Equilibrium computation of free-energy differences using non-equilibrium methods

Mathias ROUSSET[1], Gabriel STOLTZ[2,3]

[1] LSP, Université Paul Sabatier (Toulouse, France)

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

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

http://cermics.enpc.fr/∼stoltz/

References

- Strategy inspired by the Sequential Monte Carlo philosophy
- M. Rousset and G. Stoltz, *An interacting particle approach for molecular dynamics*, submitted to *M3AS* (2005)
- Presentation and preprints available at the URL `http://cermics.enpc.fr/~stoltz/`

Jarzynski’s equality: a simulated annealing strategy

References

- Jarzynski's seminal papers:
  - Nonequilibrium equality for free-energy differences, *Phys. Rev. Lett.*, 78 (14) (1997) 2690-2693.
  - Equilibrium free energy differences from nonequilibrium measurements: a master equation approach, *Phys. Rev. E* **56**(5) (1997) 5018–5035.
- A recent re-derivation (Jarzynski, *J. Stat. Mech.*, 2004)
- Improvements
  - Jarzynski's formula as a special case of a fluctuations/dissipation relation (Crooks, *PRE*, 1999 and *PRE*, 2000).
  - Path sampling method (Sun, *J. Chem. Phys.*, 2003)
- Pedagogical example: the ideal gas case (Lua/Grosberg, *J. Phys. Chem. B*, 2005)
- Many studies comparing different free-energy methods, the method is still a matter of debate among physicists!

The simulated annealing setting

- Family of Hamiltonian functions H_λ, parameter $\lambda \in [0,1]$ (change of temperature, 'alchemical transition')
- Irreducible Markovian dynamics $t \mapsto X_t^{\lambda(t)}, \quad X_0^{\lambda(0)} \sim \mu_0$
- Smooth schedule $t \mapsto \lambda(t)$ ($\lambda(1) = 0$ and $\lambda(T) = 1$)
- Canonical measure $d\mu_\lambda$ invariant
- Examples
  - Hypo-elliptic Langevin dynamics
  - Overdamped Langevin dynamics
  - Hamiltonian case (NVE) can also be considered (non irreducible)

Jarzynski's equality

- Differentiation of the *non-normalized* Boltzmann path $t \mapsto \mathrm{e}^{-\beta H_{\lambda(t)}(x)}\, dx$

$$\partial_t \Pi_{\lambda(t)}(f) = \Pi_{\lambda(t)}\left(L_{\lambda(t)}(f) - \beta \frac{\partial H_{\lambda(t)}}{\partial \lambda} \lambda'(t) f \right).$$

- Free-energy $F(\lambda) = -\beta^{-1} \ln Z_\lambda$ (partition function Z)
- Feynamn-Kac formula (with $f = 1$):

$$\mathbb{E}(\mathrm{e}^{-\beta W_t}) = \mathrm{e}^{-\beta(F(\lambda(t)) - F(0))}$$

- Work defined as $W_t = \int_0^t \frac{\partial H_{\lambda(s)}}{\partial \lambda}(X_s^{\lambda(s)}) \lambda'(s)\, ds$
- Jensen's inequality $\mathbb{E}(W_t) \geq F(\lambda(t)) - F(0) \rightarrow$ use the virtual work W_t to perform a selection between replicas

Re-interpreting the dynamics as a jump-diffusion process

Derivation of the jump-diffusion process

- Differentiation of the *normalized* Boltzmann (with $\mathcal{F}_\lambda = \mu_\lambda\left(\frac{\partial H_\lambda}{\partial \lambda}\right)$)

$$\partial_t \mu_{\lambda(t)}(f) = \mu_{\lambda(t)}\left(L_{\lambda(t)}(f) + \beta\left(\mathcal{F}_{\lambda(t)} - \frac{\partial H_{\lambda(t)}}{\partial \lambda}\right)\lambda'(t) f\right),$$

- Rewritten as a nonlinear Markovian evolution

$$\partial_t \mu_{\lambda(t)}(f) = \mu_{\lambda(t)}\left(L_{\lambda(t)}(f) + J_t(f)\right)$$

- Jump generator

$$J_t(f)(x) = \int_{\mathcal{M}} (f(y) - f(x))(\alpha_t^-(x) + \alpha_t^+(y))\mu_{\lambda(t)}(dy)$$

with transition intensities
$\alpha_t^- = \beta\lambda'(t)(\mathcal{F}_{\lambda(t)} - \frac{\partial H_{\lambda(t)}}{\partial \lambda})^-,\ \alpha_t^+ = \beta\lambda'(t)(\mathcal{F}_{\lambda(t)} - \frac{\partial H_{\lambda(t)}}{\partial \lambda})^+$

The jump-diffusion process

- Jump-diffusion process

  Process 1 $Y_0 \sim d\mu_0(x)$, $(\tau_n^b, \tau_n^d)_{n\geq 1} \sim \mathcal{E}(1)$, $T_0^d = 0, T_0^b = 0$.

  - *Between each jump time, $t \mapsto Y_t$ evolves according to the chosen dynamics;*
  - *At random times T_{n+1}^d defined by $\int_{T_n^d}^{T_{n+1}^d} \alpha_s^-(Y_s)ds = \tau_{n+1}^d$, the process jumps to a configuration y, chosen according to the probability measure $d\mu_{\lambda(T_{n+1}^d)}(y)$* [death]*;*
  - *At random times T_{n+1}^b defined by $\int_{T_n^b}^{T_{n+1}^b} \mu_{\lambda(s)}(\alpha_s^+)ds = \tau_{n+1}^b$, the process jumps to a configuration y, chosen according to the probability measure $\sim \alpha^+_{T_{n+1}^b}(y) d\mu_{\lambda(T_{n+1}^b)}(y)$* [birth]*.*

- For all $t \geq 0$, the law of Y_t is $\mu_{\lambda(t)}$.

Discretizing the process using an Interacting Particle System approach

Discretizing the process

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

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

and empirical Boltzmann distribution

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

- Selection mechanism favors replicas sampling lower virtual works $\rightarrow$ 'self-organization' to keep closer to a quasi-static transformation
- Convergence properties as $M \rightarrow \infty$

The IPS algorithm

Process 2 *Initial distribution* $(X_0^1, \dots, X_0^M) \sim d\mu_0(x)$, *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 random times* $T_{n+1}^{k,d}$ *defined by*

$$\beta \int_{T_n^{k,d}}^{T_{n+1}^{k,d}} \left(\mathcal{F}_{\lambda(s)}^M - \frac{\partial H_{\lambda(s)}}{\partial \lambda}(X_s^k) \right)^- \lambda'(s)\, ds = \tau_{n+1}^{k,d},$$

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 random times* $T_{n+1}^{k,b}$ *defined by*

$$\beta \int_{T_n^{k,b}}^{T_{n+1}^{k,b}} \left(\mathcal{F}_{\lambda(s)}^M - \frac{\partial H_{\lambda(s)}}{\partial \lambda}(X_s^k) \right)^+ \lambda'(s)\, ds = \tau_{n+1}^{k,b},$$

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].

Some numerical results

Computable quantities and estimators

- IPS generates canonical distributions for all $\lambda(t)$, and allows computation of all free-energy differences in this range
- Jarzynski's estimators (rely on the tails of the work distributions)

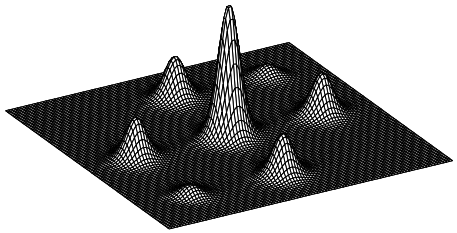
$$\Delta\hat{F}_J = -\frac{1}{\beta}\ln\left(\frac{1}{M}\sum_{k=1}^{M} e^{-\beta W_1^k}\right), \quad \Delta\hat{F}'_J = \int_0^T \mathcal{F}^{M_{\rm ind}}_{\lambda(t)}\lambda'(t)\,dt,$$

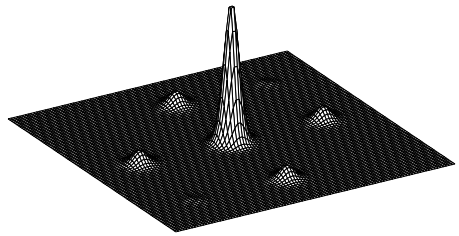
where $\mathcal{F}^{M_{\rm ind}}_{\lambda(t)} = \mu^{M_{\rm ind}}_{\lambda(t)}(\frac{\partial H_{\lambda(t)}}{\partial\lambda})$ with $\mu^{M_{\rm ind}}_{\lambda(t)}(dx) = \frac{\sum_{k=1}^{M}\delta_{X_t^k}(dx)\,\mathrm{e}^{-\beta W_t^i}}{\sum_{k=1}^{M}\mathrm{e}^{-\beta W_t^i}}$.

- IPS estimator ($\mathrm{e}^{-\beta\Delta\hat{F}_{\rm IPS}}$ unbiased estimator of $\mathrm{e}^{-\beta\Delta F}$, $\Delta\hat{F}_{\rm IPS}$ asymptotically normal with bias and variance of order M^{-1})

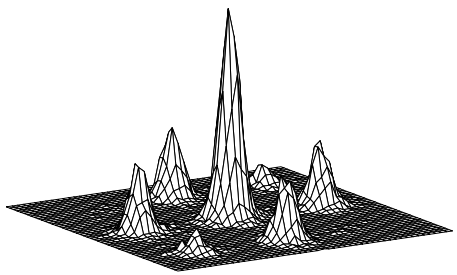
$$\Delta\hat{F}_{\rm IPS} = \int_0^T \mathcal{F}^{M}_{\lambda(t)}\lambda'(t)\,dt.$$

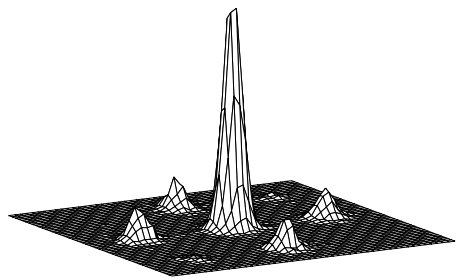
Computation of canonical distributions





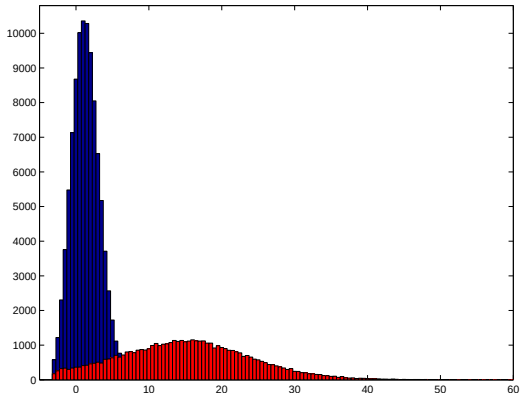
Empirical probability distribution of the dihedral angles (ϕ_1, ϕ_2) of the pentane molecule, for $T = 300$ K (Left) and $T' = 150$ K (Right) [Importance sampling, $M = 10^9$ points].

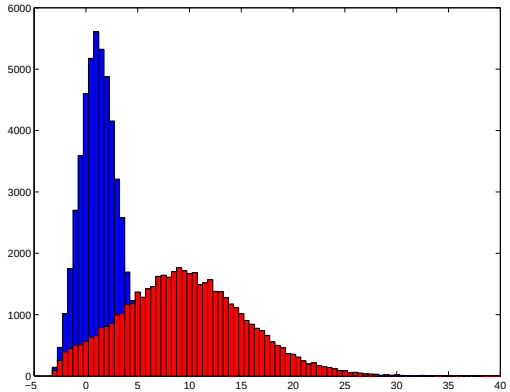




Empirical probability distributions at $T = 150$ K, generated with simulated annealing (Left), and IPS (Right), starting from $T = 300$ K, with sample size $M = 10,000$.

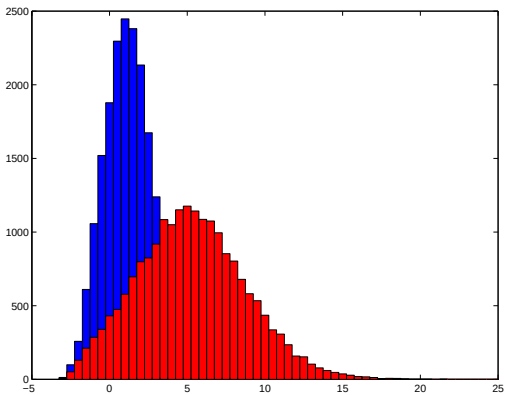
Computing free-energy differences: Widom insertion (Langevin dynamics)

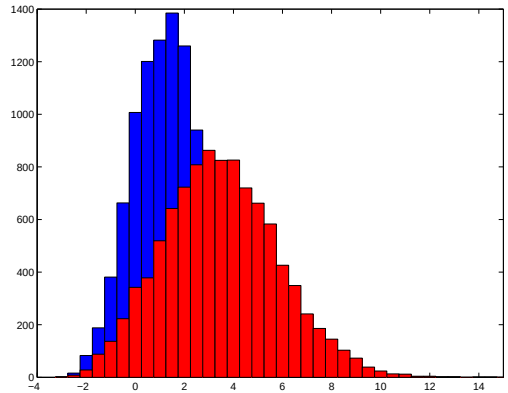




Left: $\mu_{\mathrm{IPS}} = 1.37$ (blue), $\mu_{\mathrm{SA}} = 1.32$ (red), $\tau = 1$. Right: $\mu_{\mathrm{IPS}} = 1.35, \mu_{\mathrm{SA}} = 1.29$ ($\tau = 2$).

Reference (Zwanzig, 10^8 points): $\mu_{\mathrm{ex}} = 1.31$.





Left: $\mu_{\mathrm{IPS}} = 1.29, \mu_{\mathrm{SA}} = 1.33$ ($\tau = 5$). Right: $\mu_{\mathrm{IPS}} = 1.34, \mu_{\mathrm{SA}} = 1.36$ ($\tau = 10$).