Computational Statistical Physics: A Mathematical Overview

Gabriel STOLTZ

CERMICS & MICMAC project team, ENPC (Marne-la-Vallée, France)

`http://cermics.enpc.fr/∼stoltz/`

Description of a classical system

- Positions q (configuration), momenta $p = M\dot{q}$ (M diagonal mass matrix)
- Microscopic description of a classical system (N particles):

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

- For instance, $T^*\mathcal{D} = \mathcal{D} \times \mathbb{R}^{3N}$ with $\mathcal{D} = \mathbb{R}^{3N}$ or $\mathbb{T}^{3N}$
- More complicated situations can be considered... (constraints defining submanifolds of the phase space)
- Hamiltonian

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

- All the physics is contained in V
- For instance, pair interactions $V(q_1,\ldots,q_N) = \sum_{1\leq i<j\leq N} v(|q_j - q_i|)$

Extracting macroscopic properties: Statistical physics

- Given the structure and the laws of interaction of the particles, what are the macroscopic properties of the matter composed of these particles?
- Equilibrium thermodynamic properties (pressure,...):

$$\langle A \rangle = \int_{T^*\mathcal{D}} A(q,p)\, \mu(dq\, dp)$$

- Choice of thermodynamic ensemble (probability measure $d\mu$): constrained maximisation of entropy

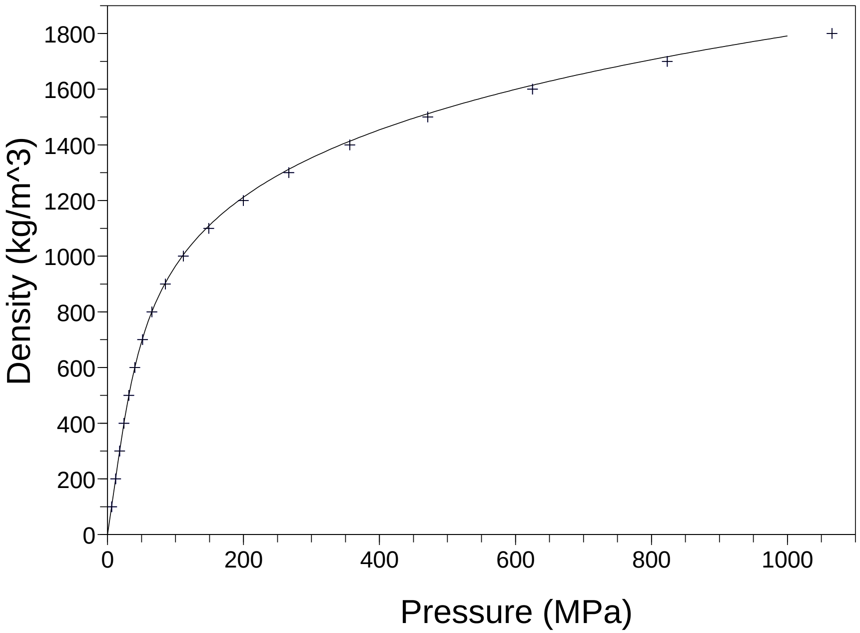
$$S(\rho) = -k_{\rm B} \int \rho \ln \rho,$$

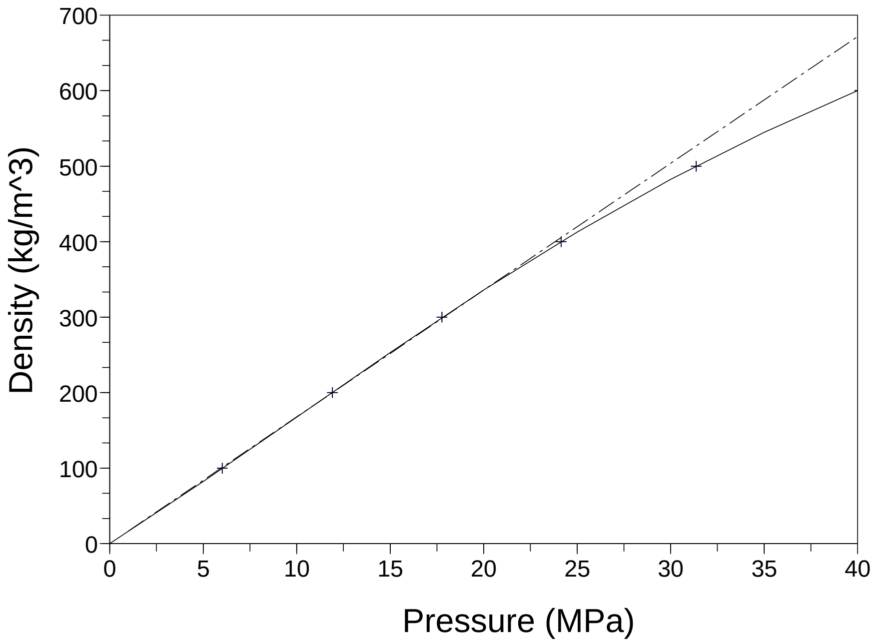
under the constraints $\rho \geq 0, \quad \int \rho = 1, \quad \int A_i\, \rho = \mathcal{A}_i$

- The choice of the variables and the observables A_i $(1 \leq i \leq m)$ determine the ensemble

An example of macroscopic data

- Pressure observable $A(q,p) = \frac{1}{3|\mathcal{D}|} \sum_{i=1}^{N} \left(\frac{p_i^2}{m_i} - q_i \cdot \nabla_{q_i} V(q) \right)$
- Lennard-Jones potential $v(r) = 4\varepsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \right]$
- Argon: $\varepsilon / k_{\mathrm{B}} =$120 K, $\sigma = 3.405$ Å



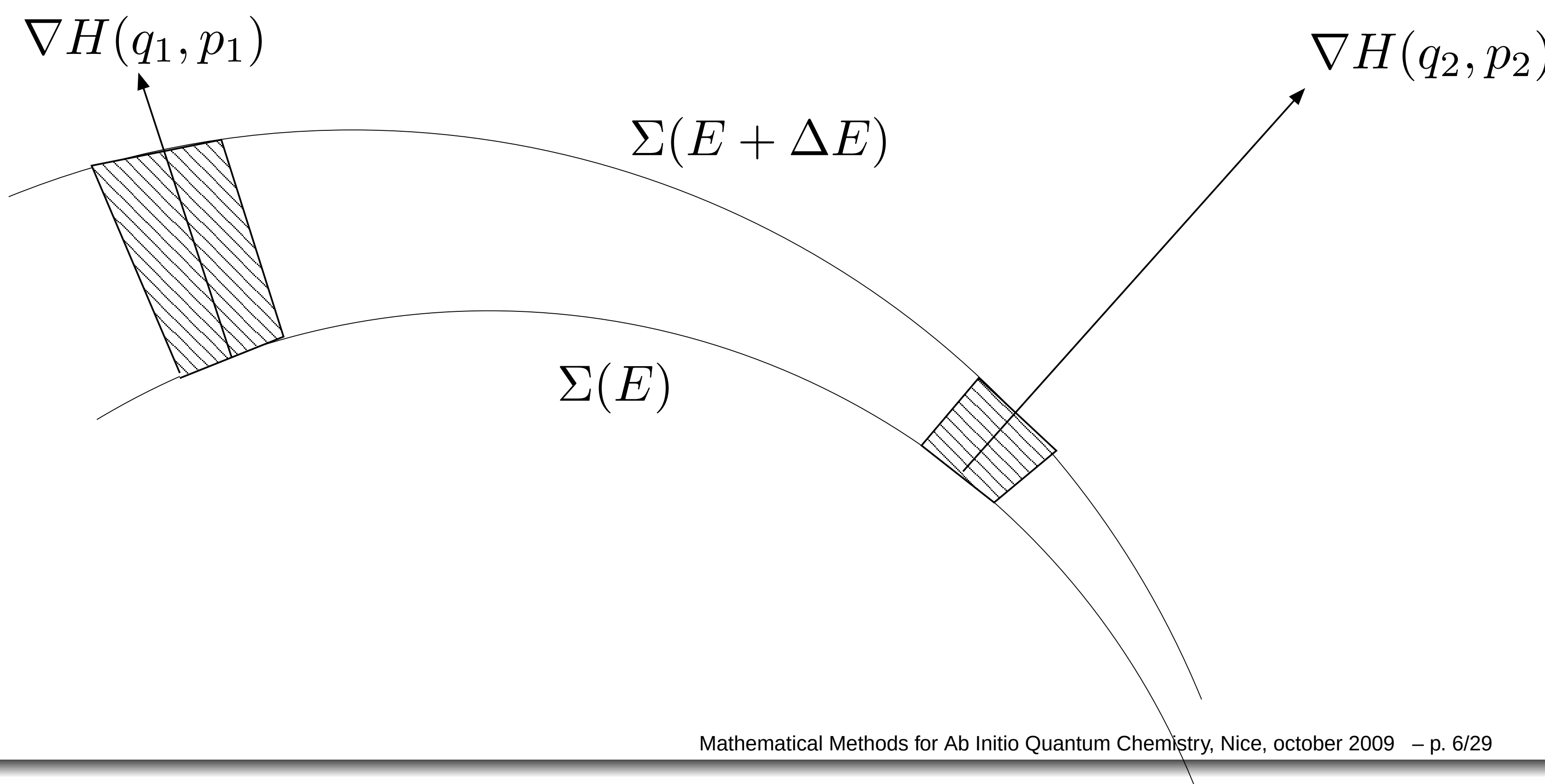


Argon state law at $T = 300$ K.

Sampling the microcanonical ensemble

The microcanonical measure

- Lebesgue measure conditioned to the set $\Sigma(E) = \{H(q,p) = E\}$
- Measure $d\mu_{\mathrm{NVE}}(q) = Z_E^{-1}\delta_{H(q,p)-E}(dq\,dp) = Z_E^{-1}\dfrac{\sigma_{\Sigma(E)}(dq\,dp)}{|\nabla H(q,p)|}$
- The partition function Z_E is a normalization constant



The Hamiltonian dynamics

- Evolution of isolated systems (Newton's law)

$$\left\{ \begin{aligned} \frac{dq(t)}{dt} &= \frac{\partial H}{\partial p}(q(t),p(t)) &&= M^{-1}p(t) \\ \frac{dp(t)}{dt} &= -\frac{\partial H}{\partial q}(q(t),p(t)) &&= -\nabla V(q(t)) \end{aligned} \right.$$

- Energy and volume preserving
- Ergodic postulate (on connected components of $\Sigma(E)$)

$$\langle A \rangle = \lim_{T \to +\infty} \frac{1}{T} \int_0^T A(q(t),p(t))\, dt$$

- Proof for integrable systems and their perturbations (KAM theory)
- Numerically interesting since it allows to replace a high-dimensional integral with an integral in dimension 1

- Verlet scheme[a] (finite difference discretization for the equation on $\ddot{q}$)

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

- Estimate the ensemble average as $\displaystyle \frac{1}{N} \sum_{n=1}^{N} A(q^n)$

- Symplectic scheme: recall that a map $(q,p) \mapsto \phi(q,p)$ is symplectic if

$$\nabla \phi(q,p) \, J \, \nabla \phi(q,p) = J, \qquad J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$$

- Backward analysis: *exact preservation of an approximate energy*, hence *approximate preservation of the exact energy*[b]

[a] L. Verlet, *Phys. Rev.* **159**(1) (1967) 98-105.

[b] E. Hairer, C. Lubich and G. Wanner (Springer, 2006)

Sampling the canonical ensemble

Classification of the methods

- Computation of $\langle A \rangle = \int_{T^*\mathcal{D}} A(q,p)\,\mu(dq\,dp)$ with

$$\mu(dq\,dp) = Z_\mu^{-1} \mathrm{e}^{-\beta H(q,p)}\,dq\,dp, \qquad \beta = \frac{1}{k_{\mathrm{B}}T}$$

- Actual issue: sampling canonical measure on configurational space

$$\nu(dq) = Z_\nu^{-1} \mathrm{e}^{-\beta V(q)}\,dq$$

- Several strategies:
  - (1) Purely stochastic methods (i.i.d sample)
  - (2) Markov chain methods and stochastic differential equations
  - (3) Deterministic methods *à la* Nosé-Hoover
- Theoretical and numerical comparison:[a] convergence for (1)-(2), in practice (2) is more convenient

[a] E. CANCÈS, F. LEGOLL ET G. STOLTZ, *M2AN*, 2007

Metropolis-Hastings algorithm

- Markov chain method (Metropolis et al. (1953), Hastings (1970))
- Given a current configuration q^n, propose $\tilde{q}^{n+1}$ according to a transition probability $T(q^n, \tilde{q})$
  - Gaussian displacement $\tilde{q}^{n+1} = q^n + \sigma\, G^n$ with $G^n \sim \mathcal{N}(0, \mathrm{Id})$, in which case $T(q,\tilde{q}) = \left(\sigma\sqrt{2\pi}\right)^{-3N} \exp\left(-\frac{|\tilde{q}-q|^2}{2\sigma^2}\right)$
  - Biased random walk $\tilde{q}^{n+1} = q^n - \alpha\nabla V(q^n) + \sqrt{\frac{2\alpha}{\beta}}\, G^n$, in which case $T(q,\tilde{q}) = \left(\frac{\beta}{4\pi\alpha}\right)^{3N/2} \exp\left(-\beta\frac{|\tilde{q}-q+\alpha\nabla V(q)|^2}{4\alpha}\right)$
- Accept the proposition with probability

$$\min\left(1,\ \frac{T(\tilde{q}^{n+1}, q^n)\,\nu(\tilde{q}^{n+1})}{T(q^n, \tilde{q}^{n+1})\,\pi(q^n)}\right),$$

and set in this case $q^{n+1} = \tilde{q}^{n+1}$; otherwise, set $q^{n+1} = q^n$.

Convergence

- Transition kernel

$$P(q, dq') = \min\Big(1, r(q,q')\Big) T(q,q')\, dq' + \Big(1 - \alpha(q)\Big)\delta_q(dq'),$$

where $\alpha(q) \in [0,1]$ is the probability to accept a move starting from q:

$$\alpha(q) = \int \min\Big(1, r(q,q')\Big) T(q,q')\, dq'.$$

- The canonical measure is reversible with respect to ν, hence invariant:

$$P(q, dq')\nu(dq) = P(q', dq)\nu(dq')$$

- Show irreducibility (properties of the proposition function): defining the n-step transition probability as $P^n(q, dq') = \int_{x \in \mathcal{D}} P(q, dx)\, P^{n-1}(x, dq')$, the condition is that, for almost all q_0 and any set A of positive measure, there exists n_0 such that $P^n(q_0, A) > 0$ when $n \geq n_0$

- Pathwise ergodicity $\lim_{N \to +\infty} \frac{1}{N} \sum_{n=1}^{N} A(q^n) = \int_{\mathcal{D}} A(q)\, \nu(dq)$

Error estimates

- Under additional assumptions, the pathwise convergence result can be refined to a central limit theorem for Markov chains:

$$\sqrt{N}\left|\frac{1}{N}\sum_{n=1}^{N} A(q^n) - \int_{\mathcal{D}} A(q)\,\nu(dq)\right| \longrightarrow \mathcal{N}(0,\sigma^2),$$

- The asymptotic variance σ^2 takes into account the correlations:

$$\sigma^2 = \mathrm{Var}_\nu(A) + 2\sum_{n=1}^{+\infty} \mathbb{E}_\nu\Big[\big(A(q^0) - \mathbb{E}_\nu(A)\big)\big(A(q^n) - \mathbb{E}_\nu(A)\big)\Big]$$

- Numerical efficiency: trade-off between acceptance and sufficiently large moves in space (rejection rate around 0.5), the aim being to reduce the sample autocorrelation
- Practical computation of error bars (confidence intervals): independent realizations or block averaging

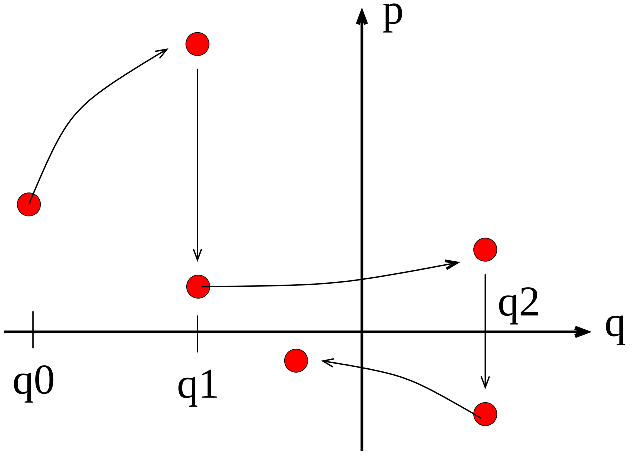
Markov chain in the configuration space (Duane et al. (1987), Schuette et al. (1999)). Starting from q^n:

- generate momenta p^n according to $Z_p^{-1}\, \mathrm{e}^{-\beta p^2/2m} dp$
- compute (an approximation of) the flow $\Phi_\tau(q^n, p^n) = (\tilde{q}^{n+1}, \tilde{p}^{n+1})$ of Newton's equations, *i.e.* integrate

$$\dot{q}_i = \frac{p_i}{m_i}, \quad \dot{p}_i = -\nabla_{q_i} V(q) \tag{1}$$

  on a time τ starting from (q^n, p^n).

- accept $\tilde{q}^{n+1}$ and set $q^{n+1} = \tilde{q}^{n+1}$ with a probability $\min\left(1, \exp -\beta(\tilde{E} - E_n)\right)$; otherwise set $q^{n+1} = q^n$.



Two parameters : τ and Δt.

Extensions: correlated momenta, random times τ, constrained dynamics, ...

Overdamped Langevin dynamics

- SDE on the configurational part only (momenta trivial to sample)

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

where $(W_t)_{t\geq 0}$ is a standard Wiener process of dimension dN

- Numerical scheme: $q^{n+1} = q^n - \Delta t\,\nabla V(q^n) + \sigma\sqrt{\Delta t}\,G^n$
- Invariance of the canonical measure $\nu(dq)$ when $\sigma = (2/\beta)^{1/2}$
- Evolution PDE for the law of the process at time t:

$$\partial_t \psi = \mathrm{div}\left(\psi_\infty \nabla\left(\frac{\psi}{\psi_\infty}\right)\right), \quad \psi_\infty = Z^{-1}\exp(-\beta V)$$

- Invariance + irreducibility (elliptic process):

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T A(q_t)\,dt = \int_{\mathcal{D}} A(q)\,\nu(dq) \quad \text{a.s.}$$

- Numerical scheme samples an approximate measure $\nu_{\Delta t}(dq)$

- Several notions of convergence: here, longtime convergence in law
- Relative entropy $\mathcal{H}(\psi(t,\cdot)\,|\,\psi_\infty) = \int \ln\left(\frac{\psi(t,\cdot)}{\psi_\infty}\right)\psi_\infty$
- It holds $\|\psi(t,\cdot) - \psi_\infty\|_{L^1} \leq \sqrt{2\mathcal{H}(\psi(t,\cdot)\,|\,\psi_\infty)}$
- Fisher information $I(\psi(t,\cdot)\,|\,\psi_\infty) = \int \left|\nabla \ln\left(\frac{\psi(t,\cdot)}{\psi_\infty}\right)\right|^2 \psi_\infty$
- A simple computation shows $\frac{d}{dt}H(\psi(t,\cdot)\,|\,\psi_\infty) = -\beta^{-1}I(\psi(t,\cdot)\,|\,\psi_\infty)$
- When a Logarithmic Sobolev Inequality holds for ψ_∞, namely $H(\phi|\,\psi_\infty) \leq \frac{1}{2R}I(\phi\,|\,\psi_\infty)$, then, by Gronwall's lemma, the relative entropy converges exponentially fast to 0, as well as the total variation distance
- Obtaining LSI: Bakry-Emery criterion (convexity), Gross (tensorization), Holley-Stroock's perturbation result

Langevin dynamics

- Stochastic perturbation of the Hamiltonian dynamics

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

- Fluctuation/dissipation relation $\sigma^2 = 2\gamma k_{\rm B} T = 2\gamma/\beta$
- Invariance of the canonical measure (stationary solution of the Fokker-Planck equation)
- Convergence of the trajectorial average, starting from a given (q^0, p^0):

$$\lim_{T \to +\infty} \frac{1}{T} \int_0^T A(q_t, p_t) \, dt = \frac{\displaystyle \int_{T^*\mathcal{D}} A(q,p) \, \mathrm{e}^{-\beta H(q,p)} \, dq \, dp}{\displaystyle \int_{T^*\mathcal{D}} \mathrm{e}^{-\beta H(q,p)} \, dq \, dp} \quad \text{a.s.}$$

- Numerical schemes obtained by a splitting strategy for instance (Verlet scheme + partial randomization of momenta)

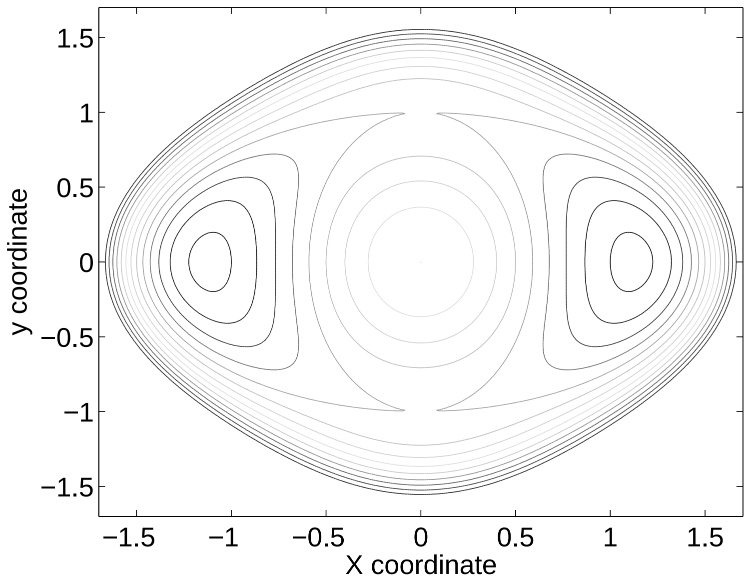
Free energy biased dynamics

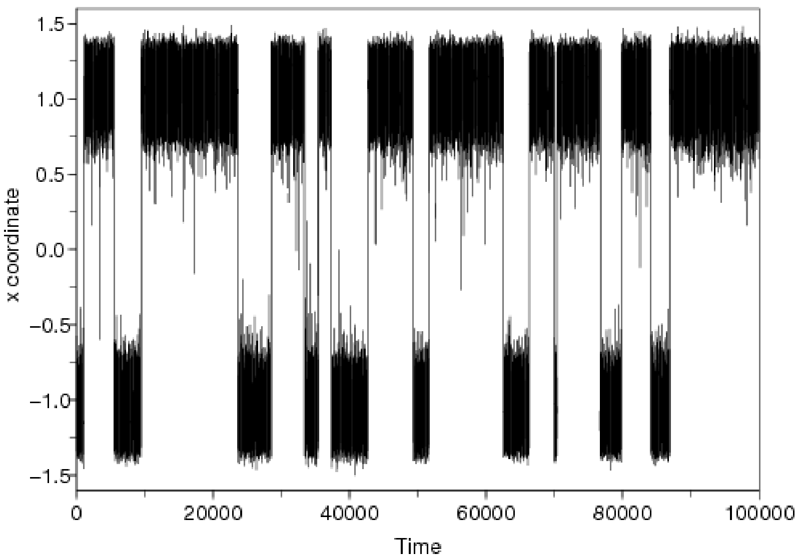
Metastability (1)

Numerical discretization of the overdamped Langevin dynamics:

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

where $G^n \sim \mathcal{N}(0,1)$ i.i.d.

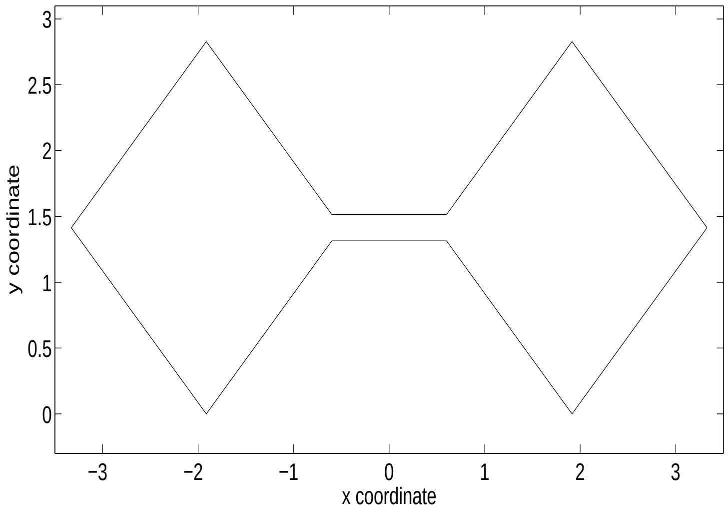




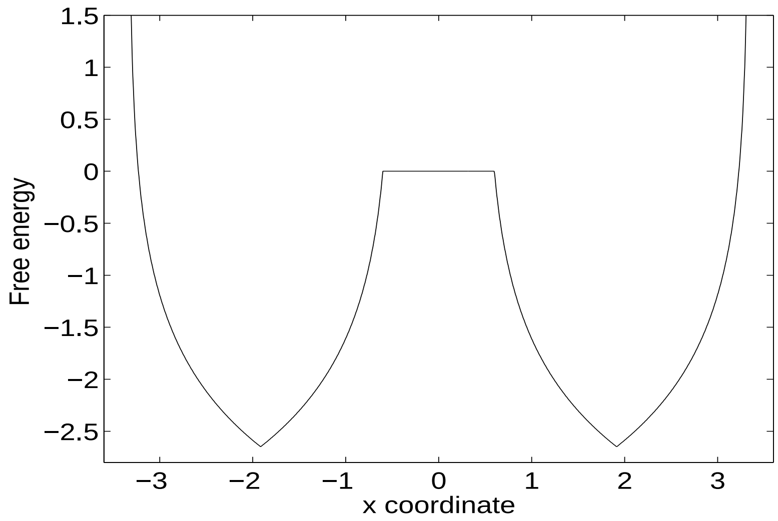
Projected trajectory in the x variable for $\Delta t = 0.01$, $\beta = 8$.

Metastability (2)

- Although the trajectorial average converges to the phase-space average, the convergence may be slow...
- Slowly evolving macroscopic function of the microscopic degrees of freedom
- Two origins : energetic or entropic barriers (in fact, free energy barrier)

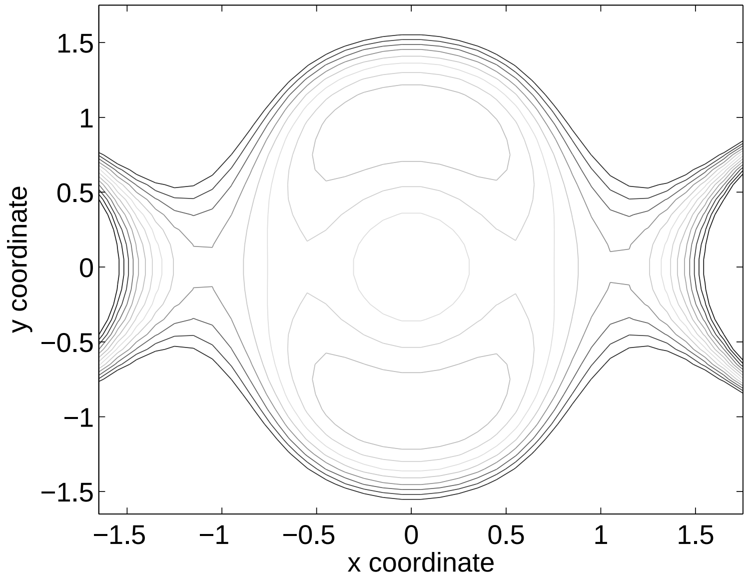


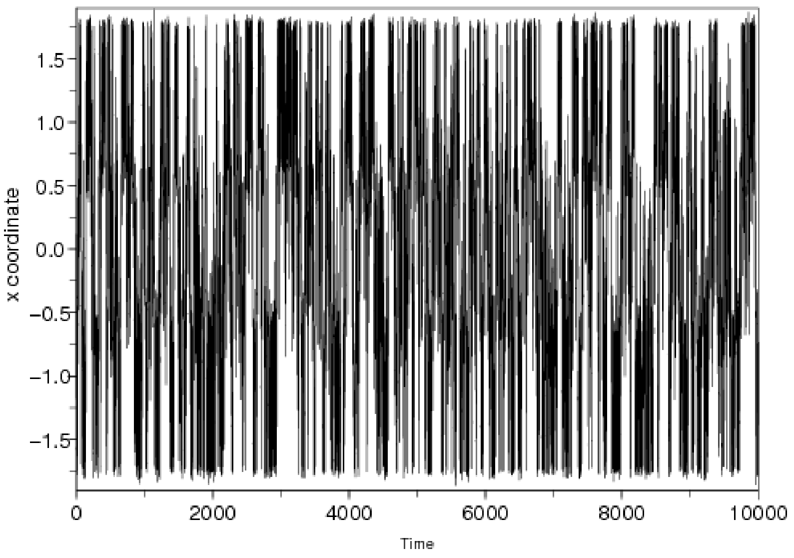
(a) Entropic barrier.



(b) Associated free energy.

- Assume the free energy F associated with the slow direction x has been computed, and sample the modified potential $\mathcal{V}(x,y) = V(x,y) - F(x)$.





Projected trajectory in the x variable for $\Delta t = 0.01$, $\beta = 8$.

- Many more transitions! The variable x is uniformly distributed.
- Estimate canonical averages through reweighting: $\dfrac{\sum_{n=1}^{N} A(x^n)\,\mathrm{e}^{-\beta F(x^n)}}{\sum_{n=1}^{N} \mathrm{e}^{-\beta F(x^n)}}$

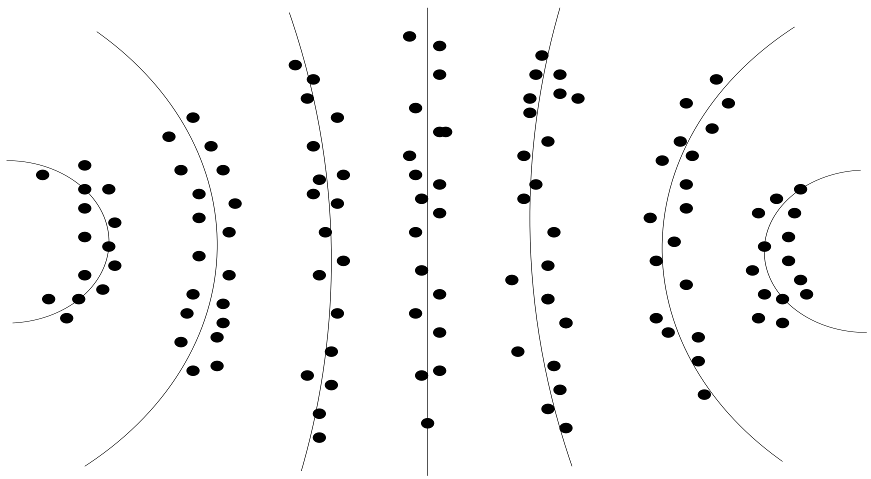
Computation of free energy differences (1)

- Absolute free energy $F = -\frac{1}{\beta} \ln \int_{T^*\mathcal{D}} \mathrm{e}^{-\beta H(q,p)} \, dq \, dp$
- Motivation (Gibbs, 1902):
  - canonical measure $\mu(q,p) = Z^{-1} \exp(-\beta H(q,p))$
  - start from the thermodynamic identity $F = U - TS$
  - average energy $U = \int H \mu$
  - entropy $S = -k_{\mathrm{B}} \int \mu \ln \mu$
- (given) reaction coordinate $\xi : \mathbb{R}^{3N} \to \mathbb{R}^m$ (angle, length,. . .):

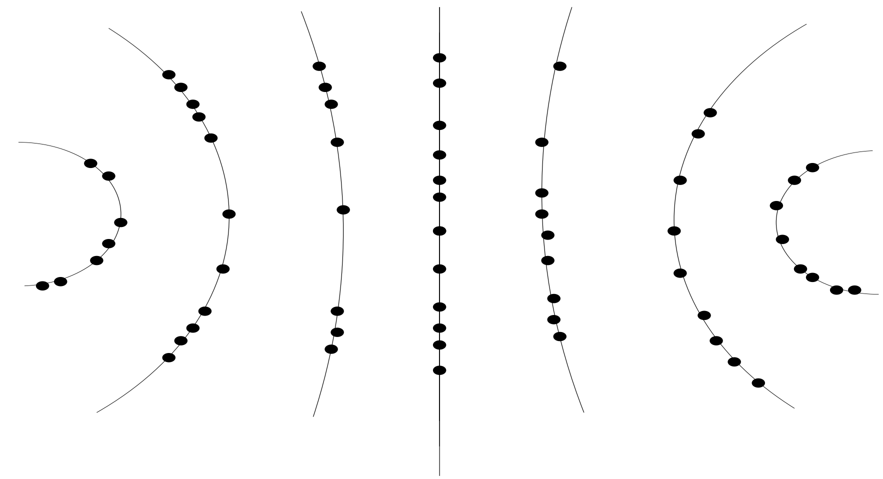
$$\Delta F = -\beta^{-1} \ln \left(\frac{\displaystyle\int_{T^*\mathcal{D}} \mathrm{e}^{-\beta H(q,p)} \, \delta_{\xi(q)-z_1} \, dq \, dp}{\displaystyle\int_{T^*\mathcal{D}} \mathrm{e}^{-\beta H(q,p)} \, \delta_{\xi(q)-z_0} \, dq \, dp} \right).$$

Recall $\delta_{\xi(q)-z} = |\nabla \xi|^{-1} d\sigma_{\Sigma_z}$ supported on $\Sigma_z = \{\xi(q) = z\}$

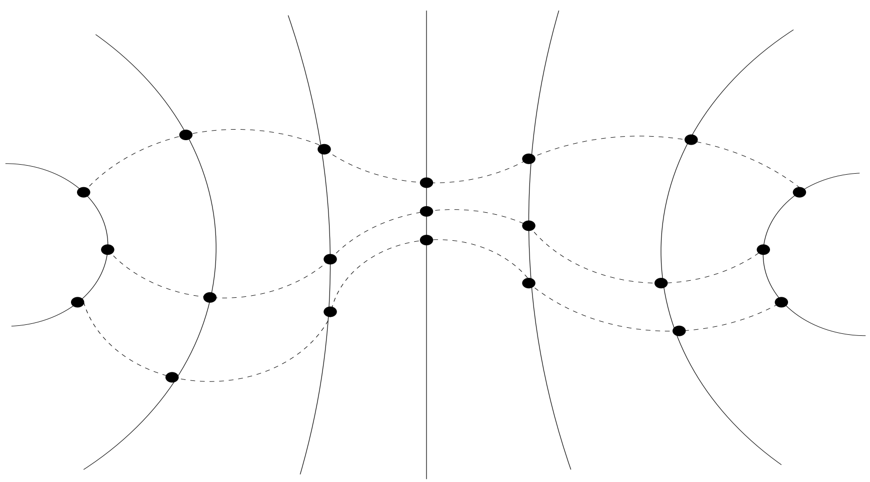
Cartoon comparison of the methods



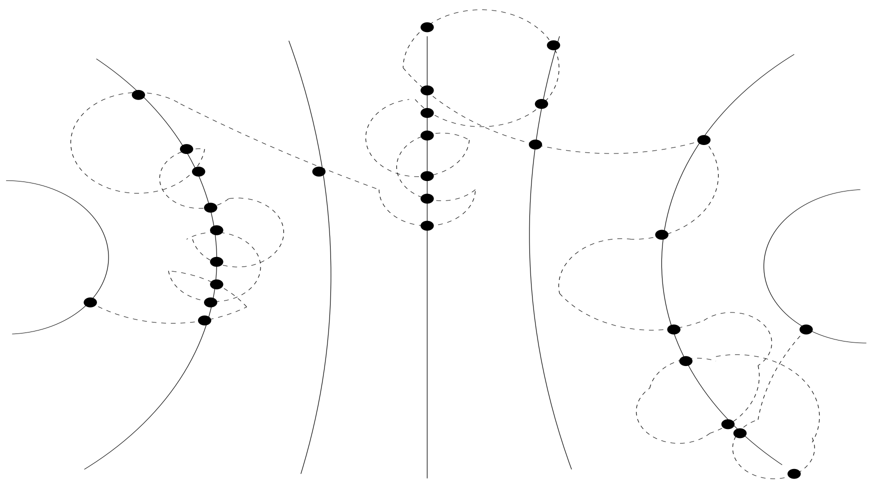
(a) Histogram method



(b) Thermodynamic integration



(c) Nonequilibrium switching dynamics



(d) Adaptive dynamics

- Simplified setting: $q = (x, y)$ and $\xi(q) = x \in \mathbb{R}$ so that

$$F(x_2) - F(x_1) = -\beta^{-1} \ln \left(\frac{\overline{\psi}_{\mathrm{eq}}(x_2)}{\overline{\psi}_{\mathrm{eq}}(x_1)} \right), \qquad \overline{\psi}_{\mathrm{eq}}(x) = \int \mathrm{e}^{-\beta V(x,y)} \, dy$$

- Notice that the mean force $F'(x) = \dfrac{\displaystyle\int \partial_x V(x,y) \, \mathrm{e}^{-\beta V(x,y)} \, dy}{\displaystyle\int \mathrm{e}^{-\beta V(x,y)} \, dy}$

- The dynamics $dq_t = -\nabla V(q_t) \, dt + \sqrt{\frac{2}{\beta}} \, dW_t$ is metastable, contrarily to

$$\left\{ \begin{array}{l} dq_t = -\nabla \Big(V(q_t) - F(\xi(q_t)) \Big) \, dt + \sqrt{\dfrac{2}{\beta}} \, dW_t \\ F'(x) = \mathbb{E}_\mu \Big(\partial_x V(q) \, \Big| \, \xi(q) = x \Big) \end{array} \right.$$

- Replace equilibrium expectation with $F'(t, x) = \mathbb{E} \Big(\partial_x V(q_t) \, \Big| \, \xi(q_t) = x \Big)$

- Adaptive Biasing Force method[a]

$$\begin{cases} dq_t = -\nabla\Big(V(q_t) - F(t,\xi(q_t))\Big)\,dt + \sqrt{\dfrac{2}{\beta}}\,dW_t \\ F'(t,x) = \mathbb{E}\Big(\partial_x V(q)\,\Big|\,\xi(q_t) = x\Big) \end{cases}$$

- Can be proved to converge as $t \to +\infty$[b]
- Replace the conditional expectation by a time-average:

$$\mathbb{E}\Big(\partial_x V(q_t)\,\Big|\,\xi(q_t) = x\Big) \simeq \frac{1}{t}\int_0^t \partial_x V(q_s)\,\mathbf{1}_{\xi(q_s)-x}\,ds$$

- Possibly use several replicas of the system, driven by independent noises and contributing to the same biasing potential
- Selection strategy[c] to enhance the diffusion

[a] See the works by Darve, Pohorille, Chipot, Hénin, ...

[b] T. Lelièvre, M. Rousset and G. Stoltz, *Nonlinearity* **21** (2008) 1155-1181

[c] T. Lelièvre, M. Rousset and G. Stoltz, *J. Chem. Phys.* **126** (2007) 134111

Adaptive dynamics: convergence

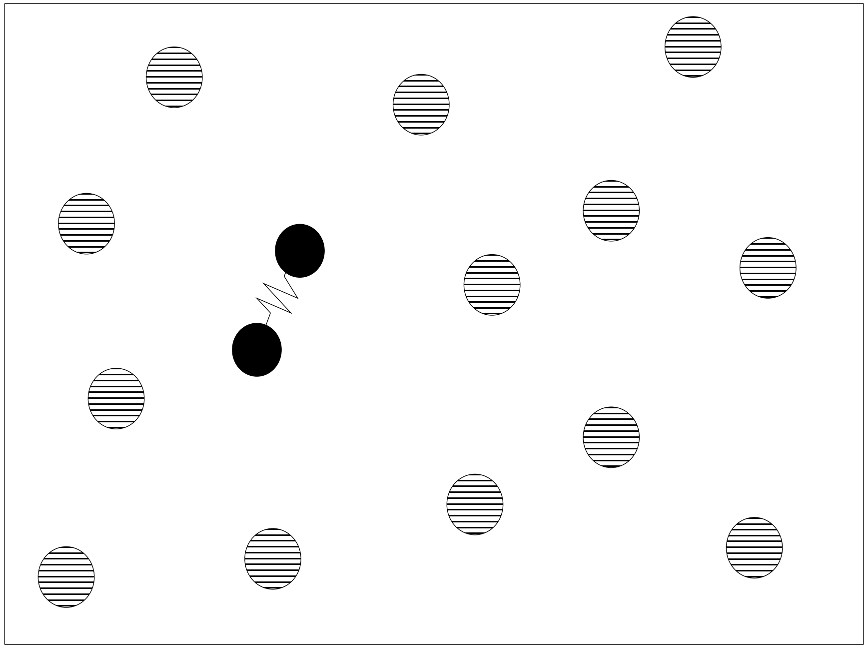
- Nonlinear PDE on the law $\psi(t,q)$:

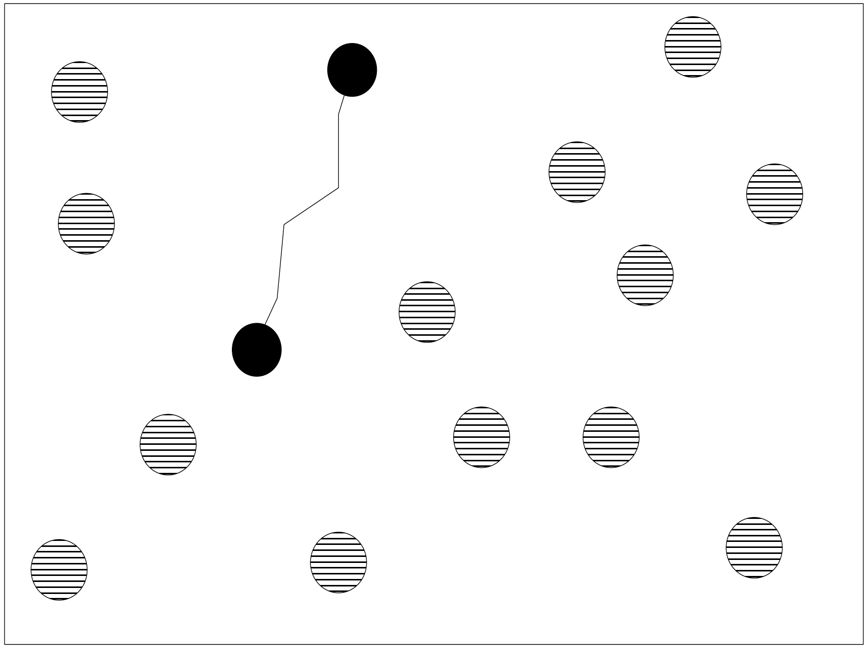
$$\left\{\begin{aligned} &\partial_t \psi = \mathrm{div}\left[\nabla\left(V - F_{\mathrm{bias}}(t,x)\right)\psi + \beta^{-1}\nabla\psi\right], \\ &F'_{\mathrm{bias}}(t,x) = \frac{\displaystyle\int \partial_x V(x,y)\,\psi(t,x,y)\,dy}{\displaystyle\int \psi(t,x,y)\,dy}. \end{aligned}\right.$$

- Stationary solution $\psi_\infty \propto \mathrm{e}^{-\beta(V - F\circ\xi)}$
- Simple diffusion for the marginals $\partial_t \overline{\psi} = \partial_{xx}\,\overline{\psi}$
- Decomposition of the total entropy $H(\psi \,|\, \psi_\infty) = \displaystyle\int_{\mathcal{D}} \ln\left(\frac{\psi}{\psi_\infty}\right)\psi$ into a macroscopic contribution (marginals in x) and a microscopic one (conditioned measures)
- Convergence of the microscopic entropy provided some uniform logarithmic Sobolev inequality holds for the conditioned measures

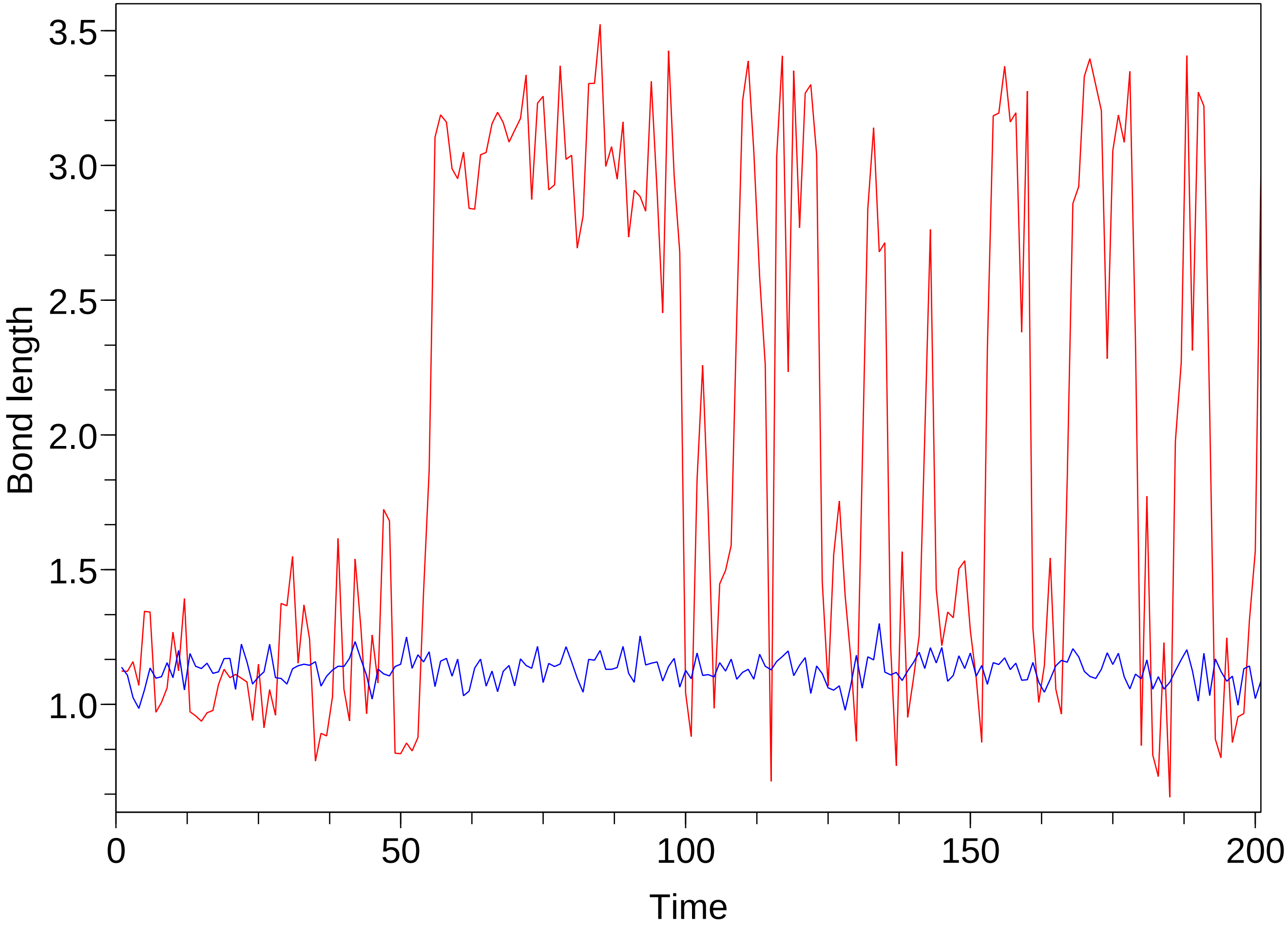
Application: Solvatation effects on conformational changes (1)

- Two particules (q_1,q_2) interacting through $V_{\rm S}(r) = h\left[1 - \frac{(r-r_0-w)^2}{w^2}\right]^2$
- Solvent: particules interacting through the purely repulsive potential $V_{\rm WCA}(r) = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] + \epsilon$ if $r \leq r_0$, 0 if $r > r_0$
- Reaction coordinate $\xi(q) = \frac{|q_1 - q_2| - r_0}{2w}$, compact state $\xi^{-1}(0)$, stretched state $\xi^{-1}(1)$





Application: Solvatation effects on conformational changes (2)



Blue: without biasing term. Red: adaptive biasing force.

Parameters: $h = 10$, density $\rho = 0.25\,\sigma^{-2}$, $w = 1$, $\beta = 3$, $\epsilon = 1$, $\tau = 0.1$

Conclusion – Mathematical classification (september 2009)

Free energy perturbation	$\longrightarrow$	Homogeneous MCs and SDEs
Thermodynamic integration	$\longrightarrow$	Projected MCs and SDEs
Nonequilibrium dynamics	$\longrightarrow$	Nonhomogenous MCs and SDEs
Adaptive dynamics	$\longrightarrow$	Nonlinear SDEs and MCs
Selection procedures	$\longrightarrow$	Particle systems and jump processes

- Which method is the most efficient in practice...?
- Some advertisement for a book to appear this year:

  T. LELIÈVRE, M. ROUSSET AND G. STOLTZ *Free energy computations: A Mathematical Perspective*, Imperial College Press.