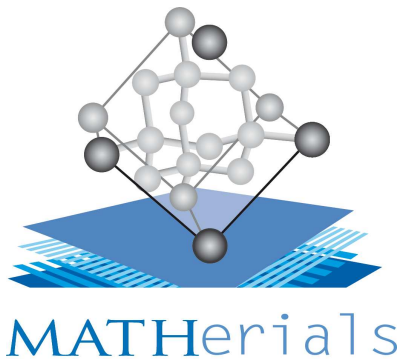
How to compute transition times?

On the Hill relation and the mean reaction time for metastable processes.

T. Lelièvre

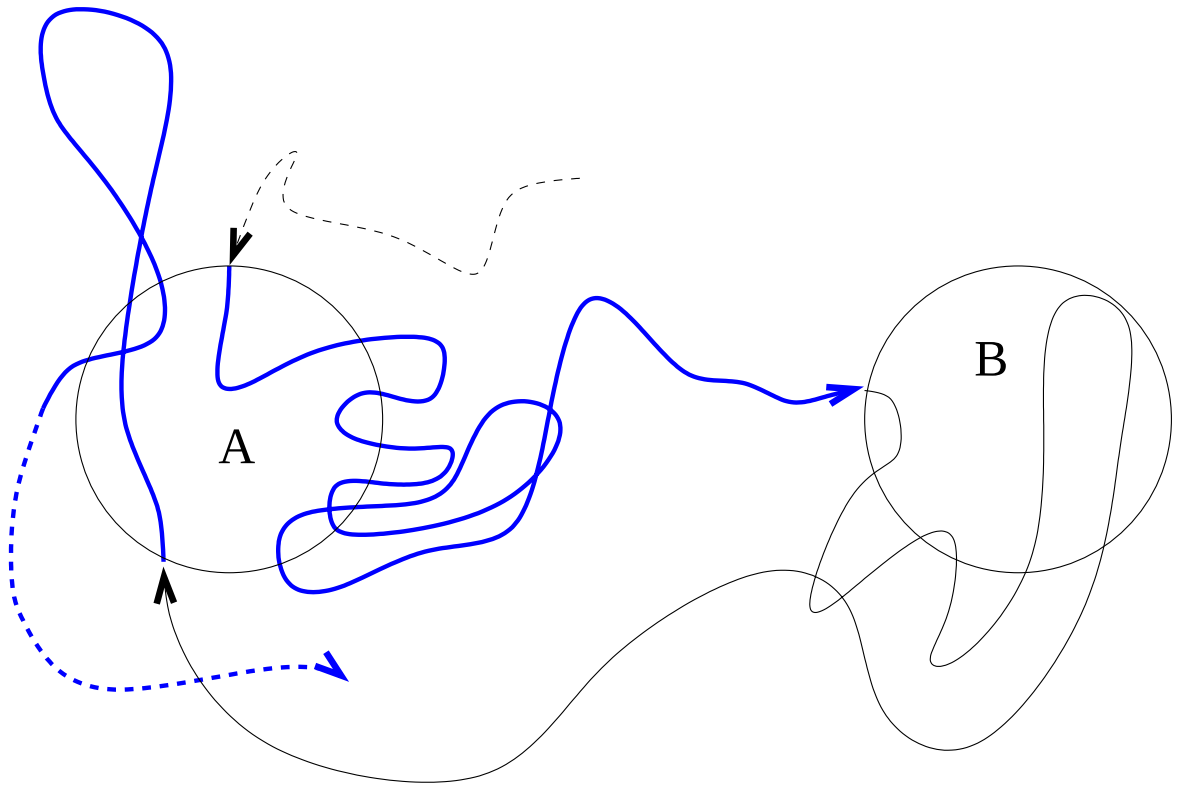
CERMICS - Ecole des Ponts ParisTech & Matherials project-team - INRIA

Joint work with M. Baudel, A. Guyader, M.Ramil and J. Reygner



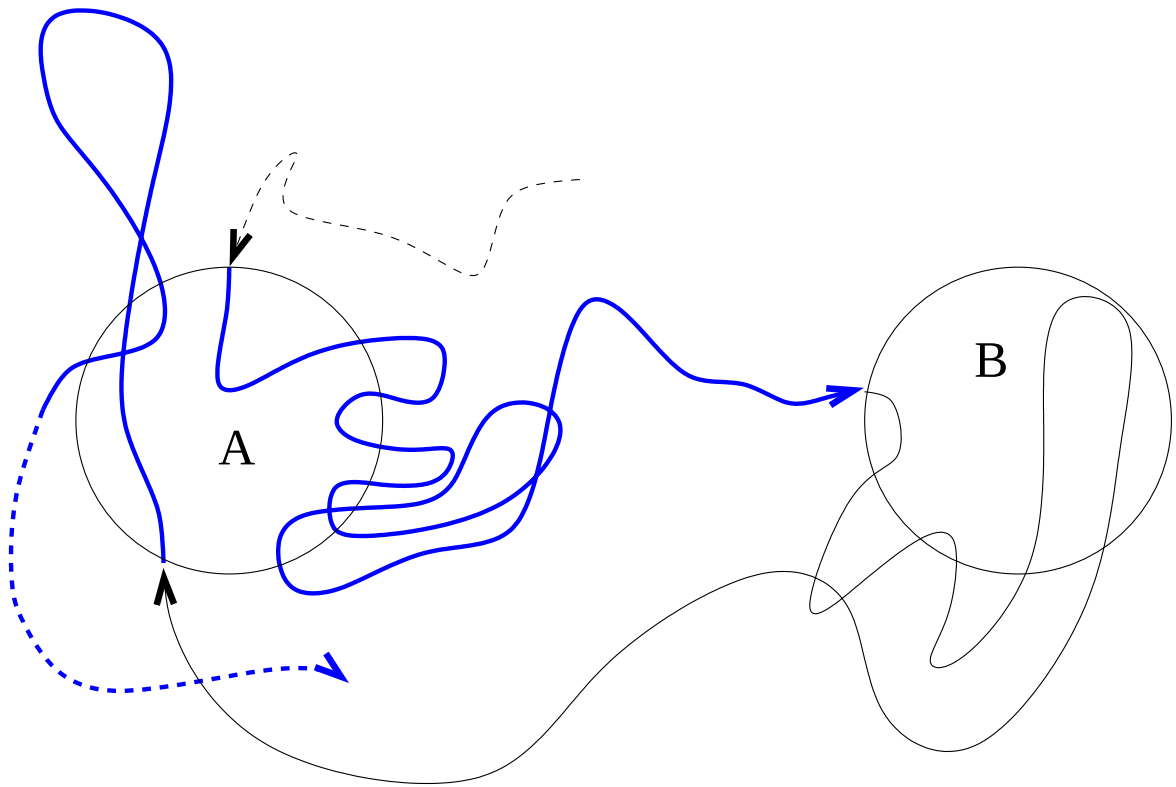
Brin Mathematics Research Center, Feb. 28th 2023

Introduction and motivation



Transition time

Let us consider an ergodic stochastic continuous in time process $(X_t)_{t\geq 0}$ in $\mathbb{R}^d$, and two disjoint subsets $A \subset \mathbb{R}^d$ and $B \subset \mathbb{R}^d$. The objective is to compute the mean transition time at equilibrium from A to B, denoted by $\Delta_{A\to B}$.



Remark: we are also interested in any statistical property of the equilibrium reactive paths from A to B.

Metastability

Examples: Molecular dynamics (A and B are defined in positions space)

- Langevin dynamics (M mass matrix, $\gamma > 0$, $\beta = (k_B T)^{-1}$)

$$\begin{cases} dQ_t = M^{-1} P_t \, dt, \\ dP_t = -\nabla V(Q_t) \, dt - \gamma M^{-1} P_t \, dt + \sqrt{2\gamma\beta^{-1}} dW_t, \end{cases}$$

  ergodic wrt $\mu(dq) \otimes Z_p^{-1} \exp\left(-\beta \frac{p^t M^{-1} p}{2}\right) dp$ with

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

  where $Z = \int \exp(-\beta V)$.

- overdamped Langevin dynamics

$$dX_t = -\nabla V(X_t) \, dt + \sqrt{2\beta^{-1}} dW_t,$$

  which is also ergodic wrt μ.

Challenge: A is typically a metastable state, so that the transition time from A to B is very large.

From continous time to discrete time: *Method 1*

Ths first step is to rewrite the problem in terms of a Markov chain.

Method 1 (only for Langevin): Consider the successive entrances in A of (Q_t, P_t):

$$Y_n = (Q_{\tau_n}, P_{\tau_n})$$

where $\tau_n = \inf\{t > \tau_{n-1}, Q_t \in A \cup B$.

The Markov chain $(Y_n)_{n\geq 0}$ is with values in $\mathcal{A} \cup \mathcal{B}$ where

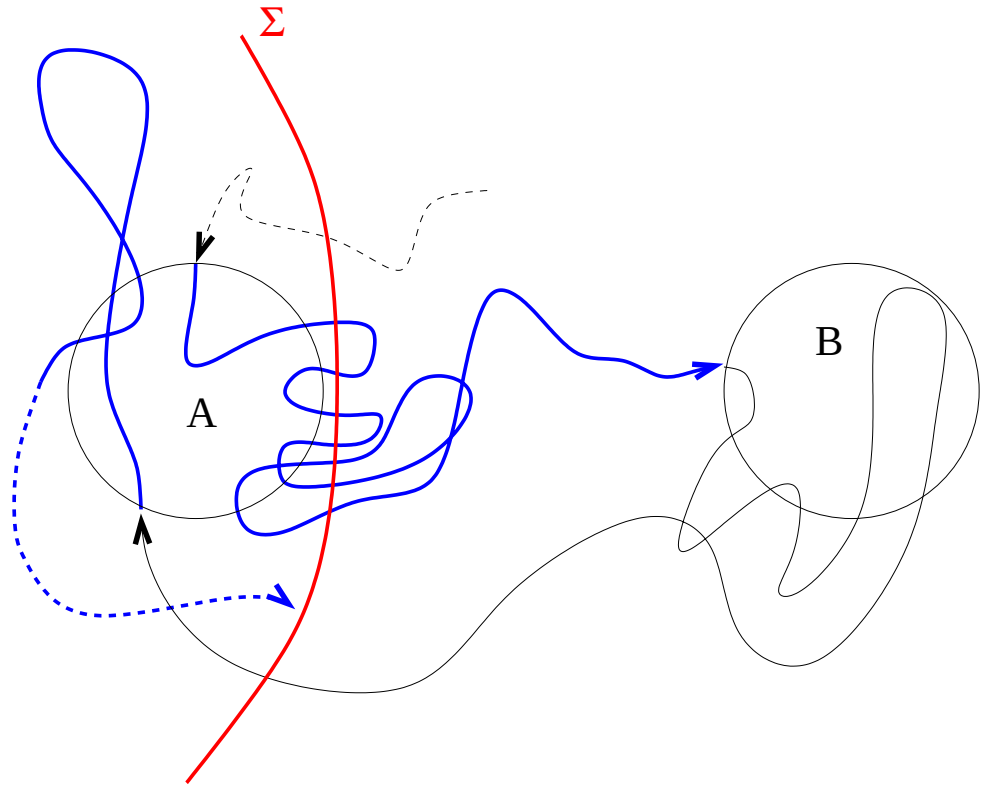
$$\mathcal{A} = \{(q,p) \in \partial A \times \mathbb{R}^d, p \cdot n_A(q) < 0\}$$

and

$$\mathcal{B} = \{(q,p) \in \partial B \times \mathbb{R}^d, p \cdot n_B(q) < 0\}.$$

From continous time to discrete time: *Method 2*

Method 2 (introduced here for overdamped Langevin, but can also been used for Langevin):



Let Σ a co-dimension 1 submanifold in-between A and B. Consider then, $(Y_n)_{n\geq 0}$ is the sequence of successive intersections of $(X_t)_{t\geq 0}$ with ∂A or ∂B, while hitting Σ in-between.

From continous time to discrete time: *Method 2*

More precisely:

$$Y_n = X_{\tau_n}$$

where

$$\tau_n^{\Sigma} = \inf\{t > \tau_{n-1}, X_t \in \Sigma\}$$

$$\tau_n = \inf\{t > \tau_n^{\Sigma}, X_t \in \mathcal{A} \cup \mathcal{B}\}.$$

The Markov chain $(Y_n)_{n\geq 0}$ is with values in $\mathcal{A} \cup \mathcal{B}$ where

$$\mathcal{A} = \partial A$$

and

$$\mathcal{B} = \partial B.$$

Reactive entrance distribution

Let us define the successive entrance times in $\mathcal{A}$ and $\mathcal{B}$ [Lu, Nolen, 2013] [E, Vanden Eijnden, 2006]:

$$T_{k+1}^{\mathcal{A}} = \inf\{n > T_k^{\mathcal{B}},\ Y_n \in \mathcal{A}\}$$

$$T_{k+1}^{\mathcal{B}} = \inf\{n > T_{k+1}^{\mathcal{A}},\ Y_n \in \mathcal{B}\}.$$

The reactive entrance distribution in $\mathcal{A}$ at equilibrium is defined by:

$$\nu_E = \lim_{K\to\infty} \hat{\nu}_{E,K}$$

where

$$\hat{\nu}_{E,K} = \frac{1}{K}\sum_{k=1}^{K} \delta_{Y_{T_k^{\mathcal{A}}}}.$$

Remark: ν_E is independent of the choice of Σ and is also the reactive entrance distribution for the original continuous time process.

Back to the mean transition time

The mean transition time at equilibrium is (strong Markov property):

$$\Delta_{A\to B} = \mathbb{E}^{\nu_E}\left(\sum_{n=0}^{T_{\mathcal{B}}-1} \Delta(Y_n)\right)$$

where

$$T_{\mathcal{B}} = \inf\{n \geq 0,\ Y_n \in \mathcal{B}\}$$

and for all $x \in \mathcal{A}$,

$$\Delta(x) = \mathbb{E}^x(\tau_1).$$

Remark: Notice that for all $x \in \mathcal{A}$,

$$\Delta(x) = \mathbb{E}^x(\tau_1 1_{Y_1\in\mathcal{A}}) + \mathbb{E}^x(\tau_1 1_{Y_1\in\mathcal{B}})$$

is the average time of loop from x back to $\mathcal{A}$ when $Y_1 \in \mathcal{A}$ plus the average time of a reactive trajectory from x to $\mathcal{B}$ when $Y_1 \in \mathcal{B}$.

Summary

Objective: Given a discrete-time Markov chain $(Y_n)_{n\geq 0}$ with values in $\mathcal{A}\cup\mathcal{B}$ and a bounded measurable function $f:\mathcal{A}\to\mathbb{R}$, estimate:

$$\mathbb{E}^{\nu_E}\left(\sum_{n=0}^{T_{\mathcal{B}}-1} f(Y_n)\right)$$

Two challenges: The sets $\mathcal{A}$ and $\mathcal{B}$ are metastable, so that (i) $T_{\mathcal{B}}$ is very large, and (ii) ν_E is difficult to sample.

Ideas: For (i), use the Hill relation to replace a longtime computation by the estimation of the probability of the event $\{Y_1\in\mathcal{B}\}$, together with a rare event sampling method (forward flux sampling -FFS- or adaptive multilevel splitting -AMS-). For (ii), on can rely on the fact that $\mathcal{A}$ is metastable: the process $(Y_n)_{n\geq 0}$ reaches some "equilibrium within $\mathcal{A}$" (quasi stationary distribution) before transitioning to $\mathcal{B}$.

Assumptions and notation

Assumptions: In the following, we assume that the Markov chain $(Y_n)_{n\geq 0}$ (with kernel K) satisfies the following hypothesis:

[A1] $(Y_n)_{n\geq 0}$ is weak-Feller meaning that $(Kf) \in \mathcal{C}(\mathcal{A}\cup\mathcal{B},\mathbb{R})$ whenever $f \in \mathcal{C}(\mathcal{A}\cup\mathcal{B},\mathbb{R})$.

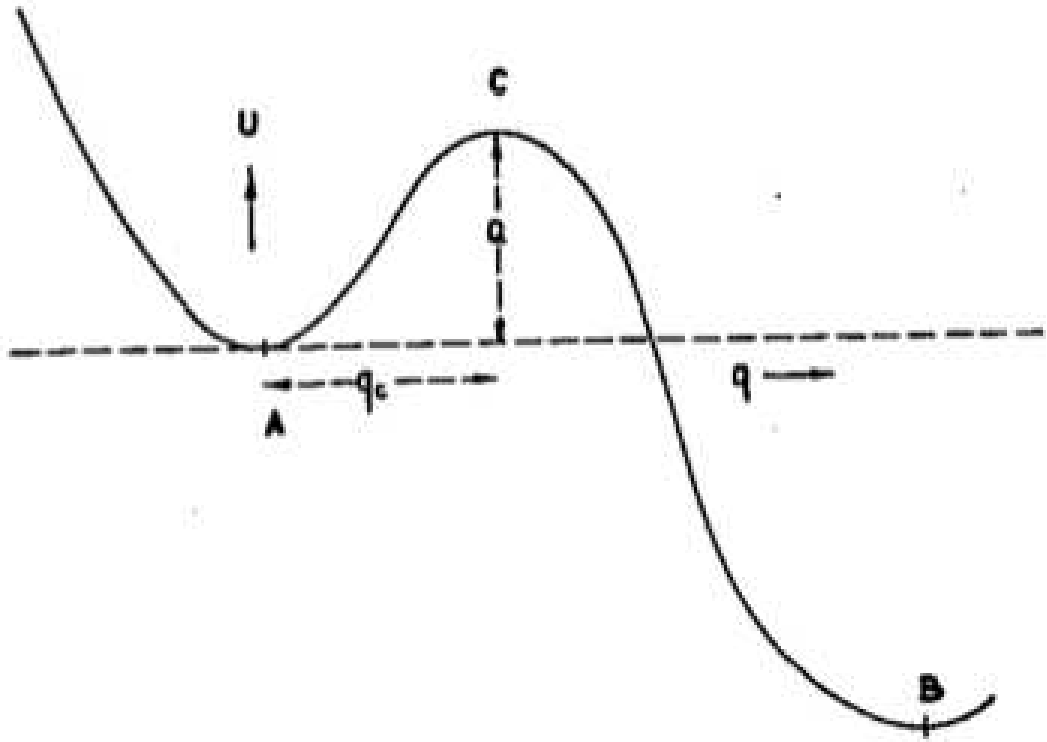
[A2] $(Y_n)_{n\geq 0}$ is positive Harris recurrent, and π_0 denotes its unique stationary probability measure.

[A3] $\pi_0(\mathcal{A}) > 0$ and $\pi_0(\mathcal{B}) > 0$.

All these assumptions are satisfied for the discrete processes built from the Langevin or overdamped Langevin dynamics.

Notation: In the following we use the block-decomposition of the kernel K of the chain $(Y_n)_{n\geq 0}$ over $\mathcal{A}\cup\mathcal{B}$: $K = \begin{bmatrix} K_{\mathcal{A}} & K_{\mathcal{A}\mathcal{B}} \\ K_{\mathcal{B}\mathcal{A}} & K_{\mathcal{B}} \end{bmatrix}$.

The Hill relation



[Kramers, 1940]

The π-return process and the Hill relation

Let π be a probability measure on $\mathcal{A}$. The π-return process $(Y_n^\pi)_{n\geq 0}$ is the Markov chain with values in $\mathcal{A}$ and transition kernel: $\forall x \in \mathcal{A}$, $\forall C \subset \mathcal{A}$,

$$K^\pi(x, C) = \mathbb{P}^x(Y_1 \in C, T_\mathcal{B} > 1) + \mathbb{P}^x(Y_1 \in \mathcal{B})\pi(C).$$

In words, $(Y_n^\pi)_{n\geq 0}$ is the chain $(Y_n)_{n\geq 0}$ "reset to π" each time Y_n enters $\mathcal{B}$.

Lemma. $(Y_n^\pi)_{n\geq 0}$ admits a unique stationary distribution, denoted by $R(\pi)$, where

$$R(\pi) = \frac{\pi(\mathrm{Id}_\mathcal{A} - K_\mathcal{A})^{-1}}{\mathbb{E}^\pi(T_\mathcal{B})}.$$

Remark: Such processes are typically used in MD when people introduce a sink in B and a source in A to create a non-equilibrium flux from A to B [Farkas, 1927] [Kramers, 1940]: Weighted Ensemble [Zuckerman, Aristoff], Milestoning [Elber, Vanden Eijnden], TIS [Bolhuis, Van Erp].

The π-return process and the Hill relation

We are now in position to state the Hill relation [Hill, 1977] [Aristoff, 2018] [Baudel, Guyader, TL, 2022].

Proposition. For any bounded measurable function $f : \mathcal{A} \to \mathbb{R}$,

$$\mathbb{E}^{\pi}\left(\sum_{n=0}^{T_{\mathcal{B}}-1} f(Y_n)\right) = \frac{R(\pi)f}{\mathbb{P}^{R(\pi)}(Y_1 \in \mathcal{B})}.$$

Application of the Hill relation to $\pi = \nu_E$

Lemma. The probability measure $R(\nu_E)$ is the stationary distribution π_0 restricted to $\mathcal{A}$:

$$R(\nu_E) = \frac{\pi_0 1_{\mathcal{A}}}{\pi_0(\mathcal{A})} =: \pi_{0|\mathcal{A}}.$$

As a consequence,

$$\mathbb{E}^{\nu_E}\left(\sum_{n=0}^{T_{\mathcal{B}}-1} f(Y_n)\right) = \frac{\pi_{0|\mathcal{A}}(f)}{\mathbb{P}^{\pi_{0|\mathcal{A}}}(Y_1 \in \mathcal{B})}.$$

The Hill relation to compute $\Delta_{A\to B}$

Back to the mean transition time:

$$\mathbb{E}^{\nu_E}\left(\sum_{n=0}^{T_{\mathcal{B}}-1}\Delta(Y_n)\right) = \Delta_{Loop}(\pi_{0|\mathcal{A}})\left(\frac{1}{\mathbb{P}^{\pi_{0|\mathcal{A}}}(Y_1\in\mathcal{B})}-1\right)+\Delta_{React}(\pi_{0|\mathcal{A}})$$

where

- $\Delta_{Loop}(\pi_{0|\mathcal{A}}) = \mathbb{E}^{\pi_{0|\mathcal{A}}}(\tau_1|Y_1\in\mathcal{A})$ is the mean time for a loop from $\pi_{0|\mathcal{A}}$ back to $\mathcal{A}$ (computed by brute force Monte Carlo)
- $\Delta_{React}(\pi_{0|\mathcal{A}}) = \mathbb{E}^{\pi_{0|\mathcal{A}}}(\tau_1|Y_1\in\mathcal{B})$ is the mean time of a reactive trajectory from $\pi_{0|\mathcal{A}}$ to $\mathcal{B}$ (computed by FFS/AMS)
- $\mathbb{P}^{\pi_{0|\mathcal{A}}}(Y_1\in\mathcal{B})$ is the probability to get a reactive traj. starting from $\pi_{0|\mathcal{A}}$ (computed by FFS/AMS)

Application to the Langevin dynamics (Method 1)

For the Langevin dynamics, one has [TL, M. Ramil, J. Reygner, 2022]

$$\pi_{0|\mathcal{A}}(dq\,dp) = Z^{-1}|p.n_A(q)|\exp(-\beta(V(q)+p^T M^{-1}p/2))\,\sigma_{\partial A}(dq)\,dp$$

which can be sampled exactly by using appropriate sampling methods (constrained sampling on ∂A).

Remark: notice that we obtain the rigidly constrained stationary measure on ∂A, not the δ measure on ∂A.

[Hartmann, Schuette] [Ciccotti] [Vanden-Eijnden] [TL, Rousset, Stoltz, Zhang]

Application to the overdamped Langevin dynamics (Method 2)

When applying Method 2 (with Σ), π_0 and, a fortiori, $\pi_{0|\mathcal{A}}$ are in general unknown and difficult to sample.

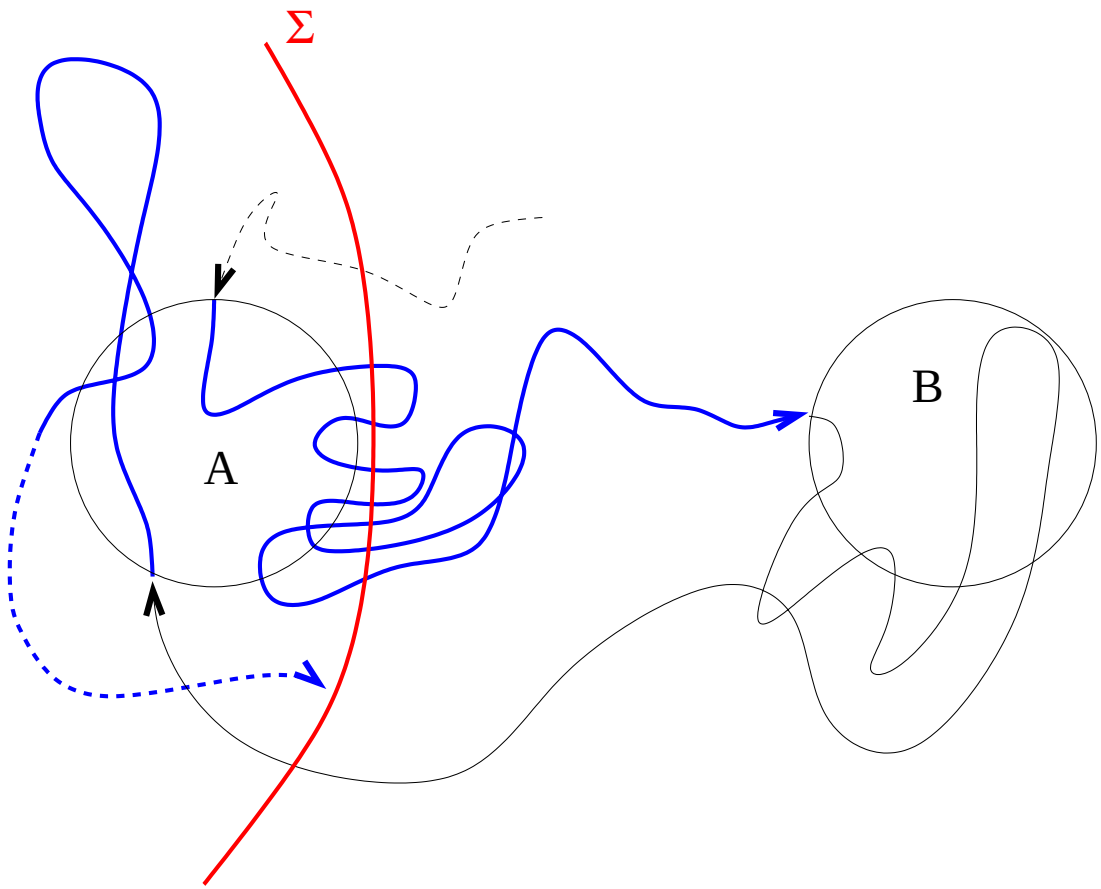
Consequently, the formula

$$\mathbb{E}^{\nu_E}\left(\sum_{n=0}^{T_{\mathcal{B}}-1} f(Y_n)\right) = \frac{\pi_{0|\mathcal{A}}(f)}{\mathbb{P}^{\pi_{0|\mathcal{A}}}(Y_1 \in \mathcal{B})}$$

is not practical since $\pi_{0|\mathcal{A}}$ is difficult to sample.

Hope: since $\mathcal{A}$ is metastable, maybe it is not needed to sample ν_E or $\pi_{0|\mathcal{A}}$ since, typically, the process will reach a local equilibrium within $\mathcal{A}$ before going to $\mathcal{B}$.

Using the QSD approximation



The quasi-stationary distribution (QSD)

Lemma. Under the assumptions above, the process $(Y_n)_{n\geq 0}$ admits a quasi-stationary distribution (QSD) ν_Q in $\mathcal{A}$, namely a probability measure ν_Q over $\mathcal{A}$ such that: $\forall C \subset \mathcal{A}$,

$$\nu_Q(C) = \mathbb{P}^{\nu_Q}(Y_1 \in C | T_{\mathcal{B}} > 1).$$

Remarks:

- QSD and Yaglom limit: if $\mathcal{L}(Y_n | T_{\mathcal{B}} > n)$ admits a limit when $n \to \infty$, this limit is a QSD.
- The ν_Q-return process admits ν_Q as an invariant distribution:

$$R(\nu_Q) = \nu_Q.$$

The Hill relation applied to $\pi = \nu_Q$

Since $R(\nu_Q) = \nu_Q$, one has:

$$\mathbb{E}^{\nu_Q}\left(\sum_{n=0}^{T_{\mathcal{B}}-1} f(Y_n)\right) = \frac{\nu_Q(f)}{\mathbb{P}^{\nu_Q}(Y_1 \in \mathcal{B})}.$$

Remark: Starting from ν_Q, $T_{\mathcal{B}}$ is geometrically distributed, with parameter $\mathbb{P}^{\nu_Q}(Y_1 \in \mathcal{B})$.

Back to the mean transition time [Cérou, Guyader, TL, Pommier, 2011]:

$$\mathbb{E}^{\nu_Q}\left(\sum_{n=0}^{T_{\mathcal{B}}-1} \Delta(Y_n)\right) = \Delta_{Loop}(\nu_Q)\left(\frac{1}{\mathbb{P}^{\nu_Q}(Y_1 \in \mathcal{B})} - 1\right) + \Delta_{React}(\nu_Q)$$

What did we gain, compared to $\pi = \nu_E$? The probability distribution ν_Q can be sampled by brute force Monte Carlo.

The algorithm to compute $\Delta_{A\to B}$

In practice:

- Simulate the process $(X_t)_{t\geq 0}$ (or $(Q_t, P_t)_{t\geq 0}$) in a neighborhood of A, registering the successive loops from A to Σ and back to A. This gives samples distributed according to ν_Q, and $\Delta_{Loop}(\nu_Q)$.
- Use AMS to simulate reactive trajectories, starting from the QSD ν_Q. This gives an estimate of $\mathbb{P}^{\nu_Q}(Y_1 \in \mathcal{B})$ and of $\Delta_{React}(\nu_Q)$.

Remark: Typically, one has $\mathbb{P}^{\nu_Q}(Y_1 \in \mathcal{B}) \ll 1$ and $\Delta_{React}(\nu_Q) \ll \frac{\Delta_{Loop}(\nu_Q)}{\mathbb{P}^{\nu_Q}(Y_1\in\mathcal{B})}$ so that

$$\mathbb{E}^{\nu_Q}\left(\sum_{n=0}^{T_{\mathcal{B}}-1} \Delta(Y_n)\right) \simeq \frac{\Delta_{Loop}(\nu_Q)}{\mathbb{P}^{\nu_Q}(Y_1 \in \mathcal{B})}.$$

This is e.g. the formula used in FFS to compute transition times

[Allen, Valeriani, ten Wolde, 2009].

Bias analysis

$$\left| \frac{\mathbb{E}^{\nu_E}\left(\sum_{n=0}^{T_{\mathcal{B}}-1} f(Y_n)\right) - \mathbb{E}^{\nu_Q}\left(\sum_{n=0}^{T_{\mathcal{B}}-1} f(Y_n)\right)}{\mathbb{E}^{\nu_E}\left(\sum_{n=0}^{T_{\mathcal{B}}-1} f(Y_n)\right)} \right| \ll 1?$$

Bias analysis

In practice, we thus compute $\mathbb{E}^{\nu_Q}\left(\sum_{n=0}^{T_{\mathcal{B}}-1} f(Y_n)\right)$ instead of the truth $\mathbb{E}^{\nu_E}\left(\sum_{n=0}^{T_{\mathcal{B}}-1} f(Y_n)\right)$.

Objective: Quantify the relative error

$$ERR = \left| \frac{\mathbb{E}^{\nu_E}\left(\sum_{n=0}^{T_{\mathcal{B}}-1} f(Y_n)\right) - \mathbb{E}^{\nu_Q}\left(\sum_{n=0}^{T_{\mathcal{B}}-1} f(Y_n)\right)}{\mathbb{E}^{\nu_E}\left(\sum_{n=0}^{T_{\mathcal{B}}-1} f(Y_n)\right)} \right|.$$

as a function of how large is the transition time wrt the convergence time to the QSD.

Transition time

The time to observe a transition to $\mathcal{B}$ is measured by

$$\frac{1}{p^+}$$

where $p^+ = \sup_{x \in \mathcal{A}} \mathbb{P}^x(Y_1 \in \mathcal{B})$.

Remark: One obviously has, for any $x \in \mathcal{A}$,

$$\frac{1}{p^+} \leq \mathbb{E}^x(T_{\mathcal{B}}).$$

Convergence time to the QSD

The convergence time to the QSD is measured by:

$$T_Q^E = \|\nu_E H_Q\|_{\mathrm{TV}}$$

where

$$H_Q f(x) = \mathbb{E}^x \left[\sum_{n=0}^{T_{\mathcal{B}}-1} (f(Y_n) - \nu_Q f) \right].$$

Why can T_Q^E be seen as a convergence time to the QSD?

$$T_Q^E \leq \sum_{n=0}^{\infty} \|\mathcal{L}^{\nu_E}(Y_n | T_{\mathcal{B}} > n) - \nu_Q\|_{\mathrm{TV}}$$

Bias analysis

Proposition. Assume that $p^+ T_Q^E < 1$. Then,

$$ERR \leq \frac{p^+ T_Q^E}{1 - p^+ T_Q^E} \left(1 + \frac{\|f\|_\infty}{|\pi_{0|\mathcal{A}} f|}\right).$$

This shows that the bias is small if the transition time is large compared to the convergence time to the QSD, i.e.

$$\frac{1}{p^+} \gg T_Q^E.$$

Remark: We have checked on examples that the upper bound is sharp in various ways. In particular, one can replace p^+ neither by $\mathbb{P}^{\nu_Q}(Y_1 \in \mathcal{B})$ nor by $\mathbb{P}^{\nu_E}(Y_1 \in \mathcal{B})$ in the RHS.

Example: the geometrically ergodic case

In the context of the overdamped Langevin dynamics, one can show that: $\exists \nu_Q$, $\exists \alpha > 0, \exists \rho \in (0,1), \forall x \in \mathcal{A}, \forall n \geq 0$,

$$\|\mathcal{L}^x(Y_n | T_\mathcal{B} > n) - \nu_Q\|_{TV} \leq \alpha \rho^n.$$

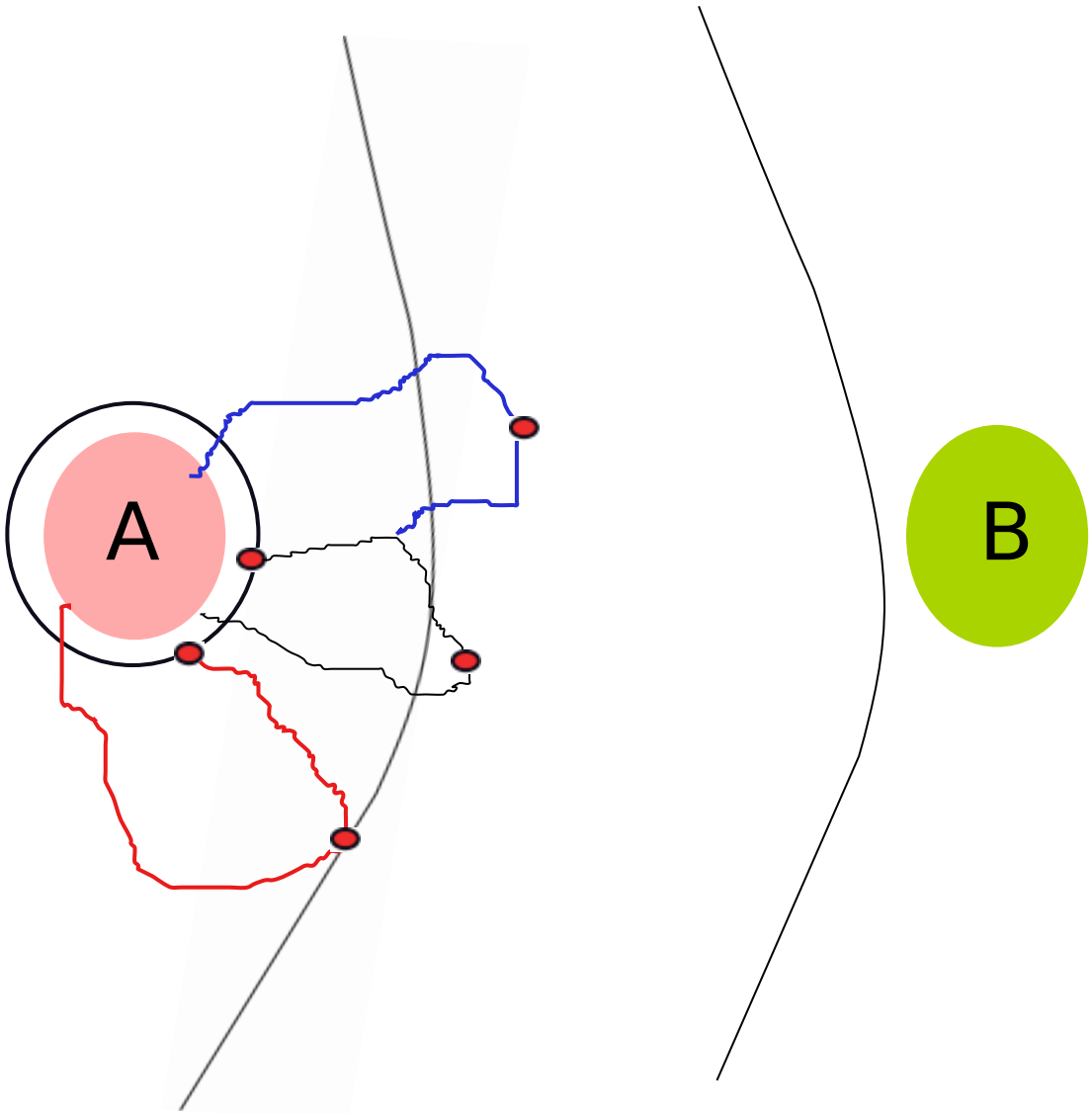
In this case,

$$T_Q^E \leq \min\left(\frac{\alpha}{1-\rho}, \inf_{c \in (0,\alpha)} \frac{2}{1-c}\left\lceil \frac{\ln(c/\alpha)}{\ln \rho}\right\rceil\right).$$

The RHS goes to $\min(\alpha, 2)$ when $\rho \to 0$.

Examples of sufficient conditions to get exponential convergence [Del Moral, Horton, Jasra, 2022]: double-sided condition [Birkhoff 1957], Dobrushin condition [Dobrushin, 1970] [Del Moral, Guionnet, 2001], Meyn-Tweedie like conditions [Champagnat, Villemonais, 2017], ...

Splitting algorithms



Multilevel splitting

Objective: sample trajectories between two given metastable states A and B.

Difficulty: A trajectory leaving A is more likely to go back to A than to reach B. This is a rare event problem.

We are interested in the rare event $\{\tau_A < \tau_B\}$ where

$$\tau_A = \inf\{t > 0,\, \boldsymbol{X}_t \in A\}, \quad \tau_B = \inf\{t > 0,\, \boldsymbol{X}_t \in B\}$$

Assume we are given a one dimensional function $\xi : \mathbb{R}^d \to \mathbb{R}$ which "indexes" the transition from A to B in the following sense:

$$A \subset \{x \in \mathbb{R}^d, \xi(x) < z_{\min}\} \text{ and } B \subset \{x \in \mathbb{R}^d, \xi(x) > z_{\max}\}.$$

Splitting algorithm: basic idea

Let us introduce

$$\tau_z = \inf\{t > 0,\, \xi(\boldsymbol{X}_t) > z\}.$$

The idea of splitting algorithms (AMS, FFS, TIS, RESTART, ...) is to write the rare event

$$\{\tau_B < \tau_A\}$$

as a sequence of nested events: for $z_{\min} < z_1 < \ldots < z_{\max}$,

$$\{\tau_{z_1} < \tau_A\} \supset \{\tau_{z_2} < \tau_A\} \supset \ldots \supset \{\tau_{z_{\max}} < \tau_A\} \supset \{\tau_B < \tau_A\}$$

and to simulate the successive *conditional events*: for $k = 1, 2, \ldots,$

$$\{\tau_{z_q} < \tau_A\} \text{ knowing that } \{\tau_{z_{q-1}} < \tau_A\}.$$

It is then easy to build an unbiased estimator of

$$\mathbb{P}(\tau_B < \tau_A) = \mathbb{P}(\tau_{z_1} < \tau_A)\mathbb{P}(\tau_{z_2} < \tau_A | \tau_{z_1} < \tau_A) \ldots \mathbb{P}(\tau_B < \tau_A | \tau_{z_{\max}} < \tau_A)$$

Splitting algorithm: adaptive level computation

Problem: How to choose the intermediate levels $(z_q)_{q\geq 1}$?

The optimum in terms of variance is attained if

$$\mathbb{P}(\tau_{z_q} < \tau_A | \tau_{z_{q-1}} < \tau_A) \text{ is constant .}$$

This naturally leads to adaptive versions (AMS, nested sampling) where the levels are determined by using *empirical quantiles*. Choose $k < n$, and given n trajectories $(\boldsymbol{X}^m_{t\wedge\tau_A})_{t>0,m=1,\ldots,n}$ in the event $\{\tau_{z_{q-1}} < \tau_A\}$, choose z_q so that

$$\mathbb{P}(\tau_{z_q} < \tau_A | \tau_{z_{q-1}} < \tau_A) \simeq \left(1 - \frac{k}{n}\right).$$

The level z_q is the k-th order statistics of $\sup_{t\geq 0} \xi(\boldsymbol{X}^m_{t\wedge\tau_A})$:

$$\sup_{t\geq 0} \xi(\boldsymbol{X}^{(1)}_{t\wedge\tau_A}) < \ldots < \sup_{t\geq 0} \xi(\boldsymbol{X}^{(k)}_{t\wedge\tau_A}) =: z_q < \ldots < \sup_{t\geq 0} \xi(\boldsymbol{X}^{(n)}_{t\wedge\tau_A}).$$

AMS: estimator of the rare event probability

Let Q_{iter} be the number of iterations to reach the level $z_{\max}$:

$$Q_{\text{iter}} = \min\{q \geq 0, z_q > z_{\max}\}$$

(where z_0 is the k-th order statistics of the n initial trajectories). Then, one obtains the estimator:

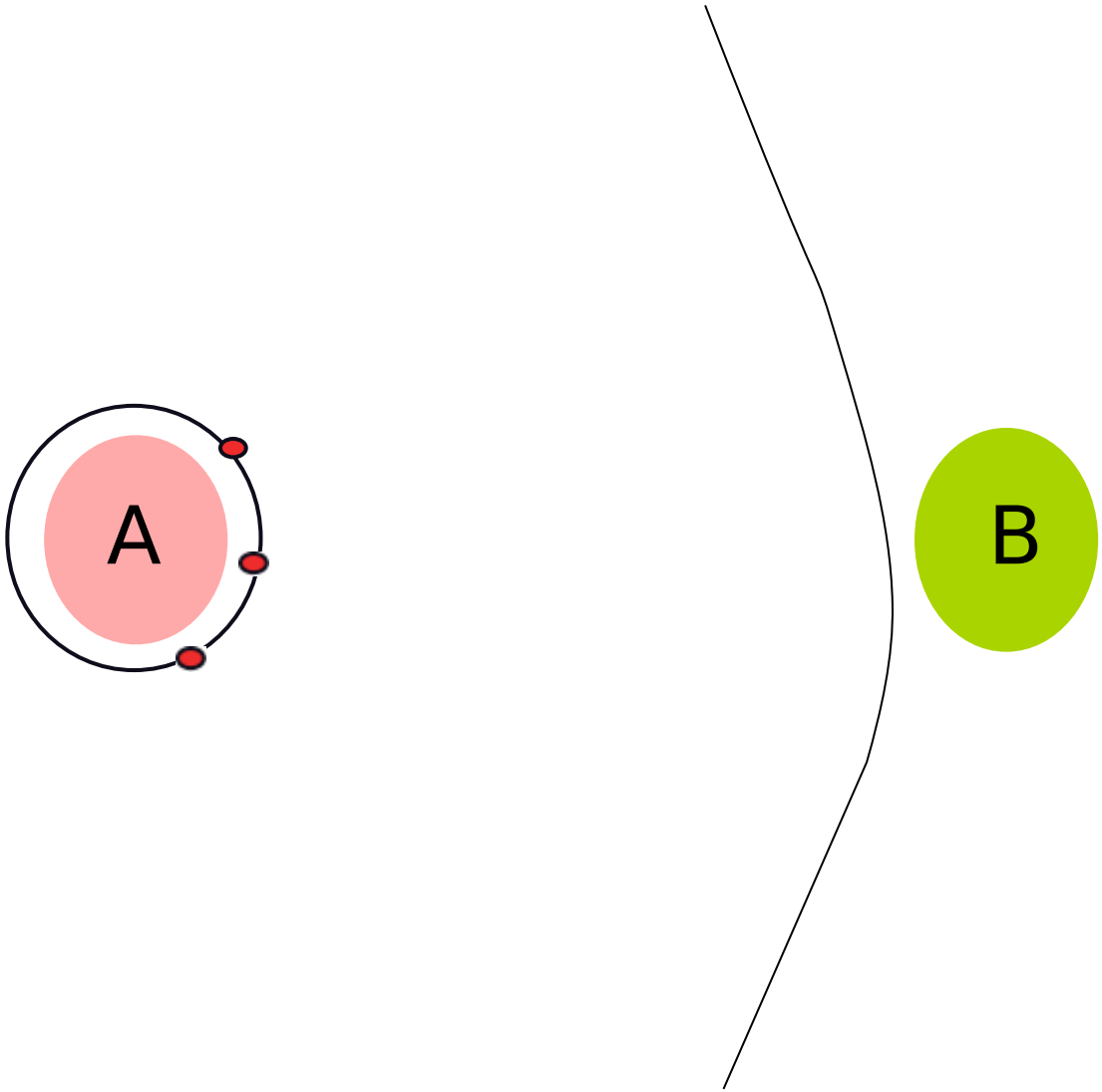
$$\left(1 - \frac{k}{n}\right)^{Q_{\text{iter}}} \simeq \mathbb{P}(\tau_{z_{\max}} < \tau_A).$$

An estimator of $\mathbb{P}(\tau_B < \tau_A)$ is

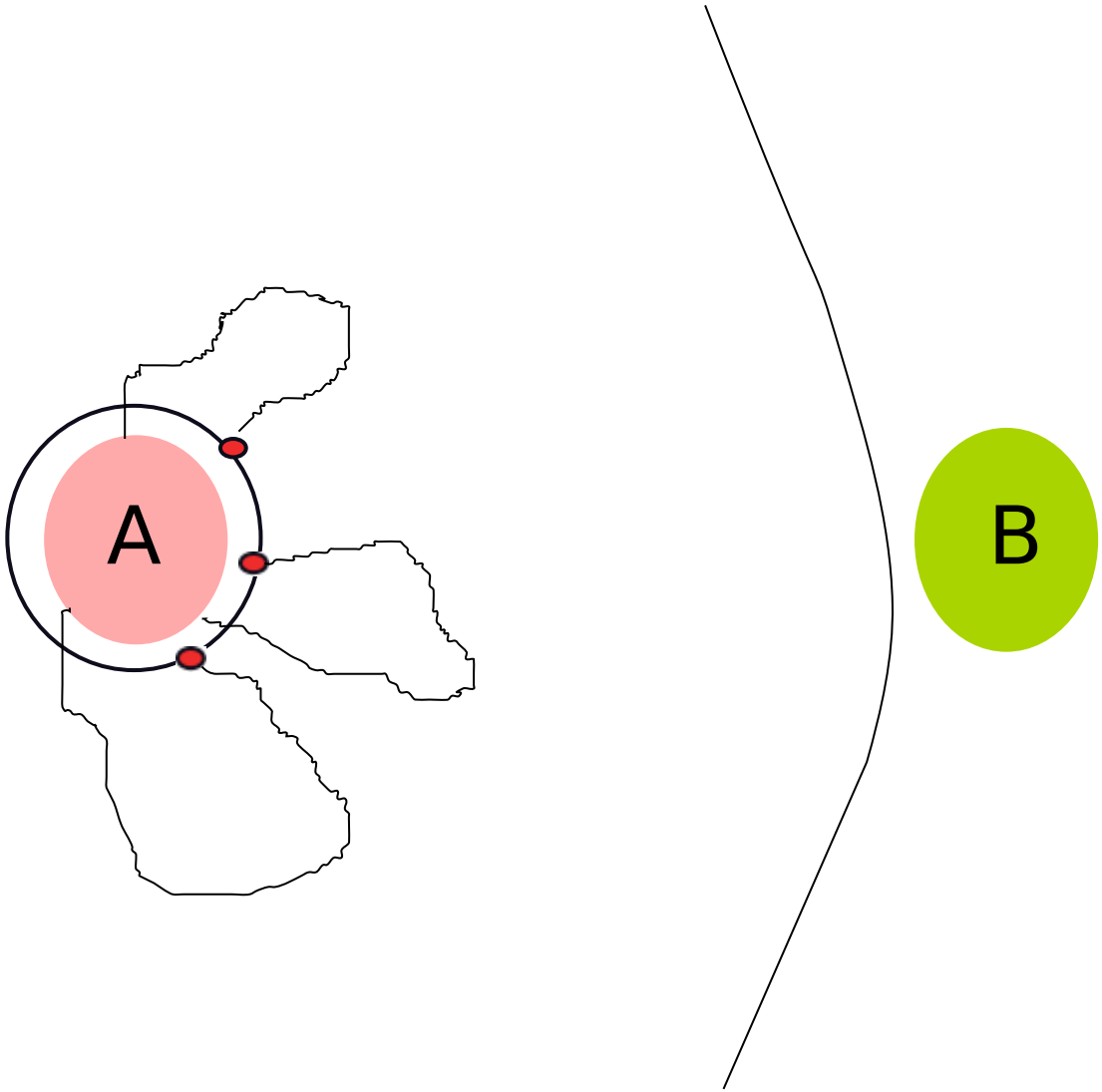
$$\hat{p} = \left(1 - \frac{k}{n}\right)^{Q_{\text{iter}}} \hat{p}_{\text{corr}}$$

where $\hat{p}_{\text{corr}}$ is the proportion of trajectories reaching B before A at the last iteration Q_{iter}.

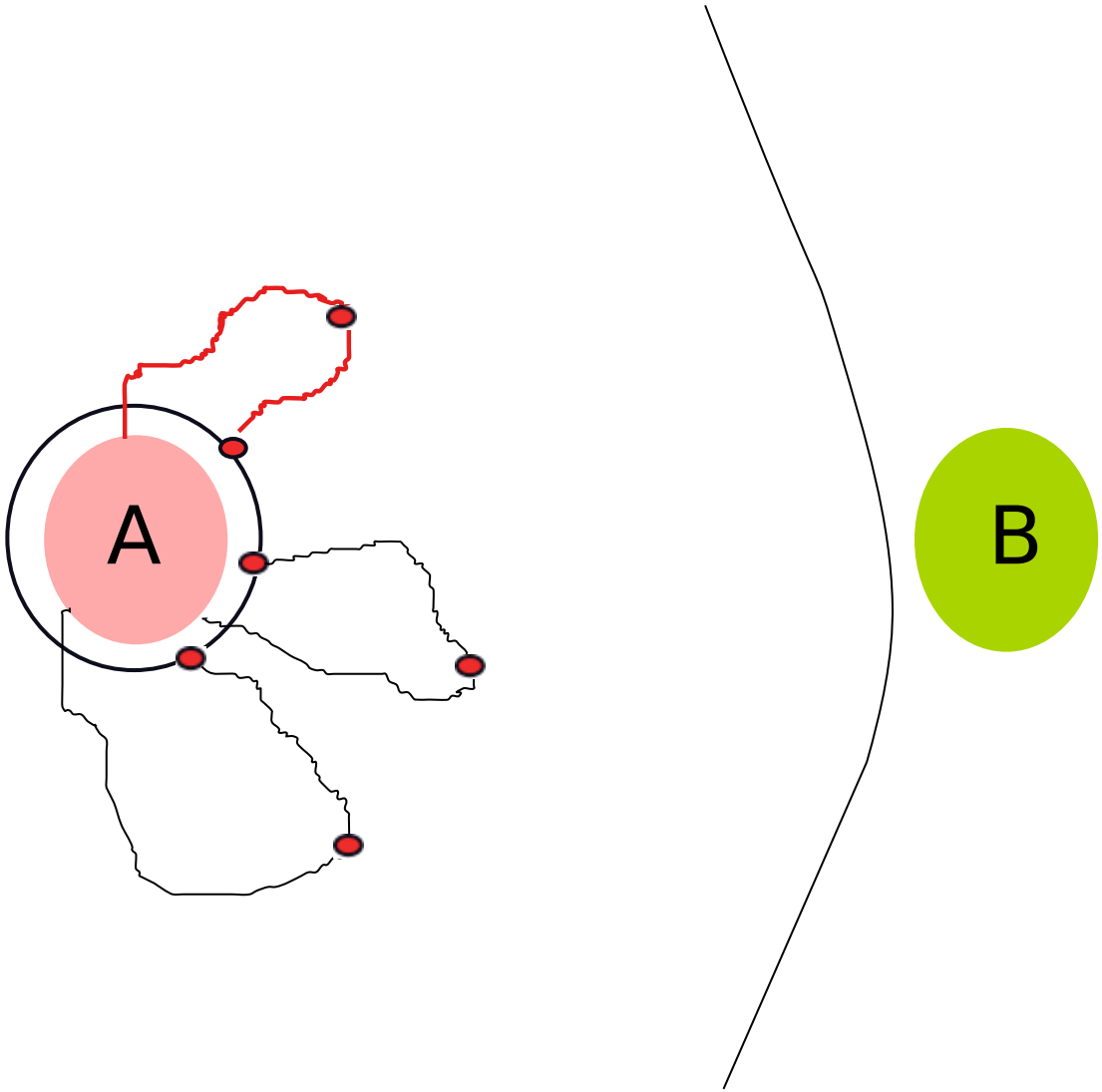
AMS Algorithm



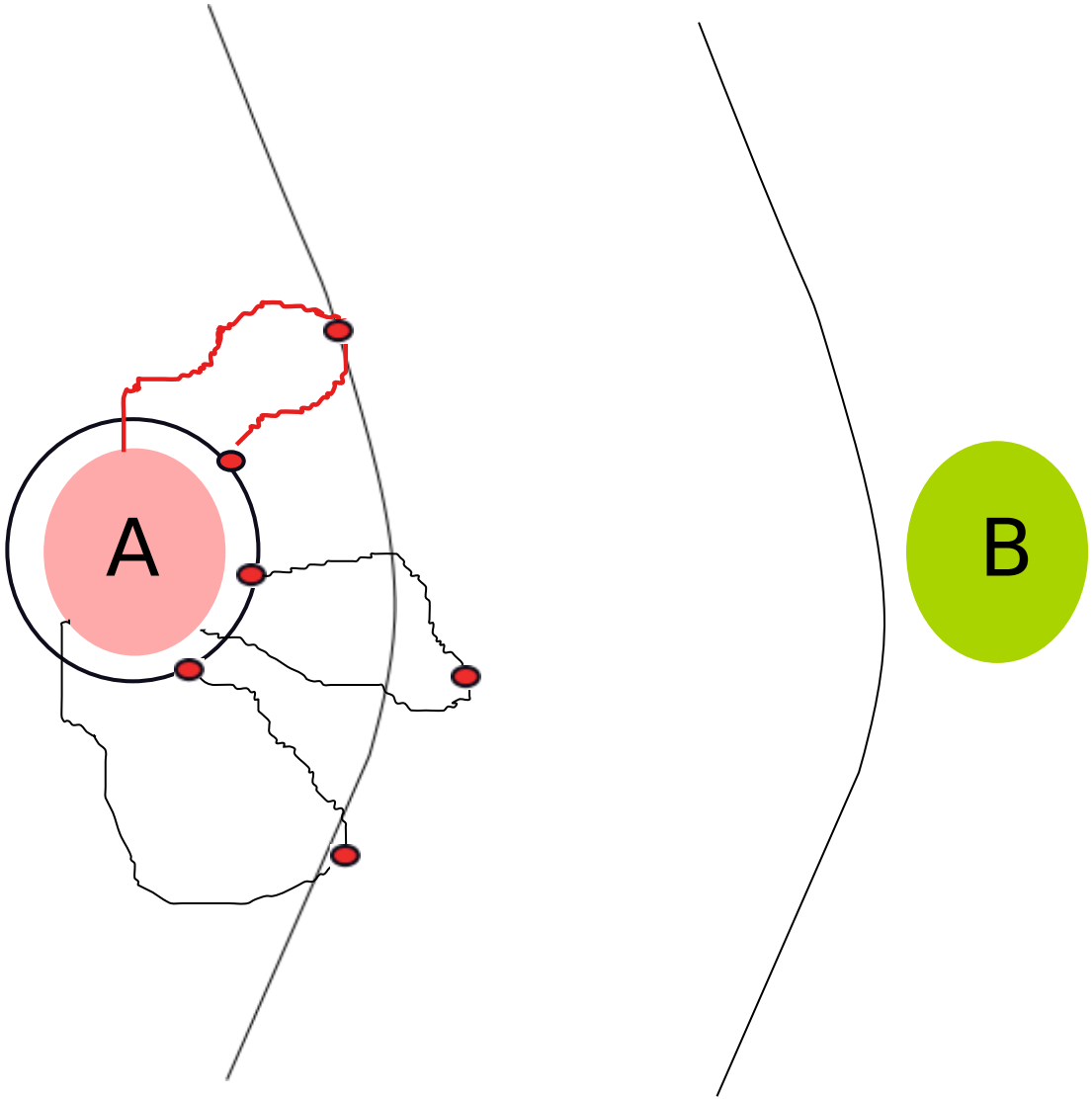
AMS Algorithm



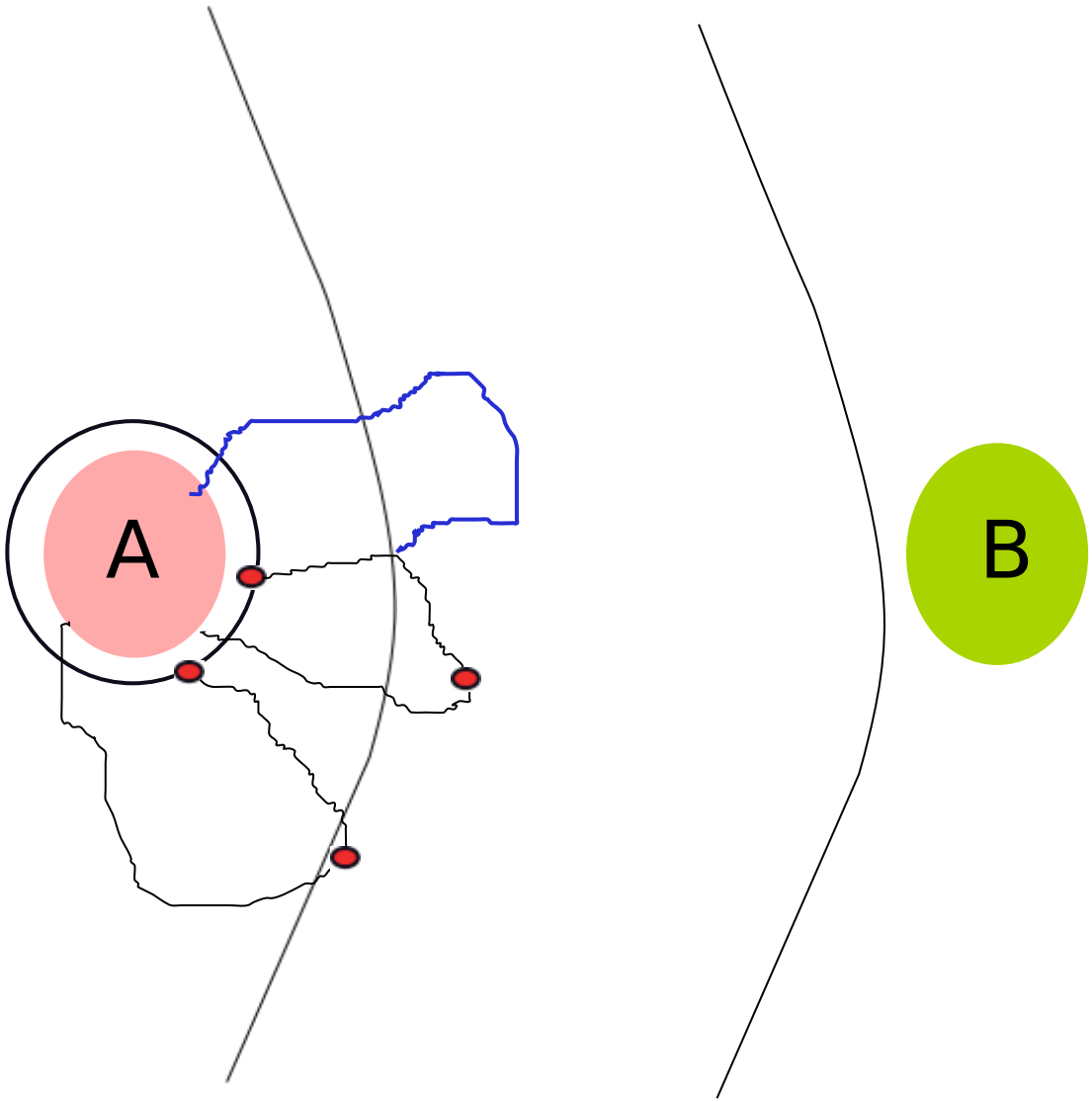
AMS Algorithm



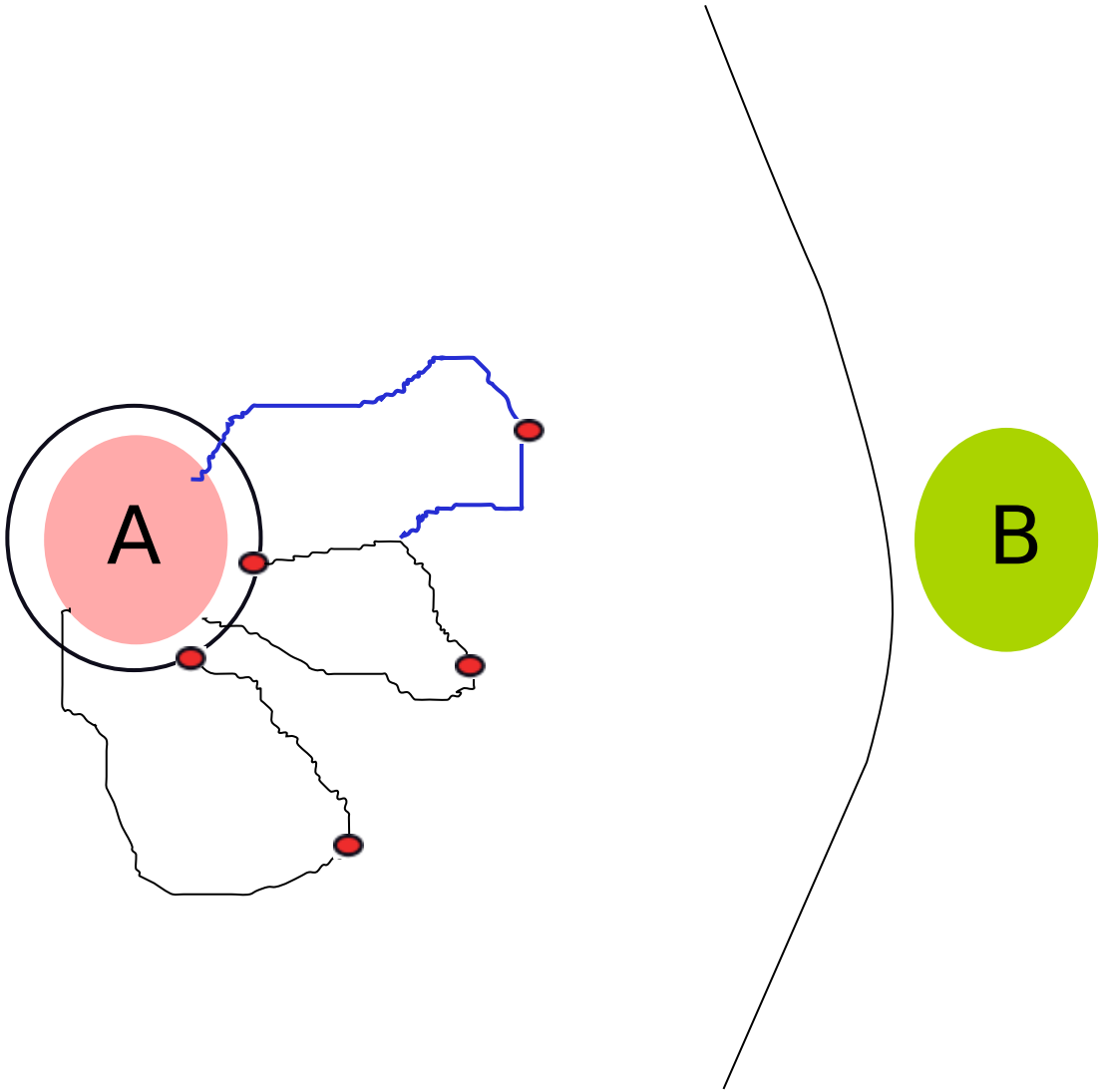
AMS Algorithm



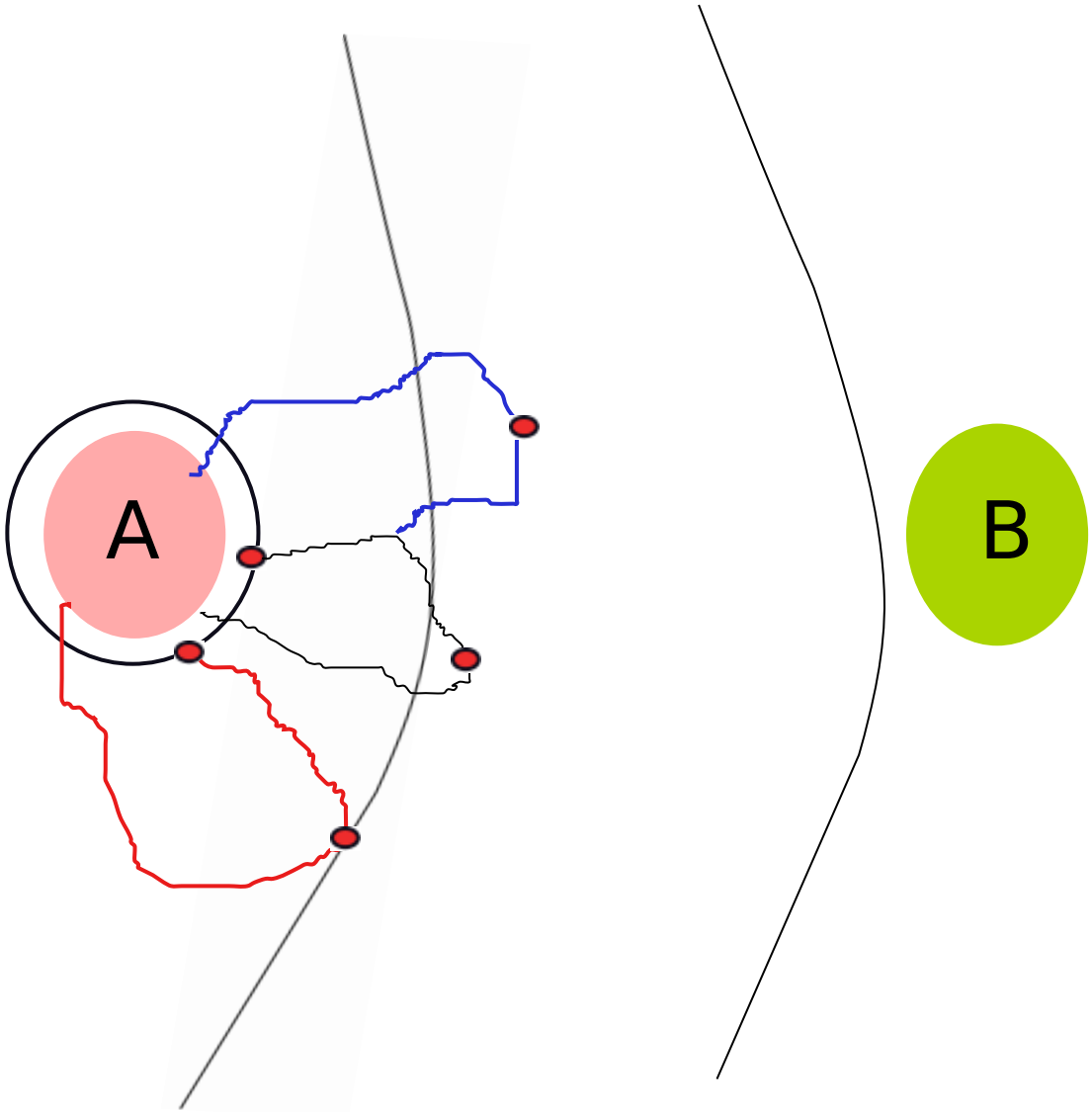
AMS Algorithm



AMS Algorithm



AMS Algorithm



AMS Algorithm: unbiasedness

In practice, the dynamics are *discrete in time* and this requires some adaptation of the AMS algorithm.

Theorem [C.-E. Bréhier, M. Gazeau, L. Goudenège, TL, M. Rousset, 2016]: For any choice of ξ, n and k,

$$\mathbb{E}(\hat{p}) = \mathbb{P}(\tau_B < \tau_A).$$

The proof is based on Doob's stopping theorem on a martingale built using filtrations indexed by the level sets of ξ. Actually, this result is proved for general path observables and in a much more general setting.

Practical counterparts:

- The algorithm is easy to parallelize.
- One can compare the results obtained with different reaction coordinates ξ to gain confidence in the results.

Results using NAMD

AMS can be used to

- analyze the ensemble of reactive paths (transition states, transition mechanisms, committor function)
- compute transition times

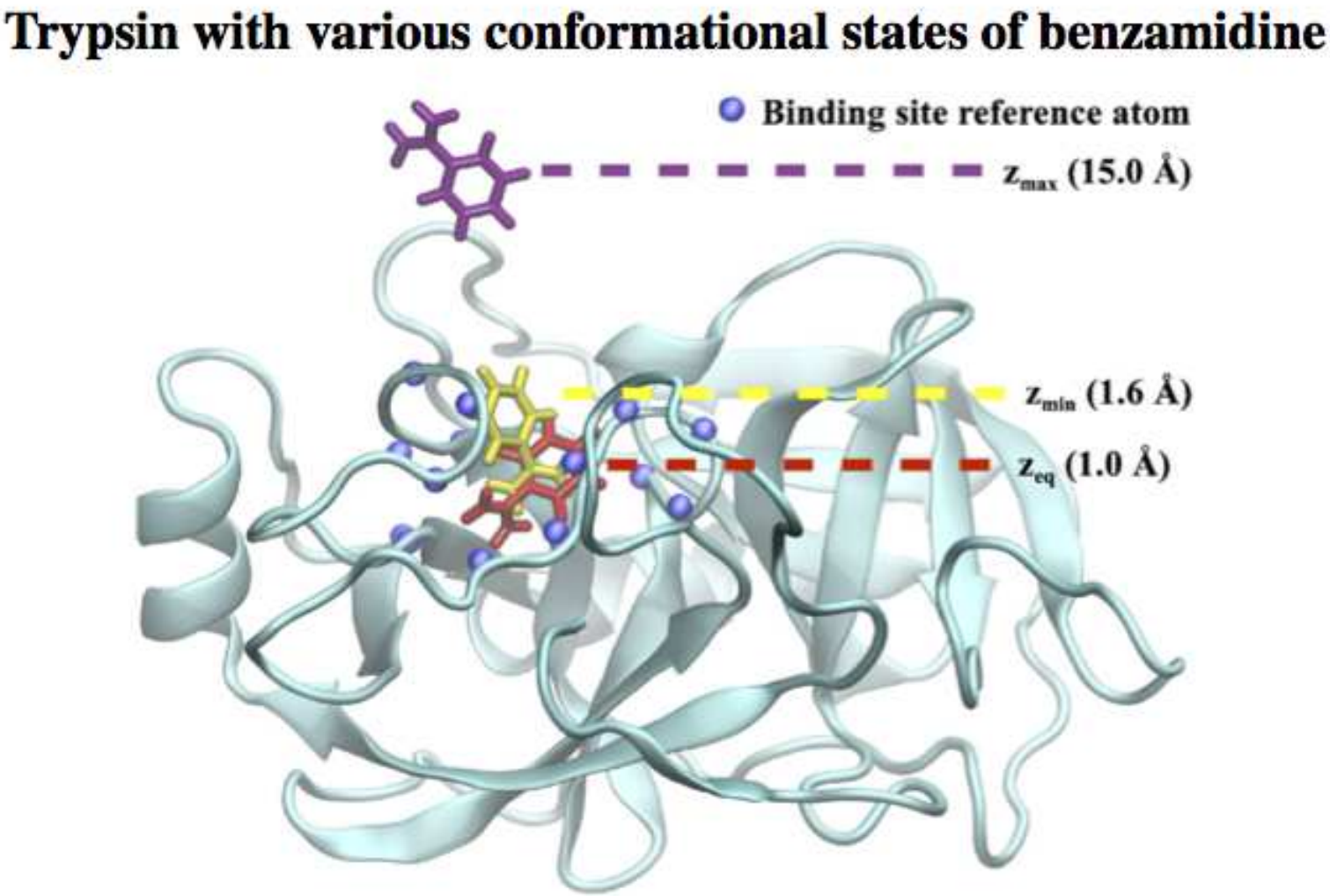
AMS is easy to implement (a pre-existing MD code can be used as a black-box). It is currently implemented in the NAMD software

(collaboration with SANOFI, C. Mayne and I. Teo, PhD of L. Lopes).

One example: computing the benzamidine-trypsin dissociation rate

Benzamidine-trypsin (1/2)

Estimation of the off rate of benzamidine from trypsin [I. Teo, C. Mayne, K. Schulten and TL, 2016].



Benzamidine-trypsin (2/2)

We obtain a dissociation rate $k_{\text{off}} = (260 \pm 240)s^{-1}$ within the same order of magnitude as the experimentally measured rate $(600 \pm 300)s^{-1}$.

The overall simulation time taken, summed over all 1000 replicas, was $2.1\mu s$ ($2.3\mu s$ after including direct MD and steered MD simulations), which is four orders of magnitude shorter than the estimated dissociation time of one event.

The main practical difficulty seems to be the determination of a 'good' domain A.

Computational setting: 68 789 atoms, with 21 800 water molecules, 62 sodium ions, and 68 chloride ions. Water: TIP3P model. CHARMM36 force field, with parameters for benzamidine obtained from the CGenFF force field. NPT conditions, at 298 K and 1 atm Langevin thermostat and barostat settings, using 2 fs time steps. AMS with $n = 1000$ replicas and $k = 1$.

Conclusion

We now have a good understanding of the formula which is used by many algorithms (FFS, AMS and the "source and sink methods": TIS, WE, milestoning) to compute the mean transition time:

- These methods are exact if the process is initialized in the initial state with the correct distribution: the reactive entrance distribution. This can be used in combination of the Hill relation to build unbiased estimators of the mean transition time for the Langevin dynamics.
- The reactive entrance distribution can be replaced by the QSD if A is metastable. This can be used in combination of the Hill relation to build estimators with controled bias if A is metastable.

Conclusion

Current research directions:

- We analyzed the bias, and not the variance. This should be possible, at least in simple prototypical cases, and maybe give some hints on good choices for some numerical parameters (e.g. position of Σ).
- In practice, it is observed that the initial conditions that indeed yield a transition to B are concentrated on some parts of the boundary ∂A. We are currently working on good sampling methods for these initial conditions [L. Lopes, T. Pigeon].

In practice, obtaining a precise result requires a throrough sampling of ν_Q.

References

- F. Cérou, A. Guyader, and D. Pommier, *A multiple replica approach to simulate reactive trajectories*, Journal of Chemical Physics, 2011.
- C.-E. Bréhier, M. Gazeau, L. Goudenège, TL , and M. Rousset, *Unbiasedness of some generalized Adaptive Multilevel Splitting algorithms*, Annals of Applied Probability, 2016.
- I. Teo, C. Mayne, K. Schulten, and TL, *Adaptive multilevel splitting method for molecular dynamics calculation of benzamidine-trypsin dissociation time*, Journal of Chemical Theory and Computation, 2016.
- M. Baudel, A. Guyader and TL, *On the Hill relation and the mean reaction time for metastable processes*, Stochastic Processes and their Applications, 2023.
- TL, M. Ramil and J. Reygner, *Estimation of statistics of transitions and Hill relation for Langevin dynamics*, to appear in Annales de l'Institut Henri Poincaré `https://arxiv.org/abs/2206.13264`