(Non)equilibrium computation of equilibrium properties

Gabriel STOLTZ[1,2]

[1] CERMICS, ENPC (Marne-la-Vallée, France)

[2] CEA/DAM (Bruyères-le-Châtel, France)

work in collaboration with Eric Cancès, Frédéric Legoll, Tony Lelièvre and Mathias Rousset

`http://cermics.enpc.fr/~stoltz/`

Outline of the talk

- Computation of static properties
  - The sampling problem
  - Presentation of some mixed stochastic/molecular dynamics methods
- Computation of free energy differences
  - Static methods (thermodynamic integration)
  - Out of equilibrium dynamics (Jarzynski)
  - Equilibration of the out of equilibrium dynamics

Sampling the canonical ensemble

Description of a classical system

- Microscopic description of a classical system (N particles):

$$(q,p) = (q_1, \ldots, q_N,\ p_1, \ldots, p_N) \in T^*\mathcal{M}$$

- Usually, $T^*\mathcal{M} = \mathbb{R}^{3N} \times \mathbb{R}^{3N}$ or $\mathbb{T}^{3N} \times \mathbb{R}^{3N}$
- More complicated situations can be considered... (submanifolds)
- Positions q (configuration)
- Momenta $p = M\dot{q}$ (M mass matrix)
- Energy

$$H(q,p) = \sum_{i=1}^{N} \frac{p_i^2}{2m_i} + V(q_1, \ldots, q_N)$$

- All the physics is contained in V

Computation of observables (static)

- Equilibrium thermodynamic properties:

$$\langle A \rangle = \int_{\mathcal{M} \times \mathbb{R}^{3N}} A(q,p)\, d\mu(q,p)$$

- Choice of a thermodynamic ensemble $\equiv$ choice of a measure $d\mu$:
  - microcanonical (NVE, constant energy) ;
  - canonical (NVT, "constant temperature") : Boltzmann measure

$$d\mu_{NVT} = \frac{1}{Z_{NVT}} \exp(-\beta H(q,p))\, dq\, dp, \quad \beta = 1/(k_B T)$$

  - Other choices are possible (grand-canonical, constant pressure,. . .)
  - Certain properties can not be computed this way (free energy, entropy)!

Problem and methods

- $A(q,p)$ and $d\mu(q,p)$ are given by physics
- Question: how to compute $\langle A \rangle = \int A(q,p)\, d\mu(q,p)$?
- Difficulty: large dimension: $(q,p) \in \mathcal{M} \times \mathbb{R}^{3N} \subset \mathbb{R}^{6N}$ with $N \geq 10^5$
- Different methods:
  - purely stochastic methods: generate random points (q_n, p_n) i.i.d. according to $d\mu$ and resort to the Law of Large Numbers:

$$\lim_{N_{\text{iter}} \to \infty} \frac{1}{N_{\text{iter}}} \sum_{n=1}^{N_{\text{iter}}} A(q^n, p^n) = \langle A \rangle$$

  - Markov chain based techniques: consider one or several realizations (q_n, p_n) of a Markov chain letting $d\mu$ invariant + "LLN".
  - (extended) molecular dynamics methods (Nosé-Hoover and beyond).

- The process $(W_t)_{t\geq 0}$ is a standard brownian motion if for $0 = t_0 < t_1 < \cdots < t_n$, the random variables $W_{t_{i+1}} - W_{t_i}$ are independent and distributed according to

$$W_{t_{i+1}} - W_{t_i} \sim \mathcal{N}(0, t_{i+1} - t_i) \sim \frac{1}{\sqrt{2\pi(t_{i+1} - t_i)}} \exp\left(-\frac{w^2}{2(t_{i+1} - t_i)}\right)$$

- Euler-Maruyama discretization of the SDE $dX_t = \sigma dW_t$:

$$x^{n+1} = x^n + \sigma\sqrt{\Delta t}\, U^n,$$

where $(U^n)_{n\geq 0}$ are i.i.d. standard gaussian random variables.

- Notice that this implies $x^n \sim \mathcal{N}(0, n\Delta t)$ (diffusive behaviour)

Overdamped Langevin dynamics

- Limit $M \to 0$ of the Langevin dynamics (see next slide!)
- SDE on the configurational part only

$$dq_t = -\nabla V(q_t)dt + \sigma dW_t,$$

  where $(W_t)_{t\geq 0}$ is a standard Wiener process of dimension dN
- Canonical measure $d\pi(q) \sim Z^{-1}\mathrm{e}^{-\beta V(q)}$ is invariant (steady solution of the associated Fokker-Planck equation)
- Fluctuation/dissipation relation $\sigma = (2/\beta)^{1/2}$
- Euler-Maruyama discretization

$$q^{n+1} = q^n - \nabla V(q^n)\,\Delta t + \sigma\sqrt{\Delta t}\,U^n$$

- Ergodicity at the continuous/numerical levels

Langevin dynamics

- Hypo-elliptic SDE (noise on p only)

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

  where $(W_t)_{t \geq 0}$ standard Wiener process of dimension dN and fluctuation/dissipation relation $\sigma = (2\xi/\beta)^{1/2}$.

- In this case, $d\mu(q,p) = \frac{1}{Z} \exp\left(-\beta H(q,p)\right) dqdp$ is an invariant measure (cf. Fokker-Planck equation)

- Ergocity along one trajectory garanteed through irreducibility (hypoelliptic process) + existence of an invariant measure

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T A(q_t^x) \, dt = \int_{\mathcal{M}} A(q) d\pi \quad \text{a.s.}$$

- Various discretizations have been proposed (BBK, Allen-Tildesley,. . .)

Computation of free energy differences

Free energy

- Free energy (not an average !) = "available phase space"

$$F = -\beta^{-1} \ln \left(\frac{1}{N!} \int_{\mathcal{M}} \mathrm{e}^{-\beta V(q)} \, dq \right)$$

- Can be computed directly only in certain cases (ideal gas, solids at low temperature,. . .)
- Free energy differences are easier to compute
  - 'Alchemical' transition (external parameter in the Hamiltonian)
  - Reaction coordinate (internal)

"Alchemical" transitions

- Potential energy $V_\lambda(q_1, \ldots, q_N)$: external parameter λ (temperature, intensity of a magnetic field, Widom insertion,...)
- Initial state: $\lambda = 0 \rightarrow$ final state: $\lambda = 1$
- Free energy differences

$$\Delta F(\lambda) = -\beta^{-1} \ln \left(\frac{\int_{\mathcal{M}} \mathrm{e}^{-\beta V_\lambda(q)}\, dq}{\int_{\mathcal{M}} \mathrm{e}^{-\beta V_0(q)}\, dq} \right)$$

- Alchemical transition = particular reaction coordinate ($Q = (\lambda, q)$)

The reaction coordinate case

- reaction coordinate $\xi(q) \simeq$ reduced description of the system in terms of a few relevant degrees of freedom
- Potential of mean force: different definitions... Most general form

$$F_f(z) = -\beta^{-1} \ln \left(Z^{-1} \int_{\Sigma_z} \mathrm{e}^{-\beta V(q)} \, f(q) \, d\sigma_{\Sigma_z}(q) \right)$$

where the manifold $\Sigma_z = \{q \in \mathcal{M} \,|\, \xi(q) = z\}$ and $Z = \int \mathrm{e}^{-\beta V(q)} \, dq$.

- Intrinsic version $f \equiv 1$
- Non-intrinsic choice: $f = |\nabla \xi(q)|^{-1}$

$$F(z) = -\beta^{-1} \ln \left(Z^{-1} \int_{\Sigma_z} \mathrm{e}^{-\beta V(q)} \, \delta_{\xi(q)-z} \, dq \right)$$

The free energy also depends on the local values of the gradient of the reaction coordinate...

Usual techniques to compute free-energy differences

- Thermodynamic integration[a]
- Free-energy perturbations[b]

$$\Delta F(\lambda) = -\beta^{-1} \ln \left\langle \mathrm{e}^{-\beta(V_\lambda(q) - V_0(q))} \right\rangle_0$$

where $\langle \cdot \rangle_0$ denotes a canonical average w.r.t. $d\mu_0(q) = Z^{-1} \mathrm{e}^{-\beta V_0(q)}\, dq$
$\rightarrow$ usual sampling techniques

- Many extensions and refinements (e.g. Umbrella sampling[c])
- Recent alternative: nonequilibrium dynamics[d]

---

[a] J.G. Kirkwood, *J. Chem. Phys.* **3**, 300 (1935)

[b] R. Zwanzig, *J. Chem. Phys.* **22**, 1420 (1954)

[c] G.M. Torrie and J.P. Valleau, *J. Comp. Phys.* **23**, 187 (1977)

[d] C. Jarzynski, *Phys. Rev. E* **56**(5) 5018 (1997)

Thermodynamic integration

- Free energy = integral

$$\Delta F = \int_0^1 F'(\lambda)\, d\lambda \simeq \sum_{i=1}^{M} (\lambda_i - \lambda_{i-1}) \left(\frac{F'(\lambda_{i-1}) + F'(\lambda_i)}{2} \right)$$

where the mean force is

$$F'(\lambda) = \frac{\int_{\mathcal{M}} \partial_\lambda V_\lambda(q)\, \mathrm{e}^{-\beta V_\lambda(q)}\, dq}{\int_{\mathcal{M}} \mathrm{e}^{-\beta V_\lambda(q)}\, dq} \simeq \frac{1}{n} \sum_{i=0}^{n} \frac{\partial V_\lambda}{\partial \lambda}(q_\lambda^n)$$

if (q_λ^n) is sampled according to the Boltzmann measure associated with V_λ

- Ergodic stochastic dynamics for a fixed value λ_i, and consider a sequence $\lambda_i \in [0, 1]$
- Alternative: average over a single long trajectory with λ varying "infinitely" slowly (quasi-static transformation)

Thermodynamic integration (2)

- Reaction coordinate $\xi(q)$
- Associated free energy

$$F_f(z) = -\beta^{-1} \ln \left(Z^{-1} \int_{\Sigma_z} \mathrm{e}^{-\beta V(q)} \, f(q) \, d\sigma_{\Sigma_z}(q) \right)$$

- Free energy difference

$$\Delta F_f = \int_0^1 F_f'(z) \, dz$$

- Mean force = canonical average on a submanifold
- Analytical expression not practical...

$$F_f'(z) = \int_{\Sigma_z} \left[\frac{\nabla V_f(q) \cdot \nabla \xi(q)}{|\nabla \xi(q)|^2} - \frac{1}{\beta |\nabla \xi(q)|} \nabla \cdot \left(\frac{\nabla \xi(\mathbf{q})}{|\nabla \xi(\mathbf{q})|} \right) \right] d\mu_{\Sigma_z, f}$$

with $V_f = V - \beta^{-1} \ln f$ and $\mu_{\Sigma_z, f} = Z_{z,f}^{-1} \, f(q) \, \mathrm{e}^{-\beta V(q)} \, dq$

Thermodynamic integration (3)

- Projected SDE using a Lagrange multiplier formulation

$$dq_{z,s} = -\nabla V(q_{z,s})\, ds + \sqrt{\frac{2}{\beta}}\, dW_s + dr_{z,s}$$

where $r_{z,s}$ is directed along $\nabla\xi(q_{z,s})$ and is such that $\xi(q_{z,s}) = z$.

- In practice, for a fixed value z, discretization (algorithmic time step Δs)

$$q_z^{n+1} = q_z^n - \nabla V_f(q_z^n)\, \Delta s + \sqrt{\frac{2\Delta s}{\beta}}\, U^n + r_z^{n+1}\, \nabla\xi(q_z^{n(+1)})$$

with r_z^{n+1} such that the constraint $\xi(q_z^{n+1}) = z$ is satisfied.

- Mean force = average over Lagrange multipliers

$$\lim_{n\to+\infty} \frac{1}{n\Delta s} \sum_{m=1}^{n} r_z^m = F_f'(z) + \mathrm{o}(\Delta s)$$

Nonequilibrium dynamics (alchemical case)

- Why not switch λ at an arbitrary rate?
- Schedule $\Lambda(t)$ such that $\Lambda(0) = 0$, $\Lambda(T) = 1$, $0 < T < +\infty$
- Start from canonical initial conditions $q_0 \sim d\mu_0(q)$
- Time inhomogeneous Markovian evolution (the potential energy changes!)

$$dq_{\Lambda,t} = -\nabla V_{\Lambda(t)}(q_{\Lambda,t})\, dt + \sqrt{\frac{2}{\beta}}\, dW_t$$

- In particular, the law of $q_{\Lambda,t}$ is not $Z^{-1}_{\Lambda(t)}\, \mathrm{e}^{-\beta V_{\Lambda(t)}}$... (distribution lags behind...) $\rightarrow$ correction through reweighting

Nonequilibrium dynamics (2)

- Definition: virtual work exerted on a trajectory

$$\mathcal{W}_T = \int_0^T \frac{\partial V_{\Lambda(t)}}{\partial \Lambda}(q_{\Lambda,t})\,\Lambda'(t)\,dt$$

- Feynman-Kac formula[a], usually known as the "Jarzynski equality"

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

- Consequence ("second law of thermodynamics"):

$$\Delta F(\Lambda(t)) \geq \mathbb{E}(\mathcal{W}_t)$$

- Remark: NVE dynamics also possible
- Remark: there are experimental validations

[a] G. Hummer and A. Szabo, *PNAS* **98**(7) (2001) 3658-3661.

Nonequilibrium dynamics (3)

- Practical implementation ($\lambda^0 = 0, \lambda^N = 1, N = T/\Delta t$) for $0 \leq i \leq M$ systems
  - update the parameter (for example $\lambda^n = \lambda^{n-1} + 1/N$)
  - one-step unconstrained diffusion (overdamped Langevin)

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

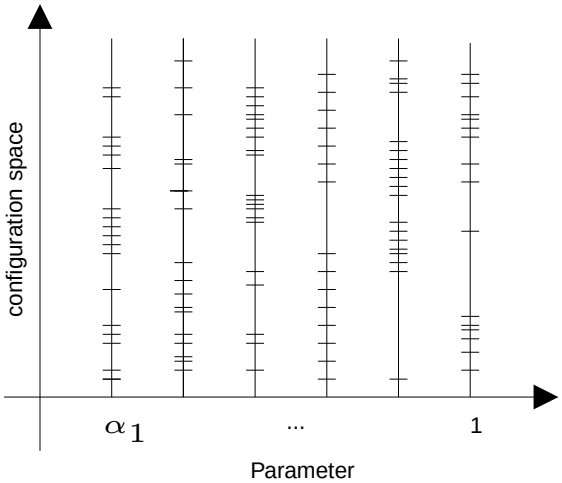
  - update the work

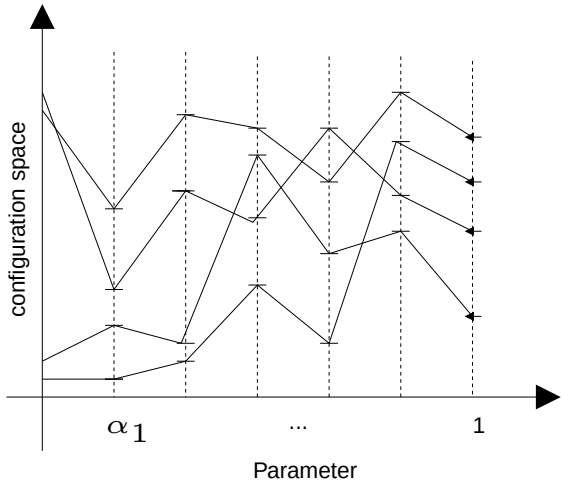
$$\mathcal{W}^{i,n+1} = \mathcal{W}^{i,n} + \frac{1}{2}\left(\frac{\partial V_{\lambda^n}}{\partial \lambda}(q^{i,n}) + \frac{\partial V_{\lambda^{n+1}}}{\partial \lambda}(q^{i,n+1})\right)(\lambda^{n+1} - \lambda^n).$$

- (Biased) free-energy estimator

$$\Delta F = -\beta^{-1} \ln\left(\frac{1}{M}\sum_{i=1}^{M} \mathcal{W}^{i,N}\right)$$

Before going on...





Left: Thermodynamic integration (adiabatic).

Right: Jarzynski nonequilibrium dynamics (canonical initial conditions, reweighting).

Nonequilibrium dynamics: The reaction coordinate case

- Geometric extension: reaction coordinate $\xi(q)$
- Variation of the constraint $z(t)$ such that $z(0) = 0, z(T) = 1$
- The "right" dynamics to consider is

$$\left\{ \begin{array}{rcl} X_0 & \sim & \mu_{\Sigma_{z(0)}} \\ dX_t & = & -\nabla V(X_t)dt + \sqrt{2\beta^{-1}}dW_t + \nabla\xi(X_t)\, d\Lambda_t \end{array} \right.$$

  with $d\Lambda_t$ such that $d[\xi(X_t)] = z'(t)dt$.

- Decomposition of the Lagrange multiplier as $\Lambda_t = \Lambda_t^{\mathrm{d}} + \Lambda_t^{\mathrm{m}} + \Lambda_t^{\mathrm{f}}$ with
  - martingale part $\nabla\xi(X_t)\, d\Lambda_t^{\mathrm{m}} = -\sqrt{2\beta^{-1}}P^{\perp}(X_t)\, dW_t,$
  - local force part $\nabla\xi(X_t)\, d\Lambda_t^{\mathrm{f}} = f(X_t)\, dt,$
  - forcing term (bias allowing a finite time switch)

$$\nabla\xi(X_t)d\Lambda_t^d = \frac{\nabla\xi(X_t)}{|\nabla\xi(X_t)|^2}\, z'(t)dt = P^{\perp}(X_t)dX_t.$$

Nonequilibrium dynamics: The reaction coordinate case (2)

- Definition of the work

$$\mathcal{W}_t = \int_0^t f(X_s)\, z'(s)\, ds = \int_0^t z'(s) d\Lambda_s^{\mathrm{f}}$$

- Feynman-Kac formula

$$\mathrm{e}^{-\beta(F(z(t))-F(z(0)))} = \mathbb{E}\left(\mathrm{e}^{-\beta \mathcal{W}_t}\right).$$

- For example, discretization of the dynamics and computation of the work according to

$$\mathcal{W}_n = \mathcal{W}_{n-1} + \frac{z(t_n) - z(t_{n-1})}{t_n - t_{n-1}} \lambda_{n+1}^{\mathrm{f}},$$

with

$$\lambda_{n+1}^{\mathrm{f}} = \lambda_{n+1} - \frac{z(t_{n+1}) - z(t_n)}{|\nabla\xi(q_n)|^2} + \sqrt{\frac{2\Delta t}{\beta}} \frac{\nabla\xi(q_n)}{|\nabla\xi(q_n)|^2} \cdot U_n.$$

Equilibrating the nonequilibrium dynamics

- Problems with the usual approach: exponential weights = only the lower tail of the work distribution counts (statistical confidence?)
- Source term in Fokker-Planck equation is reinterpreted in a probabilistic way through a birth/death process ($\simeq$ elegant continuous resampling)
- Jump/diffusion process $\bar{q}_{\Lambda,t}$ enhancing works lower than the average (and penalizing works larger than the average) in order to maintain equilibrium at all times
  - "Interacting Particle System" (genetic algorithm: mutation = diffusion)
  - the law of $\bar{q}_{\Lambda,t}$ is the canonical measure at all times
  - $\mathbb{E}(\bar{\mathcal{W}}_T) = \Delta F$
- "Simulated annealing" strategy (comparison with parallel tempering)

Equilibrating the nonequilibrium dynamics (2)

- M replicas of the sytem
- Evolution in a mean-field sense: empirical mean force

$$\mathcal{F}^M_{\Lambda(t)} = \frac{1}{M}\sum_{k=1}^{M} \frac{\partial V_{\Lambda(t)}}{\partial \Lambda}(q^k_{\Lambda,t})$$

and empirical Boltzmann distribution

$$d\mu^M_{\lambda(t)}(x) = \frac{1}{M}\sum_{k=1}^{M} \delta_{q^k_t}(dx),$$

- Excess (penalized) / deficit (enhanced) works

$$W_t^{k,\mathrm{ex/de}} = \int_0^t \left(\frac{\partial V_{\Lambda(s)}}{\partial \Lambda}(q^k_{\Lambda,t}) - \mathcal{F}^M_{\Lambda(s)} \right)^{+/-} \Lambda'(s)\, ds$$

Equilibrating the nonequilibrium dynamics (3)

Algorithm Initial distribution $(q_0^1, \dots, q_0^M) \sim d\pi_0(q)$, initial jump times $T_0^{k,d} = 0, T_0^{k,b} = 0$.

- Between each jump time, evolve independently the replicas $q_{\Lambda,t}^k$ according to the chosen dynamics;
- [Death] At random times $T_n^{k,\mathrm{ex}}$ such that

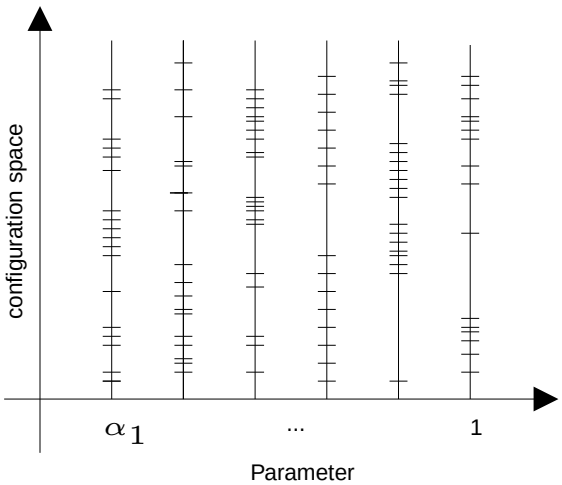
$$W^{k,\mathrm{ex}}_{T^{k,d}_{n+1}} - W^{k,\mathrm{ex}}_{T^{k,\mathrm{ex}}_n} = \tau_n^{k,\mathrm{ex}} \sim \mathcal{E}(\beta^{-1})$$

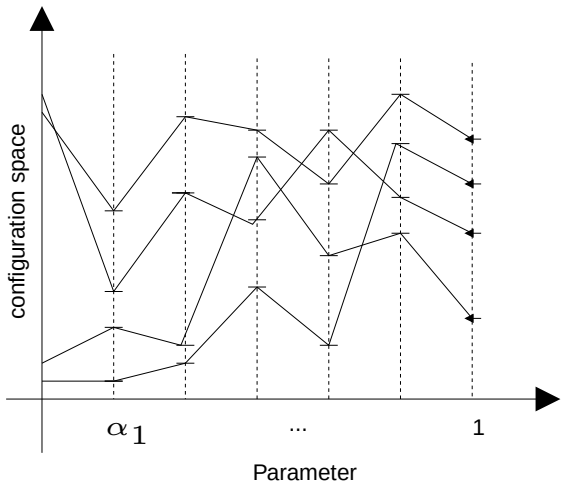
  the k-th configuration is replaced by a configuration $l \in \{1, \dots, M\}$ chosen at random;
- [Birth] At random times $T_n^{k,de}$ such that

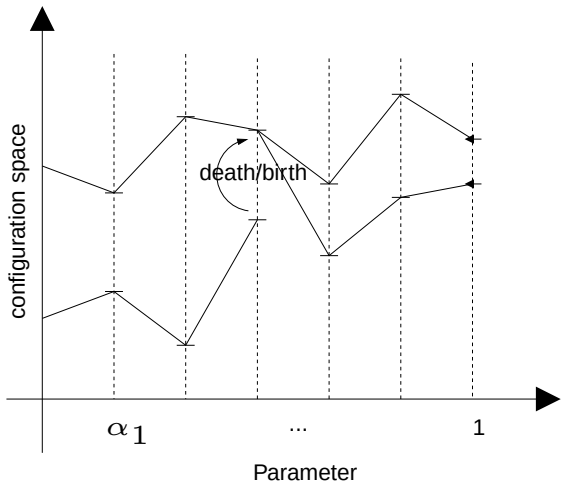
$$W^{k,\mathrm{de}}_{T^{k,\mathrm{de}}_{n+1}} - W^{k,\mathrm{de}}_{T^{k,\mathrm{de}}_n} = \tau_n^{k,\mathrm{de}} \sim \mathcal{E}(\beta^{-1})$$

  a configuration $l \in \{1, \dots, M\}$ chosen at random is replaced by the k-th configuration.

Final cartoon comparison...





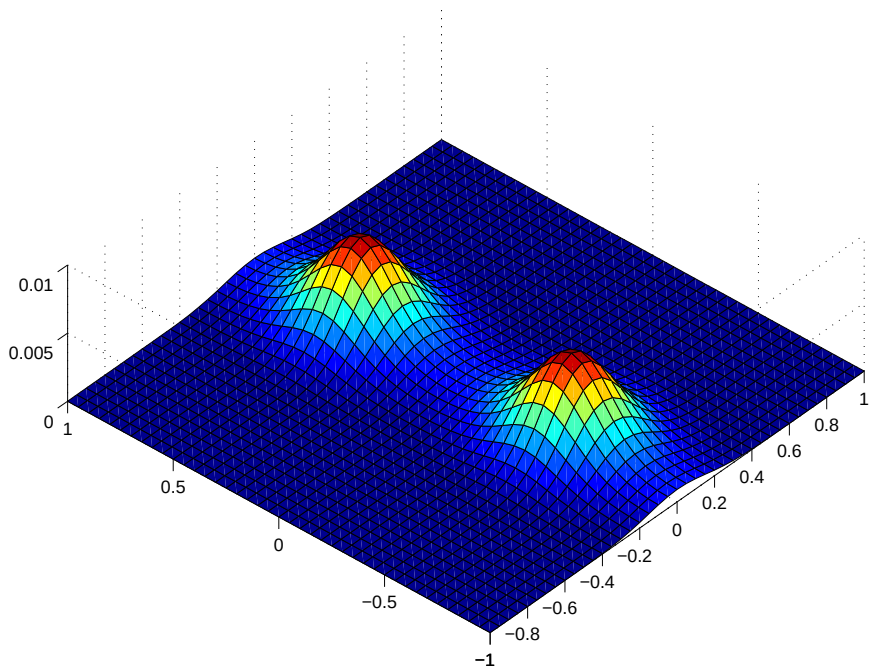


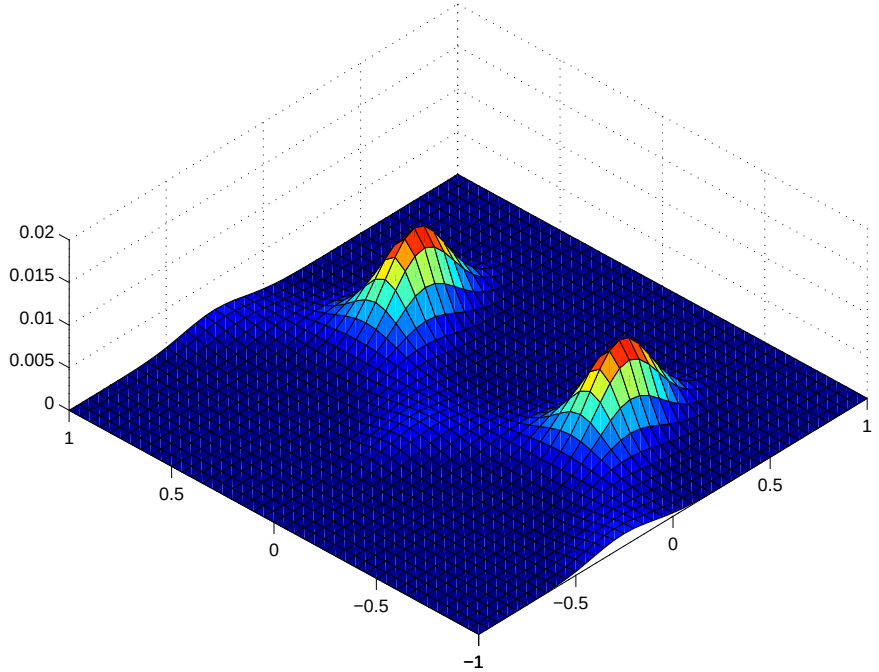
Left: Thermodynamic integration (adiabatic).

Middle: Jarzynski nonequilibrium dynamics (canonical initial conditions, reweighting).

Right: Interacting Particle Strategy (canonical initial conditions, birth/death process).

A 2D toy model





2D potential[a] $V(x,y) = \cos(2\pi x)(1 + d_1 y) + 2d_2\pi^2 y^2$.

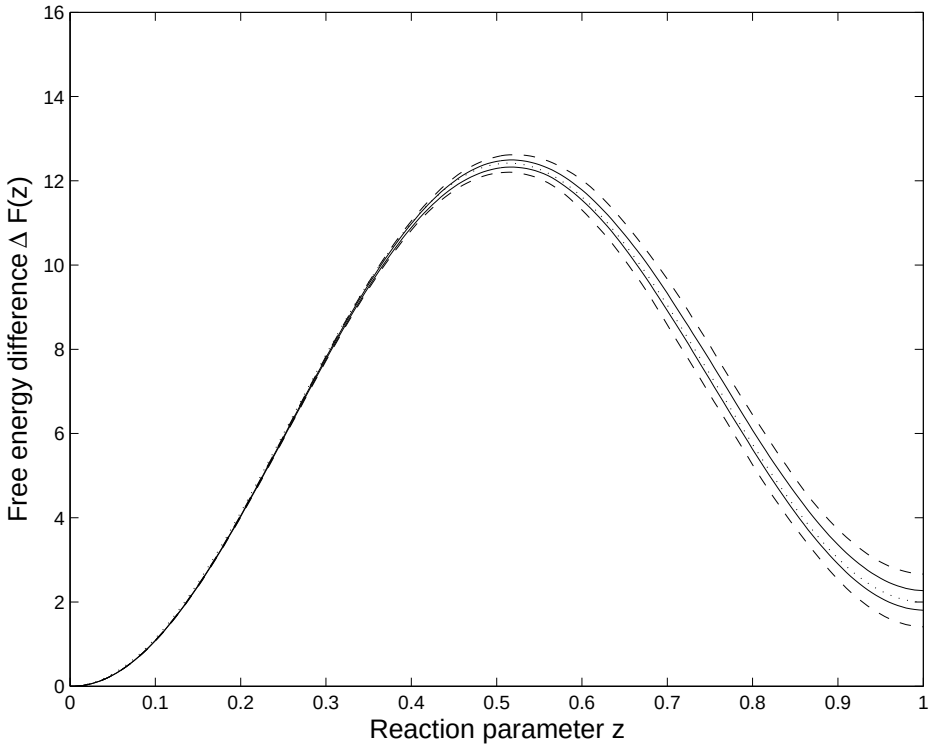
Plot of the probability density ($\beta = 1, d_2 = 1$). Left: $d_1 = 0$. Right: $d_1 = 10$.

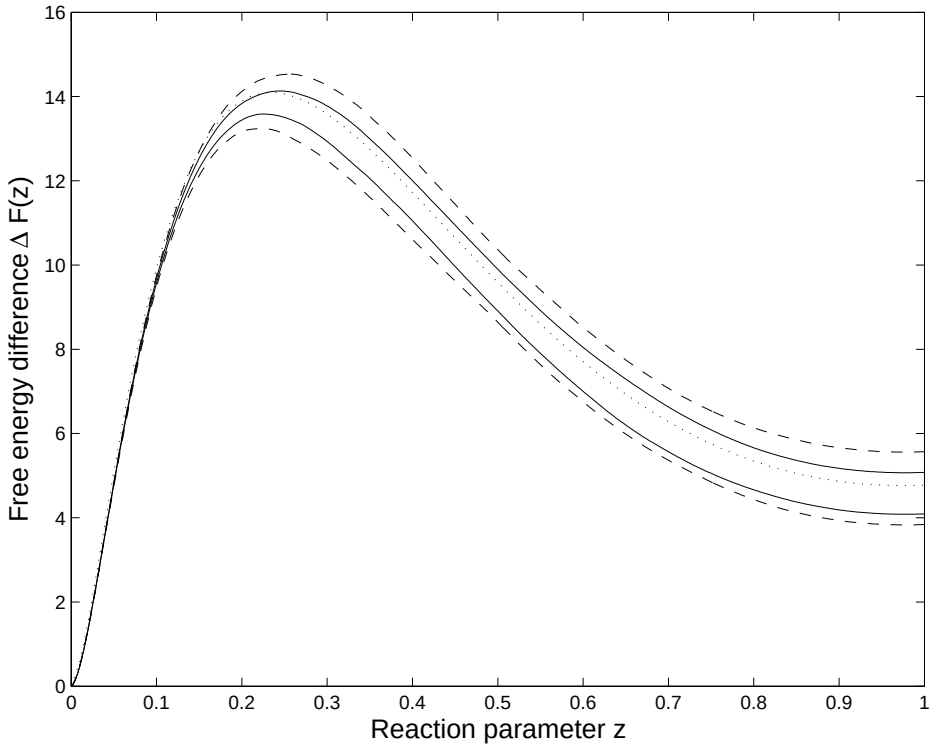
Free energy difference profiles between $x_0 = -\frac{1}{2}$ and $x_1 = 0$.

Reaction coordinate: $\xi(x,y) = \frac{1}{2^n - 1}\left[\left(1 + \frac{x - x_0}{x_1 - x_0}\right)^n - 1\right]$ $(n \geq 1)$.

[a] A. Voter (1997)

Free energy difference profiles (toy 2D case)





Free energy difference profile for $d_1 = 30, d_2 = 1$ and $\beta = 1$.
Dotted line: analytical reference. Solid lines: 95 $\%$ confidence interval (variance estimated over $K = 100$ simulations) for a nonequilibrium dynamics with $T = 1$ and $M = 10^4$. Dashed lines: id with $M = 10^3$. Left: n=1. Right: n=5.

A simplified model for solvatation effects on conformational changes

- N particles interacting through the purely repulsive WCA potential

$$V_{\mathrm{WCA}}(r) = \begin{cases} 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] + \epsilon & \text{if } r \leq \sigma, \\ 0 & \text{if } r > \sigma. \end{cases}$$

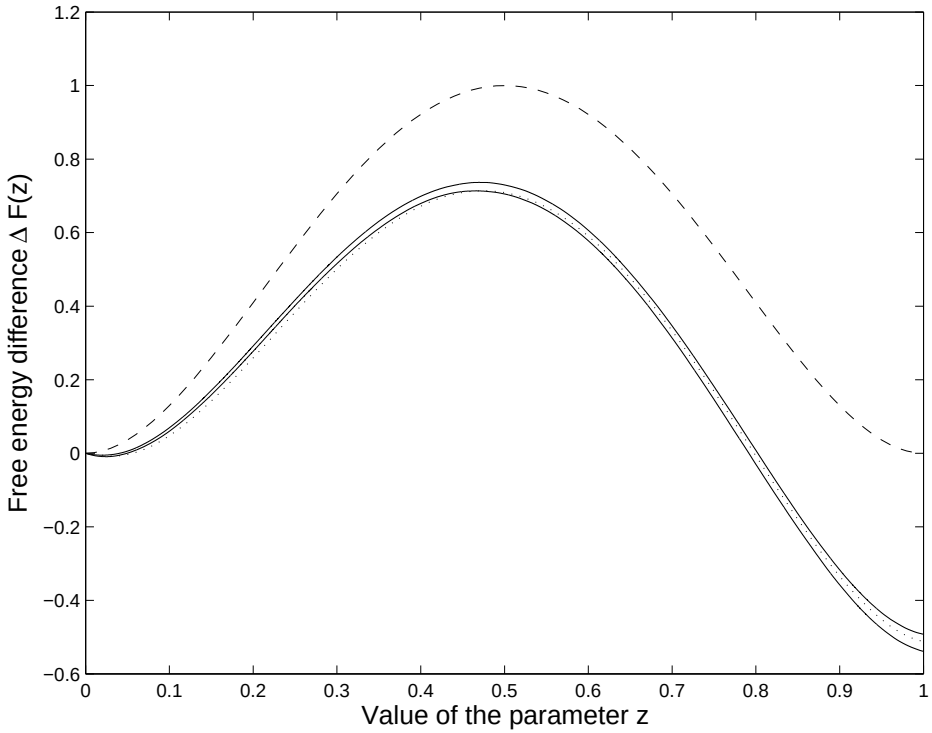
- Two solute particles interact via the double-well potential

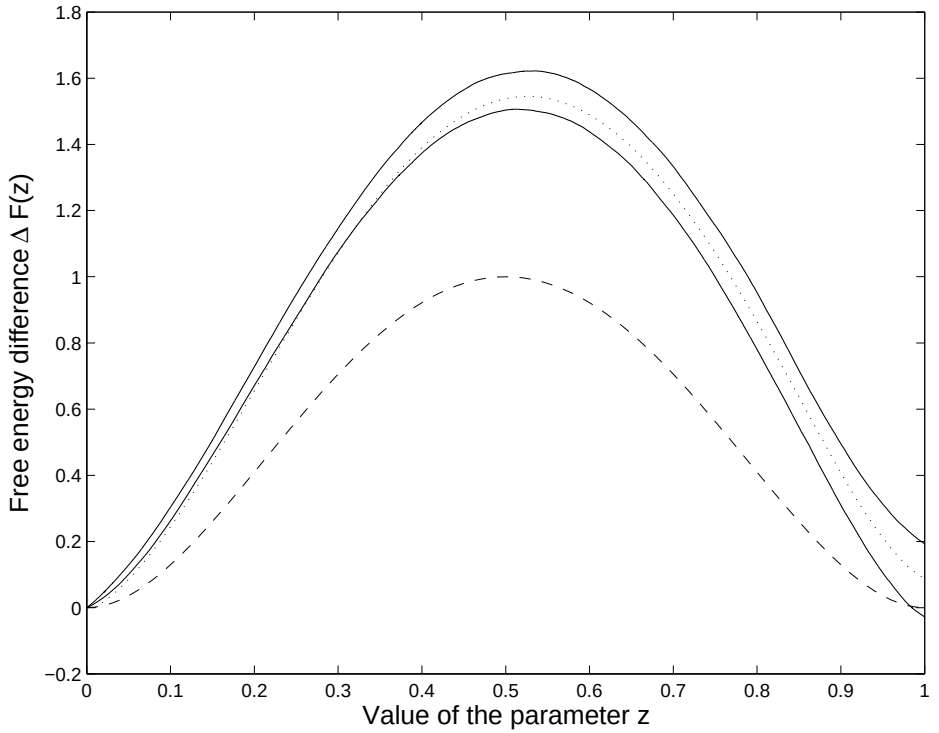
$$V_{\mathrm{S}}(r) = h\left[1 - \frac{(r-\sigma-w)^2}{w^2}\right]^2,$$

- Two energy minima (compact state $r = r_0 = 2^{1/6}\sigma$, stretched state $r = r_0 + 2w$), energy barrier h
- Reaction coordinate

$$\xi(q) = \frac{|q_1 - q_2| - r_0}{2w}$$

Influence of solvent density on free energy profiles





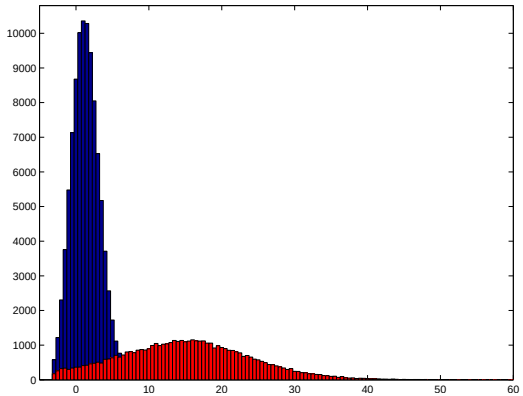
Density $\rho = a^{-2}$. Left: $a = 3$. Right: $a = 1.3$.

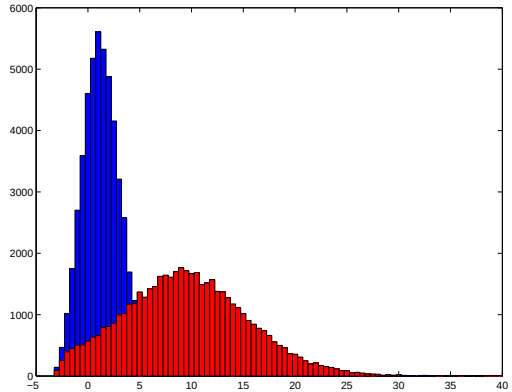
Solid line: Free-energy profiles computed using a nonequilibrium method ($T = 10$, $M = 100$, averaged over 50 realizations).

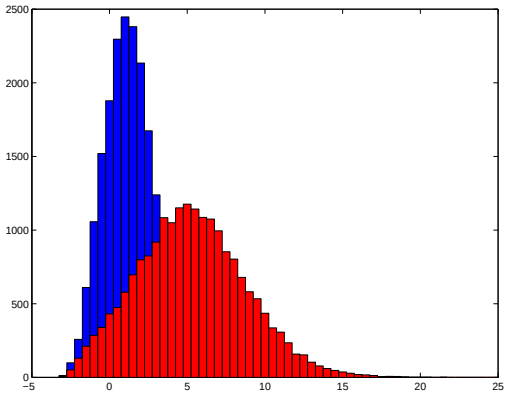
Dotted line: reference TI computation (101 points for $z \in [0, 1]$, averages taken over time $T = 2500$).

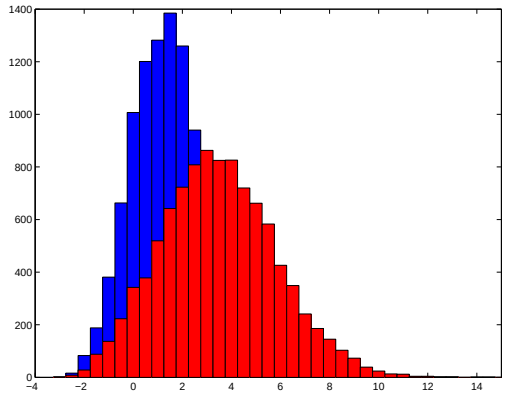
Dashed line: Energy difference for the bond (double well potential).

Computation of chemical potentiel (Widom insertion)









Work distribution for a widom insertion (blue = IPS, red = Jarzynski) for increasing switching times ($T = 1, 2, 5, 10$).

References

- **Equilibrium sampling**: E. Cancès, F. Legoll, and G. Stoltz, Theoretical and numerical comparison of some sampling methods, submitted to *M2AN* (2005)
- **Thermodynamic integration (projected SDE)**: G. Ciccotti, T. Lelièvre, and E. Vanden-Eijnden, Sampling Boltzmann-Gibbs distributions restricted on a manifold with diffusions, submitted to *CMAP* (2006)
- **Nonequilibrium computation of free energy differences**:
  - *Reaction coordinate case*: T. Lelièvre, M. Rousset, and G. Stoltz, Computation of free energy differences through nonequilibrium dynamics: The reaction coordinate case, submitted to *J. Comp. Phys.* (2006)
  - *Equilibration of the nonequilibrium dynamics*: M. Rousset and G. Stoltz, Equilibrium sampling from nonequilibrium dynamics, accepted for publication in *J. Stat. Phys.* (2006)