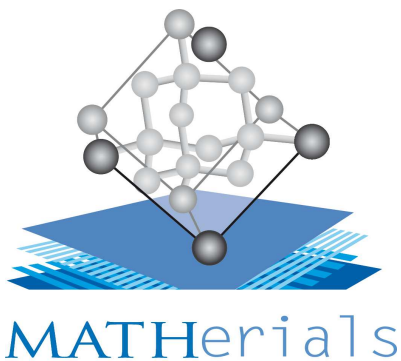
Algorithms for computational statistical physics
Part 1: Sampling the canonical distribution

T. Lelièvre

CERMICS - Ecole des Ponts ParisTech & Matherials project-team - INRIA



ICTS, August 2017

Motivation

The aim of molecular dynamics simulations is to understand the relationships between the macroscopic properties of a molecular system and its atomistic features. In particular, one would like to evaluate numerically macroscopic quantities from models at the microscopic scale.

Many applications in various fields: biology, physics, chemistry, materials science.

Various models: discrete state space (kinetic Monte Carlo, Markov State Model) or continuous state space (Langevin).

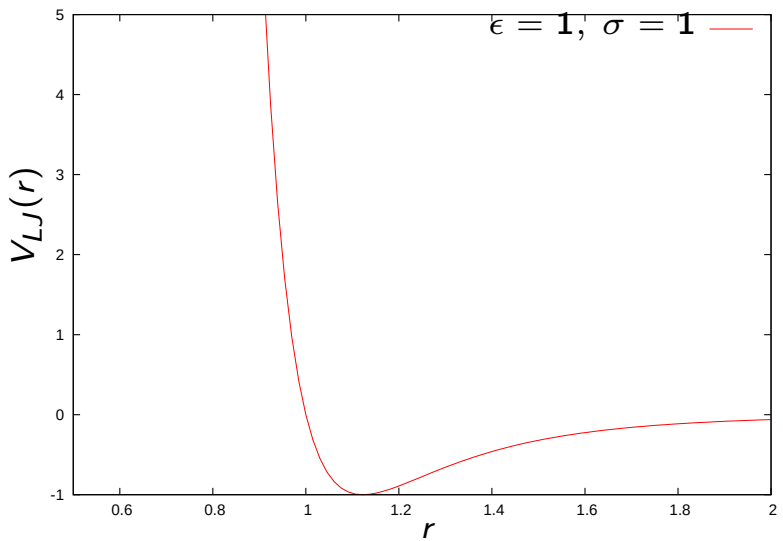
The basic ingredient: a potential V which associates to a configuration $(\boldsymbol{x}_1, ..., \boldsymbol{x}_N) = \boldsymbol{x} \in \mathbb{R}^{3N_{atom}}$ an energy $V(\boldsymbol{x}_1, ..., \boldsymbol{x}_{N_{atom}})$. The dimension $d = 3N_{atom}$ is large (a few hundred thousand to millions).

Empirical force field

Typically, V is a sum of potentials modelling interaction between two particles, three particles and four particles:

$$V = \sum_{i<j} V_1(\boldsymbol{x}_i, \boldsymbol{x}_j) + \sum_{i<j<k} V_2(\boldsymbol{x}_i, \boldsymbol{x}_j, \boldsymbol{x}_k) + \sum_{i<j<k<l} V_3(\boldsymbol{x}_i, \boldsymbol{x}_j, \boldsymbol{x}_k, \boldsymbol{x}_l).$$

For example,
$V_1(\boldsymbol{x}_i, \boldsymbol{x}_j) = V_{LJ}(|\boldsymbol{x}_i - \boldsymbol{x}_j|)$
where
$V_{LJ}(r) = 4\epsilon\left(\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right)$ is the Lennard-Jones potential.



Dynamics

Newton equations of motion:

$$\begin{cases} d\boldsymbol{X}_t = M^{-1}\boldsymbol{P}_t\,dt \\ d\boldsymbol{P}_t = -\nabla V(\boldsymbol{X}_t)\,dt \end{cases}$$

Dynamics

Newton equations of motion + thermostat: Langevin dynamics:

$$\left\{ \begin{array}{l} d\boldsymbol{X}_t = M^{-1}\boldsymbol{P}_t\, dt \\ d\boldsymbol{P}_t = -\nabla V(\boldsymbol{X}_t)\, dt - \gamma M^{-1}\boldsymbol{P}_t\, dt + \sqrt{2\gamma\beta^{-1}}\, d\boldsymbol{W}_t \end{array} \right.$$

where $\gamma > 0$. Langevin dynamics is ergodic wrt $\mu(d\boldsymbol{x}) \otimes Z_p^{-1} \exp\left(-\beta \frac{\boldsymbol{p}^t M^{-1}\boldsymbol{p}}{2}\right) d\boldsymbol{p}$ with

$$d\mu = Z^{-1} \exp(-\beta V(\boldsymbol{x}))\, d\boldsymbol{x}$$

where $Z = \int \exp(-\beta V(\boldsymbol{x}))\, d\boldsymbol{x}$ is the partition function and $\beta = (k_B T)^{-1}$ is proportional to the inverse of the temperature.

Dynamics

Newton equations of motion + thermostat: Langevin dynamics:

$$\begin{cases} d\boldsymbol{X}_t = M^{-1}\boldsymbol{P}_t\,dt \\ d\boldsymbol{P}_t = -\nabla V(\boldsymbol{X}_t)\,dt - \gamma M^{-1}\boldsymbol{P}_t\,dt + \sqrt{2\gamma\beta^{-1}}d\boldsymbol{W}_t \end{cases}$$

where $\gamma > 0$. Langevin dynamics is ergodic wrt $\mu(d\boldsymbol{x}) \otimes Z_p^{-1} \exp\left(-\beta \frac{\boldsymbol{p}^t M^{-1}\boldsymbol{p}}{2}\right) d\boldsymbol{p}$ with

$$d\mu = Z^{-1}\exp(-\beta V(\boldsymbol{x}))\,d\boldsymbol{x}$$

where $Z = \int \exp(-\beta V(\boldsymbol{x}))\,d\boldsymbol{x}$ is the partition function and $\beta = (k_B T)^{-1}$ is proportional to the inverse of the temperature.

In the following, we focus on the *overdamped Langevin* (or gradient) dynamics

$$d\boldsymbol{X}_t = -\nabla V(\boldsymbol{X}_t)\,dt + \sqrt{2\beta^{-1}}d\boldsymbol{W}_t,$$

which is also ergodic wrt μ.

Macroscopic quantities of interest

These dynamics are used to compute macroscopic quantities:

(i) Thermodynamic quantities (averages wrt μ of some observables): stress, heat capacity, free energy,...

$$\mathbb{E}_\mu(\varphi(\boldsymbol{X})) = \int_{\mathbb{R}^d} \varphi(\boldsymbol{x})\, \mu(d\boldsymbol{x}) \simeq \frac{1}{T} \int_0^T \varphi(\boldsymbol{X}_t)\, dt.$$

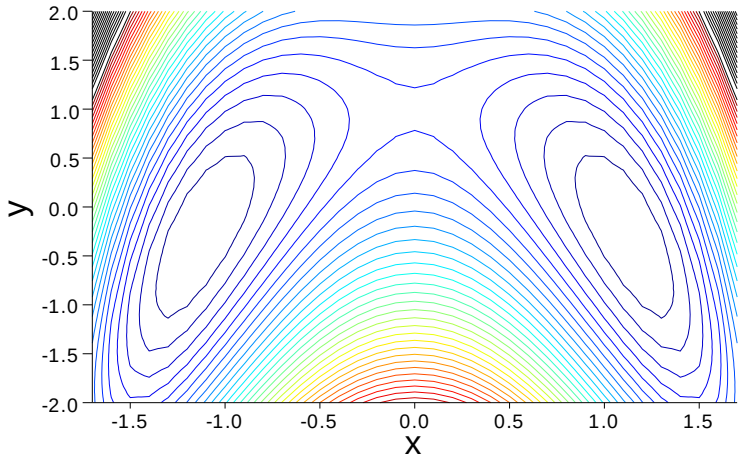
(ii) Dynamical quantities (averages over trajectories): diffusion coefficients, viscosity, transition rates,...

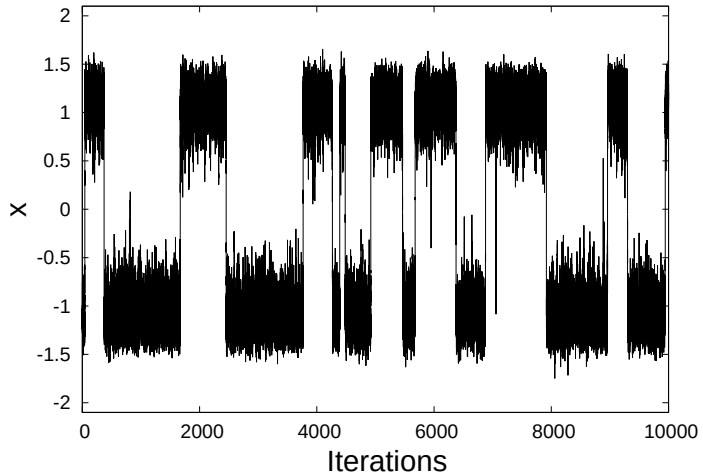
$$\mathbb{E}(\mathcal{F}((\boldsymbol{X}_t)_{t\geq 0})) \simeq \frac{1}{M} \sum_{m=1}^{M} \mathcal{F}((\boldsymbol{X}_t^m)_{t\geq 0}).$$

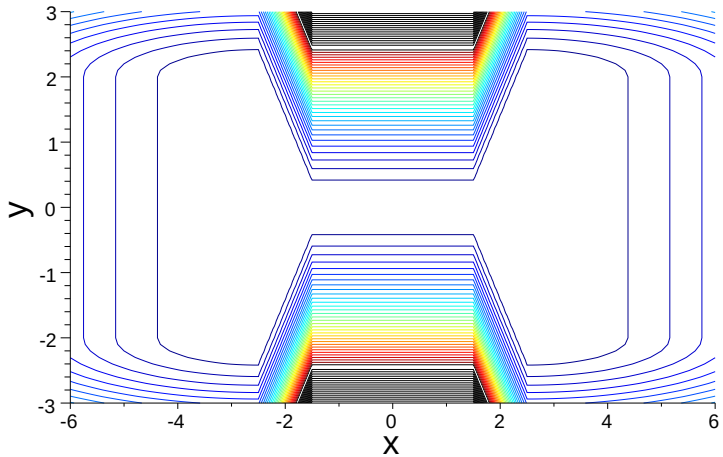
Difficulties: (i) high-dimensional problem ($N \gg 1$); (ii) $\boldsymbol{X}_t$ is a metastable process and μ is a multimodal measure.

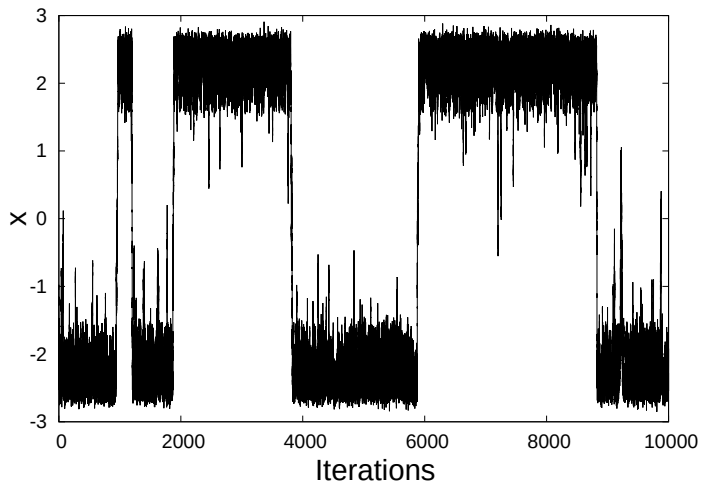
Metastability: energetic and entropic barriers

A two-dimensional schematic picture







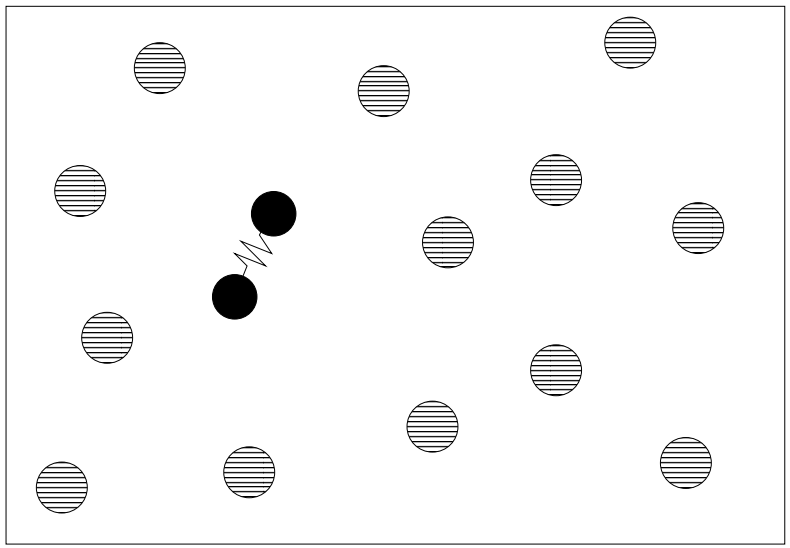


$\longrightarrow$ • Slow convergence of trajectorial averages

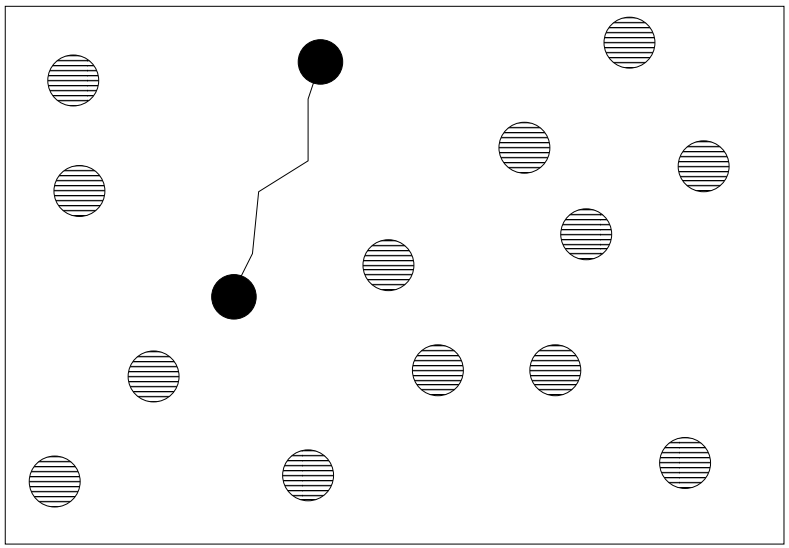
• Transitions between metastable states are rare events

A toy model for solvation

Influence of the solvation on a dimer conformation [Dellago, Geissler].



Compact state.

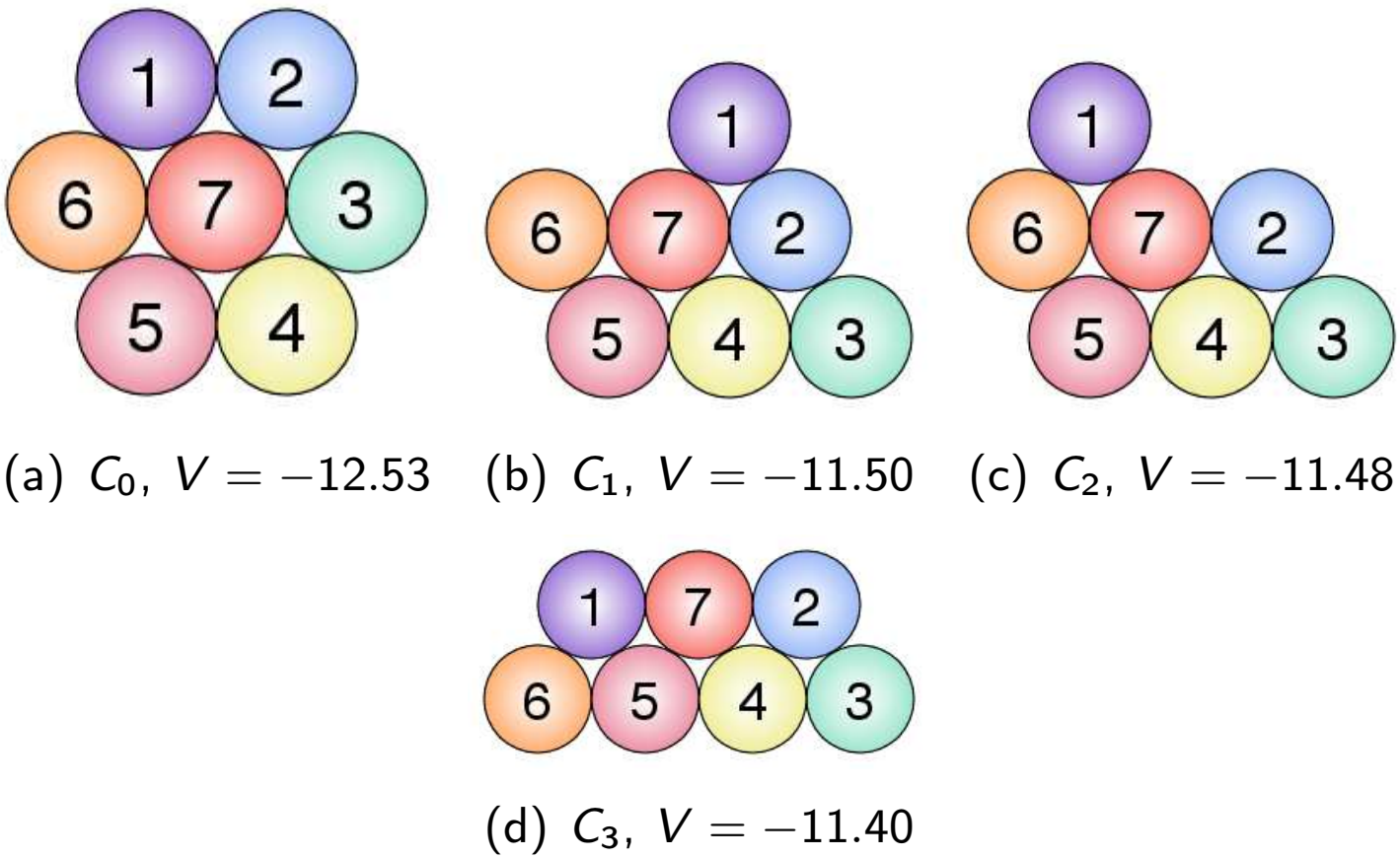


Stretched state.

The particles interact through a pair potential: truncated LJ for all particles except the two monomers (black particles) which interact through a double-well potential. A slow variable is the distance between the two monomers.

A toy example in material sciences

The 7 atoms Lennard Jones cluster in 2D.



(a) C_0, $V = -12.53$ (b) C_1, $V = -11.50$ (c) C_2, $V = -11.48$

(d) C_3, $V = -11.40$

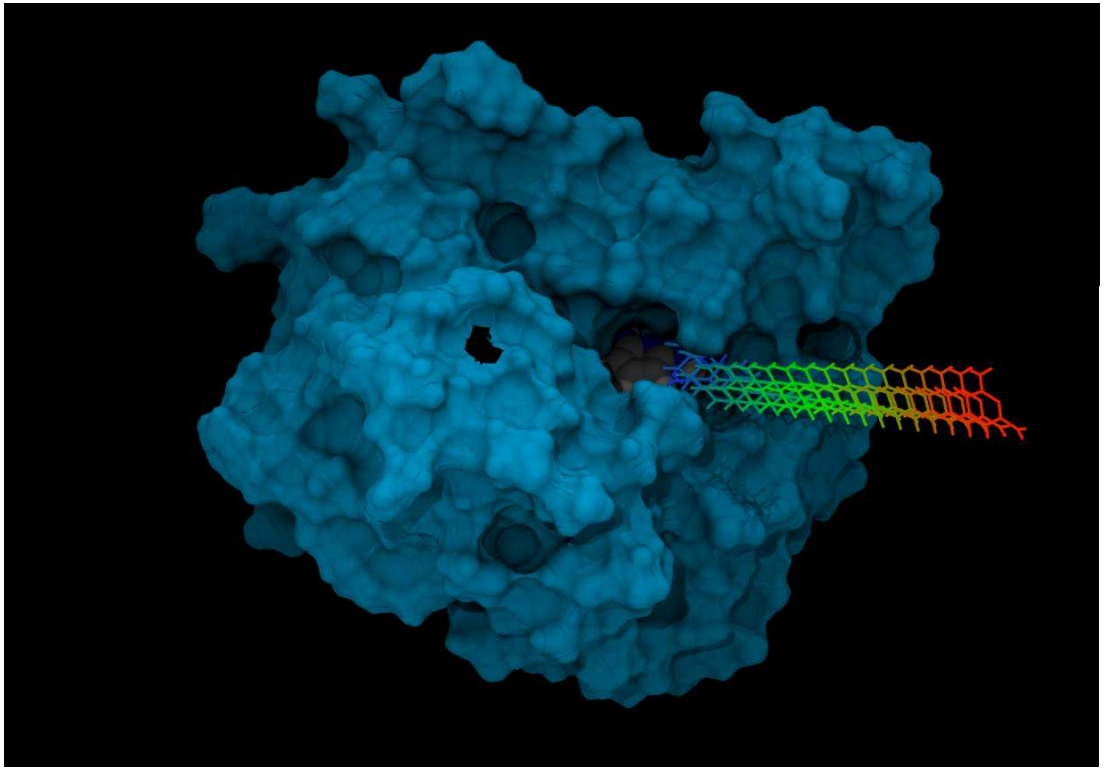
Figure: Low energy conformations of the Lennard-Jones cluster.

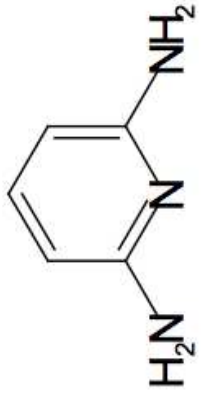
$\longrightarrow$ simulation

Simulations of biological systems

Unbinding of a ligand from a protein

(Diaminopyridine-HSP90, Courtesy of SANOFI)





Elementary time-step for the molecular dynamics $= 10^{-15}\,\mathrm{s}$

Dissociation time $= 0.5\,\mathrm{s}$

Challenge: bridge the gap between timescales

Outline

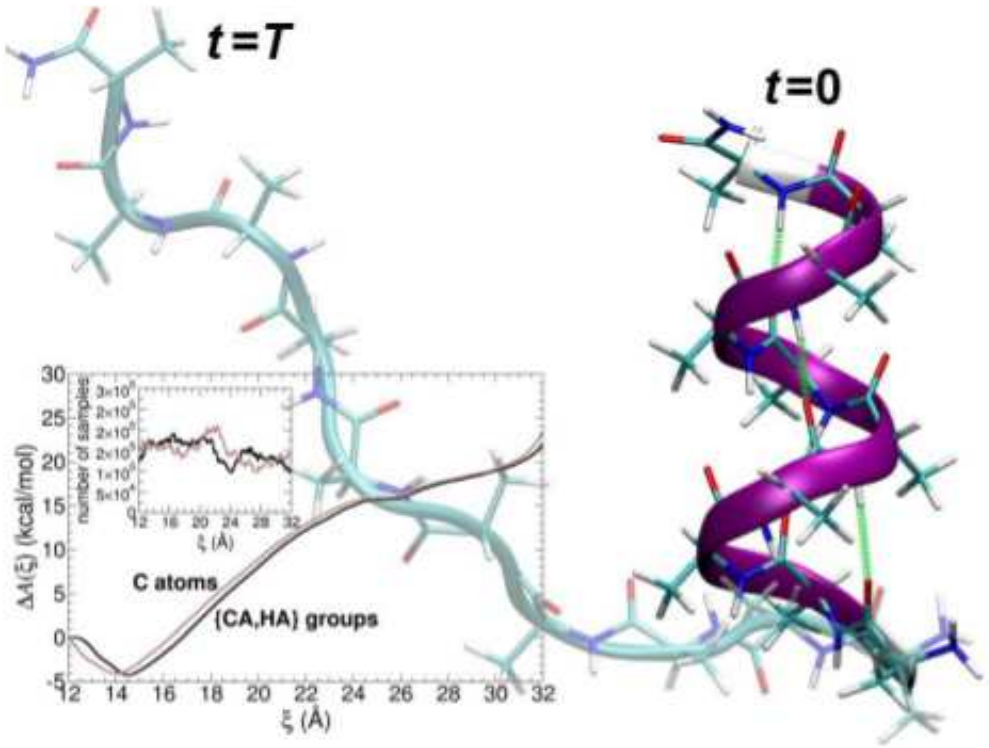
Outline of this part:

1. Definition of the free energy associated to a reaction coordinate.
2. Adaptive biasing techniques: Free energy computation methods based on biased stochastic processes.

Mathematical tools: delta measure and co-area formula, Entropy techniques and Logarithmic Sobolev Inequalities.

Underlying question: how to properly define and quantify metastability ? Various answers: (i) rate of convergence to equilibrium; (ii) exit time from metastable states; (iii) decorrelation time; (iv) asymptotic variance of estimators.

Reaction coordinate and free energy



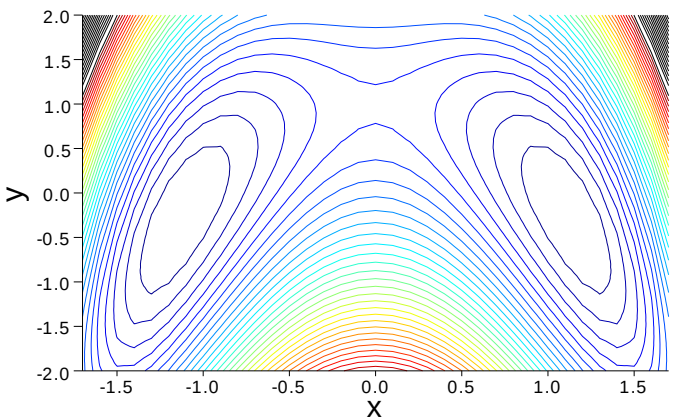
Reaction coordinate

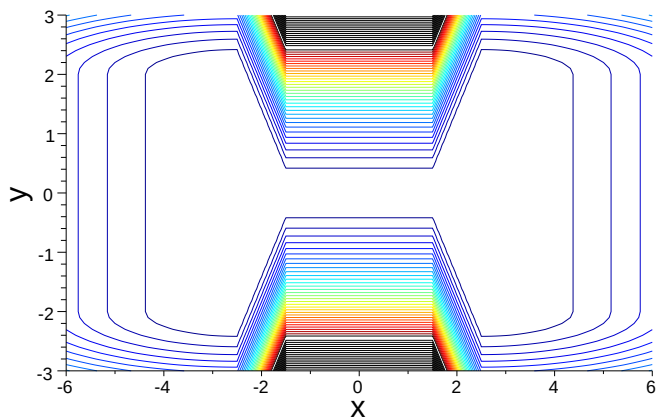
We suppose in the following that we know a slow variable of dimension 1: $\xi(\boldsymbol{X}_t)$, where $\xi : \mathbb{R}^d \to \mathbb{T}$ is a so-called reaction coordinate.

This reaction coordinate will be used to efficiently sample the canonical measure using two techniques: (i) constrained dynamics (thermodynamic integration) or (ii) biased dynamics (adaptive importance sampling technique).

Free energy will play a central role.

For example, in the 2D simple examples: $\xi(x, y) = x$.





Free energy

Let us introduce two probability measures associated to μ and ξ:

- The image of the measure μ by ξ:

$$\xi_* \mu\,(dz) = \exp(-\beta A(z))\,dz$$

where the free energy A is defined by:

$$A(z) = -\beta^{-1} \ln \left(\int_{\Sigma(z)} e^{-\beta V} \delta_{\xi(\boldsymbol{x})-z}(d\boldsymbol{x}) \right),$$

with $\Sigma(z) = \{\boldsymbol{x},\, \xi(\boldsymbol{x}) = z\}$ is a (smooth) submanifold of $\mathbb{R}^d$, and $\delta_{\xi(\boldsymbol{x})-z}(d\boldsymbol{x})\,dz = d\boldsymbol{x}$.

- The probability measure μ conditioned to $\xi(\boldsymbol{x}) = z$:

$$\mu_{\Sigma(z)}(d\boldsymbol{x}) = \frac{\exp(-\beta V(\boldsymbol{x}))\,\delta_{\xi(\boldsymbol{x})-z}(d\boldsymbol{x})}{\exp(-\beta A(z))}.$$

Free energy (2d case)

In the simple case $\xi(x, y) = x$, we have:

- The image of the measure μ by ξ:

$$\xi_* \mu\,(dx) = \exp(-\beta A(x))\, dx$$

where the free energy A is defined by:

$$A(x) = -\beta^{-1} \ln \left(\int_{\Sigma(x)} e^{-\beta V(x,y)} dy \right)$$

and $\Sigma(x) = \{(x, y),\ y \in \mathbb{R}\}$.

- The probability measure μ conditioned to $\xi(x, y) = x$:

$$\mu_{\Sigma(x)}(dy) = \frac{\exp(-\beta V(x, y))\, dy}{\exp(-\beta A(x))}.$$

The delta measure and the co-area formula

- The measure $\delta_{\xi(\boldsymbol{x})-z}$ is defined by: for all test functions $\varphi : \mathbb{T} \to \mathbb{R}$ and $\psi : \mathbb{R}^d \to \mathbb{R}$,

$$\int_{\mathbb{R}^d} \varphi \circ \xi(\boldsymbol{x}) \psi(\boldsymbol{x}) \, d\boldsymbol{x} = \int_{\mathbb{T}} \varphi(z) \left(\int_{\Sigma(z)} \psi(\boldsymbol{x}) \delta_{\xi(\boldsymbol{x})-z}(d\boldsymbol{x}) \right) dz.$$

- The measure $\delta_{\xi(\boldsymbol{x})-z}$ can be understood using a regularization procedure: for any test function $\psi : \mathbb{R}^d \to \mathbb{R}$,

$$\int_{\Sigma(z)} \psi(\boldsymbol{x}) \delta_{\xi(\boldsymbol{x})-z}(d\boldsymbol{x}) = \lim_{\epsilon \to 0} \int_{\mathbb{R}^d} \psi(\boldsymbol{x}) \delta^\epsilon(\xi(\boldsymbol{x}) - z) \, d\boldsymbol{x}$$

  where $\lim_{\epsilon \to 0} \delta^\epsilon = \delta$ (Dirac mass at zero).
- The measure $\delta_{\xi(\boldsymbol{x})-z}$ is related to the Lebesgue measure on $\Sigma(z)$ through:

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

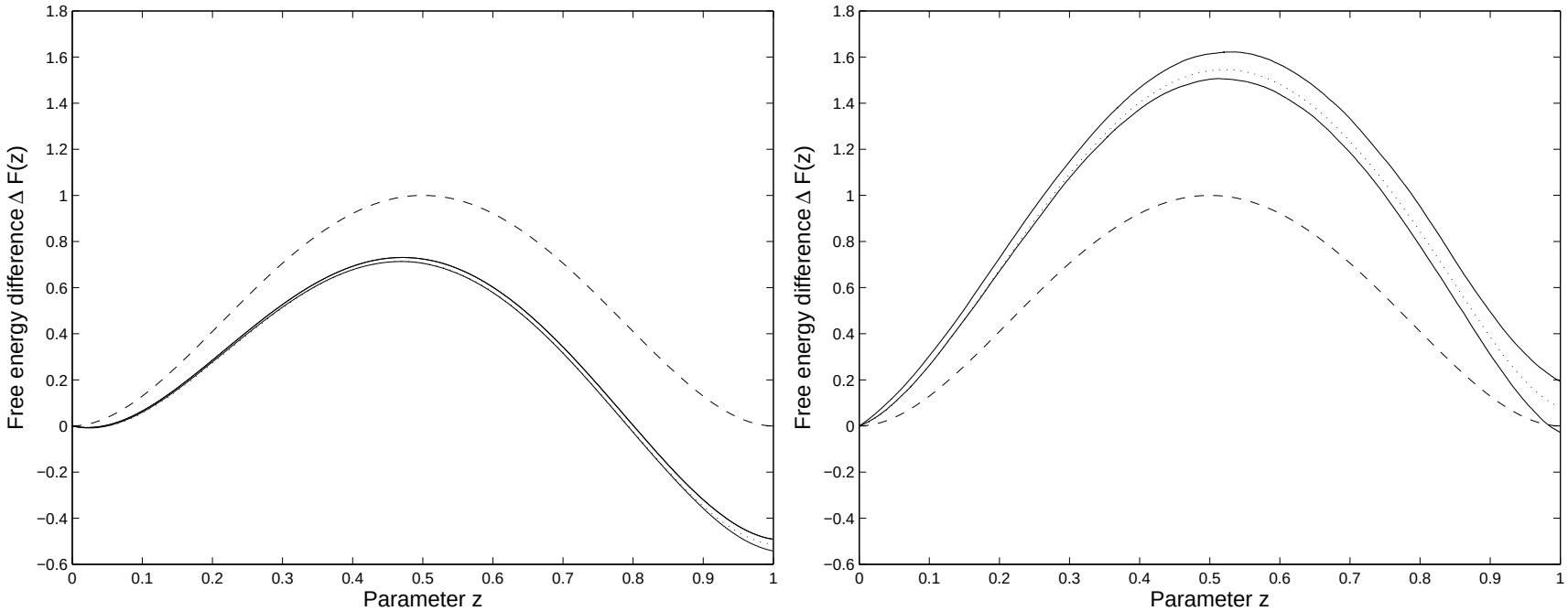
  This is the co-area formula. We thus have: $A(z) = -\beta^{-1} \ln \left(\int_{\Sigma(z)} e^{-\beta V} |\nabla \xi|^{-1} d\sigma_{\Sigma(z)} \right)$.

Free energy: Remarks

- A is the free energy associated with the reaction coordinate or collective variable ξ (angle, length, ...). The aim of many molecular dynamic simulations is to compute A.
- A is defined up to an additive constant, so that it is enough to compute free energy differences, or the derivative of A (the mean force).
- $A(z) = -\beta^{-1} \ln Z_{\Sigma(z)}$ and $Z_{\Sigma(z)}$ is the partition function associated with the conditioned probability measures: $\mu_{\Sigma(z)} = Z_{\Sigma(z)}^{-1} e^{-\beta V} |\nabla \xi|^{-1} d\sigma_{\Sigma(z)}$.
- If $U = \displaystyle\int_{\Sigma(z)} V \, Z_{\Sigma(z)}^{-1} e^{-\beta V} \delta_{\xi(\boldsymbol{x})-z}(d\boldsymbol{x})$ and $S = -k_B \displaystyle\int_{\Sigma(z)} \ln \left(Z_{\Sigma(z)}^{-1} e^{-\beta V} \right) Z_{\Sigma(z)}^{-1} e^{-\beta V} \delta_{\xi(\boldsymbol{x})-z}(d\boldsymbol{x})$, then $A = U - TS$ (since $\beta^{-1} = k_B T$).

Free energy on a simple example

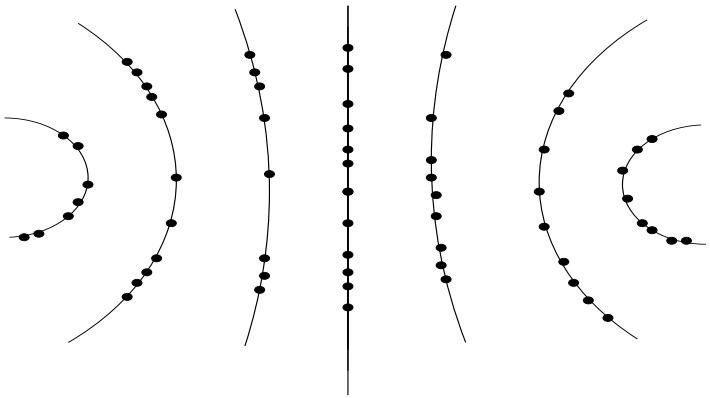
What is free energy ? The simple example of the solvation of a dimer. (Profiles computed using thermodynamic integration.)



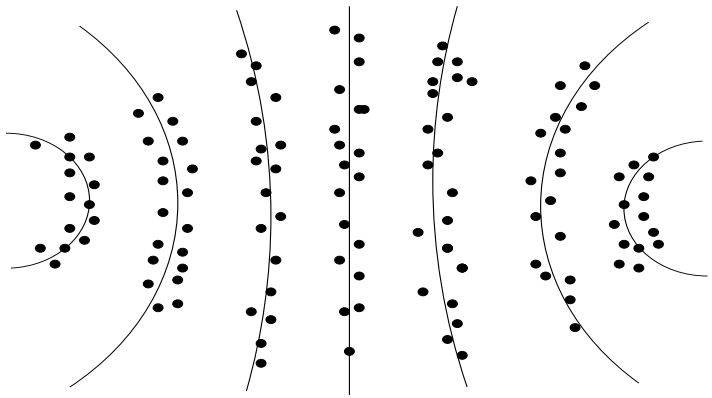
The density of the solvent molecules is lower on the left than on the right. At high (resp. low) density, the compact state is more (resp. less) likely. The "free energy barrier" is higher at high density than at low density. Related question: interpretation of the free energy barrier in terms of dynamics ?

Free energy calculation techniques

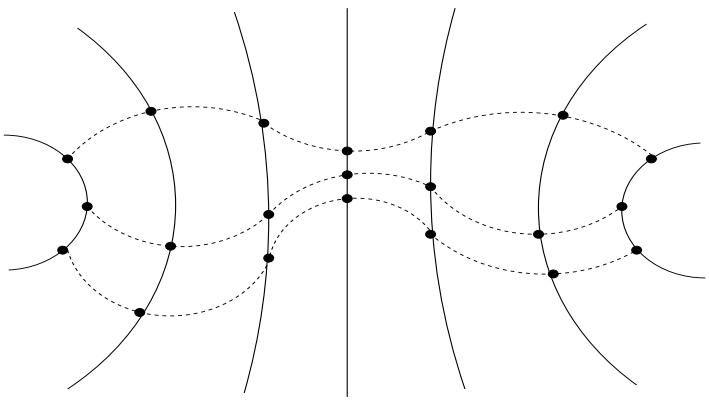
There are many free energy calculation techniques:



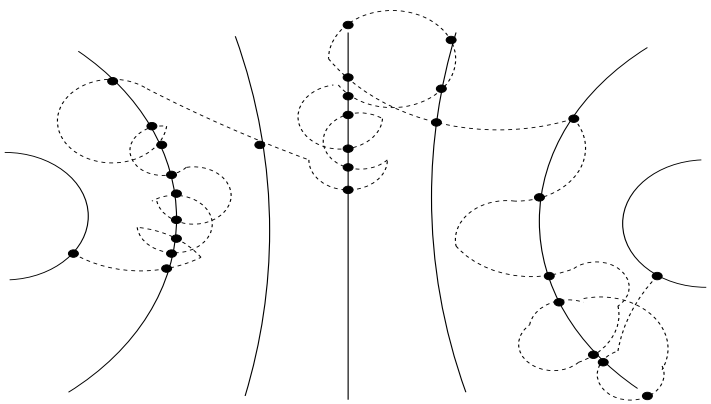
(a) Thermodynamic integration.



(b) Histogram method.

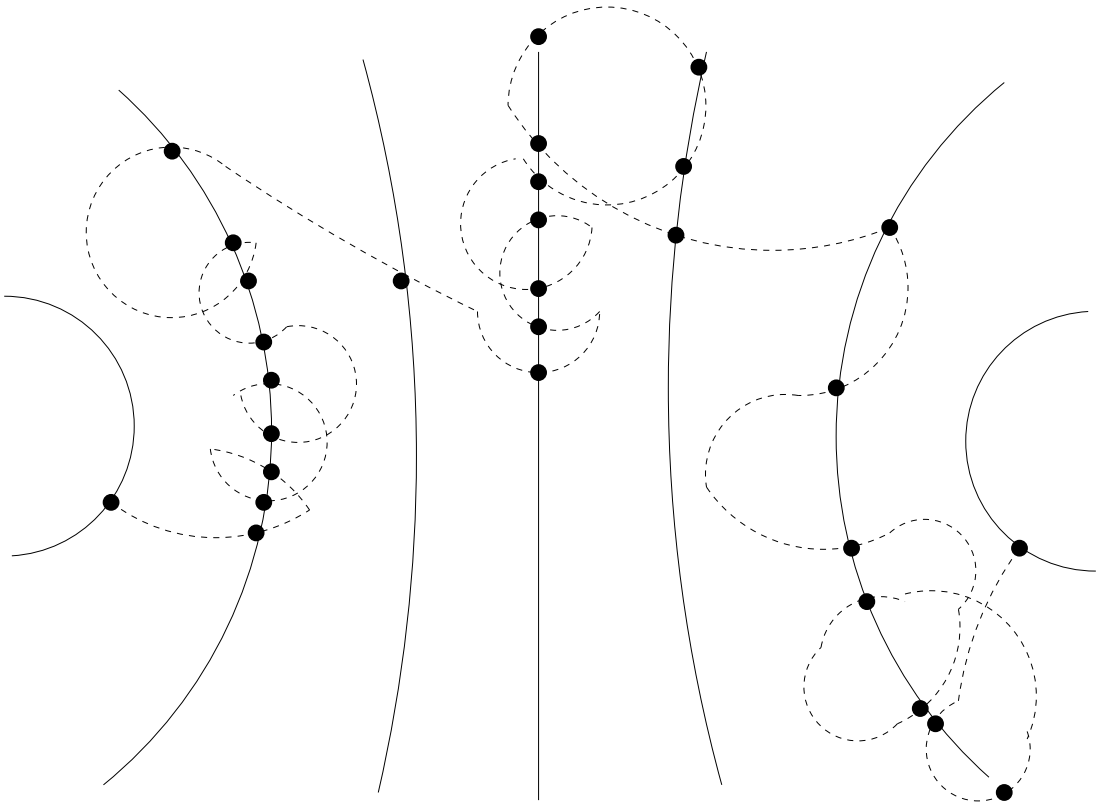


(c) Non equilibrium dynamics.



(d) Adaptive dynamics.

Adaptive biasing techniques

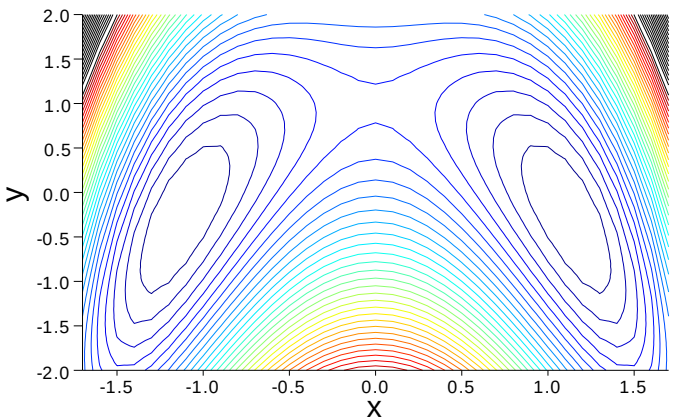


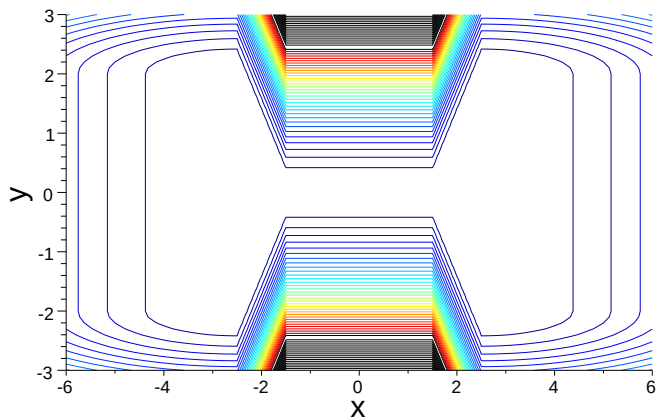
Adaptive biasing techniques

We suppose that we know a slow variable of dimension 1: $\xi(\boldsymbol{X}_t)$, where $\xi : \mathbb{R}^d \to \mathbb{T}$ is a so-called reaction coordinate.

This reaction coordinate will be used to bias the dynamics (adaptive importance sampling technique), using the free energy A associated with the reaction coordination ξ.

For example, in the 2D simple examples: $\xi(x, y) = x$.





Adaptive biasing techniques

The bottom line of adaptive methods is the following: for "well chosen" ξ the potential $V - A \circ \xi$ is less rugged than V. Indeed, by construction $\xi_* \exp(-\beta(V - A \circ \xi)) = 1_{\mathbb{T}}$.

Problem: A is unknown ! Idea: use a time dependent potential of the form

$$\mathcal{V}_t(\boldsymbol{x}) = V(\boldsymbol{x}) - A_t(\xi(\boldsymbol{x}))$$

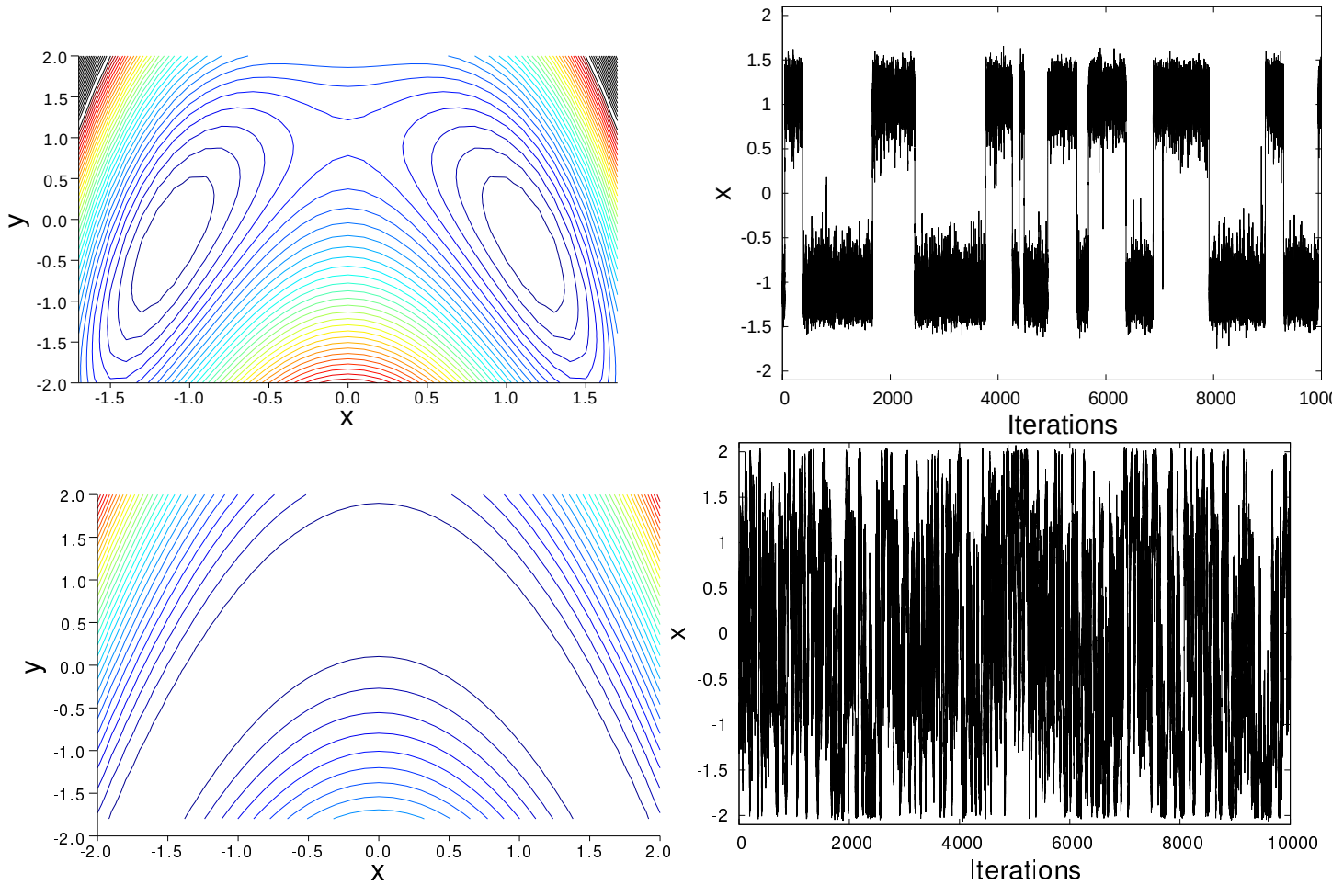
where A_t is an approximation at time t of A, given the configurations visited so far.

Hopes:

- build a dynamics which goes quickly to equilibrium,
- compute free energy profiles.

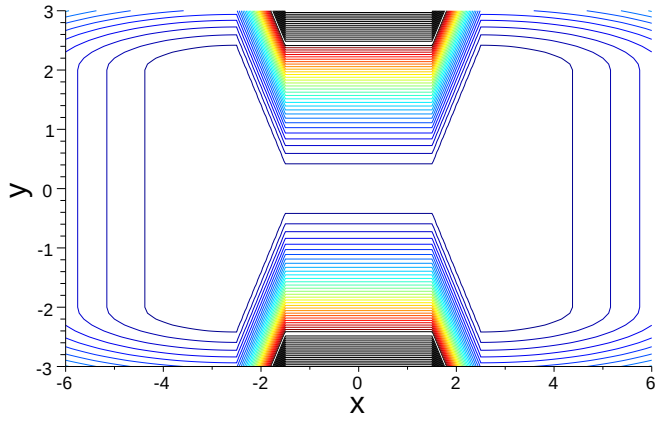
Wang-Landau, ABF, metadynamics: *Darve, Pohorille, Hénin, Chipot, Laio, Parrinello, Wang, Landau,...*

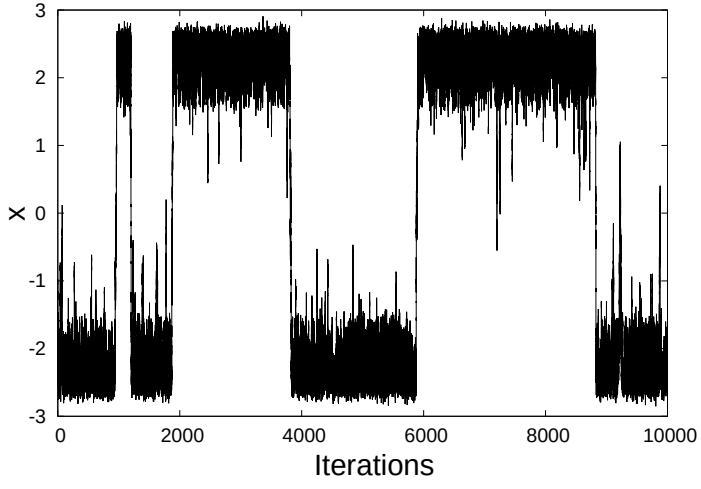
Free energy biased dynamics (1/2)

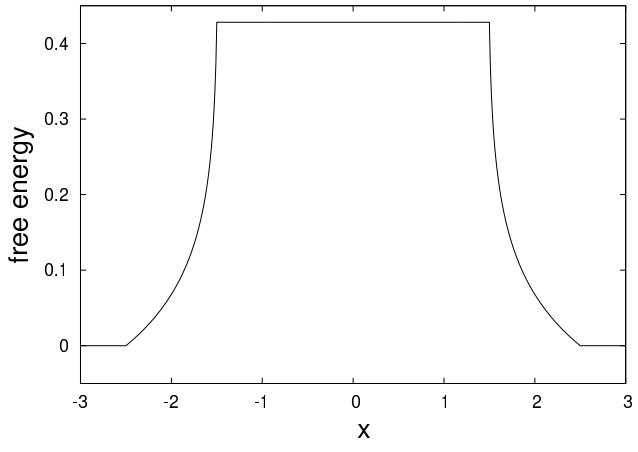


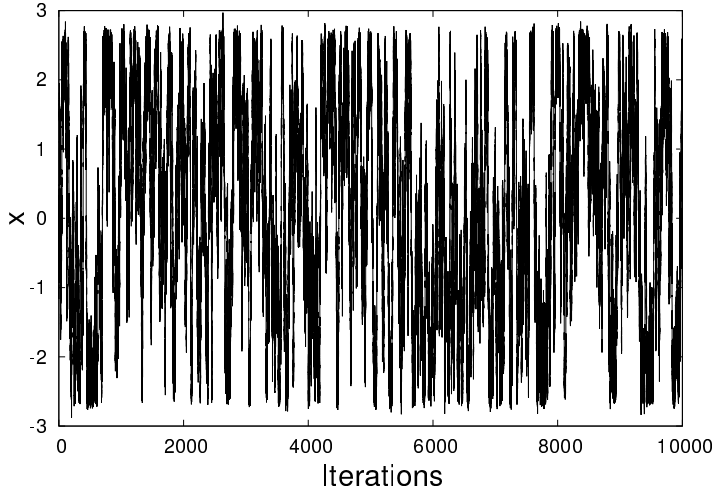
A 2D example of a free energy biased trajectory: energetic barrier.

Free energy biased dynamics (2/2)









A 2D example of a free energy biased trajectory: entropic barrier.

Updating strategies

How to update A_t ? Two methods depending on wether A'_t (Adaptive Biasing Force) or A_t (Adaptive Biasing Potential) is approximated.
To avoid geometry problem, an extended configurational space $(\boldsymbol{x}, z) \in \mathbb{R}^{n+1}$ may be considered, together with the meta-potential:

$$V^k(\boldsymbol{x}, z) = V(\boldsymbol{x}) + k(z - \xi(\boldsymbol{x}))^2.$$

Choosing $(\boldsymbol{x}, z) \mapsto z$ as a reaction coordinate, the associated free energy A^k is close to A (in the limit $k \to \infty$, up to an additive constant).

Adaptive algorithms used in molecular dynamics fall into one of these four possible combinations [TL, M. Rousset, G. Stoltz, J Chem Phys, 2007]:

	A'_t	A_t
V	ABF	Wang-Landau
V^k	...	metadynamics

The mean force

For the Adaptive Biasing Force (ABF) method, the idea is to use the formula

$$A'(z) = \frac{\displaystyle\int_{\Sigma(z)} \left(\frac{\nabla V \cdot \nabla \xi}{|\nabla \xi|^2} - \beta^{-1} \mathrm{div} \left(\frac{\nabla \xi}{|\nabla \xi|^2} \right) \right) e^{-\beta V} \, \delta_{\xi(\boldsymbol{x})-z}(d\boldsymbol{x})}{\displaystyle\int_{\Sigma(z)} e^{-\beta V} \, \delta_{\xi(\boldsymbol{x})-z}(d\boldsymbol{x})}$$

$$= \int_{\Sigma(z)} f \, d\mu_{\Sigma(z)} = \mathbb{E}_\mu(f(\boldsymbol{X}) | \xi(\boldsymbol{X}) = z).$$

The mean force $A'(z)$ is the average of f with respect to $\mu_{\Sigma(z)}$.

The mean force

In the simple case $\xi(x,y) = x$, remember that

$$A(x) = -\beta^{-1} \ln \left(\int e^{-\beta V(x,y)} dy \right),$$

so that

$$A'(x) = \frac{\displaystyle\int_{\Sigma(x)} \partial_x V \, e^{-\beta V(x,y)} \, dy}{\displaystyle\int_{\Sigma(x)} e^{-\beta V(x,y)} \, dy}$$

$$= \int \partial_x V \, d\mu_{\Sigma(x)}.$$

Notice that actually, whatever A_t is,

$$A'(z) = \frac{\displaystyle\int_{\Sigma(z)} f \, e^{-\beta(V - A_t \circ \xi)} \, \delta_{\xi(\boldsymbol{x}) - z}(d\boldsymbol{x})}{\displaystyle\int_{\Sigma(z)} e^{-\beta(V - A_t \circ \xi)} \, \delta_{\xi(\boldsymbol{x}) - z}(d\boldsymbol{x})}.$$

The mean force

Proof in the general case : $A'(z) = -\beta^{-1} \frac{\frac{d}{dz} \int_{\Sigma(z)} \exp(-\beta V) \delta_{\xi(\boldsymbol{x})-z}(d\boldsymbol{x})}{\int_{\Sigma(z)} \exp(-\beta V) \delta_{\xi(\boldsymbol{x})-z}(d\boldsymbol{x})}$

and

$$\int_{\mathbb{T}} \left(\int_{\Sigma(z)} \exp(-\beta V) \delta_{\xi(\boldsymbol{x})-z}(d\boldsymbol{x}) \right)' \phi(z)\, dz$$

$$= - \int_{\mathbb{T}} \int_{\Sigma(z)} \exp(-\beta V) \delta_{\xi(\boldsymbol{x})-z}(d\boldsymbol{x}) \phi'\, dz$$

$$= - \int_{\mathbb{T}} \int_{\Sigma(z)} \exp(-\beta V) \phi' \circ \xi\, \delta_{\xi(\boldsymbol{x})-z}(d\boldsymbol{x})\, dz$$

$$= - \int_{\mathbb{R}^d} \exp(-\beta V) \phi' \circ \xi d\boldsymbol{x} = - \int_{\mathbb{R}^d} \exp(-\beta V) \nabla(\phi \circ \xi) \cdot \frac{\nabla \xi}{|\nabla \xi|^2} d\boldsymbol{x}$$

$$= \int_{\mathbb{R}^d} \nabla \cdot \left(\exp(-\beta V) \frac{\nabla \xi}{|\nabla \xi|^2} \right) \phi \circ \xi\, d\boldsymbol{x}$$

$$= \int_{\mathbb{T}} \int_{\Sigma(z)} \left(-\beta \frac{\nabla V \cdot \nabla \xi}{|\nabla \xi|^2} + \nabla \cdot \left(\frac{\nabla \xi}{|\nabla \xi|^2} \right) \right) \exp(-\beta V) \delta_{\xi(\boldsymbol{x})-z}(d\boldsymbol{x}) \phi(z)\, d$$

The ABF method

Thus, we would like to simulate:

$$\begin{cases} d\boldsymbol{X}_t = -\nabla(V - A \circ \xi)(\boldsymbol{X}_t)\, dt + \sqrt{2\beta^{-1}} d\boldsymbol{W}_t, \\ A'(z) = \mathbb{E}_\mu\left(f(\boldsymbol{X}) | \xi(\boldsymbol{X}) = z\right) \end{cases}$$

but A is unknown...

The ABF method

The ABF dynamics is then:

$$\begin{cases} d\boldsymbol{X}_t = -\nabla(V - A_t \circ \xi)(\boldsymbol{X}_t)\, dt + \sqrt{2\beta^{-1}} d\boldsymbol{W}_t, \\ A_t'(z) = \mathbb{E}\left(f(\boldsymbol{X}_t) | \xi(\boldsymbol{X}_t) = z\right). \end{cases}$$

The ABF method

The ABF dynamics is then:

$$\begin{cases} d\boldsymbol{X}_t = -\nabla(V - A_t \circ \xi)(\boldsymbol{X}_t)\, dt + \sqrt{2\beta^{-1}} d\boldsymbol{W}_t, \\ A_t'(z) = \mathbb{E}\left(f(\boldsymbol{X}_t) | \xi(\boldsymbol{X}_t) = z\right). \end{cases}$$

The associated (nonlinear) Fokker-Planck equation writes:

$$\begin{cases} \partial_t \psi = \operatorname{div}\left(\nabla(V - A_t \circ \xi)\psi + \beta^{-1}\nabla\psi\right), \\ A_t'(z) = \dfrac{\displaystyle\int f\, \psi\, \delta_{\xi(\boldsymbol{x})-z}(d\boldsymbol{x})}{\displaystyle\int \psi\, \delta_{\xi(\boldsymbol{x})-z}(d\boldsymbol{x})}, \end{cases}$$

where $\boldsymbol{X}_t \sim \psi(t, \boldsymbol{x})\, d\boldsymbol{x}$.

The ABF method

The ABF dynamics is then:

$$\begin{cases} d\boldsymbol{X}_t = -\nabla(V - A_t \circ \xi)(\boldsymbol{X}_t)\, dt + \sqrt{2\beta^{-1}} d\boldsymbol{W}_t, \\ A_t'(z) = \mathbb{E}\left(f(\boldsymbol{X}_t) | \xi(\boldsymbol{X}_t) = z\right). \end{cases}$$

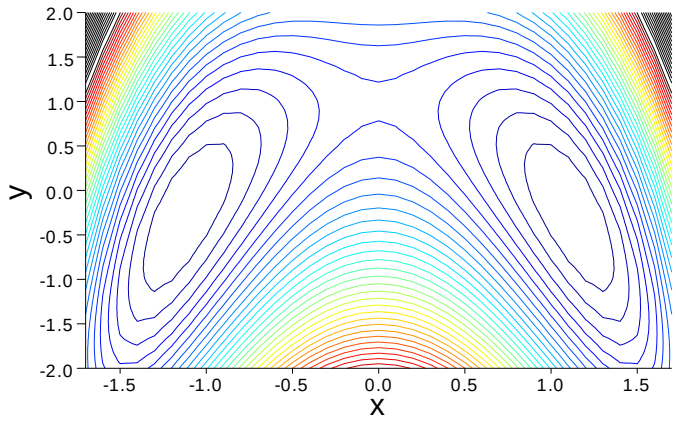
The associated (nonlinear) Fokker-Planck equation writes:

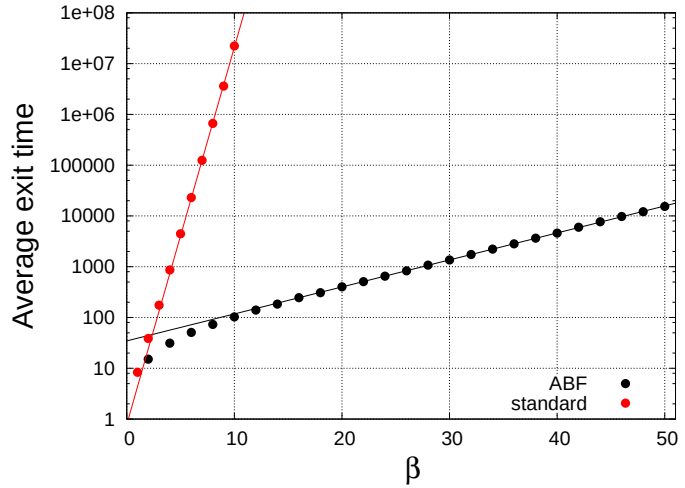
$$\begin{cases} \partial_t \psi = \operatorname{div}\left(\nabla(V - A_t \circ \xi)\psi + \beta^{-1}\nabla\psi\right), \\ A_t'(z) = \dfrac{\displaystyle\int f\, \psi\, \delta_{\xi(\boldsymbol{x})-z}(d\boldsymbol{x})}{\displaystyle\int \psi\, \delta_{\xi(\boldsymbol{x})-z}(d\boldsymbol{x})}, \end{cases}$$

where $\boldsymbol{X}_t \sim \psi(t, \boldsymbol{x})\, d\boldsymbol{x}$.

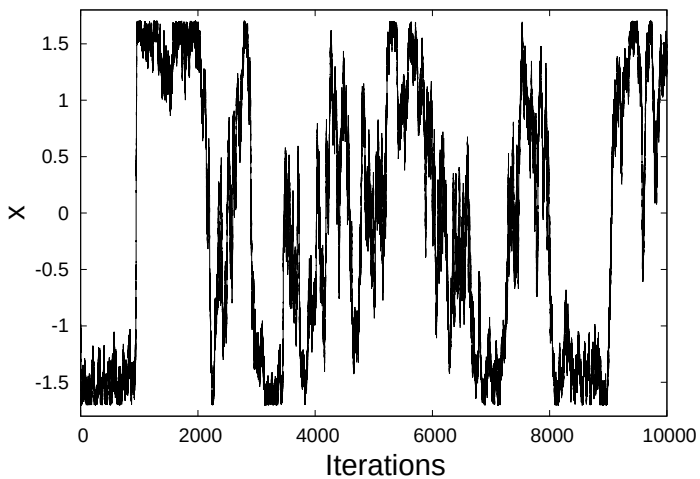
Questions: Does A_t' converge to A' ? What did we gain compared to the original gradient dynamics ?

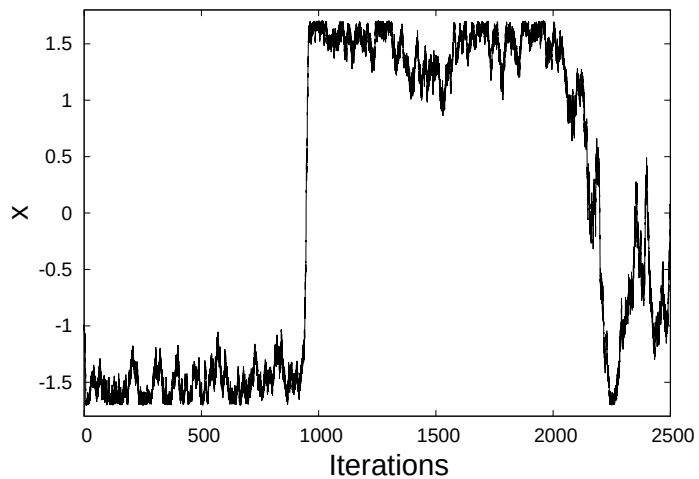
Back to the 2D example





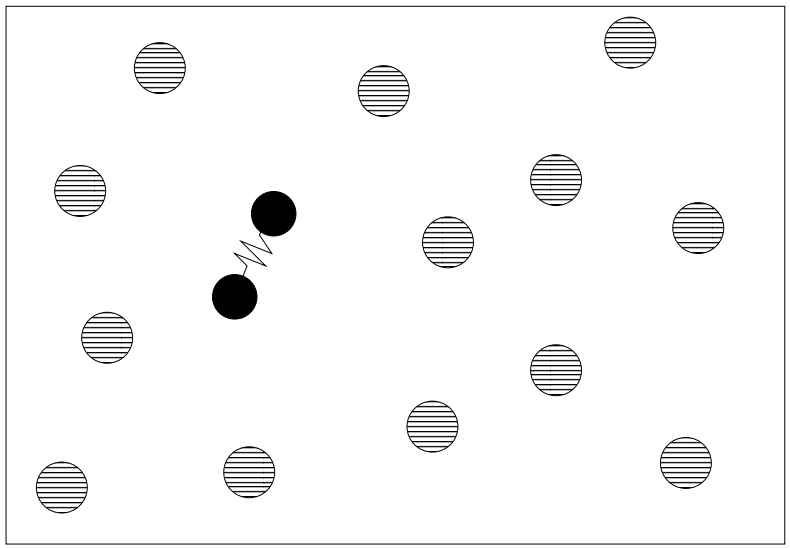
Left: the 2D potential – energetic barrier; Right: average exit time from the left well



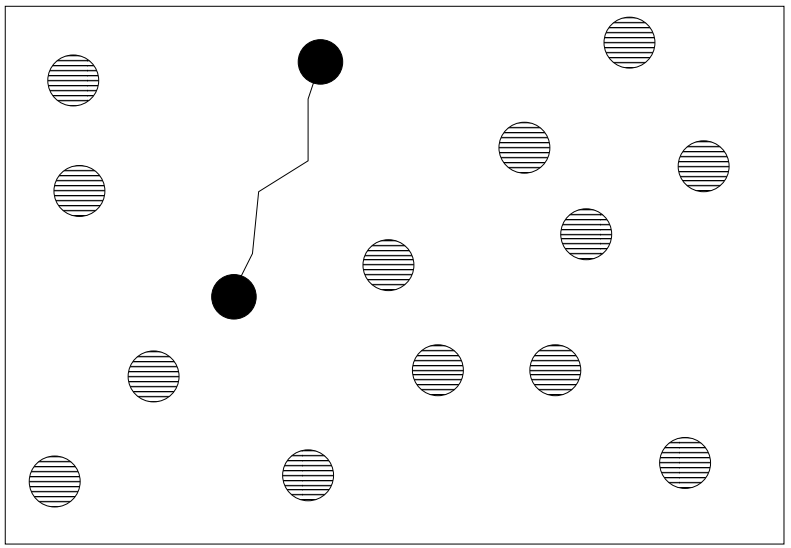


The ABF trajectory (right: zoom on the first 2500 iterations)

Back to the toy example for solvation



Compact state.



Stretched state.

The reaction coordinate ξ is the distance between the two monomers. $\longrightarrow$ simulation

Longtime convergence and entropy (1/3)

Recall the original gradient dynamics:

$$d\boldsymbol{Q}_t = -\nabla V(\boldsymbol{Q}_t)\, dt + \sqrt{2\beta^{-1}} d\boldsymbol{W}_t.$$

The associated (linear) Fokker-Planck equation writes:

$$\partial_t \phi = \mathrm{div}\, \left(\nabla V \phi + \beta^{-1} \nabla \phi \right).$$

where $\boldsymbol{Q}_t \sim \phi(t, \boldsymbol{q})\, d\boldsymbol{q}$.

The metastable behaviour of $\boldsymbol{Q}_t$ is related to the multimodality of μ, which can be quantified through the rate of convergence of ϕ to $\phi_\infty = Z^{-1} \exp(-\beta V)$.

A classical approach for partial differential equations (PDEs): entropy techniques.

Longtime convergence and entropy (2/3)

Notice that the Fokker-Planck equation rewrites

$$\partial_t \phi = \beta^{-1} \mathrm{div} \left(\phi_\infty \nabla \left(\frac{\phi}{\phi_\infty} \right) \right).$$

Let us introduce the entropy:

$$E(t) = H(\phi(t,\cdot)|\phi_\infty) = \int \ln \left(\frac{\phi}{\phi_\infty} \right) \phi.$$

We have (Csiszár-Kullback inequality):

$$\|\phi(t,\cdot) - \phi_\infty\|_{L^1} \leq \sqrt{2E(t)}.$$

Longtime convergence and entropy (3/3)

$$\begin{aligned}\frac{dE}{dt} &= \int \ln\left(\frac{\phi}{\phi_\infty}\right)\partial_t\phi \\ &= \beta^{-1}\int \ln\left(\frac{\phi}{\phi_\infty}\right)\operatorname{div}\left(\phi_\infty\nabla\left(\frac{\phi}{\phi_\infty}\right)\right) \\ &= -\beta^{-1}\int\left|\nabla\ln\left(\frac{\phi}{\phi_\infty}\right)\right|^2\phi =: -\beta^{-1}I(\phi(t,\cdot)|\phi_\infty).\end{aligned}$$

If V is such that the following Logarithmic Sobolev inequality (LSI(R)) holds: $\forall\phi$ pdf,

$$H(\phi|\phi_\infty) \leq \frac{1}{2R}I(\phi|\phi_\infty)$$

then $E(t) \leq E(0)\exp(-2\beta^{-1}Rt)$ and thus ϕ converges to ϕ_∞ exponentially fast with rate $\beta^{-1}R$.

Metastability $\Longleftrightarrow$ small R

Convergence of ABF (1/4)

A convergence result [TL, M. Rousset, G. Stoltz, Nonlinearity 2008]: Recall the ABF Fokker-Planck equation:

$$\begin{cases} \partial_t \psi = \mathrm{div}\, \left(\nabla (V - A_t \circ \xi) \psi + \beta^{-1} \nabla \psi \right), \\ A_t'(z) = \dfrac{\int f\, \psi\, \delta_{\xi(x)-z}(dx)}{\int \psi\, \delta_{\xi(x)-z}(dx)}. \end{cases}$$

Suppose:

(H1) "Ergodicity" of the microscopic variables: the conditional probability measures $\mu_{\Sigma(z)}$ satisfy a LSI(ρ),

(H2) Bounded coupling: $\left\| \nabla_{\Sigma(z)} f \right\|_{L^\infty} < \infty$,

then

$$\| A_t' - A' \|_{L^2} \leq C \exp(-\beta^{-1} \min(\rho, r) t).$$

The rate of convergence is limited by:

- the rate r of convergence of $\overline{\psi} = \int \psi\, \delta_{\xi(x)-z}(dx)$ to $\overline{\psi_\infty}$,
- the LSI constant ρ (the real limitation).

Convergence of ABF (2/4)

In summary:

- Original gradient dynamics: $\exp(-\beta^{-1}Rt)$ where R is the LSI constant for μ;
- ABF dynamics: $\exp(-\beta^{-1}\rho t)$ where ρ is the LSI constant for the conditioned probability measures $\mu_{\Sigma(z)}$.

If ξ is well chosen, $\rho \gg R$: the free energy can be computed very efficiently.

Two ingredients of the proof:

(1) The marginal $\overline{\psi}(t,z) = \int \psi(t,\boldsymbol{x})\,\delta_{\xi(\boldsymbol{x})-z}(d\boldsymbol{x})$ satisfies a closed PDE:

$$\partial_t \overline{\psi} = \beta^{-1}\partial_{z,z}\overline{\psi} \text{ on } \mathbb{T},$$

and thus, $\overline{\psi}$ converges towards $\overline{\psi_\infty} \equiv 1$, with exponential speed $C\exp(-4\pi^2\beta^{-1}t)$. (Here, $r = 4\pi^2$).

Convergence of ABF (3/4)

(2) The total entropy can be decomposed as [N. Grunewald, F. Otto, C. Villani, M. Westdickenberg, Ann. IHP, 2009]:

$$E = E_M + E_m$$

where

The total entropy is $E = H(\psi|\psi_\infty)$,

The macroscopic entropy is $E_M = H(\overline{\psi}|\overline{\psi_\infty})$,

The microscopic entropy is

$$E_m = \int H\Big(\psi(\cdot|\xi(\boldsymbol{x}) = z)\Big|\psi_\infty(\cdot|\xi(\boldsymbol{x}) = z)\Big)\,\overline{\psi}(z)\,dz.$$

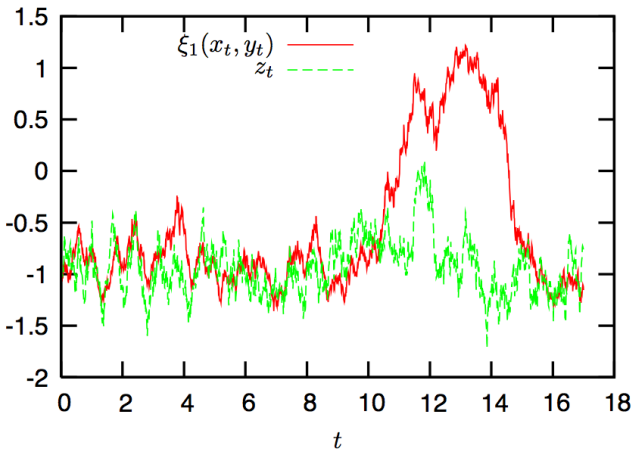
We already know that E_M goes to zero: it remains only to consider E_m...

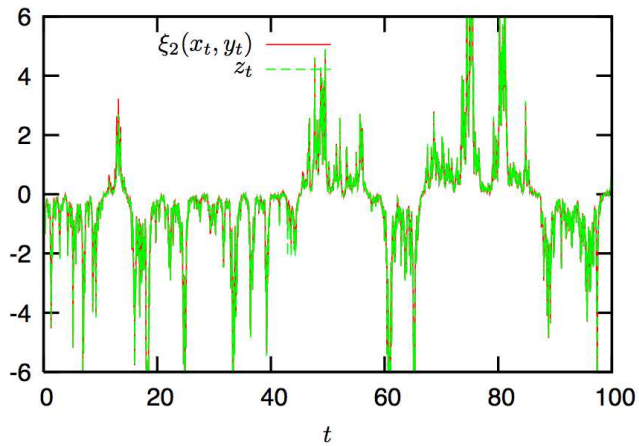
Convergence of ABF (4/4)

Other results based on this set of assumptions:

- [TL, JFA 2008] LSI for the cond. meas. $\mu_{\Sigma(z)}$
  + LSI for the marginal $\overline{\mu}(dz) = \xi_* \mu(dz)$
  + bdd coupling ($\|\nabla_{\Sigma(z)} f\|_{L^\infty} < \infty$) $\Longrightarrow$ LSI for μ.
- [F. Legoll, TL, Nonlinearity, 2010] Effective dynamics for $\xi(\boldsymbol{Q}_t)$. Uniform control in time:

$$H(\mathcal{L}(\xi(\boldsymbol{Q}_t))|\mathcal{L}(z_t)) \leq C \left(\frac{\|\nabla_{\Sigma(z)} f\|_{L^\infty}}{\rho} \right)^2 H(\mathcal{L}(\boldsymbol{Q}_0)|\mu).$$





Discretization of ABF

Discretization of adaptive methods can be done using two (complementary) approaches:

- Use empirical means over many replicas (interacting particle system):

$$\mathbb{E}(f(\boldsymbol{X}_t)|\xi(\boldsymbol{X}_t)=z) \simeq \frac{\sum_{m=1}^N f(\boldsymbol{X}_t^{m,N})\,\delta^\alpha(\xi(\boldsymbol{X}_t^{m,N})-z)}{\sum_{m=1}^N \delta^\alpha(\xi(\boldsymbol{X}_t^{m,N})-z)}.$$

  This approach is easy to parallelize, flexible (selection mechanisms) and efficient in cases with multiple reactive paths. [TL, M. Rousset, G. Stoltz, 2007; C. Chipot, TL, K. Minoukadeh, 2010 ; TL, K. Minoukadeh, 2010]

- Use trajectorial averages along a single path:

$$\mathbb{E}(f(\boldsymbol{X}_t)|\xi(\boldsymbol{X}_t)=z) \simeq \frac{\int_0^t f(\boldsymbol{X}_s)\,\delta^\alpha(\xi(\boldsymbol{X}_s)-z)\,ds}{\int_0^t \delta^\alpha(\xi(\boldsymbol{X}_s)-z)\,ds}.$$

  The longtime behavior is much more difficult to analyze.

Back to the original problem

How to use free energy to compute canonical averages $\int \varphi d\mu = Z^{-1} \int \varphi e^{-\beta V}$?

- Importance sampling:

$$\int \varphi \, d\mu = \frac{\int \varphi e^{-\beta A \circ \xi} \, Z_A^{-1} e^{-\beta(V - A \circ \xi)}}{\int e^{-\beta A \circ \xi} \, Z_A^{-1} e^{-\beta(V - A \circ \xi)}}.$$

- Conditioning:

$$\int \varphi \, d\mu = \frac{\int_z \left(\int_{\Sigma(z)} \varphi \, d\mu_{\Sigma(z)} \right) e^{-\beta A(z)} \, dz}{\int_z e^{-\beta A(z)} \, dz}.$$

This requires the sampling of the conditional probability measure $\mu_{\Sigma(z)}$ which can be done using projected Langevin dynamics [TL, M. Rousset, G. Stoltz, 2011].

ABF: Current developments and open problems

- Avoid the computation of ξ: extended-ABF
- Projection on a gradient of the mean force (Helmholtz decomposition)
- Reaction coordinates in larger dimension: exchange bias, separated representations
- Extension of the analysis to the Langevin dynamics
- Extension of the analysis to approximations of the mean force or the free energy based on time averages

Other techniques to compute thermodynamic quantities

Other algorithms which are used in MD to sample efficiently μ:

- Umbrella sampling and statistical reconstruction: Histogram methods
- Conditioning and constrained sampling: Thermodynamic integration
- Interacting replicas techniques: Parallel tempering, Replica exchange dynamics, ...
- Modify the dynamics: Metropolis Hastings algorithms with well-chosen proposals, non-reversible perturbations,...
- Out of equilibrium methods: fluctuation relations *à la* Jarzynski-Crooks

References

Monographs on numerical methods in MD:

- Allen, M., Tildesley, D.: Computer Simulation of Liquids. Oxford University Press (1989)
- Frenkel, D., Smit, B.: Understanding Molecular Simulation: From Algorithms to Applications. Oxford University Press (2001)
- Hairer, E., Lubich, C., Wanner, G.: Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations, Springer-Verlag (2006).
- Leimkuhler, B., Reich, S.: Simulating Hamiltonian Dynamics. Cambridge University Press (2005).
- Rapaport, D.: The Art of Molecular Dynamics Simulation. Cambridge University Press (2004)
- Tuckerman, M.: Statistical Mechanics: Theory and Molecular Simulation. Oxford University Press (2010).

A book on the mathematics for stochastic MD