Tackling long timescales in molecular simulation

Carsten Hartmann (FU Berlin, Matheon)

jointly with Christof Schütte, Ralf Banisch, Han Wang and Wei Zhang

SCICADE 2013, Valladolid, 16–20 September 2013

Take-home message

We can design **optimal nonequilibrium MD simulations that mimic full equilibrium statistics** of any given observable at low numerical cost (zero-variance estimators, shorter trajectories).

Outline

Molecular conformation dynamics

$1.3\mu s$ MD simulation of green tea at room temperature (visualization: Amira@ZIB).

Rare events in molecular simulation

Given a Markov process $(X_t)_{t\geq 0}$, discrete or continuous in time, we want to estimate probabilities $p \ll 1$, such as

$$p = P\left(\tau < T\right),$$

with τ the time to reach the target conformation, or rates

$$k = \frac{1}{\mathbb{E}[\tau]}\,.$$

Guiding example: Bistable system

- Overdamped Langevin equation

$$dX_t = -\nabla V(X_t)dt + \sqrt{2\epsilon}dB_t$$

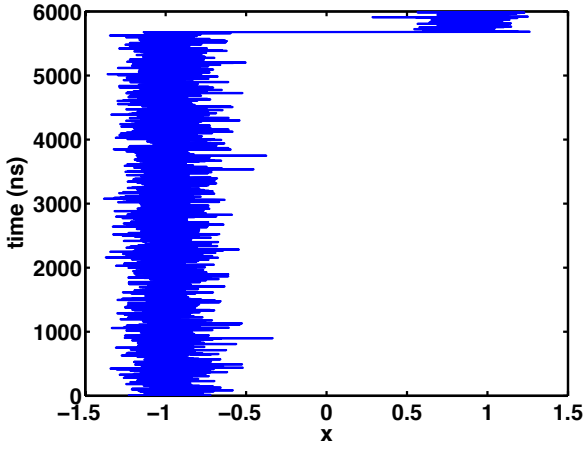
- **Exit time** asymptotics as $\epsilon \to 0$

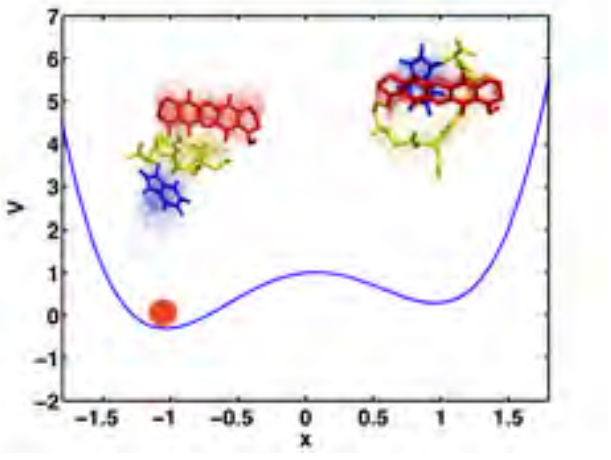
$$\mathbb{E}[\tau] \asymp e^{\Delta V/\epsilon}$$

- Hence, for moderate values of T,

$$P(\tau < T)$$

is **exponentially small** in ϵ.





Rare events in molecular simulation, cont'd

Given N **independent realizations** of $X_t = X_t(\omega)$, the simplest way to estimate probabilities, such as

$$p = P(\tau < T)$$

is by **Monte-Carlo**:

$$p \approx \frac{1}{N} \sum_{i=1}^{N} \mathbf{1}_{\{\tau(\omega_i) < T\}}$$

Probability of rare events, cont'd

Although the naïve MC estimator is unbiased with bounded variance $p(1-p)/N$, the **relative error** is not:

$$\delta_{\text{rel}} = \frac{\text{standard deviation}}{\text{mean}} = \frac{1}{p}\sqrt{\frac{p(1-p)}{N}}$$

blows up as $p \to 0$.

This is a **common feature when estimating rare events**.

Rare events in molecular dynamics

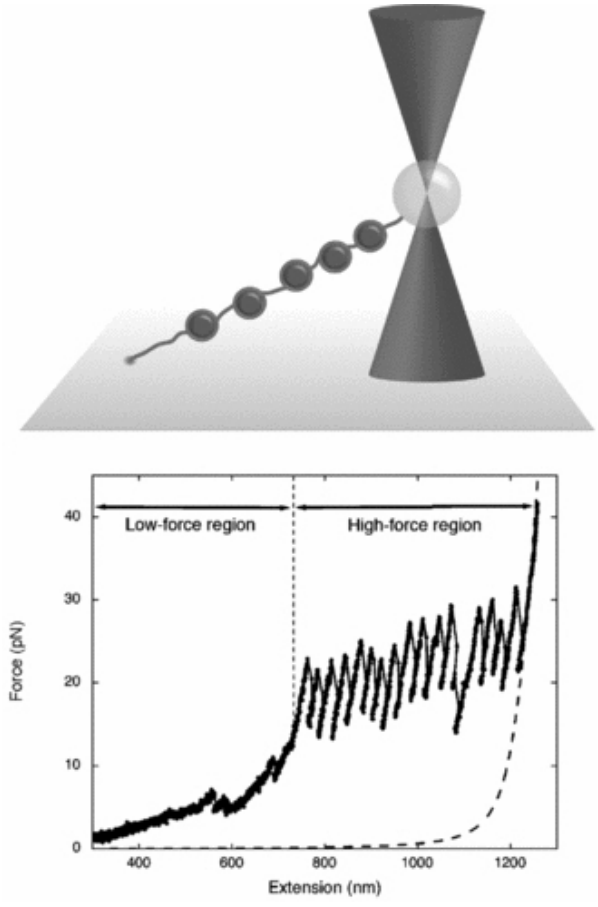
Reweighting I

Reweighting II

Algorithmic aspects and numerical examples

Conclusions and open problems

Nonequilibrium molecular dynamics



from: [Dame, Biochem Soc Trans, 2008]

- **Enhanced sampling** of rare barrier crossing events by applying a force u: umbrella sampling, TMD, SMD, ABF, ...
- **Forcing messes up the statistics**; can we force the system in a gentle way?
- Mathematically: **change of measure** from P to $Q = Q(u)$ with likelihood ratio

$$\varphi = \frac{dQ}{dP}\,.$$

[Schlitter et al, Mol Sim, 1993], [Darve & Pohorille, JCP, 2001], [Schulten & Park, JCP, 2004], ...

Guiding example

- **Mean first exit time** for small ϵ

$$\mathbb{E}[\tau] \asymp \exp(\Delta V/\epsilon)$$

- **Tilting the potential**

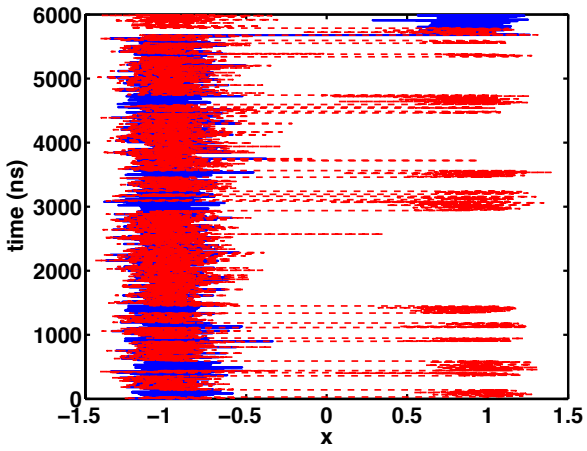
$$U(x,t) = V(x) - u_t x$$

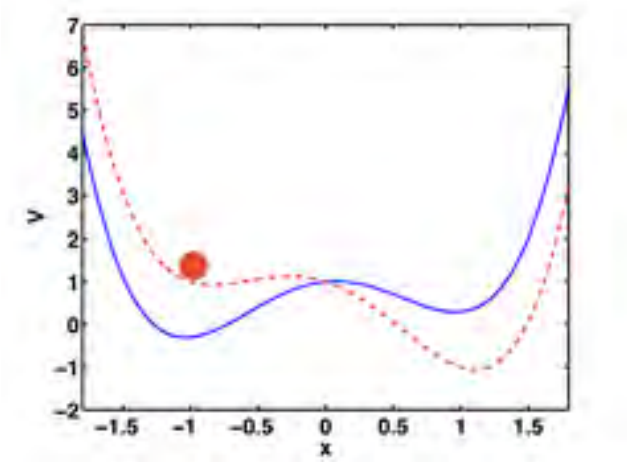
decreases the energy barrier.

- Overdamped Langevin equation

$$dX_t = (u_t - \nabla V(X_t))\, dt + \sqrt{2\epsilon} dB_t$$

with **time-dependent forcing** u_t.





Importance sampling identity I: change of measure

Let $Q \ll P$ be two probability measures with likelihood ratio

$$\varphi = \frac{dQ}{dP} > 0\,.$$

Then

$$\mathbb{E}[\tau] = \int_\Omega \tau(\omega) dP(\omega) = \int_\Omega \tau(\omega)\varphi^{-1}(\omega) dQ(\omega) =: \mathbb{E}_Q[\tau\varphi^{-1}]$$

Guiding example, cont'd

- **Unbiased MC estimator** for $\mathbb{E}[\tau]$

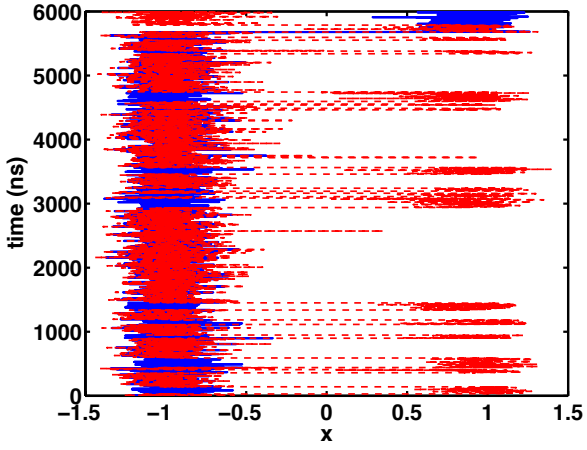
$$\mathbb{E}[\tau] \approx \frac{1}{N} \sum_{i=1}^{N} \tau_u(\omega_i) \varphi^{-1}(\omega_i) \,,$$

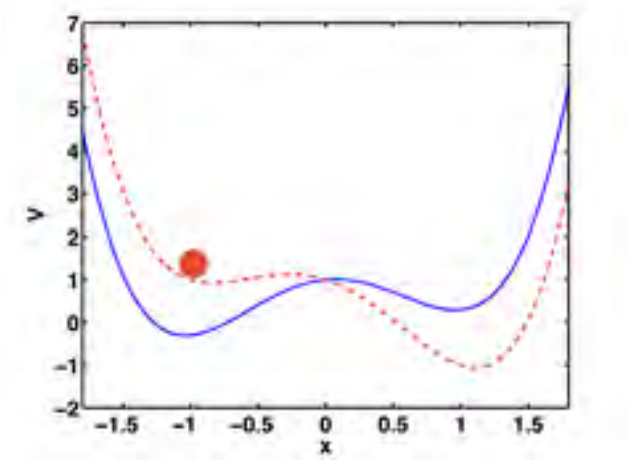
where

$$\varphi = e^{-\xi(u)/\epsilon}$$

is explicitly known.

- A naïve choice of u will **increase the variance**, with much lower efficiency than that of the vanilla MC estimator.





Rare events in molecular dynamics

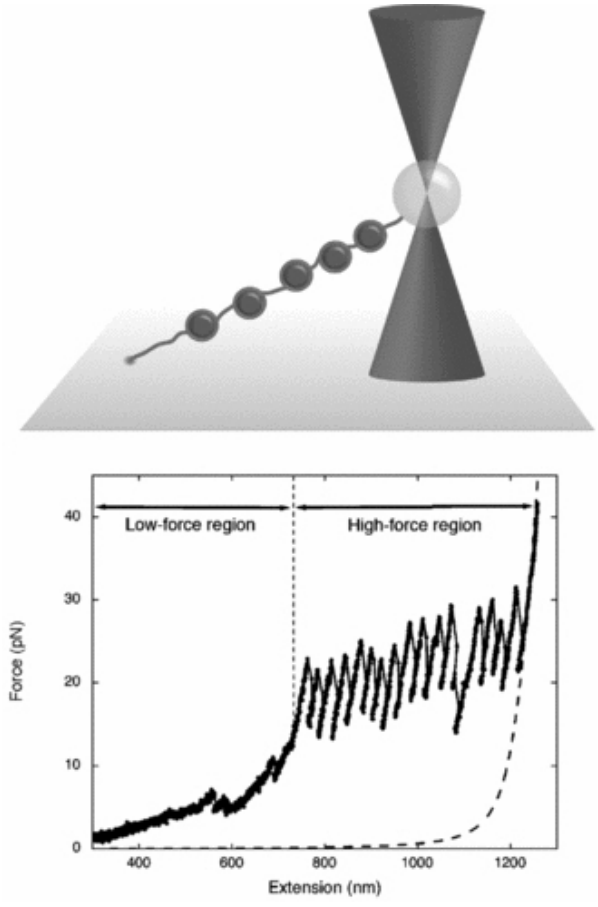
Reweighting I

Reweighting II

Algorithmic aspects and numerical examples

Conclusions and open problems

Nonequilibrium molecular dynamics, cont'd



from: [Dame, Biochem Soc Trans, 2008]

- **Aim:** extract equilibrium properties from nonequilibrium simulations.
- Calculation of **equilibrium free energies** á la Jarzynski and Crooks:

$$\Delta F = -\epsilon \log \mathbb{E}_Q[\exp(-W/\epsilon)]\,,$$

with W the work exerted under u.

- Note the **second law** like inequality

$$\Delta F \leq \mathbb{E}_Q[W]$$

[Jarzynski, PRL, 1997], [Crooks, J Stat Phys, 1998]

Nonequilibrium molecular dynamics: some observations

The free energy in Jarzynski's formula has the form of a **cumulant generating function** of the random variable W:

$$-\epsilon \log \mathbb{E}_Q[\exp(-W/\epsilon)] \approx \mathbb{E}_Q[W] - \frac{1}{2\epsilon}\mathbb{E}_Q[(W - \mathbb{E}_Q[W])^2]$$

A cumulant generating functions encodes information about moments, provided they exist. Moreover, to lowest order in $1/\epsilon$, **mean and variance are approximately decoupled**.

Nonequilibrium molecular dynamics: some observations

The free energy in Jarzynski's formula has the form of a **cumulant generating function** of the random variable W:

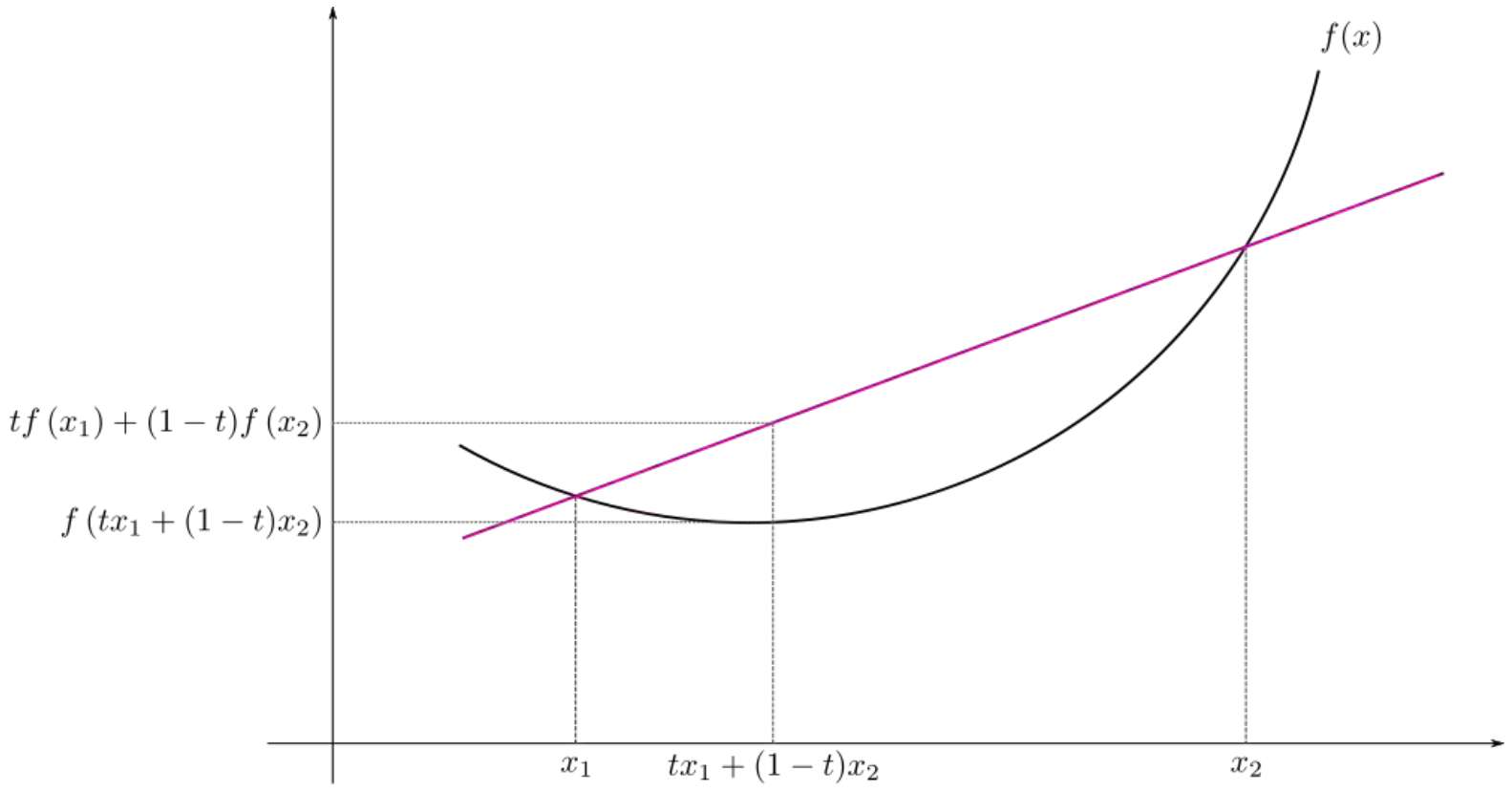
$$-\epsilon \log \mathbb{E}_Q[\exp(-W/\epsilon)] \approx \mathbb{E}_Q[W] - \frac{1}{2\epsilon}\mathbb{E}_Q[(W - \mathbb{E}_Q[W])^2]$$

A cumulant generating functions encodes information about moments, provided they exist. Moreover, to lowest order in $1/\epsilon$, **mean and variance are approximately decoupled**.

So, let's go back to our rare event sampling problem. . .

Importance sampling identity II: Jensen's inequality

$\mathbb{E}[f(X)] \geq f(\mathbb{E}[X])$ when f is a convex function, e.g. $f(x) = e^x$.



Free energy as the CGF of a random variable

Define the **generalized free energy of the random variable** Z with respect to the equilibrium distribution P as

$$F_Z(\epsilon) = -\epsilon \log \mathbb{E}[\exp(-Z/\epsilon)]\,.$$

The free energy now describes the **full equilibrium statistics** of the random variable Z, e.g., for the exit time $Z = \tau$:

$$F_\tau(\infty) = \mathbb{E}[\tau]\,, \quad F'_\tau(\infty) = \mathbb{E}[(\tau - \mathbb{E}[\tau])^2]\,, \ldots$$

Free energy as the CGF of a random variable

Define the **generalized free energy of the random variable** Z with respect to the equilibrium distribution P as

$$F_Z(\epsilon) = -\epsilon \log \mathbb{E}[\exp(-Z/\epsilon)] \,.$$

The free energy now describes the **full equilibrium statistics** of the random variable Z, e.g., for the exit time $Z = \tau$:

$$F_\tau(\infty) = \mathbb{E}[\tau] \,, \quad F'_\tau(\infty) = \mathbb{E}[(\tau - \mathbb{E}[\tau])^2] \,, \ldots$$

Importance sampling identities I & II

Change of measure & Jensen's inequality entail

$$F_Z(\epsilon) \leq \mathbb{E}_Q \left[Z + \epsilon \log \varphi \right] .$$

with **equality iff** $Z + \epsilon \log \varphi$ **is a.s. constant**.

In thermodynamic language, the inequality reads

$$F_Z(\epsilon) = \min_{Q \ll P} \{ \mathbb{E}_Q[Z] + \epsilon H(Q \| P) \}$$

with $H(Q\|P)$ the **relative entropy** between Q and P.

(The minimizer exists and is unique, but its normalization constant is the quantity that we want to compute.)

[Dai Pra et al., Math Control Signals Systems, 1996]; cf. [Vaikuntanathan & Jarzynski, PRL, 2008]

Importance sampling identities I & II

Change of measure & Jensen's inequality entail

$$F_Z(\epsilon) \leq \mathbb{E}_Q \left[Z + \epsilon \log \varphi \right] .$$

with **equality iff** $Z + \epsilon \log \varphi$ **is a.s. constant**.

In thermodynamic language, the inequality reads

$$F_Z(\epsilon) = \min_{Q \ll P} \{ \mathbb{E}_Q[Z] + \epsilon H(Q \| P) \}$$

with $H(Q\|P)$ the **relative entropy** between Q and P.

(The minimizer exists and is unique, but its normalization constant is the quantity that we want to compute.)

[Dai Pra et al., Math Control Signals Systems, 1996]; cf. [Vaikuntanathan & Jarzynski, PRL, 2008]

We have turned a difficult sampling problem into a potentially more difficult minimization problem. Now let us turn it into something more familiar and more useful. . .

Nonequilibrium diffusions

Assume **overdamped Langevin dynamics** (generating P)

$$dX_t = -\nabla V(X_t)dt + \sqrt{2\epsilon}\, dB_t\,, \quad X_0 = x$$

The requirement $Q \ll P$ **leaves only one choice** for the nonequilibrium dynamics generating Q, namely

$$dX_t^u = (u_t - \nabla V(X_t^u))dt + \sqrt{2\epsilon}\, dB_t\,, \quad X_0^u = x\,.$$

The control force u must be chosen so as to **minimize the free energy functional** (i.e. the right hand side of the inequality).

[Fleming, SIAM J Control, 1978], [Dupuis & Wang, Ann Appl Probab, 2005], [H. & Schütte, JSTAT 2012]

Nonequilibrium diffusions

Assume **overdamped Langevin dynamics** (generating P)

$$dX_t = -\nabla V(X_t)dt + \sqrt{2\epsilon}\, dB_t\,, \quad X_0 = x$$

The requirement $Q \ll P$ **leaves only one choice** for the nonequilibrium dynamics generating Q, namely

$$dX_t^u = (u_t - \nabla V(X_t^u))dt + \sqrt{2\epsilon}\, dB_t\,, \quad X_0^u = x\,.$$

The control force u must be chosen so as to **minimize the free energy functional** (i.e. the right hand side of the inequality).

[Fleming, SIAM J Control, 1978], [Dupuis & Wang, Ann Appl Probab, 2005], [H. & Schütte, JSTAT 2012]

Guiding example, cont'd

- Free energy of the **first exit time**

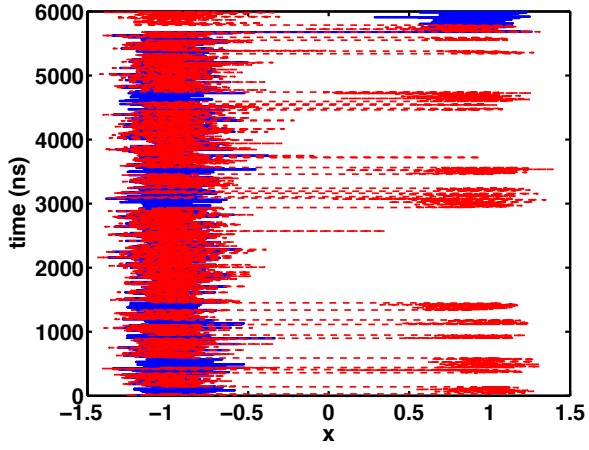
$$F_\tau(x;\epsilon) = \min_u \mathbb{E}_Q \left[\tau_u + \frac{1}{2} \int_0^{\tau_u} |u_t|^2 dt \right]$$

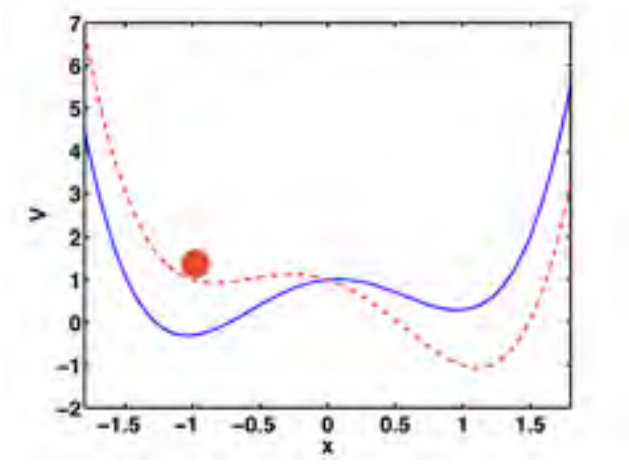
- The **optimal control is Markovian** and of gradient form:

$$u_t^* = -2\nabla F_\tau(X_t^{u^*};\epsilon)\,,$$

- **Optimally tilted potential**

$$U = V + 2F_\tau\,.$$





Some remarks

Control Penalization: The relative entropy minimization puts a quadratic penalty on the control force.

NFL Theorem: The optimal bias is the gradient of the quantity that we want to compute:

$$u_t^* = -2\nabla F_Z(X_t^{u^*}; \epsilon)\,.$$

(System is again in equilibrium—detailed balance holds.)

[H. & Schütte, JSTAT 2012], [Wang, H. & Schütte, Mol. Phys., 2013], [H. et al., submitted to Entropy, 2013]

Some remarks

Control Penalization: The relative entropy minimization puts a quadratic penalty on the control force.

NFL Theorem: The optimal bias is the gradient of the quantity that we want to compute:

$$u_t^* = -2\nabla F_Z(X_t^{u^*}; \epsilon)\,.$$

(System is again in equilibrium—detailed balance holds.)

[H. & Schütte, JSTAT 2012], [Wang, H. & Schütte, Mol. Phys., 2013], [H. et al., submitted to Entropy, 2013]

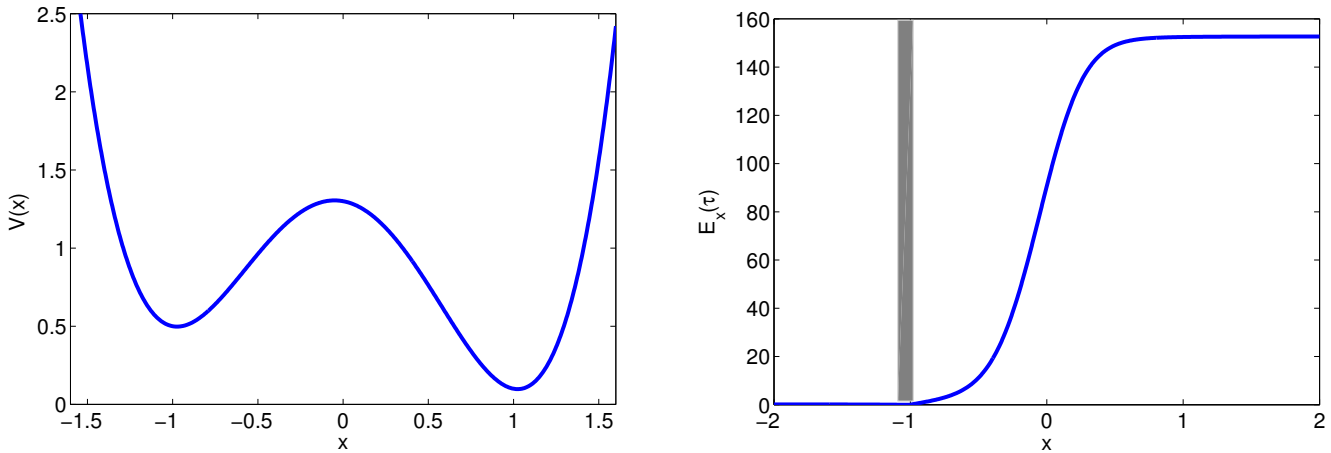
Numerical example I

Fast passage to a terminal set

$$-\log \mathbb{E}[\exp(-\sigma\tau)] = \min_u \mathbb{E}_Q \left[\sigma\tau^u + \frac{1}{2}\int_0^{\tau^u} |u_t|^2 \, dt \right]$$

where the minimization is subject to the **tilted dynamics**

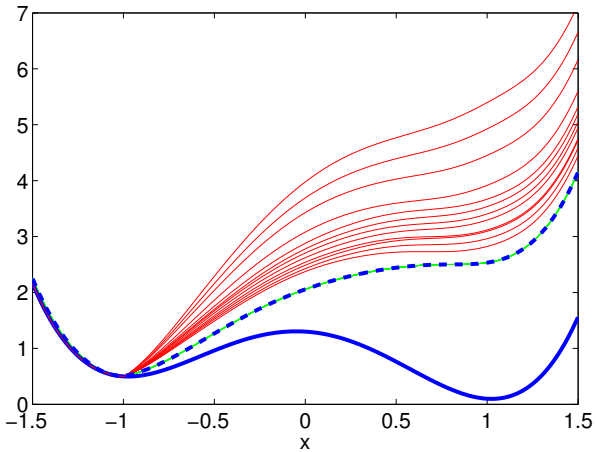
$$dX_t^u = (u_t - \nabla V(X_t^u))\, dt + \sqrt{2\epsilon} dB_t\,, \quad X_0^u = x\,.$$

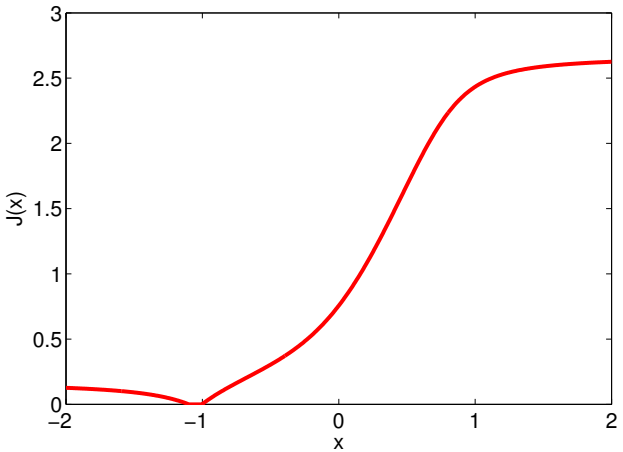


Skew double-well potential V and MFPT of the set $S = [-1.1, -1]$ for $\epsilon = 0.25$ (FEM reference solution).

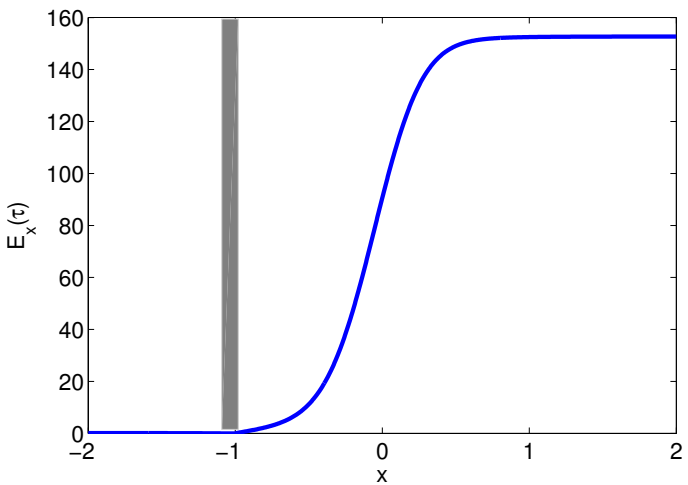
[H. & Schütte, JSTAT 2012]

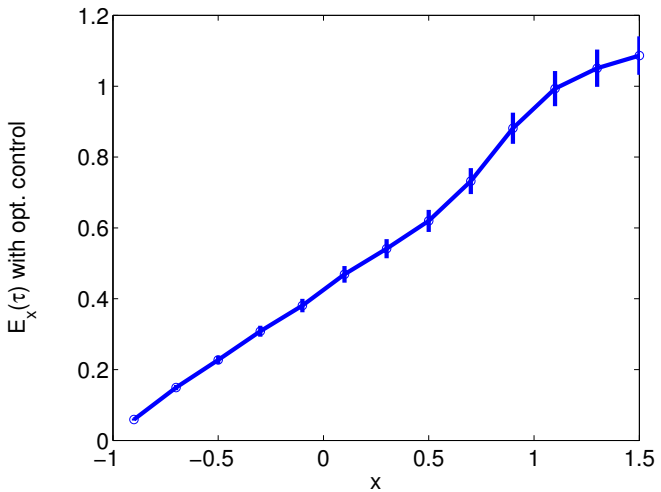
Fast passage to a terminal set, cont'd





ABF reversed: first 11 iterates of the EM algorithm using 10 Gaussians and limiting free energy.





MFPT (unbiased estimate), based on 2000 realizations of the tilted dynamics. Total CPU speed-up is $\sim$ 100.

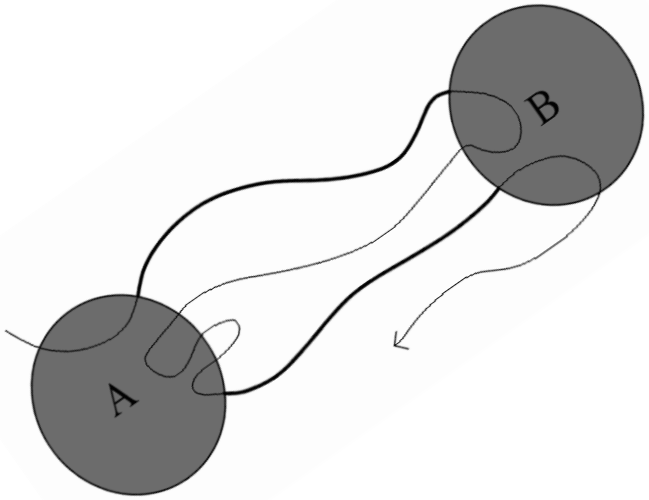
Numerical example II

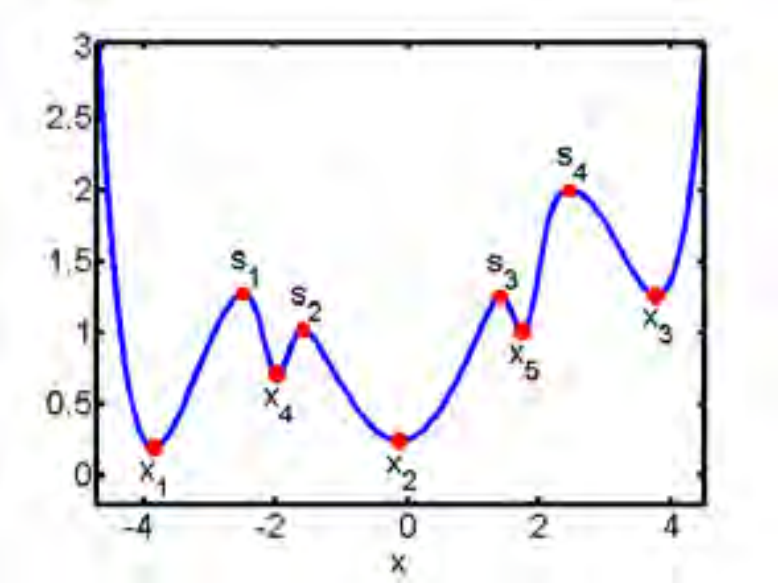
Committor probabilities

Probability to hit target set A before B, when starting at $X_0 = x$:

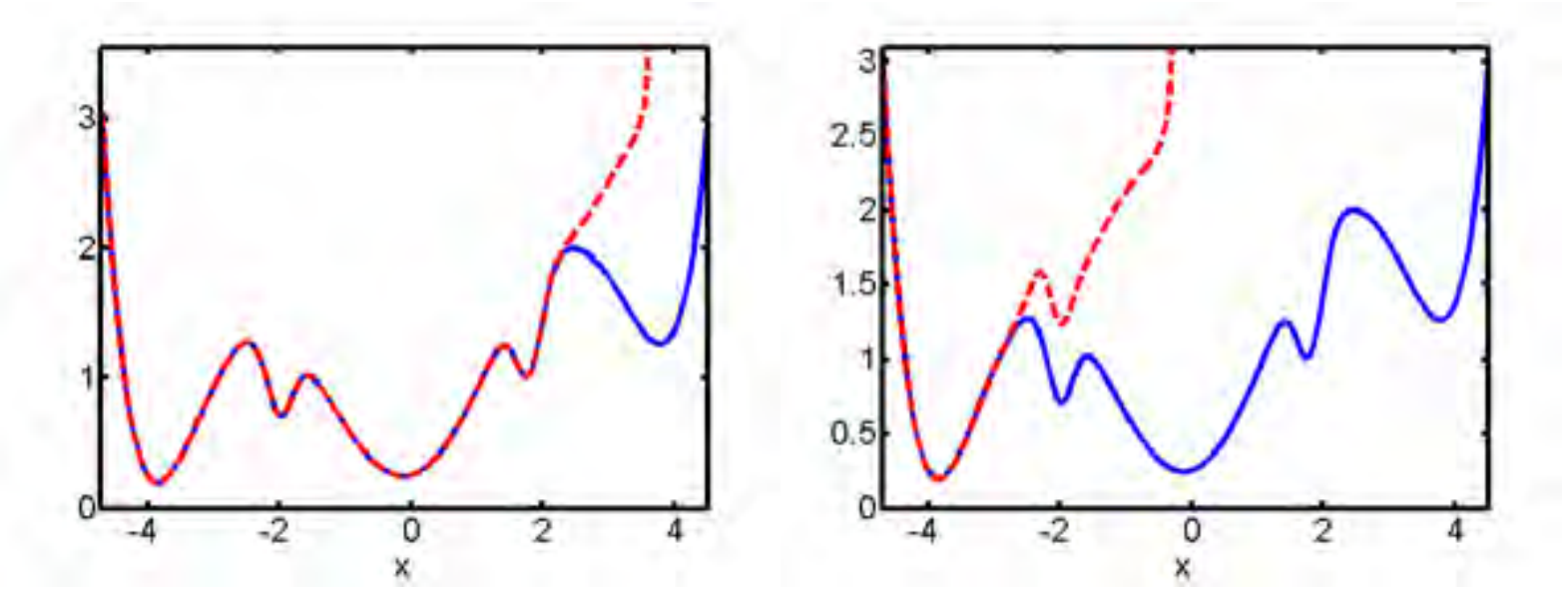
$$c(x) = \mathbb{E}[\mathbf{1}_A(X_{\tau_{A\cup B}})]$$

$$-\epsilon \log c(x) = \min_u \mathbb{E}_Q \left[\frac{1}{2} \int_0^{\tau^u_{A\cup B}} |u_s|^2 ds - \epsilon \log \mathbf{1}_A(X^u_{\tau^u_{A\cup B}}) \right]$$





Committor probabilities, cont'd



Optimal biasing potential for the committor with $A = (-4, -3.8)$ and $B = (3.8, 4)$ or $(-0.2, 0)$

[H. et al., submitted to Entropy, 2013]

Numerical example III

α-helical conformation of alanine dipeptide

Maximize

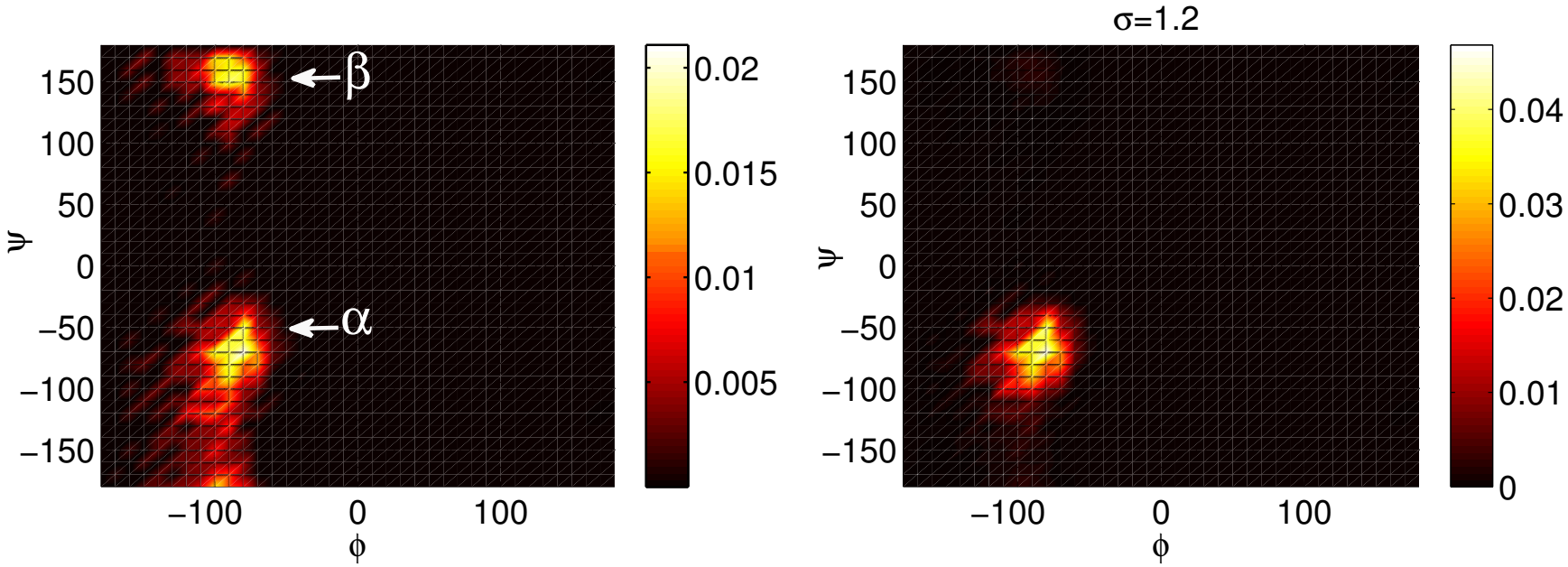
$$J(u) = \liminf_{T\to\infty} \mathbb{E}\left[\frac{1}{T}\int_0^T \left(\sigma\chi_\alpha(X_t^u) - \frac{1}{2}|u_t|^2\right) dt\right],$$

subject to $dX_t^u = (u_t - \nabla V(X_t^u))\, dt + \sqrt{2\epsilon}\, dB_t$.

MD simulation of ADP in a box of 256 water molecules (TIP3P, CHARMM force field)

α-helical conformation of alanine dipeptide, cont'd

Approximation of the continuous dynamics by a **Markov jump process** on $\mathcal{S} = \{1, \ldots, 100\}$ with generator $M = (M_{ij})_{i,j}$, based on a Galerkin discretization of the torsion angle space.



Marginal distributions in the dihedral angles of uncontrolled and tilted system (100 × 100 EVP for principal EV)

[Schütte, Winkelmann & H, Math Program, 2011]

Conclusions

Optimally driven nonequilibrium MD can **mimic equilibrium statistics**, replacing a difficult sampling problem for rare events by a potentially more difficult **minimization problem**.

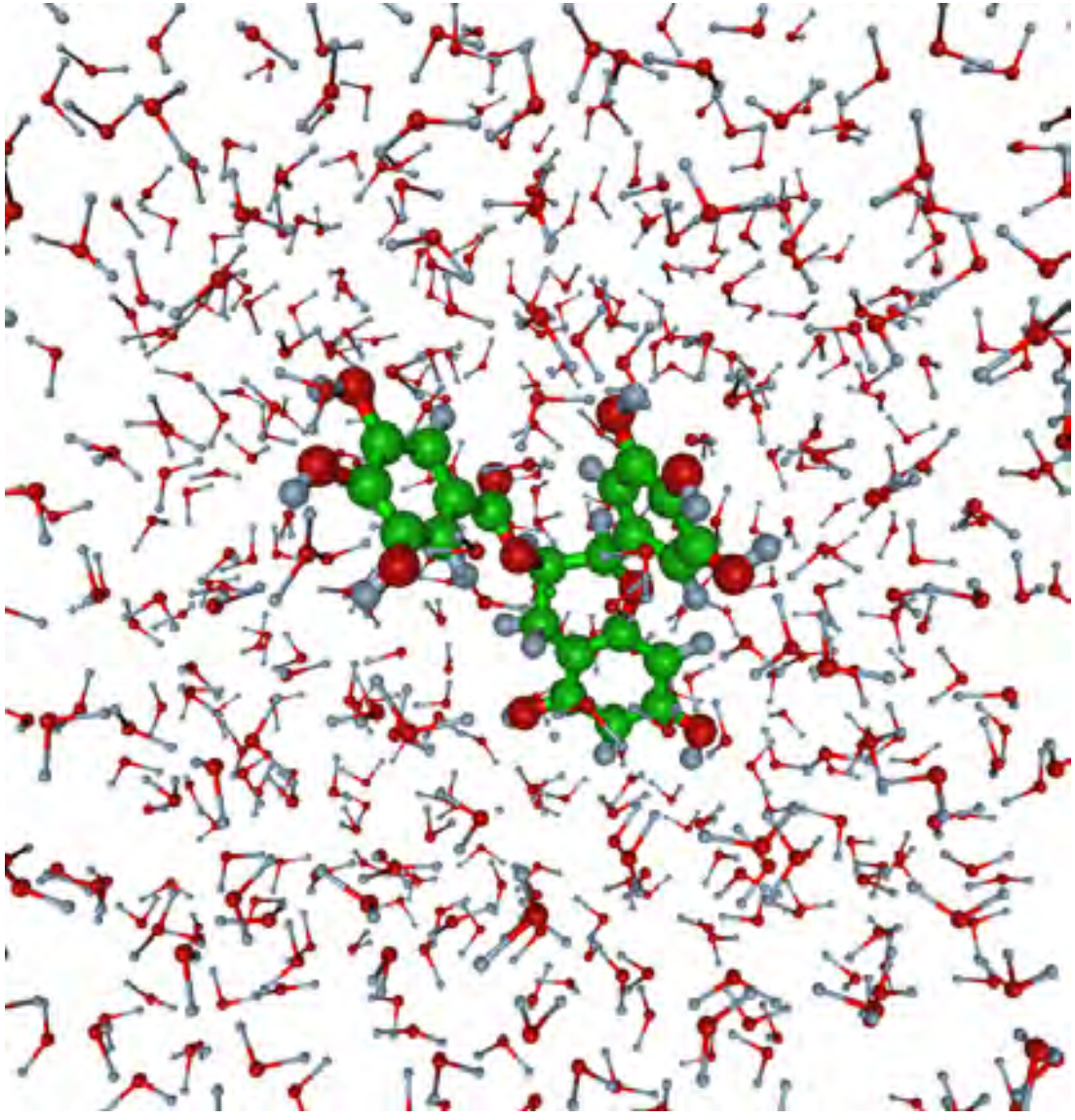
But: the approach can only be useful for sampling rare events if the underlying optimal control problems can be solved efficiently.

Conclusions

Optimally driven nonequilibrium MD can **mimic equilibrium statistics**, replacing a difficult sampling problem for rare events by a potentially more difficult **minimization problem**.

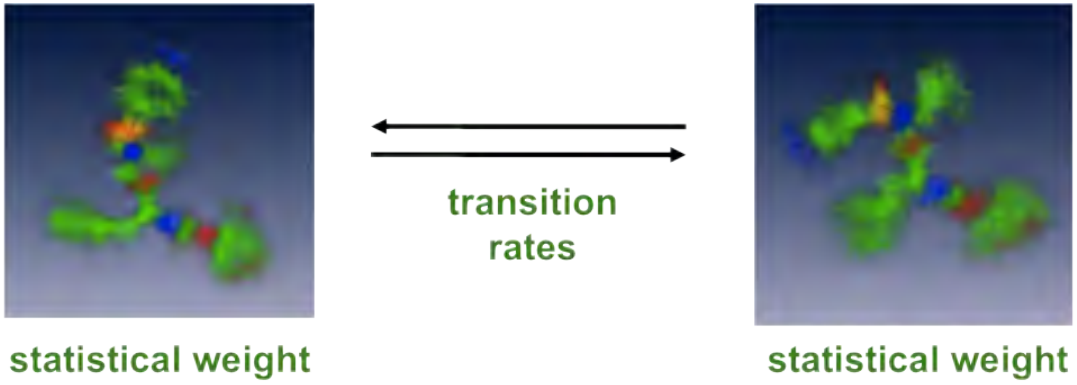
But: the approach can only be useful for sampling rare events if the underlying optimal control problems can be solved efficiently.

Open problems (the bad news)



Open problems (and some good news)

Dynamics on the relevant time scale are relatively simple . . .



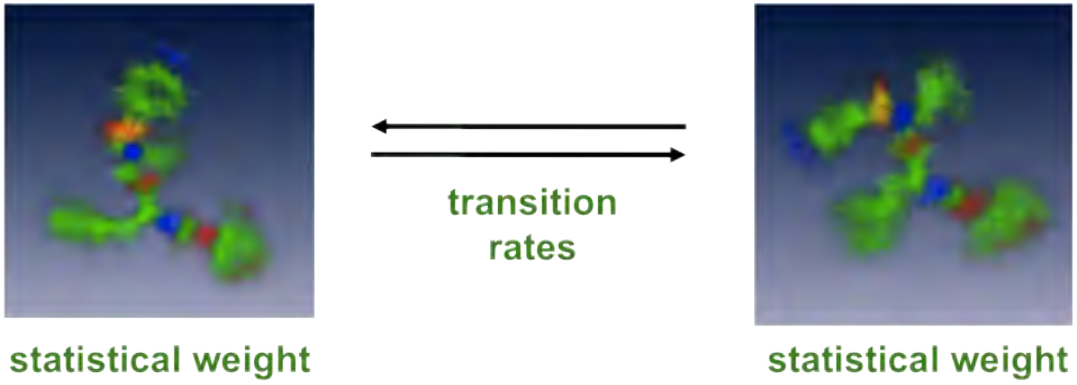
. . . and so are the controls (if one has an idea of what a sensible order parameter might be to represent the bias potential).

The functional can be minimized directly using an EM algorithm. Ideas from the machine learning community should be looked at.

The relation to Jarzysnki's approach is yet an open question.

Open problems (and some good news)

Dynamics on the relevant time scale are relatively simple . . .



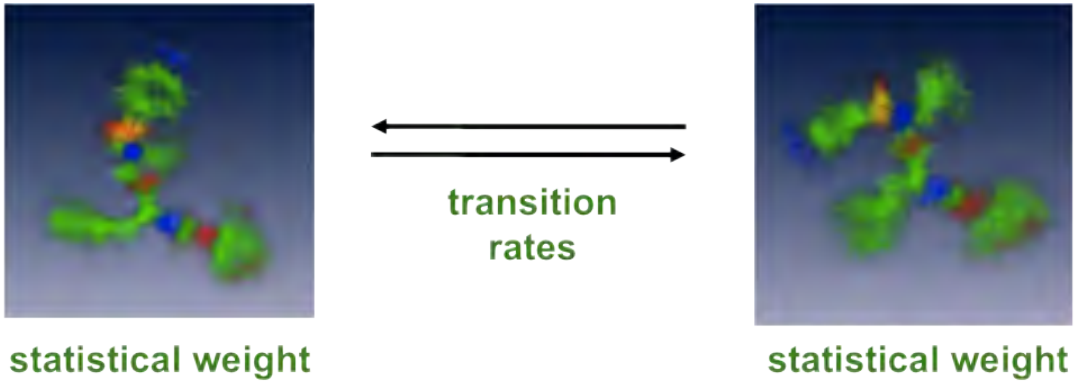
. . . and so are the controls (if one has an idea of what a sensible order parameter might be to represent the bias potential).

The functional can be minimized directly using an EM algorithm. Ideas from the machine learning community should be looked at.

The relation to Jarzysnki's approach is yet an open question.

Open problems (and some good news)

Dynamics on the relevant time scale are relatively simple . . .



. . . and so are the controls (if one has an idea of what a sensible order parameter might be to represent the bias potential).

The functional can be minimized directly using an EM algorithm. Ideas from the machine learning community should be looked at.

The relation to Jarzysnki's approach is yet an open question.

Thank you for your attention.

further information on biocomputing.mi.fu-berlin.de