Nonequilibrium computation of free energy differences: some new algorithms

Gabriel STOLTZ

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

http://cermics.enpc.fr/∼stoltz/

Joint work with Tony Lelièvre (CERMICS) and Mathias Rousset (INRIA Lille)

Free-energy

- System of N particles $(q,p) = (q_1, \ldots, q_N,\ p_1, \ldots, p_N)$
- Hamiltonian

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

- Free energy (defined up to constants)

$$F = -\beta^{-1} \ln \left(\int_{\mathcal{E}} \mathrm{e}^{-\beta H(q,p)}\, dq\, dp \right)$$

- 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 variables)

- Alchemical transition: indexed by an external parameter λ (force field parameter, magnetic field,...)

$$\Delta F = -\beta^{-1} \ln \left(\frac{\int_{\mathcal{E}} \mathrm{e}^{-\beta H_1(q,p)}\, dq\, dp}{\int_{\mathcal{E}} \mathrm{e}^{-\beta H_0(q,p)}\, dq\, dp} \right) ;$$

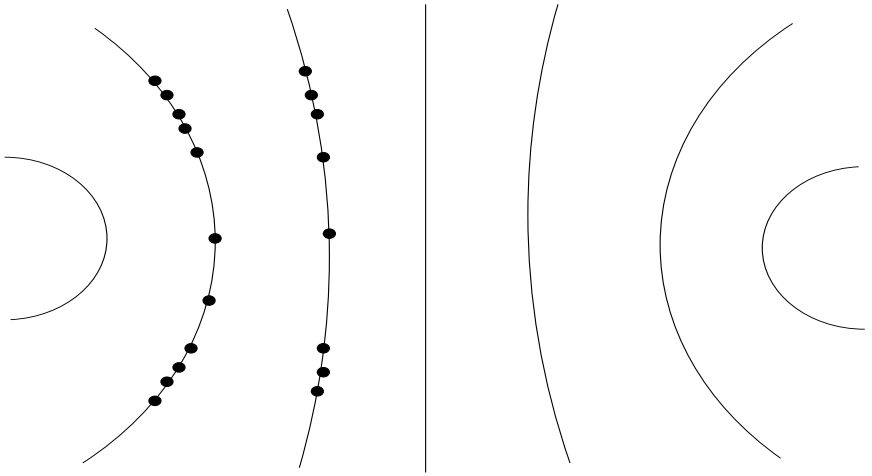
Typically, $H_\lambda = (1-\lambda)H_0 + \lambda H_1$

- (given) reaction coordinate $\xi : \mathbb{R}^{3N} \to \mathbb{R}^m$ (angle, length,...):

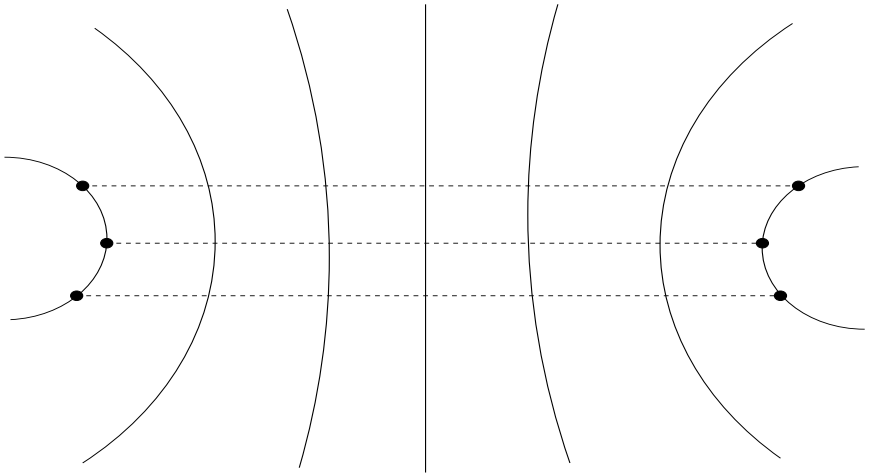
$$\Delta F = -\beta^{-1} \ln \left(\frac{\int_{\mathcal{E}} \mathrm{e}^{-\beta H(q,p)}\, \delta_{\xi(q)-z_1}\, dq\, dp}{\int_{\mathcal{E}} \mathrm{e}^{-\beta H(q,p)}\, \delta_{\xi(q)-z_0}\, dq\, dp} \right) .$$

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

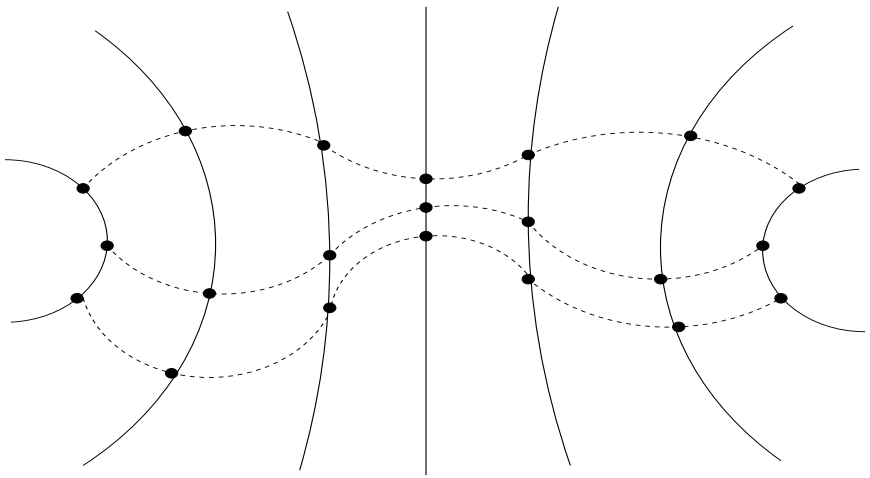
Cartoon comparison of the methods



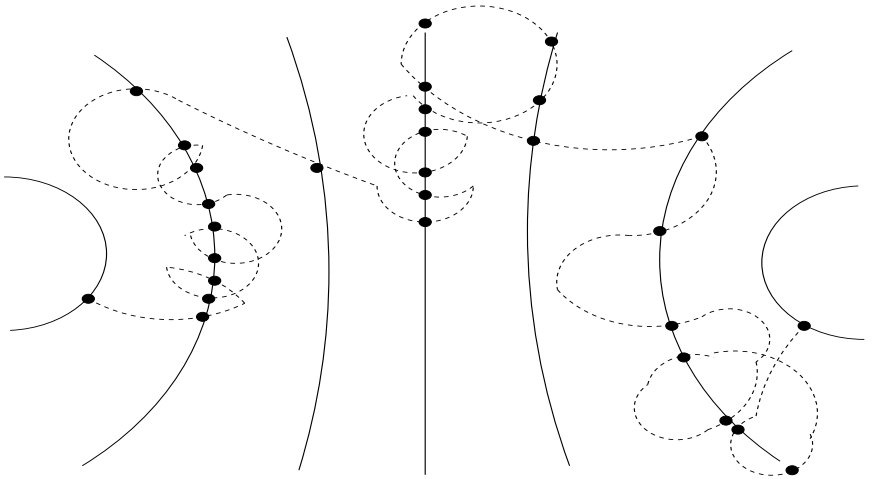
(a) Thermodynamic integration



(b) Free energy perturbation



(c) Nonequilibrium switching dynamics



(d) Adaptive dynamics

Mathematical classification (june 2008)

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

Outline of the talk

- Alchemical transitions
  - Jarzynski's equality
  - A selection process to avoid the degeneracy of weights
- The reaction coordinate case
  - Thermodynamic integration
  - The corresponding nonequilibrium dynamics
- Path sampling strategies
  - Computation of free energy differences using path sampling
  - A refined shooting algorithm

Alchemical transitions:

Selection processes

The Jarzynski equality

- Schedule $\Lambda(t)$ such that $\Lambda(0) = 0$, $\Lambda(T) = 1$, $0 < T < +\infty$, and canonical initial conditions $q_0 \sim d\mu_0(q)$
- Time inhomogeneous Markovian evolution (the potential energy changes!)

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

- Define the work along a trajectory: $\mathcal{W}_T = \displaystyle\int_0^T \frac{\partial V_{\Lambda(t)}}{\partial \Lambda}(q_{\Lambda,t})\, \Lambda'(t)\, dt$
- Jarzynski's equality[a] (Feynman-Kac formula[b])

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

- Consequence $\Delta F \geq \mathbb{E}(\mathcal{W}_t)$ (second law of thermodynamics)
- The dynamics has an influence on the accuracy of the results!

---

[a] C. Jarzynski, *Phys. Rev. Lett.* (1997)

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

Equilibrating the nonequilibrium dynamics

- Degeneracy of weights and the law of $q_{\Lambda,t}$ is not the canonical measure!
- Birth/death process (jump/diffusion) to restore equilibrium at all times:[a] process $\bar{q}_{\Lambda,t}$
  - simulate several replicas of the system in parallel [diffusion process]
  - to each replica, two exponentially distributed random times are attached (birth/death times)
  - if the current work performed on the replica is lower than the average, decrease the birth time, otherwise decrease the death time
  - if the birth time of the k-th replica is zero, select another replica l at random, and set $q^l = q^k$ (id death process) [jump process]
- The law of $\bar{q}_{\Lambda,t}$ is the canonical measure at all times and $\mathbb{E}(\overline{\mathcal{W}}_T) = \Delta F$
- Convergence can be proved when the number of particles goes to infinity[b]

[a] M. Rousset and G. Stoltz, *J. Stat. Phys.* (2005)

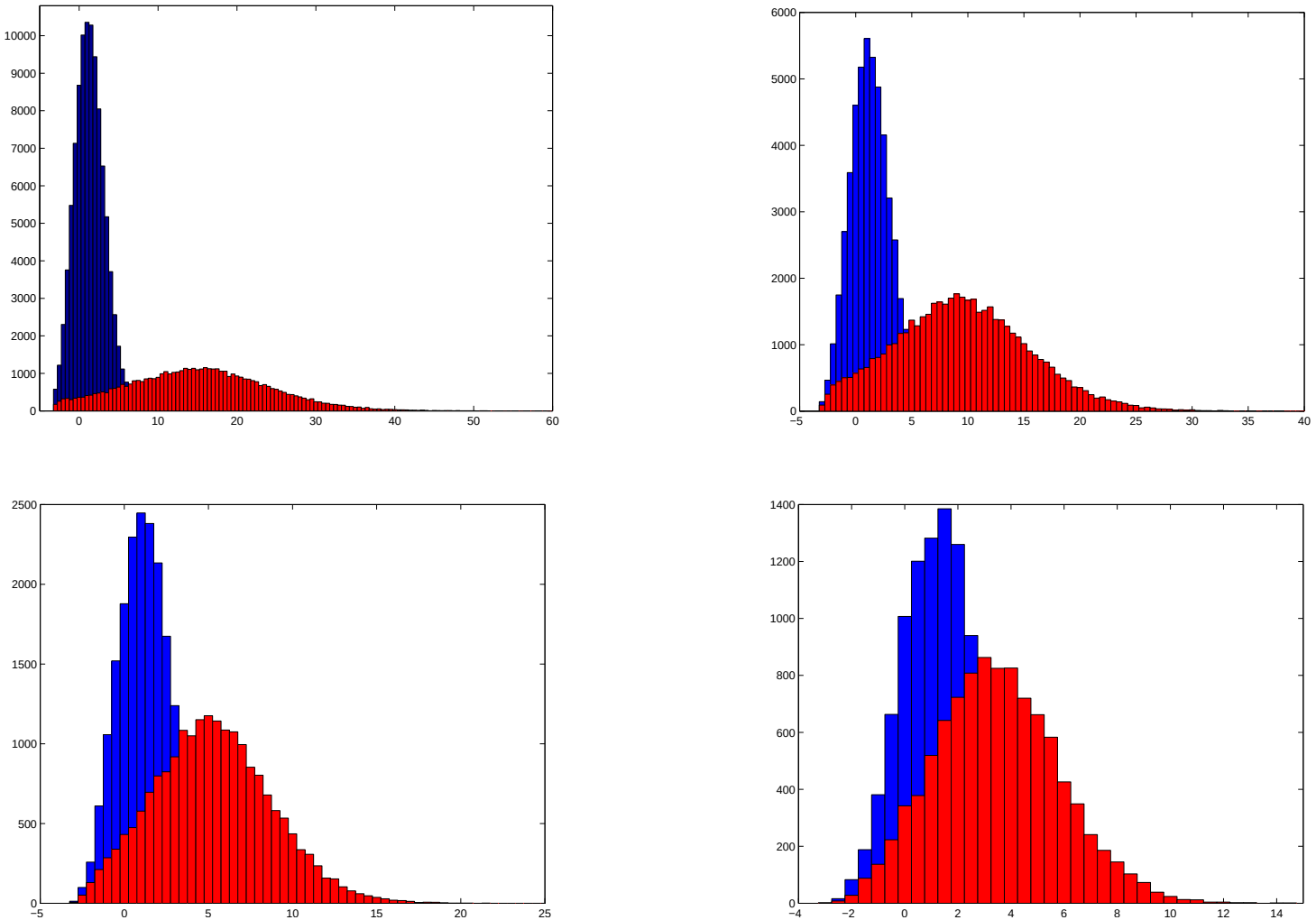
[b] M. Rousset, *SIAM J. Math. Anal.* (2006)

Initial distribution $(q_0^1, \dots, q_0^M) \sim d\mu_0(x)$, times $\tau_1^{k,b}, \tau_1^{k,d} \sim \mathcal{E}(1)$, initial jump times $T_0^{k,d} = 0, T_0^{k,b} = 0$.

- Between each jump time, evolve independently the replicas X_t^k according to the chosen dynamics;
- At times $T_{n+1}^{k,d}$ such that $\beta \int_{T_n^{k,d}}^{T_{n+1}^{k,d}} \left(\mathcal{F}_{\lambda(s)}^M - \frac{\partial H_{\lambda(s)}}{\partial \lambda}(X_s^k) \right)^- \lambda'(s)\, ds = \tau_{n+1}^{k,d}$: pick at random $l \in \{1, \dots, M\}$, replace k-th replica by the l-th. New time $\tau_{n+2}^{k,d} \sim \mathcal{E}(1)$ [death];
- At times $T_{n+1}^{k,b}$ such that $\beta \int_{T_n^{k,b}}^{T_{n+1}^{k,b}} \left(\mathcal{F}_{\lambda(s)}^M - \frac{\partial H_{\lambda(s)}}{\partial \lambda}(X_s^k) \right)^+ \lambda'(s)\, ds = \tau_{n+1}^{k,b}$: pick at random $l \in \{1, \dots, M\}$, replace l-th replica by the k-th. New time $\tau_{n+2}^{k,b} \sim \mathcal{E}(1)$ [birth].

In these expressions, $\mathcal{F}_{\lambda(t)}^M = \frac{1}{M} \sum_{k=1}^M \frac{\partial H_{\lambda(t)}}{\partial \lambda}(q_t^k)$ is the empirical average thermodynamical force

Application: Computation of chemical potentiel (Widom insertion)



Work histograms for increasing switching times ($T = 1, 2, 5, 10$).

Blue: IPS strategy. Red: usual nonequilibrium switching.

Nonequilibrium dynamics: the reaction coordinate case

Thermodynamic integration (1)

- Reaction coordinate $\xi(q)$, and free energy difference (configurational part)

$$\Delta F = -\beta^{-1} \ln \left(\frac{\displaystyle\int_{\mathcal{M}} \mathrm{e}^{-\beta V(q)}\, \delta_{\xi(q)-z_1}\, dq}{\displaystyle\int_{\mathcal{M}} \mathrm{e}^{-\beta V(q)}\, \delta_{\xi(q)-z_0}\, dq} \right) .$$

- Thermodynamic integration $\Delta F = \displaystyle\int_0^1 F'(z)\, dz$
- Mean force = canonical average on a submanifold
- Analytical expression not practical...

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

with $\widetilde{V} = V + \beta^{-1} \ln |\nabla \xi|$ and $d\mu_{\Sigma_z} = Z_z^{-1}\, \mathrm{e}^{-\beta V(q)}\, \delta_{\xi(q)-z}$

Thermodynamic integration (2)

- Projected SDE using a Lagrange multiplier formulation[a]

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

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)

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

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

- Mean force = average over Lagrange multipliers

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

[a] G. Ciccotti, T. Lelièvre and E. Vanden-Eijnden, *Comm. Pure Appl. Math.* (2007)

Nonequilibrium dynamics: The reaction coordinate case

- Extension to nonequilibrium switchings:[a] avoid steered MD
- Variation of the constraint $z(t)$ such that $z(0) = 0, z(T) = 1$
- Consider the dynamics

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

  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 $d\Lambda_t^{\mathrm{f}} = f(X_t)\, dt$,
  - forcing term (bias for finite time switching) $d\Lambda_t^{d} = \dfrac{1}{|\nabla \xi(X_t)|^2}\, z'(t)dt$

[a] T. Lelièvre, M. Rousset, and G. Stoltz, *J. Comput. Phys.* (2007)

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

$$q^n = q^{n-1} - \nabla \widetilde{V}(q^{n-1})\, \Delta t + \sqrt{\frac{2\Delta t}{\beta}}\, U^{n-1} + \lambda^n\, \nabla \xi(q^{n(-1)})$$

and computation of the work according to

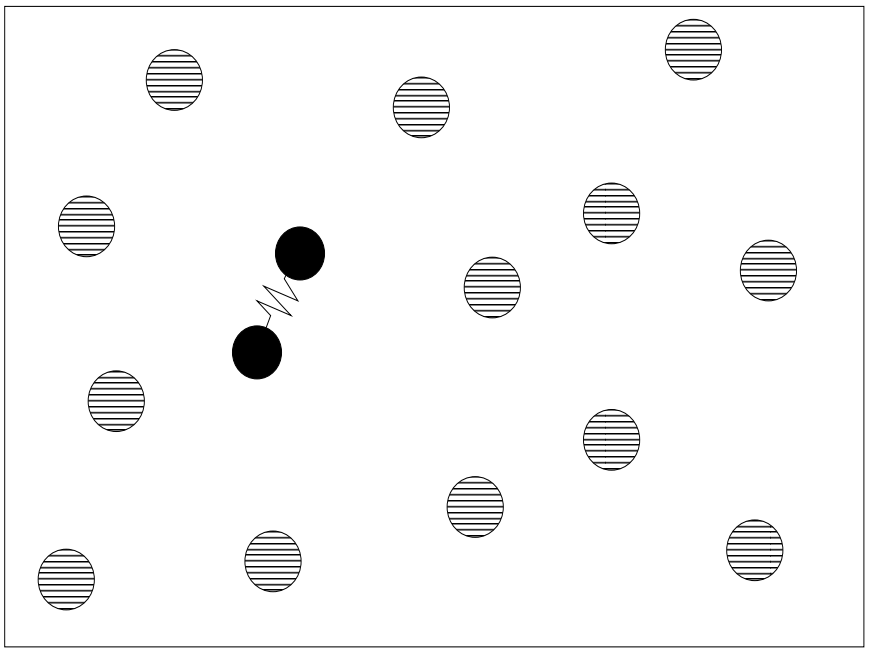
$$\mathcal{W}_n = \mathcal{W}_{n-1} + \frac{z(t_n) - z(t_{n-1})}{t_n - t_{n-1}}\, \lambda_n^{\mathrm{f}},$$

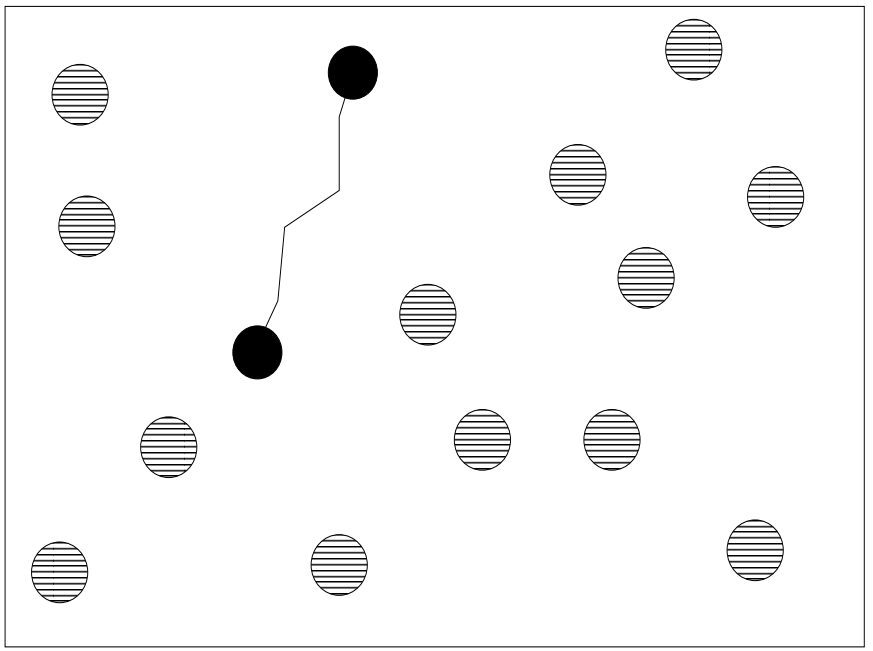
with

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

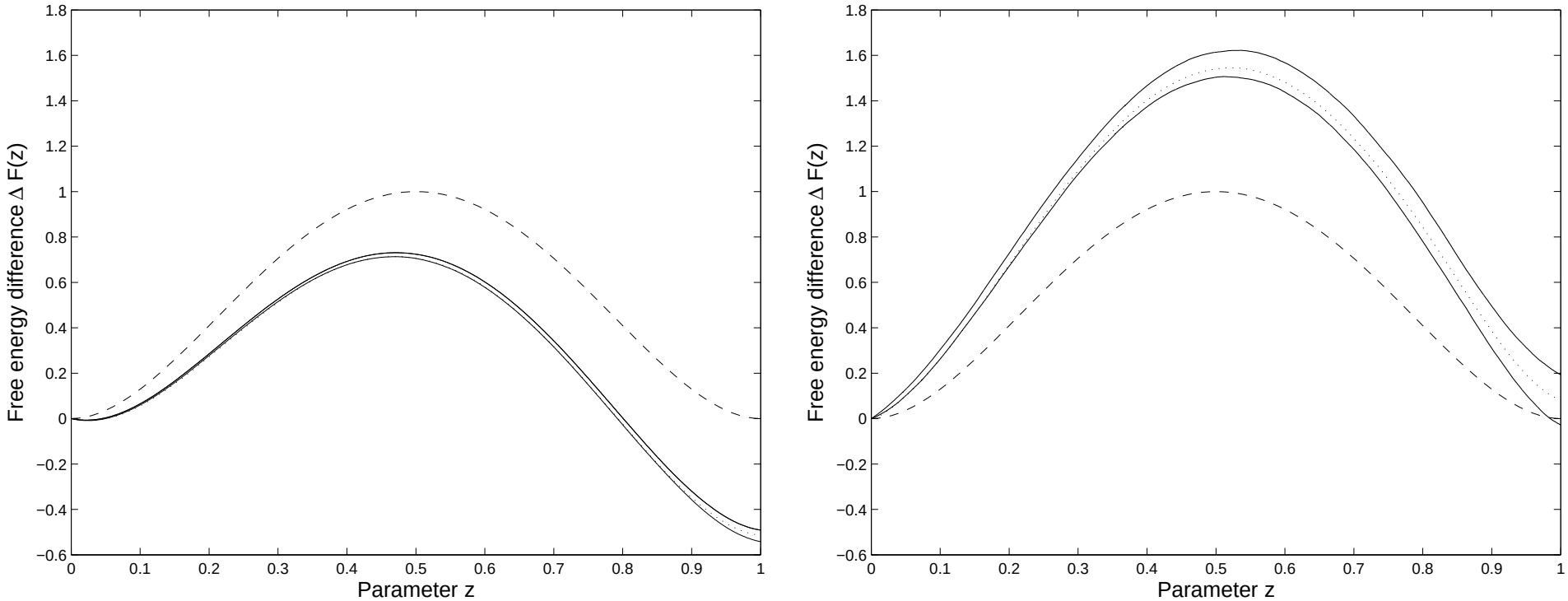
Application: Solvatation effects on conformational changes (1)

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





Application: Solvatation effects on conformational changes (2)



Comparison of free energy difference profiles using the reaction at low densities on the left, and high densities on the right.
Reference computed with TI (dashed line). Confidence interval obtained from 50 realizations of nonequilibrium process (solid lines).

Path sampling strategies

Nonequilibrium equality in the path space

- Path $x = ((x_0, \lambda_0) \ldots, (x_L, \lambda_L))$ ($x = q$ or $x = (q, p)$)
- Measure on path space $\pi(x, \lambda) = Z_L^{-1} \rho(x_0) \prod_{i=0}^{L-1} \mathrm{p}(x_i, x_{i+1}; \lambda_{i+1})$ so that

$$\mathrm{e}^{-\beta F} = \int \mathrm{e}^{-\beta W(x)} \, d\pi(x, \lambda)$$

- For overdamped Langevin dynamics with the usual Euler-Maruyama discretization, the conditional probability is

$$\mathrm{p}(q_i, q_{i+1}; \lambda_{i+1}) = \left(\frac{\beta}{4\pi\Delta t} \right)^{N/2} \exp\left(-\frac{\beta}{4\Delta t} \left| q_{i+1} - q_i + \nabla V_{\lambda_{i+1}}(q_i) \Delta t \right|^2 \right).$$

- Path sampling using a Metropolis-Hastings algorithm
- Importance sampling techniques can be used[a,b,c]

---

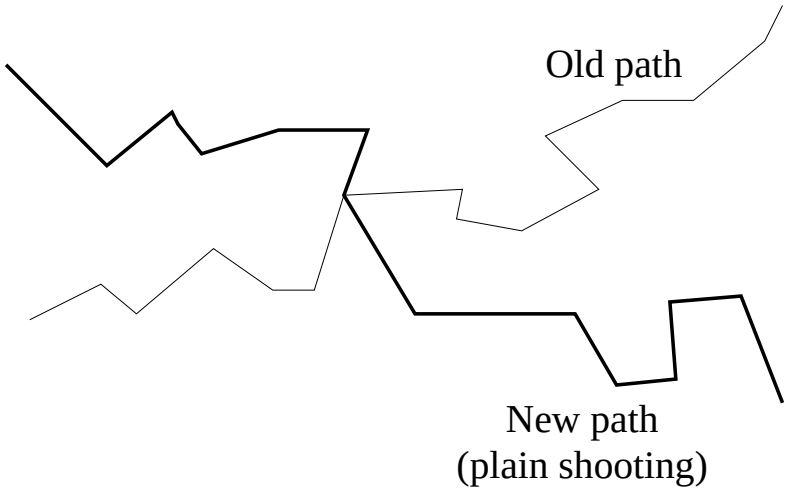
[a] Sun, *J. Chem. Phys.* (2003)

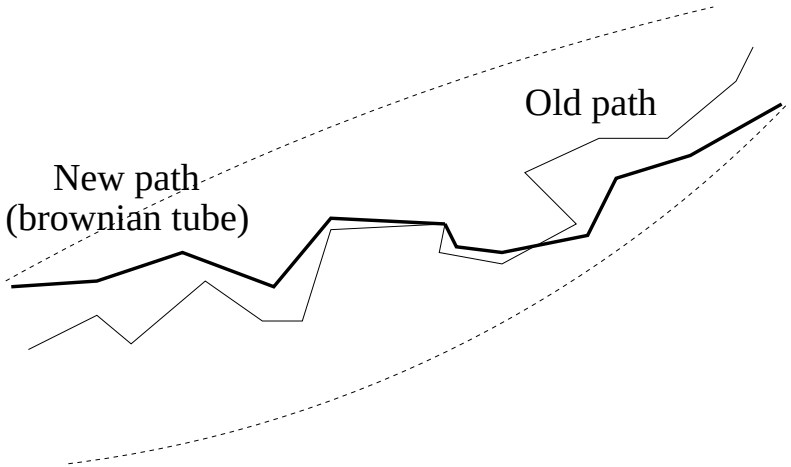
[b] Athènes, *Eur. Phys. J. B* (2004)

[c] Ytreberg and Zuckerman, *J. Chem. Phys.* (2004)

Sampling nonequilibrium paths

- Shooting algorithm vs. Brownian tube[a]





- From an old to a new trajectory, using $dq_t = -\nabla V(q_t)\,dt + \sigma\,dW_t$
  - One configuration fixed (q_{t_0} for $0 \leq t_0 \leq T$)
  - Choose a realization of the brownian correlated to the previous one:

$$\alpha(t) = \mathbb{E}(W_t^{\text{old}} W_t^{\text{new}})$$

  - Forward and backward integrations

[a] G. Stoltz, *J. Comput. Phys.* (2007)

Sampling nonequilibrium paths (2)

- Usual shooting algorithm: $\alpha(t) = 0$. This amounts in practice to redrawing new uncorrelated random increments.
- Noise history algorithm:[a] $\alpha(t) = \begin{cases} 0 & \text{if } t_0 \leq t \leq t_0 + \Delta t, \\ 1 & \text{otherwise.} \end{cases}$

  In practice, all the random numbers used to generate the trajectories are kept but one.
- Brownian tube: general $0 \leq \alpha(t) \leq 1$. Redraw new random numbers correlated to the previous ones.
- Interest:
  - Generalizes both approaches
  - Allows to obtain arbitrary acceptance rates
  - Tested for path sampling, what about free energy computations?

[a] Crooks and Chandler, *Phys. Rev. E* (2001)

Sampling nonequilibrium paths (3)

- Shooting index k chosen, decomposition of the path as a forward part $((q_k, \lambda_k), \ldots, (q_L, \lambda_L))$ and a backward part $((q_k, \lambda_k), \ldots, (q_0, \lambda_0))$.
- Associated noises (U_j forward part, $\overline{U}_j$ backward part):

$$U_j = \frac{1}{\sqrt{2\beta^{-1}}} \left(q_{j+1} - q_j + \nabla V(q_j, \lambda_{j+1})\Delta t\right) \quad (k \leq j \leq L-1),$$

$$\overline{U}_j = \frac{1}{\sqrt{2\beta^{-1}}} \left(q_j - q_{j+1} + \nabla V(q_{j+1}, \lambda_{j+1})\Delta t\right) \quad (0 \leq j \leq k-1).$$

- Modify as $U'_j = \alpha_j U_j + \sqrt{1-\alpha_j^2} G_j$, $\overline{U}'_j = \alpha_j \overline{U}_j + \sqrt{1-\alpha_j^2}\overline{G}_j$
- Integrate backward using $\{\overline{U}'_j\}$, and forward using $\{U'_j\}$
- Probability to generate $(q', \lambda) = ((q'_0, \lambda_0), \ldots, (q'_L, \lambda_L))$ from (q, λ)

$$\mathcal{P}_{\rm gen}((q,\lambda),(q',\lambda)) = w_k \prod_{0 \leq j \leq k-1} \mathrm{p}_{\alpha_j}(\overline{U}_j, \overline{U}'_j) \prod_{k \leq j \leq L-1} \mathrm{p}_{\alpha_j}(U_j, U'_j),$$