptance ($\mathcal{P}_{2i}^{2i+1} \equiv \mathcal{P}_{2i+2'}^{2i+1}$)

$$\tau_1(2i+1) = \min\left(\left[\frac{\mathcal{K}_1(\mathcal{P}_{2i}^{2i+1})}{\mathcal{K}_0(\mathcal{P}_{2i}^{2i+1})}\right]^{1-\theta}, \left[\frac{\mathcal{K}_1(\mathcal{P}_{2i+2}^{2i+1})}{\mathcal{K}_0(\mathcal{P}_{2i+2}^{2i+1})}\right]^{1-\theta}\right),$$

or, in case of rejection, $\mathcal{P}_{2i}^{2i+1} \equiv \mathcal{P}_{2i+2}^{2i+1}$

$$\tau_1(2i+1) = \left[\frac{\mathcal{K}_1(\mathcal{P}_{2i+1'}^{2i+1})}{\mathcal{K}_0(\mathcal{P}_{2i+1'}^{2i+1})}\right]^{1-\theta} - \left[\frac{\mathcal{K}_1(\mathcal{P}_{2i}^{2i+1})}{\mathcal{K}_0(\mathcal{P}_{2i}^{i+1})}\right]^{1-\theta}.$$

This residence weight is also invariant under a chain-reversible owing to the symmetric role played by the indices $2i+2$ and $2i$ One finally checks that our algorithm obeys the weighted detailed balance condition (Eq. (39)).

The residence weight ratio $\langle\tau_1\rangle/\langle\tau_0\rangle$ that gives an estimate of the desired quantity is similar in spirit to the biased occupation ratios in expanded ensemble methods [4]. In preceding articles [18,19] about the residence weight method, trial states, rather than trial paths, were considered to be generated with biased Monte Carlo schemes [37–42] obeying a super detailed balance condition [1]. Reversals to the previous state or configuration, rather than path rejections, were depicted which required to introduce the concept of non-Markovian sampling. This point is also reminiscent of the various non-Markovian preliminary procedures of expanded ensemble methods [4,8,9,17]. Nevertheless, this method could have been considered as Markovian, if one had artificially incorporated the memory kernel of the process into the sampled ensemble by extending it [43]. This would have led to the concepts of path-ensemble and path-sampling, since super detailed balance assumes detailed balance over each path connecting any two consecutive states of the chain.

Finally, note that imposing a weighted detailed balance condition between $\mathcal{Z}_0$ and $\mathcal{Z}_1$ subensembles implies the time alternation constraint. The effect of not satisfying this constraint is illustrated in Table 3 of Section 4.4 where it is shown that, in such a case, incorrect estimations are obtained resulting from the violation of detailed balance. At variance, the expanded path method, which does not account for all the rejected path information, is free from any alternation constraint. Hence, it is less accurate but more general than the residence weights method. For this reason, the method of expanded paths can be considered as a additional variant method of expanded ensembles. Expanded ensemble methods can indeed be implemented in microcanonical [44,45], thermal or path ensembles.

References

1. D. Frenkel, B. Smit, *Understanding molecular simulation* (Academic Press, San Diego, 2001)
2. S.V. Shevkunov, A. Martsinovski, P. Vorontsov-Velyaminov, *Teplofisika vysokikh temperatur* (Pervaja statja o osobennykh ansamblej, 1988), Vol. 26, p. 246
3. Y. Iba, Int. J. Mod. Phys. C **12**, 623 (2001)
4. A.P. Lyubartsev, A.A. Martsinovski, S.V. Shevkunov, P.N. Vorontsov-Velyaminov, J. Chem. Phys. **96**, 1776 (1992)
5. E. Marinari, G. Parisi, Europhys. Lett. **19**, 451 (1992)
6. A.P. Lyubartsev, A. Laaksonen, P.N. Vorontsov-Velyaminov, Mol. Phys. **82**, 455 (1994)
7. A. Brukhno, T.V. Kuznetsova, A.P. Lyubartsev, P.N. Vorontsov-Velyaminov, Vysokomolekuljarnye Soedinenija **38**, 3336 (1996)
8. F. Wang, D. Landau, Phys. Rev. Lett. **86**, 2050 (2001)
9. F. Wang, D. Landau, Phys. Rev. E **64**, 056101 (2001)
10. L.R. Pratt, J. Chem. Phys. **85**, 5045 (1986)
11. C. Dellago, P. Bolhuis, F. Csajka, D. Chandler, J. Chem. Phys. **108**, 1964 (1998)
12. C. Dellago, P. Bolhuis, F. Csajka, D. Chandler, J. Chem. Phys. **108**, 9236 (1998)
13. T. Vlugt, C. Dellago, B. Smit, J. Chem. Phys. **113**, 8791 (2000)
14. T. Vlugt, B. Smit, Phys. Chem. Comm. **2**, 1 (2001)
15. A.N. Kolmogorov, Math. Ann. **113**, 766 (1936)
16. R.L. Stratonovich, *Nonlinear Nonequilibrium Thermodynamics I* (Springer-Verlag, Berlin, 1992)
17. J. Ferkinghoff-Borg, Eur. Phys. J. B **29**, 481 (2002)
18. M. Athènes, Phys. Rev. E **66**, 016701 (2002)
19. M. Athènes, Phys. Rev. E **66**, 046705 (2002)

20. P. Langevin, Comptes Rendus **146**, 530 (1908)
21. D. Euvrard, *Résolution numérique des équations aux dérivés partielles* (Masson, Paris, 1994)
22. R. Dewar, J. Phys. A **36**, 631 (1999)
23. D.J. Evans, G.D. Cohen, G.P. Moriss, Phys. Rev. Lett. **71**, 2401 (1993)
24. J. Kurchan, J. Phys. A **31**, 3719 (1998)
25. G.E. Crooks, Phys. Rev. E **97**, 2361 (2000)
26. R. Zwanzig, J. Chem. Phys. **22**, 1420 (1954)
27. C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997)
28. C. Jarzynski, Phys. Rev. E **56**, 5018 (1997)
29. G.E. Crooks, J. Stat. Phys. **90**, 1481 (1998)
30. D.A. Hendrix, C. Jarzynski, J. Chem. Phys. **114**, 5974 (2001)
31. P.J. Rossky, J.D. Doll, H.L. Friedman, J. Chem. Phys. **69**, 4628 (1978)
32. C. Pangali, M. Rao, B.J. Berne, Chem. Phys. Lett. **55**, 413 (1978)
33. B. Mehlig, D.W. Heermann, B.M. Forrest, Mol. Phys. **76**, 1347 (1992)
34. S. Duane, A.D. Kennedy, B.J. Pendleton, D. Roweth, Phys. Lett. B **195**, 216 (1987)
35. B. Mehlig, D.W. Heermann, B.M. Forrest, Phys. Rev. B **45**, 645 (1992)
36. A.P. Omelyan, I.P. Mryglod, R. Folk, Comp. Phys. Commun. **151**, 272 (2003)
37. J.A. Siepmann, D. Frenkel, Mol. Phys. **75**, 59 (1992)
38. D. Frenkel, G.C. Mooij, B. Smit, J. Phys.: Condens. Matter **4**, 3053 (1992)
39. D. Frenkel, G.C. Mooij, B. Smit, Mol. Phys. **75**, 983 (1992)
40. T. Garel, H. Orland, J. Phys. A **23**, L621 (1990)
41. S. Consta, N.B. Wilding, D. Frenkel, Z. Alexandrowicz, J. Chem. Phys. **110**, 3220 (1999)
42. S. Consta, T.J. Vlugt, H.J.W. Hoeth, D. Smit, B. Frenkel, Mol. Phys. **97**, 1243 (1999)
43. C. Robert, *Méthodes de Monte Carlo par Chaînes de Markov* (Economica, Paris, 1996)
44. B.A. Berg, T. Neuhaus, Phys. Lett. B **267**, 249 (1991)
45. B.A. Berg, T. Neuhaus, Phys. Rev. Lett. **68**, 9 (1992)