Journal of Statistical Physics, Vol. 123, No. 6, June 2006 (© 2006)
DOI: 10.1007/s10955-006-9090-2

Equilibrium Sampling From Nonequilibrium Dynamics

Mathias Rousset[1] and Gabriel Stoltz[2,3]

Received October 25, 2005; accepted March 13, 2006
Published Online: May 19, 2006

We present some applications of an Interacting Particle System (IPS) methodology to the field of Molecular Dynamics. This IPS method allows several simulations of a nonequilibrium random process to keep closer to equilibrium at each time, thanks to a selection mechanism based on the relative virtual work induced on the system. It is therefore an efficient improvement of usual nonequilibrium simulations, which can be used to compute canonical averages, free energy differences, and typical transitions paths.

KEY WORDS: Non-equilibrium molecular dynamics, interacting particle system, genetic algorithms, free energy estimation
AMS: 65C05, 65C35, 80A10.

Phase-space integrals are widely used in Statistical Physics to relate the macroscopic properties of a system to the elementary phenomenon at the microscopic scale.[11] In constant temperature (NVT) molecular simulations, these integrals often take the form

$$\mu(A) = \langle A \rangle = \int_{T^*\mathcal{M}} A(q, p)\, d\mu(q, p). \tag{1}$$

where $\mathcal{M}$ denotes the position space (also called the *configuration space*), and $T^*\mathcal{M}$ denotes its cotangent space. A generic element of the position space $\mathcal{M}$ will be denoted by $q = (q_1, \cdots, q_{\mathrm{N}})$ and a generic element of the momentum space by $p = (p_1, \cdots, p_N)$. We will consider here that $\mathcal{M} \sim \mathbb{R}^{3N}$ or $\mathbb{T}^{3N}$ (a torus of dimension $3N$, which arises when using periodic boundary conditions), and that

[1] Laboratoire Statistique et Probabilités, Université Paul Sabatier 118, Route de Narbonne, 31062 TOULOUSE Cedex 4, France; e-mail: mathias.rousset@polytechnique.org
[2] CERMICS, Ecole Nationale des Ponts et Chaussées, 6/8 avenue Blaise Pascal, 77455 Marne-la-Vallée, France
[3] CEA/DAM Ile-de-France, BP 12, 91680 Bruyères-le-Châtel, France; e-mail: stoltz@cermics.enpc.fr

$T^*\mathcal{M} \sim \mathbb{R}^{3N} \times \mathbb{R}^{3N}$ or $\mathbb{T}^{3N} \times \mathbb{R}^{3N}$, though in general more complicated situations should be considered, when performing Blue Moon sampling[4] for example.

The measure μ is the canonical probability measure

$$d\mu(q, p) = Z^{-1} \exp(-\beta H(q, p))\, dq\, dp, \tag{2}$$

where $\beta = 1/k_{\mathrm{B}}T$ (T denotes the temperature and k_{B} the Boltzmann constant) and where H denotes the Hamiltonian of the molecular system:

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

In the above expression, V is the potential experienced the N particles, and $M = \mathrm{Diag}(m_1, \cdots, m_{\mathrm{N}})$ where m_i is the mass of the i-th particle. The constant Z in (2) is the normalization constant defined as

$$Z = \int_{T^*\mathcal{M}} \exp(-\beta H(q, p))\, dq\, dp.$$

Some quantities can not be expressed through relations such as (1). One important example is the free energy of a system, defined (up to a constant) as

$$F = -\beta^{-1} \ln Z.$$

It is though often the case in practice that a straightforward sampling of μ is difficult. Indeed, high dimensional systems exhibit many local minima in which the system remains trapped, especially when the temperature is low. In those cases, alternative approaches have to be used, such as those built on the *simulated annealing*[18] paradigm, and its extension to more general *switching* dynamics. The idea is to switch slowly from an initial simple sampling problem, to the target sampling problem, through a well chosen interpolation. This allows to attain deeper local minima, but, due to its nonequilibrium nature, is not efficient as such to sample accurately the target measure.

We mention that variations have been proposed, especially *tempering* methods (see ref. 15 for a review), the most famous being *parallel tempering*.[21] These methods consider an additional parameter describing the configuration system (e.g. the temperature), and sample those extended configurations according to some stochastic rules. However, these methods ask for a prior distribution of the additional parameters (for example a temperature ladder in parallel tempering method), which are usually estimated through some preliminary runs.[15]

As noted by many authors, switching strategies can be used to compute exactly ratios of partition function (free energy differences), through an explicit computation of importance weights of nonequilibrium simulations, often referred to as Jarzynski's equality[16,17] (see also ref. 22).

We present here a complementary approach to the above nonequilibrium strategies. It is similar to the method of,[13] known as "population Monte-Carlo,"

in which multiple replicas are used to represent the distribution under study. A weight is associated to each replica, and resamplings are performed at discrete fixed times to avoid degeneracy of the weights. This methodology is widely used in the fields of Quantum Monte Carlo[1,25] or Bayesian Statistics, where it is referred to as Sequential Monte Carlo.[8,6] Note that in the probability and statistics fields, each simulation is called a 'walker' or 'particle'; we use here the name 'replica,' which is more apppropriate to the Molecular Dynamics context.

Our method extends the population Monte-Carlo method to the time-continuous case. It consists in running M replicas of the system in parallel, resorting typically to a stochastic dynamics, and considering exchanges between them, according to a certain probabilistic rule depending on the work done on each system. This procedure can be seen as automatic time continuous resampling, and all replicas have the same weight at any time of the simulation. This method drastically increases the number of significative transitions paths in nonequilibrium simulations. These heuristic explanations are precised in Sec. 2. The set of all replicas (or walkers) is called an 'Interacting Particle System' (IPS),[7] and can be seen as a genetic algorithm where the mutation step is the stochastic dynamics considered.

The article is organized as follows. We first precise classical simulated annealing type methods in Sec. 1. We then describe the associated IPS method in Sec. 2, as well as its numerical implementation. Possible applications and some numerical results are then presented in Secs. 3 and 4.

1. NONEQUILIBRIUM METHODS

Consider a family of Hamiltonian functions $H_\lambda : T^*\mathcal{M} \to \mathbb{R}$ indexed by a parameter $\lambda \in (0, 1)$. The corresponding Hamiltonian dynamics are

$$\begin{cases} \dfrac{dq}{dt} = \dfrac{\partial H_\lambda}{\partial p}, \\ \dfrac{dp}{dt} = -\dfrac{\partial H_\lambda}{\partial q}. \end{cases} \tag{3}$$

The family $(H_\lambda)_{\lambda\in(0,1)}$ indexes a path between the original state described by a Hamiltonian H_0 and the final state charaterized by a Hamiltonian H_1. A canonical probability measure μ_λ can be associated to each Hamiltonian H_λ:

$$d\mu_\lambda(q, p) = \frac{1}{Z_\lambda} \mathrm{e}^{-\beta H_\lambda(q,p)}\, dq\, dp,$$

where the normalizing constant Z_λ is

$$Z_\lambda = \int_{T^*\mathcal{M}} \mathrm{e}^{-\beta H_\lambda(q,p)}\, dq\, dp.$$

We will also denote by $F(\lambda) = -\beta^{-1} \ln Z_\lambda$ the corresponding free energy.

We wish to sample according to $d\mu_1$, from an initial, easily obtained sample of $d\mu_0$. For each replica of the previous sample, the corresponding configuration of the system is brought slowly to the end state along a path $(\lambda(t))_{t\in(0,T)}$ for a time $T > 0$. The final sample of configurations is hopefully close to $d\mu_1$.

Typically, we can consider $H_\lambda(q, p) = (1-\lambda)H_0(q, p) + \lambda H_1(q, p)$ (for example when performing a change of temperature from T to $T' : H_0(q, p) = H(q, p)$, $H_1(q, p) = \frac{\beta'}{\beta} H(q, p)$). It can also represent a modification of the potential, sometimes called 'alchemical transition' in the physics and chemistry litterature. The folding of a protein could be studied this way for example, by setting initially all the long-range interactions to zero, whereas the final state corresponds to a Hamiltonian were all interactions are set on. In this case,

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

The nonequilibrium-like strategies can also be extended to the reaction coordinate case.[20] In this case, the initial and the final state are indexed by some order parameter $\xi(q)$.

1.1. Markovian Nonequilibrium Simulations

The usual way to achieve this method is to perform a time inhomogeneous irreducible Markovian dynamics

$$t \mapsto X_t^{\lambda(t)}, \qquad X_0^{\lambda(0)} \sim \mu_0, \tag{4}$$

for $t \in (0, T)$, and a smooth schedule $t \mapsto \lambda(t)$ verifying $\lambda(1) = 0$ and $\lambda(T) = 1$, and such that for all *given* $\lambda \in (0, 1)$, the Boltzmann distribution $d\mu_\lambda$ is invariant under the dynamics $t \mapsto X_t^\lambda$.

The variable x can represent the whole degrees of freedom (q, p) of the system, or only the configuration part q. Depending on the context, the invariant measure μ will therefore be the canonical measure (2), or its marginal with respect to the momenta, which reads

$$d\tilde{\mu}_\lambda(q) = \frac{1}{\tilde{Z}_\lambda} \mathrm{e}^{-\beta V_\lambda(q)}\, dq,$$

with

$$\tilde{Z}_\lambda = \int_{\mathcal{M}} \mathrm{e}^{-\beta V_\lambda(q)}\, dq.$$

When we do not wish to precise further the dynamics, we simply call x the configuration of the system, $H_\lambda(x)$ its energy and $d\mu_\lambda(x)$ the invariant measure. The actual invariant measure should be clear from the context.

Denoting by $p_{s,t}(x, y)dy = \mathbb{E}(X_t^{\lambda(t)} \in dy | X_s^{\lambda(s)} = x)$ the density kernel of the process, the evolution of the process law is characterized by the backward Kolmogorov equation (t and y being given):

$$\partial_s p_{s,t}(., y) = -L_{\lambda(s)}(p_{s,t}(., y)),$$

or its forward version (s and x being given):

$$\partial_t p_{s,t}(x, .) = L^*_{\lambda(t)}(p_{s,t}(x, .)).$$

The operator $L_{\lambda(t)}$ is called the infinitesimal generator of the dynamics, and its dual $L^*_{\lambda(t)}$ is defined by

$$\int \varphi_1 L_{\lambda(t)}(\varphi_2)\, dx = \int \varphi_2 L^*_{\lambda(t)}(\varphi_1)\, dx. \tag{5}$$

The invariance of $\mu_{\lambda(t)}$ under the instantaneous dynamic can then be expressed through the balance condition:

$$\forall \varphi, \qquad \int L_{\lambda(t)}(\varphi) d\mu_{\lambda(t)} = 0. \tag{6}$$

The dynamics we have in mind are (for a fixed $\lambda \in [0, 1]$):

- The hypo-elliptic Langevin dynamics on $T^*\mathcal{M}$

$$\begin{cases} dq_t^\lambda = M^{-1} p_t^\lambda dt, \\ dp_t^\lambda = -\nabla V(q_t^\lambda)\, dt - \xi M^{-1} p_t^\lambda\, dt + \sigma\, dW_t, \end{cases} \tag{7}$$

  where W_t denotes a standard $3N$-dimensional Brownian motion. The paradigm of Langevin dynamics is to introduce in the Newton equations of motion (3) some fictitious brownian forces modelling fluctuations, balanced by viscous damping forces modelling dissipation. The parameters $\sigma, \xi > 0$ represent the magnitude of the fluctuations and of the dissipation respectively, and are linked by the fluctuation-dissipation relation:

$$\sigma = (2\xi/\beta)^{1/2}. \tag{8}$$

  Therefore, there remains one adjustable parameter in the model. The infinitesimal generator is given by:

$$L_\lambda \varphi = M^{-1} p \cdot \nabla_q \varphi - \nabla V(q) \cdot \nabla_p \varphi - \xi M^{-1} p \cdot \nabla_p \varphi + \frac{\xi}{\beta} \Delta_p \varphi.$$

- The elliptic overdamped Langevin dynamics[4] in the configuration space $\mathcal{M}$:

$$dq_t^\lambda = -\nabla V_\lambda\left(q_t^\lambda\right) dt + \sigma\, dW_t, \tag{9}$$

where the magnitude of the random forcing is given here by

$$\sigma = \sqrt{\frac{2}{\beta}}.$$

The corresponding infinitesimal generator is given by:

$$L_\lambda \varphi = \frac{1}{\beta}\Delta_q \varphi - \nabla V_\lambda(q) \cdot \nabla_q \varphi.$$

Let us remark that the overdamped Langevin dynamics (9) is obtained from the Langevin dynamics (7) by letting the mass matrix M go to zero and by setting $\xi = 1$, which amounts here to rescaling the time.

It is well known that, for a fixed $\lambda \in (0, 1)$, these dynamics are ergodic under mild assumptions on the potential V[3].

When the schedule is sufficiently slow, the dynamics is said quasi-static, and the law of the process $X_t^{\lambda(t)}$ is assumed to stay close to its local steady state throughout the transformation. As said before, this is out of reach at low temperature (more precisely, large deviation results[10] ensure that the typical escape time from metastable states grows exponentially fast with β, which compells quasi-static transformations to being exponentially slow with β).

It is therefore interesting to consider approaches built on switched Markovian dynamics, but able to deal with reasonably fast transition schemes.

1.2. Importance Weights of Non Equilibrium Simulations

For a given nonequilibrium run $X_t^{\lambda(t)}$ we denote by

$$\mathcal{W}_t = \int_0^t \frac{\partial H_{\lambda(s)}}{\partial \lambda}(X_s^{\lambda(s)})\lambda'(s)\, ds$$

the out of equilibrium virtual work induced on the system during the time interval $(0, t)$. The quantity $\mathcal{W}_t$ gives the importance weights of nonequilibrium simulations with respect to the target equilibrium distribution. Indeed, it was shown in ref. 16 that

$$\mathbb{E}(e^{-\beta \mathcal{W}_t}) = e^{-\beta(F(\lambda(t))-F(0))}. \tag{10}$$

[4] This dynamics is actually known as the 'Langevin dynamic' in the probability and statistics fields. We adopt here the physical names of these stochastic processes, which are more natural when dealing with molecular dynamics.

This fluctuation equality is known as the Jarzynski's equality, and can be derived through a Feynman-Kac formula,[14] as follows.

Consider the Feynman–Kac density kernel defined by

$$\int \varphi(y) p^w_{s,t}(x,y)dy = \mathbb{E}\big(\varphi\big(X^{\lambda(t)}_t\big)\mathrm{e}^{-\beta(\mathcal{W}_t-\mathcal{W}_s)}|X^{\lambda(s)}_s = x\big), \tag{11}$$

and characterized by the following extended backward Komogorov evolution:

$$\partial_s p^w_{s,t}(.,y) = -L_{\lambda(s)}\left(p^w_{s,t}(.,y)\right) + \beta\frac{\partial H_{\lambda(s)}}{\partial\lambda}\lambda'(s)p^w_{s,t}(.,y).$$

Using this identity and the balance Eq. (6) gives:

$$\partial_s \int p^w_{s,t}(x,y)e^{-\beta H_{\lambda(s)}(x)}dx = 0$$

and thus after integration on $(0,t)$, we get the fundamental Feynamn-Kac fluctuation equality:

$$\frac{Z_t}{Z_0}\int \varphi d\mu_{\lambda(t)} = \mathbb{E}\big(\varphi\big(X^{\lambda(t)}_t\big)e^{-\beta\mathcal{W}_t}\big). \tag{12}$$

Therefore, taking $\varphi = 1$, it follows

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

and Jensen's inequality then gives

$$\mathbb{E}(\mathcal{W}_t) \geq F(\lambda(t)) - F(0).$$

This inequality is an equality if and only if the transformation is quasi-static on $(0,t)$; in this case the random variable $\mathcal{W}_t$ is actually constant and equal to ΔF. When the evolution is reversible, this means that equilibrium is maintained at all times.

As an improvement, we will replace the exponential importance weights of the nonequilibrium paths by a selection rule between replicas.

2. THE INTERACTING PARTICLE SYSTEM METHOD

Our strategy is inspired by the re-sampling methods in Sequential Monte Carlo (SMC) literature.[6,9] In this time continuous context, the idea is to replace importance weights of simulations performed in parallel, by a selection operation between replicas.

We first present the IPS approximation in Sec. 2.1, as well as convergence results of the discretized measure to the target measure. A justification through a mean-field interpretation is then proposed in Sec. 2.2. The numerical implementation of the IPS method is eventually discussed in Sec. 2.3.

2.1. The IPS and its Statistical Properties

Recall that the potential of mean force defined by

$$\mathcal{F}_{\lambda(t)} = \int \frac{\partial H_\lambda}{\partial \lambda}(x)\, d\mu_{\lambda(t)}(x)$$

is the average force applied to the system during an infinitely slow transformation. It can be used in a thermodynamic integration to compute free energy differences:

$$F(1) - F(0) = \int_0^T \mathcal{F}_{\lambda(t)} \lambda'(t)\, dt. \tag{13}$$

The first step is to rewrite the Feynman-Kac formula (12) by introducing a dichotomy when a replica is receiving either excess or deficit work compared to the potential of mean force.

To this end, we define respectively the excess and deficit force, and the excess and deficit work as

$$f_t^{\mathrm{ex}}(x) = \left(\frac{\partial H_{\lambda(t)}}{\partial \lambda} - \mathcal{F}_{\lambda(t)}\right)^+(x), \quad f_t^{\mathrm{de}}(x) = \left(\frac{\partial H_{\lambda(t)}}{\partial \lambda} - \mathcal{F}_{\lambda(t)}\right)^-(x)$$

$$\mathcal{W}_t^{\mathrm{ex}} = \int_0^t f_s^{\mathrm{ex}}(X_s^{\lambda(s)})\lambda'(s)\, ds, \quad \mathcal{W}_t^{\mathrm{de}} = \int_0^t f_s^{\mathrm{de}}\big(X_s^{\lambda(s)}\big)\lambda'(s)\, ds, \tag{14}$$

where $x^+ = \max\{x, 0\}$ and $x^- = \max\{-x, 0\}$ (so that $x = x^+ - x^-$). We then rewrite

$$\int \phi\, d\mu_{\lambda(t)} = \frac{\mathbb{E}\big(\varphi\big(X_t^{\lambda(t)}\big)\mathrm{e}^{-\beta\left(\mathcal{W}_t^{\mathrm{ex}} - \mathcal{W}_t^{\mathrm{de}}\right)}\big)}{\mathbb{E}\left(\mathrm{e}^{-\beta(\mathcal{W}_t^{\mathrm{ex}} - \mathcal{W}_t^{\mathrm{de}})}\right)}. \tag{15}$$

We now present the particle interpretation of (15) enabling a numerical computation through the use of empirical distributions. Consider M Markovian systems described by variables $X_t^k (1 \le k \le M)$. We approximate the virtual force by

$$\mathcal{F}_{\lambda(t)}^M = \frac{1}{M}\sum_{k=1}^M \frac{\partial H_{\lambda(t)}}{\partial \lambda}\big(X_t^k\big),$$

and the Boltzmann distribution by

$$d\mu_{\lambda(t)}^M(x) = \frac{1}{M}\sum_{k=1}^M \delta_{X_t^k}(dx),$$

which are their empirical versions. This naturally gives from definitions (14) empirical approximations of excess/deficit forces $f_t^{M,\mathrm{ex/de}}$ and works $\mathcal{W}_t^{k,\mathrm{ex/de}}$.

The replicas evolve according to a branching process with the following stochastic rules (see refs. 25 and 27 for further details):

Process 1. *Consider an initial distribution* $(X_0^1, \cdots, X_0^{\mathrm{M}})$ *generated from* $d\mu_0(x)$. *Generate idependent times* $\tau_1^{k,b}$, $\tau_1^{k,d}$ *from an exponential law of mean* β^{-1} *(the upperscripts b and d refer to 'birth' and 'death' respectively), and initialize the jump times* $T^{b/d}$ *as* $T_0^{k,d} = 0$, $T_0^{k,b} = 0$.

For $0 \le t \le T$.

- *Between each jump time, evolve independently the replicas* X_t^k *according to the dynamics (4);*
- *At random times* $T_{n+1}^{k,d}$ *defined by*

$$\mathcal{W}^{k,\mathrm{ex}}_{T_{n+1}^{k,d}} - \mathcal{W}^{k,\mathrm{ex}}_{T_n^{k,d}} = \tau_{n+1}^{k,d},$$

  an index $l \in \{1, \cdots, M\}$ *is picked at random, and the configuration of the k-th replica is replaced by the configuration of the l-th replica. A time* $\tau_{n+2}^{k,d}$ *is generated from an exponential law of mean* β^{-1}*;*
- *At random times* $T_{n+1}^{k,b}$ *defined by*

$$\mathcal{W}^{k,\mathrm{de}}_{T_{n+1}^{k,d}} - \mathcal{W}^{k,\mathrm{de}}_{T_n^{k,d}} = \tau_{n+1}^{k,b},$$

  an index $l \in \{1, \cdots, M\}$ *is picked at random, and the configuration of the l-th replica is replaced by the configuration of the k-th replica. A time* $\tau_{n+2}^{k,b}$ *is generated from an exponential law of mean* β^{-1}.

The selection mechanism therefore favors replicas which are sampling values of the virtual work $\mathcal{W}_t$ lower than the empirical average. The system of replicas is 'self-organizing' to keep closer to a quasi-static transformation.

In refs. 7 and 25, several convergence results and statistical properties of the replicas distribution are proven. They are summarized in the following proposition:

Proposition 2.1. *Assume that* $(t, x) \mapsto \frac{\partial H_{\lambda(t)}}{\partial \lambda}(x)$ *is a continuous bounded function on* $(0, T) \times T^*\mathcal{M}$ *(or* $(0, T) \times \mathcal{M}$ *in the case of overdamped Langevin dynamics), and that the dynamic* (4) *is ergodic. Then for any* $t \in (0, T)$,

- *The estimator*

$$\exp\left(-\beta \int_0^t \mathcal{F}^M_{\lambda(s)} \lambda'(s)\, ds\right) \tag{16}$$

  is an unbiased estimator of $e^{-\beta(F(\lambda(t))-F(0))}$*;*
- *for all test function* φ, *the estimator* $\int \varphi \, d\mu^M_{\lambda(t)}$ *is an asymptotically normal estimator of* $\int \varphi \, d\mu_{\lambda(t)}$, *with bias and variance of order* M^{-1}.

Let us emphasize that the sample $(X_t^k)_{1 \le k \le M}$ is in particular an empirical approximation of the canonical measure $d\mu_{\lambda(t)}$ for all t.

The proof follows from Lemma 3.20, Proposition 3.25 and Theorem 3.28 of[7] (see also refs. 25 and 27 for further details).

The unbiased estimation of un-normalized quantities is a very usual property in particle system methods. It comes from the fundamental property that at each "time step," each replica may branch with a number of offsprings equal in average to its relative importance weight.

2.2. Consistency Through a Mean-Field Limit

In order to prove the consistency of the IPS approximation, we consider the ideal setting where the number of replicas goes to infinity ($M \to +\infty$). This point of view is equivalent to a mean-field or Mc Kean interpretation of the IPS (denoted by the superscript 'mf'). In this limit, the behavior of any single replica, denoted by X_t^{mf}, is then independent from any finite number of other ones. We shall consider the mean field distribution and force:

$$\mathrm{Law}\big(X_t^{\mathrm{mf}}\big) = d\mu_t^{mf} = \mu_t^{\mathrm{mf}}(x)\,dx,$$
$$\mathcal{F}_t^{\mathrm{mf}} = \int \frac{\partial H_{\lambda(t)}}{\partial \lambda} d\mu_t^{\mathrm{mf}}.$$

The associated mean field excess/deficit forces $f_t^{\mathrm{mf,ex/de}}$ and works $\mathcal{W}_t^{\mathrm{mf,ex/de}}$ are defined as in (14).

In view of Process 1, the stochastic process X_t^{mf} is a jump-diffusion process which evolves according to the following stochastic rules (some facts about pure Markov jump processes are recalled in the Appendix):

Process 2. *Generate X_0^{mf} from $d\mu_0(x)$. Generate idependent clocks $(\tau_n^b, \tau_n^d)_{n\geq 1}$ from an exponential law of mean β^{-1}, and initialize the jump times $T^{b/d}$ as $T_0^d = 0$, $T_0^b = 0$.*

For $0 \leq t \leq T$,

- *Between each jump time, $t \mapsto X_t^{\mathrm{mf}}$ evolves according to the dynamics* (4)*;*
- *At random times T_{n+1}^d defined by*

$$\mathcal{W}_{T_{n+1}^d}^{\mathrm{mf,ex}} - \mathcal{W}_{T_n^d}^{\mathrm{mf,ex}} = \tau_{n+1}^d,$$

the process jumps to a configuration x, chosen according to the probability measure $d\mu_{T_{n+1}^d}^{\mathrm{mf}}(x)$;

- *At random times T_{n+1}^b defined by*

$$\mathbb{E}\big(\mathcal{W}_t^{\mathrm{mf,de}}\big)|_{t=T_{n+1}^b} - \mathbb{E}\big(\mathcal{W}_t^{\mathrm{mf,de}}\big)|_{t=T_n^b} = \tau_{n+1}^b,$$

the process jumps to a configuration x, chosen according to the probability measure proportional to $f^{\mathrm{mf,de}}_{T^b_{n+1}}(x)d\mu^{mf}_{T^b_{n+1}}$.

Remark 1. *Note that, in the treatment of the deficit work, we take in Process 2 the point of view of the jumping replica; whereas in Process 1, we take the point of view of the attracting replica which induces a branching.*

From the above probabilistic description, we can derive the Markov generator of the mean-field process, given by the sum of a diffusion and a jump generator (see the Appendix):

$$L^{\mathrm{mf}}_t = L_{\lambda(t)} + J_{t,\mu^{\mathrm{mf}}_t},$$

where the jump generator $J_{t,\mu^{\mathrm{mf}}_t}$ is defined as

$$J_{t,\mu^{\mathrm{mf}}_t}(\varphi)(x) = \beta\lambda'(t)\int(\varphi(y)-\varphi(x))\big(f^{\mathrm{mf,ex}}_t(x) + f^{\mathrm{mf,de}}_t(y)\big)d\mu^{\mathrm{mf}}_t(y).$$

A straitforward integration (using (5)) gives the fundamental balance identity of the jump generator:

$$J^*_{t,\mu^{\mathrm{mf}}_t}\big(\mu^{mf}_t\big) = \beta\left(\mathcal{F}^{\mathrm{mf}}_t - \frac{\partial H_{\lambda(t)}}{\partial\lambda}\right)\lambda'(t)\mu^{mf}_t$$

which implies, by forward Kolmogorov,

$$\partial_t\mu^{\mathrm{mf}}_t = L^*_{\lambda(t)}\big(\mu^{\mathrm{mf}}_t\big) + \beta\left(\mathcal{F}^{\mathrm{mf}}_t - \frac{\partial H_{\lambda(t)}}{\partial\lambda}\right)\lambda'(t)\mu^{\mathrm{mf}}_t$$

so that finally

$$\partial_t\left(\mu^{\mathrm{mf}}_t \mathrm{e}^{-\beta\int_0^t\mathcal{F}^{\mathrm{mf}}_s ds}\right) = L^*_{\lambda(t)}\left(\mu^{\mathrm{mf}}_t \mathrm{e}^{-\beta\int_0^t\mathcal{F}^{\mathrm{mf}}_s ds}\right) - \beta\frac{\partial H_{\lambda(t)}}{\partial\lambda}\lambda'(t)\mu^{\mathrm{mf}}_t \mathrm{e}^{-\beta\int_0^t\mathcal{F}^{\mathrm{mf}}_s\, ds}.$$

The latter is exactly the forward evolution equation of the Feynamn-Kac kernel $p^w_{0,t}$ defined in (11), and thus $\int p^w_{0,t}(x,.)d\mu_0(x) = \mu^{mf}_t e^{-\beta\int_0^t\mathcal{F}^{\mathrm{mf}}_s ds}$. Using (12), this gives the identities:

$$\mu^{\mathrm{mf}}_t = \mu_{\lambda(t)}, \quad \mathcal{F}^{\mathrm{mf}}_t = \mathcal{F}_{\lambda(t)}, \quad f^{\mathrm{mf,ex/de}}_t = f^{\mathrm{ex/de}}_{\lambda(t)}.$$

and proves the consistency of the IPS approximation scheme.

2.3. Numerical Implementation

In the previous section, we discretized the measure by considering an empirical approximation. For a numerical implementation to be tractable, it remains to discretize the time evolution. Notice already that the IPS method induces no extra computation of the forces, and is therefore unexpensive to implement. However,

although the IPS can be parallelized, the processors have to exchange informations at the end of each time step, which can slow down the simulation.

There are several ways to discretize the dynamics (7) or (9). Some common schemes used in molecular dynamics are the Euler-Maruyama discretization for (9), and the BBK scheme[2] for (7). We refer to ref. 3 for alternative approaches in the field of molecular dynamics. In the sequel, we will denote by $x^{n,k}$ a numerical approximation of a realization of $X^k_{n\Delta t}$, with $x = q$ or $x = (q, p)$ depending on the context.

Euler Discretization of the Overdamped Langevin Dynamics

The Euler-Maruyama numerical scheme associated to (9) reads, when taking integration time steps Δt,

$$q^{n+1} = q^n - \Delta t \nabla V_\lambda(q^n) + \sqrt{\frac{2\Delta t}{\beta}} R^n,$$

where $(R^n)_{n\in\mathbb{N}}$ is a sequence of independent and identically distributed (i.i.d.) $3N$-dimensional standard Gaussian random vectors. The numerical convergence of this scheme can be ensured in some circumstances.[3]

Discretization of the Langevin Dynamics

When considering an integration time step Δt, the BBK discretization of (7) reads componentwise (recall that the underscripts j refer here to the components of a given $x^k = x = (q, p)$),

$$\left\{ \begin{array}{l} p_j^{n+1/2} = p_j^n + \frac{\Delta t}{2}\left(-\nabla_{qj} V(q^n) - \xi \frac{p_j^n}{m_j} + \frac{\sigma_j}{\sqrt{\Delta t}} R_j^n\right) \\ q_j^{n+1} = q_j^n + \Delta t \frac{p_j^{n+1/2}}{m_j} \\ p_j^{n+1} = \frac{1}{1+\frac{\xi\Delta t}{2m_j}}\left(P_j^{n+1/2} - \frac{\Delta t}{2}\nabla_{q_j} V(q^{n+1}) + \sigma_j \frac{\sqrt{\Delta t}}{2} R_j^{n+1}\right) \end{array} \right.$$

where the random forcing terms $R_j^n (\in \{1, \cdots, N\}$ is the label of the particles, n is the iteration index) are standard i.i.d. Gaussian random variables. The fluctuation/dissipation relation (8) must be corrected so that the kinetic temperature is correct in the simulations.[3] To this end, we set

$$\sigma_j^2 = \frac{2\xi}{\beta}\left(1 + \frac{\xi\Delta t}{2m_j}\right).$$

Notice that the relation (8) is recovered in the limit $\Delta t \to 0$.

Discretization of the Selection Operation

We consider for example the following discretization of the force exerted on the k-th replica on the time interval $[n\Delta t, (n+1)\Delta t]$:

$$\frac{\partial H^{k,\Delta t}_{\lambda_{n+1/2}}}{\partial \lambda} = \frac{1}{2}\left(\frac{\partial H_{\lambda(n\Delta t)}}{\partial \lambda}(x^{n,k}) + \frac{\partial H_{\lambda((n+1)\Delta t)}}{\partial \lambda}(x^{n+1,k})\right).$$

The mean force is then approximated by

$$\mathcal{F}^{M,\Delta t}_{\lambda_{n+1/2}} = \frac{1}{M}\sum_{k=1}^{M} \frac{\partial H^{k,\Delta t}_{\lambda_{n+1/2}}}{\partial \lambda}.$$

To get a time dicretization of the IPS, Process 1 is mimicked using the following rules:

- the time integrals are changed into sums;
- the selection times are defined as the first discrete times exceeding the exponential clocks $\tau^{b/d}$.

Further details about the numerical implementation can be found in ref. 26. Note that one can find more elaborate methods of discretization of the IPS (see ref. 27), but this one seems to be sufficient in view of the intrinsic errors introduced by the discretization of the dynamics.

3. APPLICATIONS OF THE IPS METHOD

3.1. Computation of Canonical Averages

The most obvious application of the IPS method is the computation of phase-space integrals, since an unweighted sample of all Boltzmann distributions $(\mu_{\lambda(t)})_{t\in(0,T)}$ is generated. The sample obtained can of course be improved by some additional sampling process (according to a dynamics leaving the target canonical measure invariant). This will decorrelate the replicas and may increase the quality of the sample. We refer to ref. 3 for further precisions on sampling methods, to ref. 28 for some numerical tests assessing the quality of the sample generated as a function the transition time T and the number of replicas M, and to ref. 26 for an application to pentane.

3.2. Estimation of Free Energy Difference

The free energy of a system is defined (up to a constant) as

$$F = -\beta^{-1} \ln Z$$

where Z is the partition function $Z = \int \exp(-\beta H(q,p))\, dq\, dp$. It cannot be computed with a single sample of μ_λ. Only free energy differences can be computed easily. Since the free energy of certain states is known (This is the case for perfect gases, or for solids at low temperature[24]), the free energy of any state can in principle be obtained by an integration between a state whose free energy is known, and the state of interest. Usual methods to this end are Thermodynamic integration,[19] the free energy perturbation method[32] and the related Umbrella sampling technique,[30] or Jarzynski's non equilibrium dynamics (also called 'fast growth').[16] It is still a matter of debate which method is the most efficient. While some results show that fast growth methods can be competitive in some situations,[12] other studies disagree.[23] However, a fair comparison is difficult since the dynamics used may differ (in, ref. 23 Hamiltonian dynamics are used during the switching process), and more efficient fast growth methods techniques (using, e.g. path sampling[29,31]) are still under investigation.

In the work of Jarzynski,[16] M independent realizations $(X_t^1, \cdots, X_t^M)$ of a bare out of equilibrium dynamic (4) are used to compute free energy differences through (10), with the estimator

$$\Delta \hat{F}_J = -\frac{1}{\beta}\, \ln\left(\frac{1}{M}\sum_{k=1}^{M} e^{-\beta \mathcal{W}_1^k}\right).$$

An alternative (yet similar) estimator relying on the thermodynamical integration 13, which is also considered in ref. 29 (and is implicit in ref. 16) is

$$\Delta \hat{F}'_J = \int_0^T \mathcal{F}^{M_{\rm ind}}_{\lambda(t)} \lambda'(t)\, dt,$$

where

$$\mathcal{F}^{M_{\rm ind}}_{\lambda(t)} = \mu^{M_{\rm ind}}_{\lambda(t)}\left(\frac{\partial H_{\lambda(t)}}{\partial \lambda}\right), \quad \text{with} \quad \mu^{M_{\rm ind}}_{\lambda(t)}(dx) = \frac{\Sigma_{k=1}^M \delta_{X_t^k}(dx) \mathrm{e}^{-\beta \mathcal{W}_t^k}}{\Sigma_{k=1}^M \mathrm{e}^{-\beta \mathcal{W}_t^k}}.$$

However, both estimators $\Delta \hat{F}_J$ and $\Delta \hat{F}'_J$ suffer from the fact that only a few values of $\mathcal{W}_t^i$ are really important. Indeed, because of the exponential weighting, only the lower tail of the work distribution is taken into account. The quality of the estimation then relies on those rare values, which may be a problem in practice (see e.g. ref. 23.

In the case of interacting replicas, we use similarly

$$\Delta \hat{F}_{\rm IPS} = \int_0^T \mathcal{F}^{M}_{\lambda(t)} \lambda'(t)\, dt,$$

which shares by Proposition 2.1 the same statistical properties as $\Delta \hat{F}_J$: $\Delta \hat{F}_{\rm IPS}$ is asymptotically normal with bias and variance of order M^{-1}, and the estimator $e^{-\beta \Delta \hat{F}_{\rm IPS}}$ is unbiased estimator of $e^{\beta \Delta F}$.

Our numerical comparisons often turned out to give similar free energy estimations for the IPS method and the standard Jarzynski method. However, we have mostly considered the issue of pure energetic barriers, where the difficulty of sampling comes from overcoming a *single high* barrier. The observed numerical equivalence may be explained by the fact that the selection mechanism in the IPS method does not really help to *explore* those regions of high potential energy.

When the sampling difficulties also come from barriers of more *entropic* nature (e.g. a succession of very many transition states separated by low energy barriers), the IPS may improve the estimation. Indeed, the selection mechanism helps keeping a statistical amount of replica in the areas of high probability with respect to the local Boltzmann distribution μ_λ throughout the switching process (see the numerical example in Sec. 4.1). This relaxation property may be crucial to ensure at each time a meaningful exploration ability. These points are still under investigation.

3.3. Initial Guesses for Path Sampling

The problem of free energy estimation is deeply linked with the problem of sampling meaningful transition path. In the IPS method, one can associate to each replica X_t^k a *genealogical continuous* path $(X_s^{k,\mathrm{gen}})_{s\in(0,t)}$. The latter is constructed recursively as follows for a replica k (for $0 \le t \le T$):

- at each time t, set $X_t^{k,\mathrm{gen}} = X_t^k$;
- at each random time T_n when the replica jumps and adopts a new configuration (say of replica l), set $(X_s^{k,\mathrm{gen}})_{(0,T_n)} = (X_s^{l,\mathrm{gen}})_{(0,T_n)}$.

This path represents the ancestor line of the replica, and is composed of the past paths selected for their low work values. The study of the set of genealogical paths lies outside the scope of this article (see ref. 5 for a discussion in the discrete time case). However, let us mention that for a given $t \in (0, T)$, the set of genealogical paths is sampled, in the limit $M \to \infty$, according to the law of the non-equilibrium paths $(X_s^{\lambda(s)})_{s\in(0,t)}$ weighted by the factor $e^{-\beta \mathcal{W}_t}$ (with statistical properties analogous to those of proposition 2.1). These paths are thus typical among non-equilibrium dynamics of those with non-degenerate work. Therefore, they might be fruitfully used as non-trivial initial conditions for more specialized path sampling techniques (as e.g. ref. 31).

4. NUMERICAL SIMULATIONS

4.1. A Toy Example

Consider the following family of Hamiltonians $(H_\lambda)_{\lambda\in(0,1)}$:

$$H_\lambda(x) = \frac{x^2}{2} + \lambda Q_1(x) + \frac{\lambda^2}{2} Q_2(x) + \frac{\lambda^3}{6} Q_3(x) + \frac{\lambda^4}{24} Q_4(x) \qquad (17)$$

with

$$Q_1(x) = \frac{-1}{8x^2+1}, \quad Q_2(x) = \frac{-4}{8(x-1)^2+1},$$
$$Q_3(x) = \frac{-18}{32(x-3/2)^2+1}, \quad Q_4(x) = \frac{-84}{64(x-7/4)^2+1}.$$

Some of those functions are plotted in Fig. 1. This toy one-dimensional model is reminiscent of the typical difficulties encountered when μ_0 is very different from μ_1. Notice indeed that several transitional metastable states (denoted A and B in Fig. 1) occur in the canonical distribution when going from $\lambda = 0$ to $\lambda = 1$. The probability of presence in the basins of attraction of the main stable states of H_1 (C and D in Fig. 1) is only effective when λ is close to 1.

Simulations were performed at $\beta = 13$ with the overdamped Langevin dynamics (9), and the above Hamiltonian family (17). The number of replicas was $M = 1000$, the time step $\Delta t = 0.003$, and λ is considered to be linear: $\lambda(t) = t/T$. Figure 2 presents the distribution of replicas during a slow out of equilibrium plain

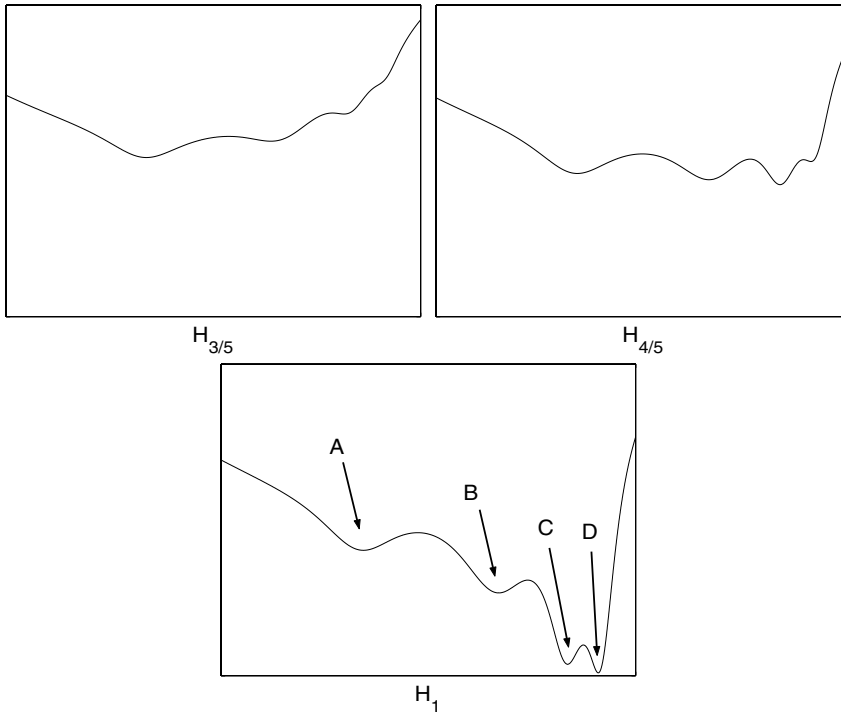

Fig. 1. Plot of some Hamiltonian functions, as defined by (17).

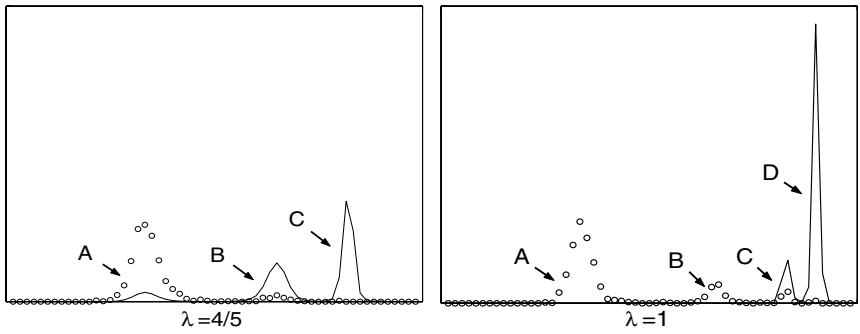

Fig. 2. Empirical densities (in dots) obtained using independant replicas.

dynamic: $T = 30$. Figure 3 presents the distribution of replicas during a faster dynamics with interaction: $T = 15$.

When performing a plain out of equilibrium dynamics (even 'slow') from $\lambda = 0$ to $\lambda = 1$, almost all replicas are trapped by the energy barrier of these transitional metastable states (see Fig. 2). In the end, a very small (almost null) proportion of replicas have performed interesting paths associated with low values of virtual work $\mathcal{W}$. When using (12) to compute thermodynamical quantities, these replicas bear almost all the weight of the degenerate sample, in view of the exponential weighting. The quality of the result therefore depends crucially on these rare values.

On the contrary, in the interacting version, the replicas can perform jumps in the configuration space thanks to the selection mechanism, and go from one metastable basin to another. In our example, as new transition states appear, only few clever replicas are necessary to attract the others in good areas (see Fig. 3). In the end, all replicas have the same weight, and the sample is not degenerate. Notice also that the final empirical distribution is fairly close to the theoretical one.

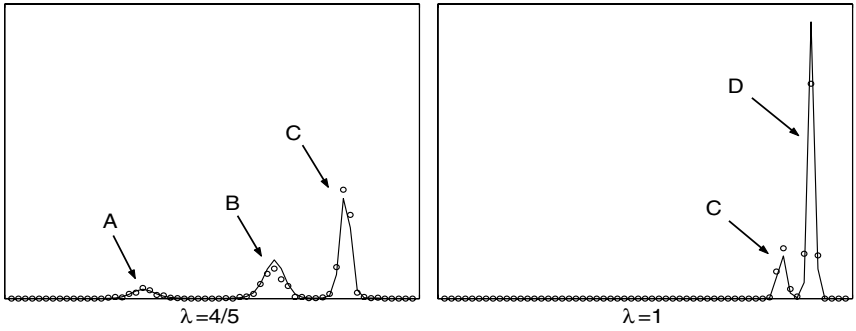

Fig. 3. Empirical densities (in dots) obtained using interacting replicas.

Table I. Error in Free Energy Estimation

Method	Bias	Variance
Plain	+0.25	0.19
Interacting	+0.15	0.10

We have also made a numerical estimation of the error of the free energy estimation, with 40 realizations of the above simulation. The results are presented in Table I, and show an important reduction of standard deviation and bias up to a factor 2 when using the IPS method.

4.2. Gradual Widom Insertion

We present here an application to the computation of the chemical potential of a soft sphere fluid. This example was considered in refs. 12 and 23 for example. We consider a two-dimensional (2D) fluid of volume $|\Omega|$, simulated with periodic boundary conditions, and formed of N particles interacting via a pairwise potential V. The chemical potential is defined, in the NVT ensemble, as

$$\mu = \frac{\partial F}{\partial N},$$

where F is the free-energy of the system. Actually, the kinetic part of the partition function Z can be straightforwardly computed, and accounts for the ideal gas contribution μ_{id}. In the large N limit, the chemical potential can be rewritten as[11]

$$\mu = \mu_{\mathrm{id}} + \mu_{\mathrm{ex}},$$

with

$$\mu_{\mathrm{id}} = -\beta^{-1} \ln \left(\frac{|\Omega|}{(N+1)\Lambda^3} \right),$$

where Λ is the "thermal de Broglie wavelength" $\Lambda = h(2\pi m \beta^{-1})^{-1/2}$ (with h Planck's constant). The excess part μ_{ex} is

$$\mu_{\mathrm{ex}} = -\beta^{-1} \ln \left(\frac{\int_{\Omega^{N+1}} \exp(-\beta V(q^{N+1}))\, dp^{N+1}}{|\Omega| \int_{\Omega} \exp(-\beta V(q^N))\, dq^N} \right),$$

where $V(q^N)$ is the potential energy of a fluid compose of N particles. We restrict ourselves to pairwise interactions, with an interaction potential Φ. Then, $V(q^N) = \sum_{1 \le i < j \le N} \Phi(|q_i - q_j|)$. Setting $\pi(q^N) = Z^{-1} \exp(-\beta V(q^N))$with $Z = \int_{\Omega^N} \exp[-\beta V(q^N)] dq^N)$ and $\Delta V(q^N, q) = V(q^{N+1}) - V(q^N)$ with $q^{N+1}(q^N, q)$, it follows

$$\mu_{\mathrm{ex}} = -\beta^{-1} \ln \left(\frac{1}{|\Omega|} \int_{\Omega} e^{-\beta \Delta V(q, q^N)} d\pi(q^N)\, dq \right). \tag{18}$$

The formula (18) can be used to compute the value of chemical potential using stochastic methods such as the free energy perturbation (FEP) method.[32] In this case, we first generate a sample of configurations of the system according to π, and then evaluate the integration in the remaining q variable by drawing positions q of the remaining variable uniformly in Ω.

Another possibility is to use fast growth methods, resorting to the following parametrization

$$H_\lambda(q^{N+1}, p^{N+1}) = \sum_{i=1}^{N+1} \frac{p_i^2}{2m} + V_\lambda(q^{N+1}) = \sum_{i=1}^{N+1} \frac{p_i^2}{2m} + V(q^N) + \lambda \Delta V(q^N, q).$$

In this case, the interactions of the remaining particle with the N first ones are progressively turned on.

As in refs. 12 and 23 we use a smoothed Lennard-Jones potential in order to avoid the singularity at the origin (Let us however note that, once the particle is inserted, it is still possible to change all the potentials to Lennard-Jones potentials, and compute the correponding free-energy difference). The Lennard Jones potential reads

$$\Phi_{\mathrm{LJ}}(r) = 2\epsilon \left(\frac{1}{2} \left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^{6} \right),$$

and the modified potential is

$$\Phi(r) = \begin{cases} a - br^2, & 0 \le r \le 0.8\,\sigma, \\ \Phi_{\mathrm{LJ}}(r) + c(r - r_c) - d, & 0.8\,\sigma \le r \le r_c, \\ 0, & r \ge r_c. \end{cases}$$

The values *a, b, c* are chosen so that the potential is C^1. The distance r_c is a prescribed cut-off radius. We consider the insertion of a particle in a 2D fluid of 25 particles, at a density $\rho\sigma^3 = 0.8$, with $r_c = 2.5\,\sigma$, $\beta\epsilon = 1$, $\Delta t = 0.0005$, and a schedule $\lambda(t) = t/T$ where T is the transition time. The results are presented in Table II, for different transitions times, but at a fixed computational cost, since

Table II. Free Energy Estimation for one Realization of Each Method, Depending on the Switching time *T* and the Number of Replicas *M* Used, Keeping *MT* Constant

Method	$M = 10^5$ $T = 1$	$M = 5 \times 10^4$ $T = 2$	$M = 2 \times 10^4$ $T = 5$	$M = 10^4$ $T = 10$
Plain noneq. dyn.	1.31 (0.015)	1.33 (0.017)	1.32 (0.023)	1.32 (0.038)
IPS	1.37 (0.025)	1.35 (0.040)	1.33 (0.033)	1.32 (0.037)

Note. The results are averaged over 10 realizations, and are presented under the form $\langle \mu \rangle > (\sqrt{\mathrm{Var}(\mu)})$. The reference value obtained through FEP is $\mu_{\mathrm{ex}} = 1.32\,k_{\mathrm{B}}T(\pm 0.01\,k_{\mathrm{B}}T)$. Notice that the results are quite comparable.

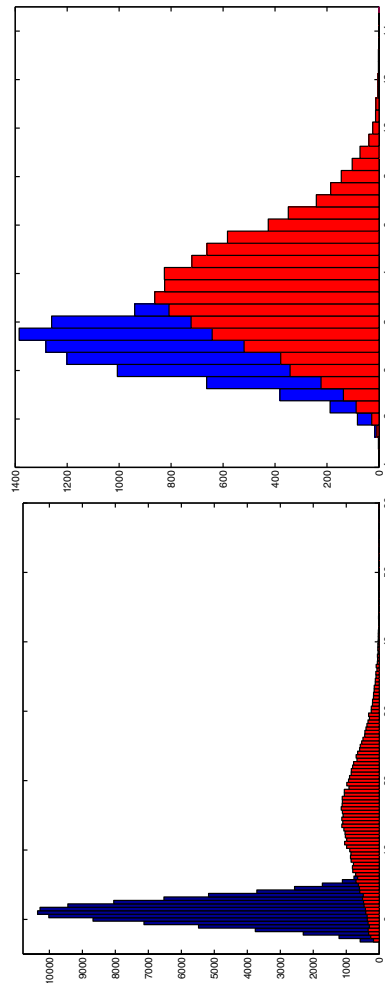

Fig. 4. Left: Comparison of the work distribution for $T = 1$. Right: Comparison of the work distributions for $T = 10$. The IPS results appear in darker colors. The target value is $1.32\, k_{\mathrm{B}} T$. Notice that the IPS work distribution is Gaussian with low variance even for the fast switching simulation.

MT is constant. Some work distributions are also depicted in Fig. 4. A reference value was computed using FEP, with 10^8 insertions, done by running $M = 10^3$ independent Langevins dynamics for the system composed of N particles, for a time $t_{\mathrm{FEP}} = 50$ (after an initial thermalization time to decorrelate the systems), and inserting one particle at random after each time-step. The reference value obtained is $\mu_{\mathrm{ex}} = 1.32 k_{\mathrm{B}} T (\pm 0.01\, k_{\mathrm{B}} T)$.

As can be seen from the results in Table II, the IPS algorithm has a comparable accuracy to Jarzynski's estimates provided the switching time is long enough. However, the work distribution is very different, and has a stable gaussian shape for all switching rates considered, whereas the work distribution obtained through the fast growth method are much wider (see in particular Fig. 4 (Left)), so that the relevant part of the work distribution (the lower tail) is only of small relative importance.

APPENDIX: PURE JUMP PROCESSES

Consider a Markov process X_t of infinitesimal generator

$$J_t(f)(x) = \int (f(y) - f(x)) \alpha_t(x) d\mu_t(y),$$

where for each time t, α_t is a bounded positive function and μ_t a probability measure. Denote by $(T_n)_{n \geq 1}$ the jump times (with $T_0 = 0$), and $(\tau_n)_{n \geq 1}$ independant clocks of exponential law of mean 1. Then the system evolves according to the following stochastic rules:

- the jump times are defined by

$$\int_{T_n}^{T_{n+1}} \alpha_t(X_t)\, dt = \tau_{n+1};$$

- at jump times, the process jumps to a configuration chosen according to the probability measure $d\mu_{T_{n+1}}(y)$.

REFERENCES

1. R. Assaraf, M. Caffarel and A. Khelif, Diffusion Monte Carlo with a fixed number of walkers. *Phys. Rev. E* **61**(4):4566–4575 (2000).
2. A. Brünger, C. B. Brooks and M. Karplus, Stochastic boundary conditions for molecular dynamics simulations of ST2 water. *Chem. Phys. Lett.* **105**:495–500 (1983).
3. E. Cancès, F. Legoll and G. Stoltz, Theoretical and numerical comparison of sampling methods for molecular dynamics. *IMA Preprint* **2040** (2005).
4. E. A. Carter, G. Ciccotti, J. T. Hynes and R. Kapral, Constrained reaction coordinate dynamics for the simulation of rare events. *Chem. Phys. Lett.* **156**:472–477 (1989).
5. P. Del Moral, *Feynman-Kac Formulae, Genealogical and Interacting Particle Systems with Applications*, Springer Series Probability and its Applications (Springer, 2004).

6. P. Del Moral, A. Doucet and G. W. Peters, Sequential Monte Carlo samplers, *Preprint version*, available on request at the URL http://www.cs.ubc.ca/∼arnaud/ (2004).
7. P. Del Moral and L. Miclo, Branching and Interacting Particle Systems approximations of Feynman -Kac formulae with applications to nonlinear filtering, Séminaire de Probabilités XXXI. *Lecture notes in Mathematics* **1729**:1–145 (2000).
8. A. Doucet, N. de Freitas and N. J. Gordon, *Sequential Monte Carlo Methods in Practice*, Series Statistics for Engineering and Information Science (Springer, 2001).
9. A. Doucet, M. Rousset, Time continuous limit of sequential samplers, *in preparation.*
10. M. I. Freidlin and A. D. Wentzell, *Random perturbations of dynamical systems*, second edition (Springer-Verlag, 1998).
11. D. Frenkel and B. Smit, *Understanding Molecular Simulation* (Academic Press, 2002).
12. D. A. Hendrix and C. Jarzynski, A "fast growth" method of computing free energy differences. *J. Chem. Phys.* **114**(14):5974–5981 (2001).
13. K. Hukushima and Y. Iba, Population annealing and its application to a spin glass. *AIP Conference Proceedings* **690**(1):200–206 (2003).
14. G. Hummer. and A. Szabo, Free energy reconstruction from nonequilibrium single-molecule pulling experiments. *PNAS* **98**(7):3658–3661 (2001).
15. Y. Iba, Extended ensemble Monte Carlo. *Int. J. Mod. Phys. C***12**:623–656 (2001).
16. C. Jarzynski, Equilibrium free energy differences from nonequilibrium measurements: A master equation approach. *Phys. Rev. E* **56**(5):5018–5035 (1997).
17. C. Jarzynski, Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.* **78**(14):2690–2693 (1997).
18. S. Kirkpatrick, C. G. Gelatt and M. P. Vecchi, Optimization by simulated annealing. *Science.* **220**:671–680 (1983).
19. J. G. Kirkwood, Statistical mechanics of fluid mixtures. *J. Chem. Phys.* **3**:300–313 (1935).
20. T. Lelièvre, M. Rousset and G. Stoltz, Computation of free energy differences through nonequilibrium stochastic dynamics: the reaction coordinate case, arXiv preprint (condmat/0603426) (2006).
21. E. Marinari and G. Parisi, Simulated tempering: A new Monte Carlo scheme. *Europhys. Lett.***19**:451–458 (1992).
22. R. Neal, Annealing importance sampling. *Statistics and Computing* **11**(2):125–139 (2001).
23. H. Oberhofer, C. Dellago and P. L. Geissler, Biased sampling of non-equilibrium trajectories: Can fast switching simulations outperform conventional free energy calculation methods? *J. Chem. Phys. B* **109**:6902–6915 (2005).
24. J. M. Rickman and R. LeSar, Free-energy calculations in materials research. *Annu. Rev. Matter. Res.* **32**:195–217 (2002).
25. M. Rousset, On the control of an Interacting Particle estimation of Schrödinger groundstates, *Preprint*, available at the URL www.Isp.ups-tlse.fr/Fp/Rousset/article3.pdf
26. M. Rousset and G. Stoltz, An interacting particle system approach for molecular dynamics. *Rapport de recherche CERMICS* **283** (2005).
27. M. Rousset *Phd Thesis*, in preparation.
28. G. Stoltz *Phd Thesis*, in preparation.
29. S. Sun, Equilibrium free energies from path sampling of nonequilibrium trajectories. *J. Chem. Phys.* **118**(13):5769–5775 (2003).
30. G. M. Torrie and J. P. Valleau, Nonphysical sampling distributions in Monte-Carlo free energy estimation: Umbrella sampling. *J. Comp. Phys.* **23**:187–199 (1977).
31. F. M. Ytreberg and D. M. Zuckerman, Single-ensemble nonequilibrium path sampling estimates of free energy differences. *J. Chem. Phys.* **120**(3):10876–10879 (2004).
32. R. Zwanzig, High-temperature equation of state by a perturbation method: I. Nonpolar gases. *J. Chem. Phys.* **22**:1420–1426 (1954).