

SOME REMARKS ON SAMPLING METHODS IN MOLECULAR DYNAMICS *

FRÉDÉRIC LEGOLL^{1,2}, TONY LELIÈVRE^{3,2} AND GABRIEL STOLTZ^{3,2}

Abstract. We present an overview of sampling techniques in molecular dynamics. We start with phase-space sampling techniques, and recall that deterministic methods may suffer from non-ergodicity problems. Then, we focus on methods relying on stochastic perturbations of the usual Hamiltonian dynamics. We also consider the problem of free energy computations, where the measure to sample is supported by a submanifold which is a level set of a so-called reaction coordinate.

Résumé. Cet article présente une rapide revue des méthodes d'échantillonnage en dynamique moléculaire. On commence par présenter les méthodes d'échantillonnage de mesures définies sur tout l'espace des phases, et plus particulièrement des méthodes stochastiques fondées sur des perturbations de la dynamique hamiltonienne. En effet, les méthodes déterministes ne sont pas toujours efficaces, voire ne sont pas ergodiques. On évoque ensuite différentes manières de calculer l'énergie libre d'un système. Dans ce cas, il s'agit d'échantillonner des mesures à support sur une sous-variété, qui est la ligne de niveau de ce qu'on appelle une coordonnée de réaction.

1. INTRODUCTION TO MOLECULAR SIMULATION

Statistical physics allows to bridge the gap between physical systems simulated at the microscopic level, and macroscopic quantities of interest, through averages over thermodynamic ensembles:

$$\langle A \rangle = \int A(q, p) d\mu(q, p). \quad (1)$$

In this expression, the function $A \equiv A(q, p)$ is an observable, and μ is a probability measure depending on the thermodynamic ensemble used.

Microscopic systems composed of N particles (typically atoms) are described by the position of the particles $q = (q_1, \dots, q_N) \in \mathbb{R}^{3N}$ and the associated momenta $p = (p_1, \dots, p_N) \in \mathbb{R}^{3N}$. The interaction between the particles is taken into account through a potential $V \equiv V(q)$, and the total energy of the system is given by the Hamiltonian

$$H(q, p) = \frac{1}{2} p^T M^{-1} p + V(q), \quad (2)$$

* The authors acknowledge the financial support provided by the ANR through the non thematic program *INGEMOL*.

¹ LAMI, Ecole Nationale des Ponts et Chaussées, 77455 Marne-la-Vallée, France ; e-mail: legoll@lami.enpc.fr.

² MICMAC, INRIA, Domaine de Voluceau-Rocquencourt, B.P. 105, 78153 Le Chesnay Cedex, France.

³ CERMICS, Ecole Nationale des Ponts et Chaussées, 77455 Marne-la-Vallée, France ; e-mail: lelievre@cermics.enpc.fr, stoltz@cermics.enpc.fr.

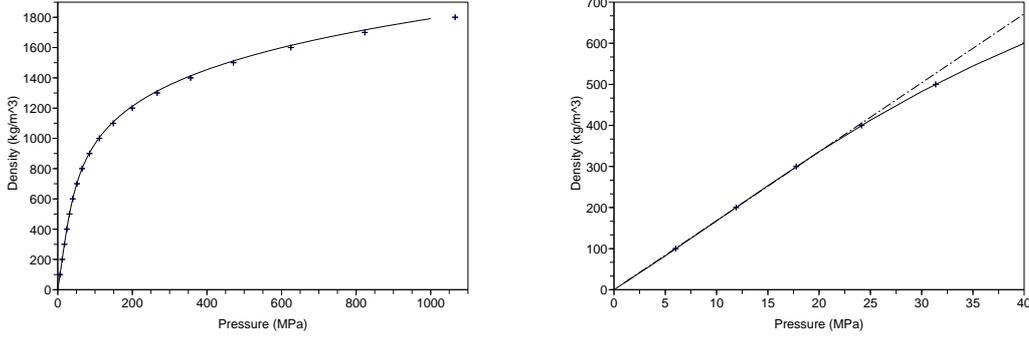


FIGURE 1. Numerical equation of state of argon at $T = 300$ K ('+') and experimental reference curve (solid line). The ideal gas regime is plotted in dash-dotted line (right picture).

where $M = \text{Diag}(m_1, \dots, m_N)$ is the mass matrix. An example of observable is the bulk pressure P in a liquid, given by $P = \langle A \rangle$ with

$$A(q, p) = \frac{1}{3|\mathcal{M}|} \sum_{i=1}^N \left(\frac{|p_i|^2}{m_i} - q_i \cdot \frac{\partial V}{\partial q_i}(q) \right),$$

where $|\mathcal{M}|$ is the volume occupied by the system. In practice, such averages can be computed with very small systems compared to the actual sizes of macroscopic systems. For example, the equation of state of Figure 1 has been computed with a system of few thousands of particles only, orders of magnitude below the Avogadro number.

1.1. The microcanonical ensemble

The most simple thermodynamic ensemble is the microcanonical ensemble which describes isolated systems. It amounts to considering microscopic configurations of equal energy:

$$\mu_{\text{mc}}(dq, dp) = \delta_{H(q,p)-E} = \frac{d\sigma_E}{|\nabla H|}, \quad (3)$$

where $d\sigma_E$ is the area measure induced by the Lebesgue measure on the manifold $\mathcal{M}(E) = \{(q, p) \mid H(q, p) = E\}$. Thermodynamic integrals (1) are computed in practice resorting to some ergodicity assumption:

$$\langle A \rangle = \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T A(\Phi_t(q, p)) dt, \quad (4)$$

where, in the microcanonical ensemble, the flow Φ_t is the flow of the Hamiltonian dynamics associated with (2):

$$\begin{cases} \dot{q}_i(t) = \frac{\partial H}{\partial p_i}(q(t), p(t)) = \frac{p_i(t)}{m_i}, \\ \dot{p}_i(t) = -\frac{\partial H}{\partial q_i}(q(t), p(t)) = -\nabla_{q_i} V(q(t)). \end{cases} \quad (5)$$

The dynamics (5) is an ordinary differential equation (ODE) which is often numerically integrated by the celebrated velocity-Verlet algorithm¹ [77]

$$\begin{cases} p^{n+1/2} = p^n - \frac{\Delta t}{2} \nabla V(q^n), \\ q^{n+1} = q^n + \Delta t M^{-1} p^{n+1/2}, \\ p^{n+1} = p^{n+1/2} - \frac{\Delta t}{2} \nabla V(q^{n+1}), \end{cases} \quad (6)$$

where Δt is the time step. The numerical flow associated with the velocity-Verlet algorithm shares two qualitative properties with the exact flow Φ_t of (5): it is *time reversible* and *symplectic*, which are very important properties for the long time numerical integration of Hamiltonian dynamics (see [28, Chap. VIII and IX] and [45]). The numerical analysis of sampling methods based on those lines (in the very particular case of completely integrable systems) can be read in [12, 13, 41].

We now focus on systems where the energy includes stiff terms, with the aim to develop algorithms that approximate the solution of the Newton equations of motion (5). A typical situation in molecular dynamics is when the potential energy is the sum of a stiff (or fast) term and a non-stiff (or slow) term:

$$V(q) = V_{\text{slow}}(q) + V_{\text{fast}}(q).$$

In general, the stiff term is much cheaper to evaluate than the non-stiff term. Indeed, the fast term usually corresponds to close range interactions, and its evaluation cost scales linearly with the system size. On the contrary, the slow term corresponds to long-range interactions, whose cost scales quadratically with the system size. It is therefore interesting to develop algorithms for which the slow forces $-\nabla V_{\text{slow}}$ can be evaluated less often than the fast forces $-\nabla V_{\text{fast}}$. One such algorithm is the so-called Impulse method [26, 76], which corresponds to a Strang splitting of the original Hamiltonian (2) in two terms, $H = H_{\text{slow}} + H_{\text{fast}}$ with

$$H_{\text{slow}}(q, p) = V_{\text{slow}}(q), \quad H_{\text{fast}}(p, q) = V_{\text{fast}}(q) + \frac{1}{2} p^T M^{-1} p.$$

The fast forces are evaluated with a time step Δt close to the time step used in the standard velocity-Verlet algorithm (6), whereas the slow forces are evaluated with a larger time step Δt_{slow} . However, there appear numerical resonances that require the slow force evaluation time step Δt_{slow} to be smaller than half the period of the fast movement [6, 23]. So, Δt_{slow} is still restricted by the highest frequency modes.

Many other methods have been developed to address this problem: the Mollify Impulse method [23], Equilibrium [34], Reversible Averaging [44], and a method based on shadow variables [36]. Some other methods have been developed to address the specific case of a harmonic stiff term, that is when

$$V_{\text{fast}}(q) = \frac{q^T \Omega^2 q}{2\varepsilon^2},$$

where Ω is a constant matrix and ε is small (see [28, Chap. XIII] for a comprehensive review, and also [25]). We also mention [38, 39], in which an approach based on a double-scale expansion is proposed to build efficient symplectic algorithms for highly oscillatory Hamiltonian systems.

1.2. The canonical ensemble

However, many systems are not isolated, but rather in contact with some thermal reservoir. In this case, the thermodynamic ensemble to describe the system is the so-called canonical ensemble (also termed NVT in view of the conservation of the number of particles, volume and temperature). It is characterized by the probability measure

$$d\mu(q, p) = Z^{-1} \exp(-\beta H(q, p)) dq dp = \Pi(q, p) dq dp, \quad (7)$$

¹See also [27] for more historic precisions: the Verlet algorithm was already known by Störmer at the beginning of the 20th century, and even by Newton!

where $\beta^{-1} = k_B T$ (T denotes the temperature and k_B the Boltzmann constant) and the partition function

$$Z = \int \exp(-\beta H(q, p)) dqdp. \quad (8)$$

There is no natural dynamics associated with this thermodynamical ensemble, in contrast to the microcanonical case. There are many (deterministic or stochastic) dynamics which preserve the canonical measure. For deterministic dynamics, it is an open question to prove the ergodicity, and there are even some theoretical and numerical counter-examples to ergodicity (see Section 2.1). On the other hand, it is usually possible to prove ergodicity for stochastic dynamics (see Section 2.2).

In some cases (especially the computation of time correlation functions), it may be desirable to consider a system governed by a Hamiltonian dynamics in its inner part, and subjected to thermal exchanges with its environment only in regions close to the boundaries. In this case a mixed thermodynamic ensemble can be used (see Section 2.3).

1.3. Free energy computations

Even if the ergodicity of the dynamics is theoretically ensured, it is often the case in practice that the long-time limit $T \rightarrow +\infty$ is difficult to reach because of the presence of very different time scales in the system. The fast time scales are typically related to some stiff components in the energy, which require very small integration time steps. For instance, bond lengths in a molecule typically oscillate with a period of the order of a femtosecond (10^{-15} s), while other quantities (such as dihedral angles) evolve on a much slower time scale. The slow time scales usually come from the presence of metastable wells in the potential or free energy. The system is trapped in such wells for a long time. Interesting events (protein folding, or conformational changes) only occur when the system has explored several such metastable wells, and this requires times at least of the order of a microsecond (10^{-6} s) and often more [66].

A possible route to tackle metastable features of the potential energy surface is to reduce the number of degrees of freedom in the system by resorting to free-energy computations (see Section 3). In this case, the information on the system is projected onto some lower dimensional manifold, indexed by a so-called reaction coordinate $\xi(q)$ describing the system at a macroscopic level. The free-energy of a system is of the general form

$$F(z) = -\beta^{-1} \ln Z(z), \quad (9)$$

where the partition function $Z(z)$ is the volume of the phase space at a given value of the reaction coordinate. Since $\xi(q)$ (if well chosen) is expected to be a slow variable of the system, it is hoped that dynamics restricted to the submanifold $Z(z)$ quickly reach their equilibrium states, so that the global metastability can be avoided by first performing ergodic averages on each submanifold $Z(z)$, and conveniently gathering all these averages in the end. This method is thermodynamic integration (see Section 3.1), and can be extended to the nonequilibrium case (see Sections 3.2 and 3.3).

Let us mention at this point that, besides the problem of computing the free energy of a system for a given reaction coordinate, a more difficult problem is the very choice of the reaction coordinate (see [21] for more precisions). A current challenge is to find a systematic way to define the right reaction coordinates (nowadays given by some physical intuition), which may change during the evolution of the system. Another challenge is to go further than the static view arising from sampling, and to study reduced or effective dynamics, that is, the *dynamics* of the reaction coordinate (see [42] for some preliminary experiments on model problems).

2. PHASE-SPACE SAMPLING METHODS

This section mainly summarizes results from [14], giving convergence results from both theoretical and numerical viewpoints for canonical averages, *i.e.* when (and if so, at which rate) the

configurations $(q^n, p^n)_{n \geq 0}$ generated by some numerical method are such that

$$\lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{n=0}^{N-1} A(q^n, p^n) = \int A(q, p) d\mu(q, p). \tag{10}$$

From a theoretical viewpoint, deterministic sampling methods based on the Nosé-Hoover paradigm may be non-ergodic (see Section 2.1). In [14] we have numerically compared several methods on a particular physically relevant example. As expected, naive Monte-Carlo methods are inefficient, in particular when the dimension of the system increases. Markov Chain Monte-Carlo techniques have been found to be particularly efficient, especially convenient stochastic perturbations of Hamiltonian dynamics such as the Langevin dynamics or the Hybrid Monte Carlo algorithm (see Section 2.2). On the other hand, deterministic sampling methods have been found to be not as efficient.

2.1. Deterministic methods

2.1.1. The Nosé-Hoover dynamics and its generalizations

As described above, the canonical ensemble is used to describe systems in contact with an external heat bath. The interaction between the physical system and the bath is such that, at equilibrium, the physical system variables are distributed according to the canonical measure (7).

The bottom line for deterministic sampling methods is to consider an extended system, composed of the physical variables (q, p) and some additional variables x describing the heat bath, and to postulate a deterministic dynamics on this extended set (q, p, x) . Ideally,

- (i) this dynamics should be ergodic for some measure $d\mu_{\text{ext}}(q, p, x)$,
- (ii) the measure $d\mu_{\text{ext}}(q, p, x) = \phi(q, p, x) dq dp dx$ should be such that

$$\int \phi(q, p, x) dx = Z^{-1} \exp(-\beta H(q, p)), \tag{11}$$

where Z is the normalization constant of the canonical measure (7).

If these two assumptions hold, then, for an observable A depending only on the physical variables (q, p) ,

$$\lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T A(q(t), p(t)) dt = \int A(q, p) d\mu_{\text{ext}}(q, p, x) = \int A(q, p) d\mu(q, p). \tag{12}$$

Hence, canonical averages can be recovered as time averages along the dynamics of the extended system.

We now describe the Nosé-Hoover dynamics, which is historically the first dynamics proposed in this vein [31, 56]. It consists in describing the heat bath by one scalar variable ξ , which can be considered as the thermostat momentum, and assuming the following dynamics:

$$\begin{cases} \dot{q}_i(t) = \frac{p_i(t)}{m_i}, \\ \dot{p}_i(t) = -\nabla_{q_i} V(q) - \frac{p_i \xi}{Q}, \\ \dot{\xi} = \sum_{i=1}^N \frac{p_i^2}{m_i} - g k_B T, \end{cases} \tag{13}$$

where g is the dimension of the manifold where p evolves (usually $3N$), T is the target temperature, and Q is a free parameter representing the mass of the thermostat. Here, the additional variable x is simply² ξ . We refer to [22] for details on the origin of this dynamics. The dynamics (13)

²Usually, a second additional variable, that represents the position of the thermostat, is also introduced. Since it is decoupled from all the other variables, we omit it here.

preserves the probability measure

$$d\mu_{\text{ext}}(q, p, \xi) = Z_{\text{ext}}^{-1} \exp\left(-\beta \left[H(q, p) + \frac{\xi^2}{2Q} \right]\right) dq dp d\xi, \quad (14)$$

where Z_{ext} is a normalization constant, and it satisfies the assumption (ii) above (see (11)). Unfortunately, proving (or disproving) that $d\mu_{\text{ext}}(q, p, \xi)$ satisfies the assumption (i) above is still mainly an open question. We come back to this question in Section 2.1.2.

Many generalizations of the Nosé-Hoover dynamics have been proposed (see [14] for more details). First, the bath can be described by many variables instead of a single one, such as in the Nosé-Hoover chain dynamics [31, 51, 56, 75]. Another idea is to develop extended *Hamiltonian* dynamics, like the Nosé-Poincaré method [8] and the Recursive Multiple Thermostats [46]. The driving idea for this approach is that there is an important literature on how to design numerical integration algorithms for Hamiltonian dynamics with good long-time properties [28]. These long-time properties are of paramount importance in molecular dynamics, since the dynamics is eventually used to compute long time averages (see (10)). All these dynamics preserve an invariant measure $d\mu_{\text{ext}}(q, p, x)$ which satisfies assumption (ii) above. Again, proving (or disproving) that these dynamics satisfy the assumption (i) above is still an open question.

2.1.2. Non ergodicity of the Nosé-Hoover dynamics in a simple case

The ergodicity of the Nosé-Hoover dynamics has never been rigorously proved. On the contrary, there is numerical evidence showing that the Nosé-Hoover method is not ergodic for some systems [31, 51, 75], including the one-dimensional harmonic oscillator. We now mainly summarize results from [43], where we prove that the Nosé-Hoover dynamics (13) is not ergodic in the case of a one-dimensional harmonic oscillator, if the thermostat mass Q is large enough.

To simplify the notations, we set $m = 1$, $\beta = 1$, and $V(q) = q^2/2$. The number of degrees of freedom to be thermalized is $g = 1$, and the Nosé-Hoover dynamics (13) reads

$$\begin{cases} \dot{q} = p, \\ \dot{p} = -q - \varepsilon\alpha p, \\ \dot{\alpha} = \varepsilon(p^2 - 1), \end{cases} \quad (15)$$

with $\varepsilon = 1/\sqrt{Q}$ and $\alpha = \varepsilon\xi$. We work in the regime $Q \gg 1$, that is $\varepsilon \ll 1$. Hence, the dynamics (15) is a perturbation of the Newton equations of motion and it is natural to introduce action-angle variables (τ, θ) for the oscillator:

$$q = \sqrt{2\tau} \cos \theta, \quad p = -\sqrt{2\tau} \sin \theta.$$

The Nosé-Hoover dynamics (15) now reads

$$\begin{cases} \dot{\theta} = 1 - \varepsilon\alpha \sin \theta \cos \theta, \\ \dot{\tau} = -2\varepsilon\tau\alpha \sin^2 \theta, \\ \dot{\alpha} = \varepsilon(2\tau \sin^2 \theta - 1). \end{cases} \quad (16)$$

We observe that τ and α are slow variables compared to θ . We thus make a change of variables as in the averaging method [70], and introduce $\hat{\tau}$ and $\hat{\alpha}$ such that

$$\tau = \hat{\tau} + \varepsilon\hat{\tau}\hat{\alpha} \sin \theta \cos \theta, \quad \alpha = \hat{\alpha} - \varepsilon\hat{\tau} \sin \theta \cos \theta.$$

The Nosé-Hoover dynamics (16) then reads

$$\begin{cases} \dot{\theta} = 1 - \varepsilon\hat{\alpha} \sin \theta \cos \theta + O(\varepsilon^2), \\ \dot{\hat{\tau}} = -\varepsilon\hat{\tau}\hat{\alpha} + O(\varepsilon^2), \\ \dot{\hat{\alpha}} = \varepsilon(\hat{\tau} - 1) + O(\varepsilon^2). \end{cases} \quad (17)$$

As a consequence, in these variables, the Nosé-Hoover dynamics is a perturbation of the dynamics

$$\begin{cases} \dot{\theta} = 1 - \varepsilon \hat{\alpha} \sin \theta \cos \theta, \\ \dot{\hat{\tau}} = -\varepsilon \hat{\tau} \hat{\alpha}, \\ \dot{\hat{\alpha}} = \varepsilon (\hat{\tau} - 1). \end{cases} \quad (18)$$

The key point of our argument is to observe that the system (18) has a first integral which is

$$G(\hat{\tau}, \hat{\alpha}) = \hat{\tau} - \ln \hat{\tau} + \frac{1}{2} \hat{\alpha}^2 - 1. \quad (19)$$

As a consequence, the Poincaré return map $Q_\varepsilon(\hat{\tau}, \hat{\alpha})$ of the plane $\Sigma = \{(\theta, \hat{\tau}, \hat{\alpha}) : \theta = 0 \bmod 2\pi\}$ for the ODE (18) has invariant curves. Let $P_\varepsilon(\hat{\tau}, \hat{\alpha})$ be the Poincaré return map of the same plane for the ODE (17). The map P_ε is a perturbation of the map Q_ε . Using KAM theory [70], we show that P_ε also has invariant curves [43]. This prevents the system (17) from being ergodic for the measure $d\mu_{\text{ext}}$ (see (14)), and so, (12) does not hold.

2.2. Stochastically perturbed Molecular dynamics methods

The dynamics (5) cannot be used to generate points according to the canonical measure, because the energy (2) is preserved by the flow. To generate points according to the canonical measure (7), there is a need for stochastic perturbations to ensure that different energy levels will be explored, and eventually all of them. These perturbations can be small and continuous (continuous stochastic processes such as in the Langevin dynamics), or large and applied at some given times (jump processes such as in the Hybrid Monte-Carlo dynamics).

2.2.1. Hybrid Monte-Carlo

The Hybrid Monte Carlo method generates points in the position space distributed according to the canonical measure. It aims at combining the advantages of Hamiltonian dynamics (that approximates the physical dynamics of the system) and of Monte Carlo methods (that explore the position space more globally). It is in fact a Metropolis-Hastings algorithm, in which proposals are constructed using the NVE Hamiltonian flow of the system. This method has been first introduced by Duane et al. in [20] and partially analyzed from a mathematical viewpoint by Schütte in [67]. It can be seen as a generalization of the Andersen thermostat method [3], and has been used in [68, 69] to identify the metastable conformations of some biological systems. There exist several refinements of the standard HMC scheme (see [14] for more details).

In the standard HMC setting, the sequence of generated positions forms a Markov chain of order one (see Figure 2). Let us emphasize that the proposal $q^{n,\tau}$ would always be accepted at step (3) if the NVE equations of motion, that are energy conserving, were integrated exactly. In practice, the time-step Δt used in the numerical integrator (6) can be chosen larger than in standard applications of MD since the dynamics of the system used to generate proposals is not constrained to accurately reproduce the physical dynamics of the system. On the other hand, it should not be too large; otherwise, the rejection rate would be large and the efficiency of the method would be low.

Convergence results rely on the theory of Markov chains. A fundamental reference is the book by Meyn and Tweedie [54] (see also the review in [24]). Roughly speaking, the convergence (10) of a single realization of a Markov chain is guaranteed provided

- (i) the canonical probability measure is preserved;
- (ii) an accessibility property is satisfied (irreducibility): denoting by P the kernel of the Markov chain, $P(q, B) > 0$ for all q and borel sets B of positive measure.

These conditions are satisfied when the potential V is bounded from above [14]. The proof of this fact relies on a least-action principle. When the potential is not bounded from above, it may be the case that the HMC algorithm is not ergodic - actually, the simplest counter-example is the

HYBRID MONTE CARLO

Algorithm 2.1. For an initial configuration q^0 and $\tau > 0$, and $n \geq 0$,

- (1) generate momenta p^n according to the canonical distribution $Z_p^{-1} \exp(-\beta p^T M^{-1} p/2)$ and compute the energy $E^n = H(q^n, p^n)$ of the configuration (q^n, p^n) ;
- (2) compute $\Phi_\tau(q^n, p^n) = (p^{n,\tau}, q^{n,\tau})$, that is, integrate the Hamiltonian equations of motion on the time interval $[0, \tau]$ starting from the initial data (q^n, p^n) ;
- (3) compute the energy $E^{n,\tau} = H(q^{n,\tau}, p^{n,\tau})$ of the new phase-space configuration. Accept the proposal $q^{n,\tau}$ with probability

$$\alpha^n = \min\left(1, e^{-\beta(E^{n,\tau} - E^n)}\right);$$

more precisely, generate a random number $U^n \sim \mathcal{U}[0, 1]$ (uniform law on $[0, 1]$), and set $q^{n+1} = q^{n,\tau}$ if $U^n \leq \alpha^n$ and $q^{n+1} = q^n$ otherwise;

FIGURE 2. The hybrid Monte-Carlo algorithm.

harmonic oscillator with τ equal to its period. In those cases, random integration times τ should be used.

2.2.2. Langevin dynamics

The paradigm of Langevin dynamics is to introduce in the Newton equations of motion (5) some fictitious brownian forces modelling fluctuations, balanced by viscous damping forces modelling dissipation. More precisely, the dynamics of the system is given by the following stochastic differential equation (SDE):

$$\begin{cases} dq_t = M^{-1} p_t dt, \\ dp_t = -\nabla V(q_t) dt - \gamma M^{-1} p_t dt + \sigma dW_t, \end{cases} \quad (20)$$

where $(W_t)_{t \geq 0}$ is a standard $3N$ -dimensional Wiener process. The parameters γ and σ represent the magnitude of the dissipation and of the fluctuations respectively, and are linked by the fluctuation-dissipation relation

$$\sigma = (2\gamma/\beta)^{1/2}. \quad (21)$$

Therefore, there remains one adjustable parameter in the model (say, γ). This dynamics is hypoelliptic (under some mild conditions on the potential), so that the average along one trajectory of the stochastic process converges to the ensemble average (1).

There are several ways to discretize the Langevin dynamics:

- (i) with convenient discretizations and a step-size Δt sufficiently small ensuring the sampling from an invariant measure close to the canonical measure (see *e.g.* [52, 72]);
- (ii) by extending usual NVE schemes used in deterministic MD simulations to the case of the Langevin dynamics (the quasi-symplectic schemes of [55]);
- (iii) by using splitting ideas borrowed from integration methods for deterministic flows (see *e.g.* [28]);
- (iv) with a Metropolized scheme (see [65] for an application to Variational Monte-Carlo).

The measure sampled by the discretized schemes (i)-(iii) is the canonical measure only up to $O(\Delta t)$ terms. The discretization strategy (iv) is therefore particularly interesting since it avoids time step errors. However, in order to obtain an efficient Metropolis-Hastings scheme [29, 53] some microscopic reversibility must be ensured, otherwise the rejection rate will be too high. For this purpose, a momentum reversion is introduced. Such a procedure was already considered for Hybrid Monte Carlo algorithms (see for instance [1]). Besides, since the Metropolis-Hastings algorithm requires transition probability densities, the discrete Langevin scheme must be elliptic. This asks

LANGEVIN METROPOLIZED ALGORITHM (SEE [65])

Algorithm 2.2. Starting from some initial configuration (q^0, p^0) , and $n \geq 0$,

- (1) Propose a move from (q^n, p^n) to $(\tilde{q}^{n+1}, \tilde{p}^{n+1})$ by performing one step of the modified Ricci-Ciccotti [60] algorithm

$$\begin{cases} q_*^{n+1} &= q^n + \frac{\Delta t}{m} p^n e^{-\gamma \Delta t / 2} - \frac{\Delta t^2}{2m} \nabla V(q^n) + e^{-\gamma \Delta t / 4} + G_1^n, \\ p_*^{n+1} &= p^n e^{-\gamma \Delta t} - \frac{\Delta t}{2} [\nabla V(q^n) + \nabla V(q_*^{n+1})] e^{-\gamma \Delta t / 2} + G_2^n, \end{cases}$$

where the identically and independently distributed (i.i.d.) gaussian random vectors (G_1^n, G_2^n) are such that

$$\langle (G_1^n)^2 \rangle = \frac{\Delta t}{\beta m \gamma} \left(2 - \frac{3 - 4e^{-\gamma \Delta t} + e^{-2\gamma \Delta t}}{\gamma \Delta t} \right) I_{3N},$$

$$\langle (G_2^n)^2 \rangle = \frac{m}{\beta} (1 - e^{-2\gamma \Delta t}) I_{3N}, \quad \langle G_1^n G_2^n \rangle = \frac{(1 - e^{-\gamma \Delta t})^2}{\beta \gamma} I_{3N}.$$

Set $(\tilde{q}^{n+1}, \tilde{p}^{n+1}) = (q_*^{n+1}, -p_*^{n+1})$ and denote by $\mathcal{P}_{\Delta t}$ the associated transition density;

- (2) Compute the acceptance rate

$$\alpha^n = \min \left(\frac{\Pi(\tilde{q}^{n+1}, \tilde{p}^{n+1}) \mathcal{P}_{\Delta t}((\tilde{q}^{n+1}, \tilde{p}^{n+1}), (q^n, p^n))}{\Pi(q^n, p^n) \mathcal{P}_{\Delta t}((q^n, p^n), (\tilde{q}^{n+1}, \tilde{p}^{n+1}))}, 1 \right);$$

- (3) Draw a random variable $U^n \sim \mathcal{U}(0, 1)$:

- if $U^n \leq \alpha^n$, accept the proposal and set $(\bar{q}^{n+1}, \bar{p}^{n+1}) = (\tilde{q}^{n+1}, \tilde{p}^{n+1})$;
- if $U^n > \alpha^n$, reject the proposal, and set $(\bar{q}^{n+1}, \bar{p}^{n+1}) = (q^n, p^n)$;

- (4) Reverse the momenta: $(q^{n+1}, p^{n+1}) = (\bar{q}^{n+1}, -\bar{p}^{n+1})$.

FIGURE 3. The Langevin Metropolized algorithm.

for integration schemes in the vein of those presented in [2]. This is summarized in the algorithm presented in Figure 3.

2.3. Thermal boundary conditions

Energy exchanges of a system with its environment can be modelled

- (i) by constraining the kinetic temperature at a fixed value in the regions close to the boundaries through some velocity rescaling, following the pioneering work [4];
- (ii) by using “thermal walls”, which are, mathematically speaking, jump processes (perturbations of the momenta of the impacting particles). Following a work of Lebowitz and Spohn [40], who showed that the momentum of the entering particles should not be drawn according to a maxwellian distribution of momenta, Ciccotti and Tenenbaum introduce thermal walls modelling the contact of impacting particles with a heat reservoir [16]. The system has free boundary conditions, but when a particle leaves the simulation domain, another one enters at the same place where the leaving particle went out, with a momentum generated from the probability distribution $C^{-1}(e \cdot p) f_T(p) \mathbf{1}_{e \cdot p > 0}$, where e is the local normal vector, f_T the distribution of the momenta at equilibrium at the temperature T (maxwellian distribution) and C is a normalization constant;
- (iii) by using a Langevin dynamics for the region of the system close to the boundary, and the usual Hamiltonian dynamics elsewhere, so that the resulting process is a diffusive process,

which is (hopefully) ergodic. The first study where only a part of the system is governed by a Langevin dynamics whereas the remaining part obeys Hamiltonian dynamics was proposed by Berkowitz and MacCammon [5] and further refined by Brooks and Karplus [9, 10].

Let us precise this last approach, which allows a seamless coupling between the Hamiltonian dynamics and an external thermostatted dynamics. The simulation domain Ω is decomposed into two disjoint domains Ω_i (inner part) and Ω_e (external part). The particles that are located in Ω_i are only subjected to the forces that derive from the interaction potential V , whereas the particles that are located in Ω_e also experience some random forcing. More precisely, we consider the dynamics:

$$\begin{cases} dq_t = M^{-1}p_t dt, \\ dp_t = -\nabla V(q_t) dt - \Gamma(q_t)M^{-1}p_t dt + \Sigma(q_t) dW_t, \end{cases} \quad (22)$$

where $(W_t)_{t \geq 0}$ is a standard $3N$ -dimensional Wiener process, and where the matrices Σ and Γ represent the magnitude of the fluctuations and of the dissipation respectively. They are linked by the fluctuation-dissipation relation:

$$\Sigma(q_t)\Sigma(q_t)^T = \frac{2}{\beta}\Gamma(q_t). \quad (23)$$

It is easy to check that the canonical probability measure is an invariant probability measure for (22) since it is a stationary solution of the associated Fokker-Planck equation.

A simple choice is a diagonal matrix for $\Gamma(q)$:

$$\Gamma(q) = \text{Diag}(\gamma(q_1), \dots, \gamma(q_N)),$$

where the function γ is taken to be a smooth decreasing function of $d(x, \partial\Omega)$ such that $\gamma(x) = 0$ in Ω_i and $\gamma(x) > 0$ in Ω_e . We then set

$$\Sigma(q) = \text{Diag}(\sigma(q_1), \dots, \sigma(q_N)), \quad \text{with } \sigma(\cdot) = \sqrt{\frac{2\gamma(\cdot)}{\beta}}. \quad (24)$$

An application to the computation of thermal conductivities with this choice can be found in [72].

3. FREE ENERGY COMPUTATIONS

The absolute free energy (9) of a system can be computed only for certain systems, such as ideal gases, or solids at low temperature (resorting to the phonon spectrum) [22, 61]. However, in many applications, the quantity of interest is the free energy *difference* between an initial and a final state. These differences indeed give information about the relative stabilities of several species, as well as their transition kinetics. These differences are expressed as

$$\Delta F = -\beta^{-1} \ln \left(\frac{\int e^{-\beta V_1(q)} dq}{\int e^{-\beta V_0(q)} dq} \right),$$

in the so-called alchemical case (that is, when the transition is indexed by an external parameter λ in the potential V_λ). In the reaction coordinate case, the transition is indexed through some level set function $\xi(q)$ indexing submanifolds of the configuration space, and

$$\Delta F = -\beta^{-1} \ln \left(\frac{\int e^{-\beta V(q)} \delta_{\xi(q)-z_1} dq}{\int e^{-\beta V(q)} \delta_{\xi(q)-z_0} dq} \right).$$

Recall that $\delta_{\xi(q)-z}$ is a measure defined on $\Sigma(z) = \{q, \xi(q) = z\}$ by

$$\delta_{\xi(q)-z} = |\nabla \xi|^{-1} d\sigma_{\Sigma(z)}.$$

Free energy differences are much more amenable to computations than the absolute free energy. Classical techniques to this end fall within four main classes:

- (i) The first one, dating back to Kirkwood [37], is *thermodynamic integration*, which mimics the quasi-static evolution of a system as a succession of equilibrium samplings, which amounts to an infinitely slow switching between the initial and final states (see Section 3.1);
- (ii) The second one, the *free energy perturbation method*, was introduced by Zwanzig [80], and is suited to the alchemical case only. It recasts free energy differences as $\Delta F = -\beta^{-1} \ln \langle e^{-\beta(V_1 - V_0)} \rangle_0$, where $\langle \cdot \rangle_0$ is the canonical average with respect to the measure $Z_0^{-1} e^{-\beta V_0(q)} dq$, so that usual sampling techniques can be employed. Notice also that there exist many refinements for those two classes of techniques, such as umbrella sampling [74] (which is an importance sampling technique);
- (iii) A more recent class of methods uses dynamics arising from a switching at a finite rate, using *nonequilibrium dynamics* with a suitable exponential reweighting, as introduced by Jarzynski in [35] (see Section 3.2);
- (iv) Finally, *adaptive dynamics* can also be used. The switching then arises as the result of an equilibrium sampling, using for example the Adaptive Biasing Force technique [17, 30] or metadynamics [33]. In those cases, the system is progressively forced to leave regions where the sampling of the reaction coordinate has been completed (Section 3.3).

In the sequel, we will present the different techniques using the overdamped Langevin dynamics

$$dq_t = -\nabla V(q_t)dt + \sigma dW_t, \tag{25}$$

where $(W_t)_{t \geq 0}$ is a standard $3N$ -dimensional Wiener process and $\sigma = (2/\beta)^{1/2}$. This dynamics is formally obtained from the Langevin dynamics (20) by letting the mass matrix M go to zero and by setting $\gamma = 1$, which amounts to rescaling the time. This dynamics is ergodic (under mild assumptions on the potential V) with respect to the canonical probability measure in the position space, which we will denote $d\pi(q) = Z_q^{-1} e^{-\beta V(q)} dq$.

3.1. Thermodynamic integration

The principle of thermodynamic integration is to write ΔF as an integral over the derivative of F with respect to the parameter λ or the reaction coordinate value z . Let us focus on the latter, with a *scalar* reaction coordinate $\xi : \mathbb{R}^{3N} \rightarrow \mathbb{R}$. Recall that, in this case, the free energy is defined (up to an additive constant) as

$$F(z) = -\beta^{-1} \ln \int e^{-\beta V} \delta_{\xi(q)-z}.$$

Using the co-area formula, it is possible to show that (see [15])

$$F'(z) = \frac{\int f^V(q) e^{-\beta V} \delta_{\xi(q)-z}}{\int e^{-\beta V} \delta_{\xi(q)-z}}, \tag{26}$$

with

$$f^V = \frac{\nabla V \cdot \nabla \xi}{|\nabla \xi|^2} - \beta^{-1} \operatorname{div} \left(\frac{\nabla \xi}{|\nabla \xi|^2} \right).$$

The free energy difference is then obtained as

$$F(z) - F(0) = \int_0^z F'(s) ds.$$

The method thus consists in computing (26), which requires to sample the probability measure $\pi_z = \left(\int e^{-\beta V} \delta_{\xi(q)-z}\right)^{-1} e^{-\beta V} \delta_{\xi(q)-z}$ with support on $\Sigma(z)$. Since this is an integral in a high-dimensional space, stochastic methods are very natural. The idea is to project the dynamics (25) onto $\Sigma(z)$ to sample π_z :

$$dq_t = P(q_t) \left(-\nabla \tilde{V}(q_t) dt + \sqrt{\frac{2}{\beta}} dW_t \right) + \frac{1}{\beta} K(q_t) dt, \quad (27)$$

where we have introduced the orthogonal projector $P(q)$ on $\Sigma(z)$ at point q :

$$P(q) = \text{Id} - \frac{\nabla \xi \otimes \nabla \xi}{|\nabla \xi|^2}(q),$$

the modified potential \tilde{V} :

$$\tilde{V} = V + \frac{1}{\beta} \ln |\nabla \xi|, \quad (28)$$

and the mean curvature vector $K(q)$ of $\Sigma(z)$ at point q :

$$K(q) = -\text{div} \left(\frac{\nabla \xi}{|\nabla \xi|} \right) \frac{\nabla \xi}{|\nabla \xi|}(q). \quad (29)$$

It can be shown that the process q_t indeed samples π_z . One interesting remark is that this dynamics can also be obtained using a Lagrange multiplier approach to impose the constraint $\xi(q_t) = z$. Indeed, q_t satisfies:

$$\begin{cases} dq_t = -\nabla \tilde{V}(q_t) dt + \sqrt{2\beta^{-1}} dW_t + \nabla \xi(q_t) d\Lambda_t, \\ \text{with } d\Lambda_t \text{ such that } \xi(q_t) = z. \end{cases} \quad (30)$$

Besides, the Lagrange multiplier $d\Lambda_t$ can be decomposed as $d\Lambda_t = d\Lambda_t^m + d\Lambda_t^f$ with

$$d\Lambda_t^m = -\sqrt{2\beta^{-1}} \frac{\nabla \xi}{|\nabla \xi|^2}(q_t) \cdot dW_t, \quad d\Lambda_t^f = f^V(q_t) dt.$$

Thus, since q_t samples π_z , it is possible to compute $F'(z)$ by simply averaging the Lagrange multiplier over a trajectory (see (26)).

In practice a natural strategy is to discretize the SDE (30) to approximate $F'(z)$. One possible scheme is for example

$$\begin{cases} q^{n+1} = q^n - \nabla \tilde{V}(q^n) \Delta t + \sqrt{2\beta^{-1} \Delta t} U^n + \nabla \xi(q^n) \Delta \Lambda_{n+1}, \\ \text{with } \Delta \Lambda_{n+1} \text{ such that } \xi(q^{n+1}) = z. \end{cases} \quad (31)$$

The random variables U^n are i.i.d. standard gaussian random vectors. It is possible to show [15] that (31) is a consistent discretization of (30), so that $\frac{1}{N} \sum_{n=1}^N \Delta \Lambda_n$ is an approximation of $F'(z)$. We finally refer to [15] for more details on the practical aspects (like variance reduction schemes or efficient computation of $\nabla \tilde{V}$), and for generalization to the case of a multi-dimensional reaction coordinate.

3.2. Nonequilibrium switching dynamics

We present in this section the Jarzynski equality to compute free energy differences using nonequilibrium dynamics. Let us focus on the alchemical case. We consider a switching schedule $\lambda(t)$ such that $\lambda(0) = 0$ and $\lambda(T) = 1$, and the non-homogeneous markovian dynamics

$$dq_t = -\nabla V_{\lambda(t)}(q_t) dt + \sqrt{2\beta^{-1}} dW_t, \quad q_0 \sim Z_0^{-1} e^{-\beta V_0(q)} dq.$$

Since the switching is done at finite rate, the system does not remain at equilibrium during the simulation. There is therefore a need for a reweighting of the final sample of configurations. Defining the out of equilibrium virtual work induced on the system during the time interval $[0, t]$ as

$$\mathcal{W}_t = \int_0^t \frac{\partial V_{\lambda(s)}}{\partial \lambda}(q_s) \lambda'(s) ds, \tag{32}$$

it holds [35]

$$\mathbb{E}(e^{-\beta \mathcal{W}_t}) = e^{-\beta(F(\lambda(t)) - F(0))}. \tag{33}$$

This equality can be derived through a Feynman-Kac formula [32]. The interest of this method is apparent from (33): realizations of the switching process can be straightforwardly parallelized resorting to many independent trajectories, so that natural *a posteriori* error bounds are provided *via* the central limit theorem. However, in practice, this method is often not efficient as such since, due to the exponential reweighting, many trajectories are discarded at the end, and some rare trajectories rule the average.

To avoid the degeneracy of weights, a selection mechanism on the replicas can be performed. This was done in [63] for the computation of free energy differences, using a so-called interacting particle system (IPS). Each particle is endowed with two random times, namely a birth and a death time. The death time is decreased when the work exerted on the replica is higher than the average work update, otherwise the birth time is decreased. When the death time is zero, the replica is replaced by another replica chosen at random. The birth process is handled in a similar way: when the birth time is zero, a replica chosen at random is replaced by the replica whose birth time is zero. This way, particles with lower work values are enhanced. We refer for example to [19] for more precisions on the practical implementation of such a procedure as well as other strategies to handle birth/death processes, and to [62] for convergence results.

The equality (33) can be seen as an average over paths. Path sampling techniques can therefore be used to compute (33) efficiently [57, 73, 79] by precisely enhancing the paths that have the larger weights (which correspond to the unlikely lower work values). More precisely,

$$e^{-\beta \Delta F} = \frac{\int e^{-\beta \mathcal{W}(x)} d\pi_L(x)}{\int d\pi_L(x)},$$

where $d\pi_L$ is a measure on a discrete path of length L , and $\mathcal{W}(x)$ is the work along a given path x . In the case of the overdamped Langevin dynamics (25) with $\lambda(t) = t/(L\Delta t)$, the probability to observe the path $x = (q_0, q_{\Delta t}, \dots, q_{L\Delta t})$ is

$$d\pi_L(x) = Z_L^{-1} e^{-\beta V_0(q_0)} \prod_{i=0}^{L-1} \exp\left(-\frac{\beta}{4\Delta t^2} |q_{(i+1)\Delta t} - q_{i\Delta t} - \Delta t \nabla V_{i/L}(q_{i\Delta t})|^2\right) dx,$$

and the work is approximated by

$$\mathcal{W}_{\Delta t}(x) = \frac{1}{L} \sum_{i=1}^L \frac{\partial V_{\lambda}}{\partial \lambda} \Big|_{\lambda=i/L} (q_{i\Delta t}).$$

Importance sampling techniques can then be used, such as rewriting

$$e^{-\beta \Delta F} = \frac{\int e^{-\beta \mathcal{W}(x)/2} d\Pi_L(x)}{\int e^{\beta \mathcal{W}(x)/2} d\Pi_L(x)},$$

where the paths are sampled according to the modified measure $d\Pi_L(x) = e^{-\beta \mathcal{W}(x)/2} d\pi_L(x)$, which enhances the paths with lower work values. Methods to sample paths can be found in [7, 18, 71].

The Jarzynski equality cannot be used as such in the reaction coordinate case, unless maybe resorting to soft constraints to switch between the initial state centered on the submanifold $\Sigma(z_0) = \{q \mid \xi(q) = z_0\}$ and the final state centered on $\Sigma(z_1)$. Steered molecular dynamics techniques use for example a penalty term $K(\xi(q) - z)^2$ in the energy of the system [58] (with K large) to 'softly' constrain the system to remain close to the submanifold $\Sigma(z)$, and varying the value z from 0 to 1 in a finite time T . It is shown in [32] how to use such a biasing potential to exactly compute free energy differences (even for a finite K), which is of particular interest for experimental studies. From a computational viewpoint however, it is expected that large values of K require small integration time steps. Moreover, it is observed in practice that the statistical fluctuations increase with larger K (see [58]). Instead, it is possible to replace the stiff constraining potential $K(\xi(q) - z)^2$ by a projection onto the submanifold $\Sigma(z)$. This situation is reminiscent of the case of molecular constraints, that can be enforced using a stiff penalty term, or more elegantly and often more efficiently, using some projection of the dynamics involving Lagrange multipliers. This is the spirit of the well known SHAKE algorithm [64]. Such a projected dynamics, extending the projected dynamics of Section 3.1 to the nonequilibrium case, has been proposed and studied in [48].

3.3. Adaptive dynamics

Instead of forcing the evolution of the system as is done in nonequilibrium simulations, it could be interesting to let the system evolve at its own pace, and maybe spend longer times in regions where it has to (because of large free-energy barriers), whereas it could rapidly leave regions where the free-energy profile is flat.

The corresponding methods are of adaptive type, and use the whole history of the exploration process to bias the current dynamics in order to force the escape from metastable sets. This is done by simultaneously estimating the free energy from an evolving ensemble of configurations of the dynamics, and using this estimate to bias the dynamics, so that the effective explored free energy surface is flattened. In the long time limit, the bias exactly gives the actual free energy profile and the metastabilities in the direction of the reaction coordinate are removed. Adaptive methods could therefore be seen as umbrella sampling with an evolving potential. This was already noticed in a previous study presenting an adaptive dynamics as a 'self-healing umbrella sampling' [50].

Recently, adaptive methods were presented in a general and unifying framework [49]. To this end, it is convenient, as is done in [11], to consider ensembles of infinitely many realizations. The system is then described by the distribution of the configurations of this ensemble at time t , denoted by $\psi_t(q)$ in the sequel. Adaptive dynamics use a time-dependent potential

$$\mathcal{V}_t(q) = V(q) - F_{\text{bias}}(t, \xi(q)), \quad (34)$$

where F_{bias} is a biasing potential depending on the history of the system. Nonequilibrium meta-dynamics [11] and Wang-Landau dynamics [78] use

$$\partial_t F_{\text{bias}}(t, z) = G_t \left(-\beta^{-1} \ln \int \psi_t(q) \delta_{\xi(q)-z} \right), \quad (35)$$

while the adaptive biasing force method (ABF, [17]) updates the bias on the mean force (26) as

$$\partial_t F'_{\text{bias}}(t, z) = \mathcal{G}_t \left(\frac{\int f^V(q) \psi_t(q) \delta_{\xi(q)-z}}{\int \psi_t(q) \delta_{\xi(q)-z}} - F'_{\text{bias}}(t, z) \right). \quad (36)$$

Derivatives with respect to the z variable are denoted by ' in this section. The functions G_t and \mathcal{G}_t are increasing functions, with $\mathcal{G}_t(0) = 0$. In both cases, the stationary state of the dynamics is unique and such that $F'_{\text{bias}}(\infty)$ is equal to the mean force F' given by (26) (see [49]).

The analysis using ensemble of realizations suggests a parallel implementation, which was already proposed in [59] in the case of metadynamics. An additional selection process on the replicas, in the IPS spirit, can enhance the sampling of the reaction coordinates in comparison with a straightforward parallel implementation [49].

A rigorous convergence result of the ABF algorithm with the update

$$F'_{\text{bias}}(t, z) = \frac{\int f^V(q) \psi_t(q) \delta_{\xi(q)-z}}{\int \psi_t(q) \delta_{\xi(q)-z}}$$

can be shown in the case of an overdamped Langevin dynamics with an infinite number of replicas [47]. Using a modified overdamped dynamics, it can be proved [47] that (i) the distribution ψ_t^ξ of the reaction coordinate satisfies

$$\partial_t \psi_t^\xi = \beta^{-1} \partial_{zz} \psi_t^\xi,$$

which is a pure diffusive behavior for the marginal law of the reaction coordinate. In particular, the original metastabilities of the free energy landscape are eliminated; (ii) the observed mean force $F'_{\text{bias}}(t)$ converges to F' when $t \rightarrow +\infty$. The convergence is exponential, the rate of convergence being limited by the minimum between the rate of convergence in each submanifold $\xi(q) = z$ (see [47] for more details), and the rate of convergence in the reaction coordinate space. The proof of convergence uses an entropy inequality, and a splitting of the total entropy between a macroscopic entropy (associated with the reaction coordinate), and the remaining microscopic entropy.

REFERENCES

- [1] E. Akhmatskaya and S. Reich. *The targetted shadowing hybrid Monte Carlo (TSHMC) method*, volume 49 of *Lecture Notes in Computational Science and Engineering*. Springer Verlag, Berlin and New York, 2006.
- [2] M. P. Allen and D. J. Tildesley. *Computer simulation of liquids*. Oxford University Press, 1987.
- [3] H. C. Andersen. Molecular-dynamics simulations at constant pressure and-or temperature. *J. Chem. Phys.*, 72(4):2384–2393, 1980.
- [4] W. T. Ashurst and W. G. Hoover. Dense-fluid shear viscosity *via* nonequilibrium molecular-dynamics. *Phys. Rev. A*, 11(2):658–678, 1975.
- [5] M. Berkowitz and J. A. McCammon. Molecular-dynamics with stochastic boundary-conditions. *Chem. Phys. Lett.*, 90(3):215–217, 1982.
- [6] J. J. Biesiadecki and R. D. Skeel. Dangers of multiple time step methods. *J. Comput. Phys.*, 109:318–328, 1993.
- [7] P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler. Transition path sampling: Throwing ropes over rough mountain passes, in the dark. *Ann. Rev. Phys. Chem.*, 53:291–318, 2002.
- [8] S. D. Bond, B. J. Leimkuhler, and B. B. Laird. The Nosé-Poincaré method for constant temperature molecular dynamics. *J. Comput. Phys.*, 151(1):114–134, 1999.
- [9] C. L. Brooks and M. Karplus. Deformable stochastic boundaries in molecular-dynamics. *J. Chem. Phys.*, 79(12):6312–6325, 1983.
- [10] A. Brünger, C. L. Brooks, and M. Karplus. Stochastic boundary-conditions for molecular-dynamics simulations of ST2 water. *Chem. Phys. Lett.*, 105(5):495–500, 1984.
- [11] G. Bussi, A. Laio, and M. Parrinello. Equilibrium free energies from nonequilibrium metadynamics. *Phys. Rev. Lett.*, 96(9):090601, 2006.
- [12] E. Cancès, F. Castella, P. Chartier, E. Faou, C. Le Bris, F. Legoll, and G. Turinici. High-order averaging schemes with error bounds for thermodynamical properties calculations by molecular dynamics simulations. *J. Chem. Phys.*, 121(21):10346–10355, 2004.
- [13] E. Cancès, F. Castella, P. Chartier, E. Faou, C. Le Bris, F. Legoll, and G. Turinici. Long-time averaging for integrable Hamiltonian dynamics. *Numerische Mathematik*, 100(2):211–232, 2005.
- [14] E. Cancès, F. Legoll, and G. Stoltz. Theoretical and numerical comparison of sampling methods for molecular dynamics. *accepted for publication in M2AN*, 2007.
- [15] G. Ciccotti, T. Lelièvre, and E. Vanden-Eijnden. Sampling Boltzmann-Gibbs distributions restricted on a manifold with diffusions. *to appear in Comm. Pure Appl. Math.*, 2007.
- [16] G. Ciccotti and A. Tenenbaum. Canonical ensemble and non-equilibrium states by molecular-dynamics. *J. Stat. Phys.*, 23(6):767–772, 1980.
- [17] E. Darve and A. Pohorille. Calculating free energies using average force. *J. Chem. Phys.*, 115(20):9169–9183, 2001.

- [18] C. Dellago, P. G. Bolhuis, and P. L. Geissler. Transition path sampling. *Adv. In Chem. Physics, Vol 123*, 123:1–78, 2002.
- [19] A. Doucet, N. de Freitas, and N.J. Gordon. *Sequential Monte Carlo Methods in Practice*. Series Statistics for Engineering and Information Science. Springer, 2001.
- [20] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth. Hybrid Monte-Carlo. *Phys. Lett. B*, 195(2):216–222, 1987.
- [21] W. E and E. Vanden-Eijnden. *Metastability, conformation dynamics, and transition pathways in complex systems. Multiscale modelling and simulation*, volume 39 of *Lect. Notes Comput. Sci. Eng.* Springer, Berlin, 2004.
- [22] D. Frenkel and B. Smit. *Understanding Molecular Simulation, From Algorithms to Applications (2nd ed.)*. Academic Press, 2002.
- [23] B. Garcia-Archilla, J. M. Sanz-Serna, and R. D. Skeel. Long time step methods for oscillatory differential equations. *SIAM J. Sci. Comput.*, 20(3):930–963, 1998.
- [24] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter. *Markov Chain Monte Carlo in practice*. Chapman and Hall, 1996.
- [25] V. Grimm and M. Hochbruck. Error analysis of exponential integrators for oscillatory second-order differential equations. *J. Phys. A*, 39:5495–5507, 2006.
- [26] H. Grubmüller, H. Heller, A. Windemuth, and K. Schulten. Generalized Verlet algorithm for efficient molecular dynamics simulations with long range interaction. *Mol. Sim.*, 6:121–142, 1991.
- [27] E. Hairer, C. Lubich, and G. Wanner. Geometric numerical integration illustrated by the Störmer-Verlet method. *Acta Numerica*, 12:399–450, 2003.
- [28] E. Hairer, C. Lubich, and G. Wanner. *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*, volume 31 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, Heidelberg, 2006.
- [29] W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57:97–109, 1970.
- [30] J. Hénin and C. Chipot. Overcoming free energy barriers using unconstrained molecular dynamics simulations. *J. Chem. Phys.*, 121(7):2904–2914, 2004.
- [31] W. G. Hoover. Canonical dynamics - Equilibrium phase-space distributions. *Phys. Rev. A*, 31(3):1695–1697, 1985.
- [32] G. Hummer and A. Szabo. Free energy reconstruction from nonequilibrium single-molecule pulling experiments. *Proc. Nat. Acad. Sci. USA*, 98(7):3658–3661, 2001.
- [33] M. Iannuzzi, A. Laio, and M. Parrinello. Efficient exploration of reactive potential energy surfaces using Car-Parrinello molecular dynamics. *Phys. Rev. Lett.*, 90(23):238302, 2003.
- [34] J. A. Izaguirre, S. Reich, and R. D. Skeel. Longer time steps for molecular dynamics. *J. Chem. Phys.*, 110(20):9853–9864, 1999.
- [35] C. Jarzynski. Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.*, 78(14):2690–2693, 1997.
- [36] Z. Jia and B. J. Leimkuhler. Geometric integrators for multiple time-scale simulation. *J. Phys. A*, 39(16):5379–5403, 2006.
- [37] J. G. Kirkwood. Statistical mechanics of fluid mixtures. *J. Chem. Phys.*, 3(5):300–313, 1935.
- [38] C. Le Bris and F. Legoll. Dérivation de schémas numériques symplectiques pour des systèmes hamiltoniens hautement oscillants (derivation of symplectic numerical schemes for highly oscillatory hamiltonian systems). *C. R. Acad. Sci. Paris, Série I*, 344(4):277–282, 2007.
- [39] C. Le Bris and F. Legoll. Homogenization approach for integrators of highly oscillatory hamiltonian systems: some examples. *in preparation*, 2007.
- [40] J. L. Lebowitz and H. Spohn. Transport properties of the Lorentz gas - Fourier’s law. *J. Stat. Phys.*, 19(6):633–654, 1978.
- [41] F. Legoll. *PhD thesis*. Université Paris VI, 2004.
- [42] F. Legoll and T. Lelièvre. Effective dynamics using conditional expectations. *in preparation*, 2007.
- [43] F. Legoll, M. Luskin, and R. Moeckel. Non-ergodicity of the Nosé-Hoover thermostatted harmonic oscillator. *Arch. Rational Mech. Anal.*, 184:449–463, 2007.
- [44] B. J. Leimkuhler and S. Reich. A reversible averaging integrator for multiple time-scale dynamics. *J. Comput. Phys.*, 171:95–114, 2001.
- [45] B. J. Leimkuhler and S. Reich. *Simulating Hamiltonian dynamics*, volume 14 of *Cambridge monographs on applied and computational mathematics*. Cambridge University Press, 2005.
- [46] B. J. Leimkuhler and C. R. Sweet. A Hamiltonian formulation for recursive multiple thermostats in a common timescale. *Siam J. Appl. Dyn. Syst.*, 4(1):187–216, 2005.
- [47] T. Lelièvre, F. Otto, M. Rousset, and G. Stoltz. Long-time convergence of the Adaptive Biasing Force method. *in preparation*, 2007.
- [48] T. Lelièvre, M. Rousset, and G. Stoltz. Computation of free energy differences through nonequilibrium stochastic dynamics: the reaction coordinate case. *J. Comput. Phys.*, 222(2):624–643, 2006.
- [49] T. Lelièvre, M. Rousset, and G. Stoltz. Computation of free energy profiles with parallel adaptive dynamics. *J. Chem. Phys.*, 126:134111, 2007.

- [50] S. Marsili, A. Barducci, R. Chelli, P. Procacci, and V. Schettino. Self-healing umbrella sampling: A non-equilibrium approach for quantitative free energy calculations. *J. Phys. Chem. B*, 110(29):14011–14013, 2006.
- [51] G. J. Martyna, M. L. Klein, and M. Tuckerman. Nosé-Hoover chains - the canonical ensemble via continuous dynamics. *J. Chem. Phys.*, 97(4):2635–2643, 1992.
- [52] J. C. Mattingly, A. M. Stuart, and D. J. Higham. Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise. *Stoch. Proc. Appl.*, 101(2):185–232, 2002.
- [53] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equations of state calculations by fast computing machines. *J. Chem. Phys.*, 21(6):1087–1091, 1953.
- [54] S. P. Meyn and R. L. Tweedie. *Markov chains and stochastic stability*. Communications and control engineering series. Springer-Verlag, London, New York, 1993.
- [55] G. N. Milstein and M. V. Tretyakov. Quasi-symplectic methods for Langevin-type equations. *IMA J. Numer. Anal.*, 23(4):593–626, 2003.
- [56] S. Nosé. A unified formulation of the constant temperature molecular-dynamics methods. *J. Chem. Phys.*, 81(1):511–519, 1984.
- [57] H. Oberhofer, C. Dellago, and P. L. Geissler. Biased sampling of nonequilibrium trajectories: Can fast switching simulations outperform conventional free energy calculation methods? *J. Phys. Chem. B*, 109(14):6902–6915, 2005.
- [58] S. Park, F. Khalili-Araghi, E. Tajkhorshid, and K. Schulten. Free energy calculation from steered molecular dynamics simulations using Jarzynski’s equality. *J. Chem. Phys.*, 119(6):3559–3566, 2003.
- [59] 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(8):3533–3539, 2006.
- [60] A. Ricci and G. Ciccotti. Algorithms for Brownian dynamics. *Mol. Phys.*, 101(12):1927–1931, 2003.
- [61] J. M. Rickman and R. LeSar. Free-energy calculations in materials research. *Ann. Rev. Materials Research*, 32:195–217, 2002.
- [62] M. Rousset. On the control of an interacting particle estimation of Schrödinger groundstates. *SIAM J. Math. Anal.*, 38:824–844, 2006.
- [63] M. Rousset and G. Stoltz. Equilibrium sampling from nonequilibrium dynamics. *J. Stat. Phys.*, 123(6):1251–1272, 2006.
- [64] J. P. Ryckaert, G. Ciccotti, and H. J. C. Berendsen. Numerical-integration of cartesian equations of motion of a system with constraints - molecular-dynamics of n-alkanes. *J. Comput. Phys.*, 23(3):327–341, 1977.
- [65] A. Scemama, T. Lelièvre, G. Stoltz, E. Cancès, and M. Caffarel. An efficient sampling algorithm for variational Monte Carlo. *J. Chem. Phys.*, 125:114105, 2006.
- [66] T. Schlick. *Molecular Modeling and Simulation*. Springer, 2002.
- [67] C. Schütte. *Habilitation Thesis*. Freie Universität Berlin, 1999.
- [68] C. Schütte, A. Fischer, W. Huisinga, and P. Deuffhard. A direct approach to conformational dynamics based on hybrid Monte Carlo. *J. Comput. Phys.*, 151(1):146–168, 1999.
- [69] C. Schütte and W. Huisinga. *Biomolecular conformations can be identified as metastable sets of molecular dynamics*, volume X of *Handbook of Numerical Analysis (Special volume on computational chemistry)*. Elsevier, 2003.
- [70] C. L. Siegel and J. K. Moser. *Lecture on Celestial Mechanics*. Springer, New York, 1971.
- [71] G. Stoltz. Path sampling with stochastic dynamics: Some new algorithms. *to appear in J. Comput. Phys.*, 2007.
- [72] G. Stoltz. *PhD thesis*. Ecole Nationale des Ponts et Chaussées, 2007.
- [73] S. X. Sun. Equilibrium free energies from path sampling of nonequilibrium trajectories. *J. Chem. Phys.*, 118(13):5769–5775, 2003.
- [74] G. M. Torrie and J. P. Valleau. Non-physical sampling distributions in Monte-Carlo free-energy estimation - Umbrella sampling. *J. Comput. Phys.*, 23(2):187–199, 1977.
- [75] M. E. Tuckerman and G. J. Martyna. Understanding modern molecular dynamics: Techniques and applications. *J. Phys. Chem. B*, 104(2):159–178, 2000.
- [76] M. E. Tuckerman, B. J. Berne, and G. J. Martyna. Reversible multiple time scale molecular dynamics. *J. Chem. Phys.*, 97:1990–2001, 1992.
- [77] L. Verlet. Computer “experiments” on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. *Phys. Rev.*, 159:98–103, 1967.
- [78] F. G. 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(5):056101, 2001.
- [79] F. M. Ytreberg and D. M. Zuckerman. Single-ensemble nonequilibrium path-sampling estimates of free energy differences. *J. Chem. Phys.*, 120(23):10876–10879, 2004.
- [80] R. W. Zwanzig. High-temperature equation of state by a perturbation method I. Nonpolar gases. *J. Chem. Phys.*, 22(8):1420–1426, 1954.