Numerical methods to overcome metastability in molecular dynamics

T. Lelièvre

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

Joint work with F. Cérou, A. Guyader, C. Le Bris, M. Luskin, D. Perez and D. Pommier. Special thanks to A. Voter.

Beyond Molecular Dynamics:
Long Time Atomic-Scale Simulations
March 2012

Introduction

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 to evaluate numerically macroscopic quantities from models at the microscopic scale.

Some examples of macroscopic quantities:

(i) Thermodynamics quantities (average of some observable wrt an equilibrium measure): stress, heat capacity, free energy,...

$$\mathbb{E}_\mu(\varphi(\boldsymbol{X})) = \int_{\mathbb{R}^d} \varphi(\boldsymbol{x})\, \mu(d\boldsymbol{x}).$$

(ii) Dynamical quantities (average over trajectories at equilibrium): diffusion coefficients, viscosity, transition rates,...

$$\mathbb{E}(\mathcal{F}((\boldsymbol{X}_t)_{t\geq 0})) = \int_{\mathcal{C}^0(\mathbb{R}_+,\mathbb{R}^d)} \mathcal{F}((\boldsymbol{x}_t)_{t\geq 0}))\, \mathcal{W}(d((\boldsymbol{x}_t)_{t\geq 0})).$$

Introduction

A molecular dynamics model amounts essentially in choosing a potential V which associates to a configuration $(x_1, ..., x_N) = x \in \mathbb{R}^{3N}$ an energy $V(x_1, ..., x_N)$.

In the canonical (NVT) ensemble, configurations are distributed according to the Boltzmann-Gibbs probability measure:

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

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

Difficulties: (i) high-dimensional problem ($N \gg 1$) ; (ii) μ is a multimodal measure.

Introduction

To sample μ, ergodic dynamics wrt to μ are used. A typical example is the *over-damped Langevin* (or gradient) dynamics:

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

It is the limit (when the mass goes to zero or the damping parameter to infinity) of the *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 M is the mass tensor and γ is the friction coefficient.

To compute dynamical quantities, these are also typically the dynamics of interest. Thus,

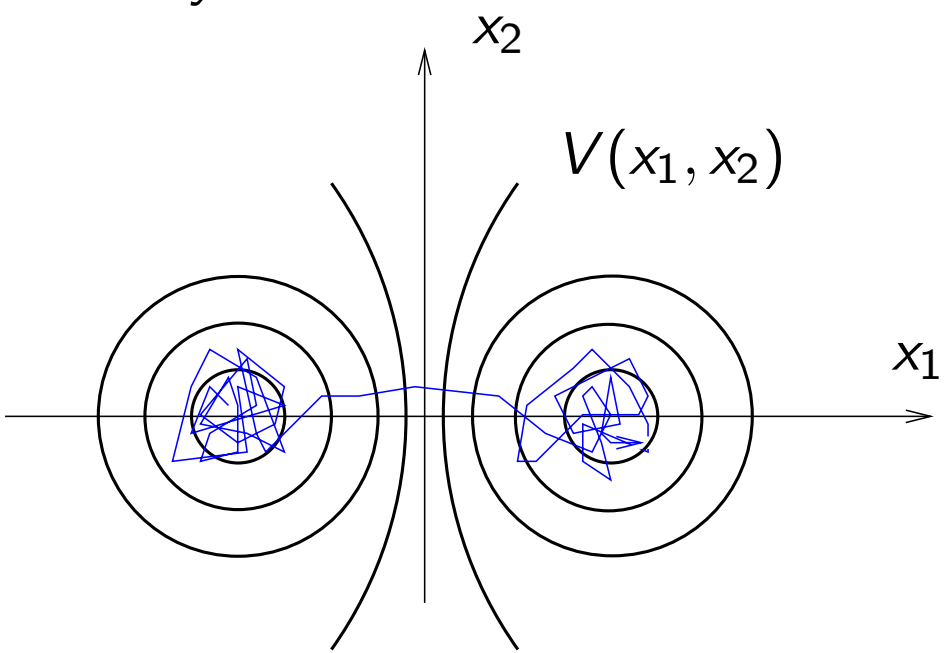
$$\mathbb{E}_\mu(\varphi(\boldsymbol{X})) \simeq \frac{1}{T}\int_0^T \varphi(\boldsymbol{X}_t)\, dt \text{ and } \mathbb{E}(\mathcal{F}((\boldsymbol{X}_t)_{t\geq 0})) \simeq \frac{1}{N}\sum_{m=1}^{N} \mathcal{F}((\boldsymbol{X}_t^m)_{t\geq 0})$$

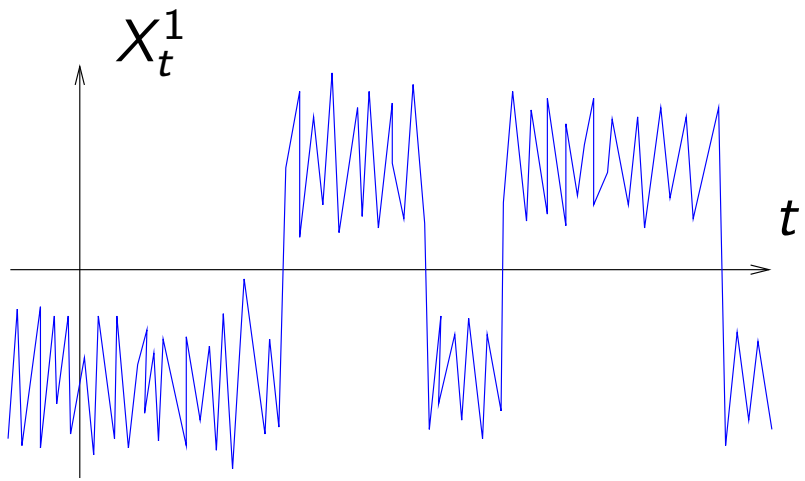
In the following, we mainly consider the over-damped Langevin dynamics.

Introduction

Difficulty: In practice, $\boldsymbol{X}_t$ is a metastable process, so that the convergence to equilibrium is very slow, and sampling metastable trajectories is very difficult.

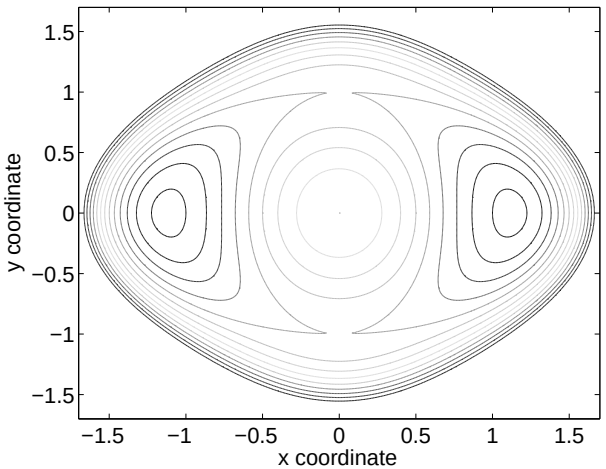
A 2d schematic picture: X_t^1 is a slow variable (a metastable dof) of the system.

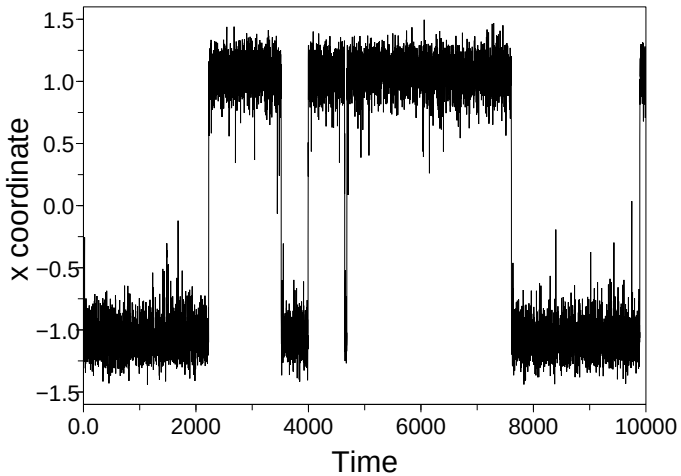




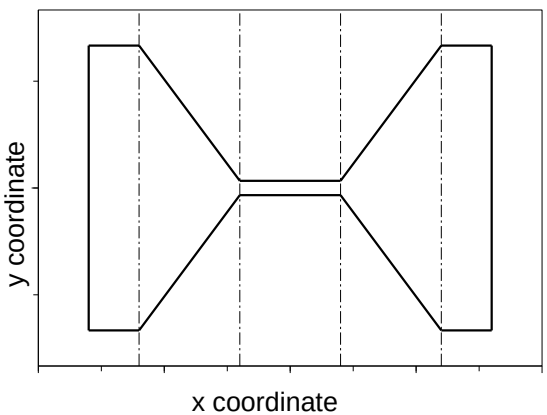
Introduction

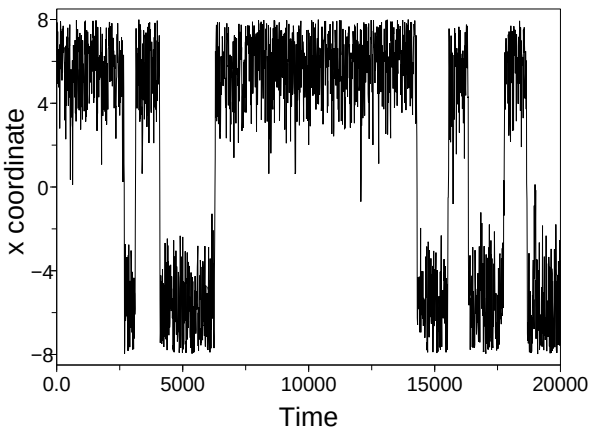
Where does metastability come from ?





Energetic barrier.





Entropic barrier.

Introduction

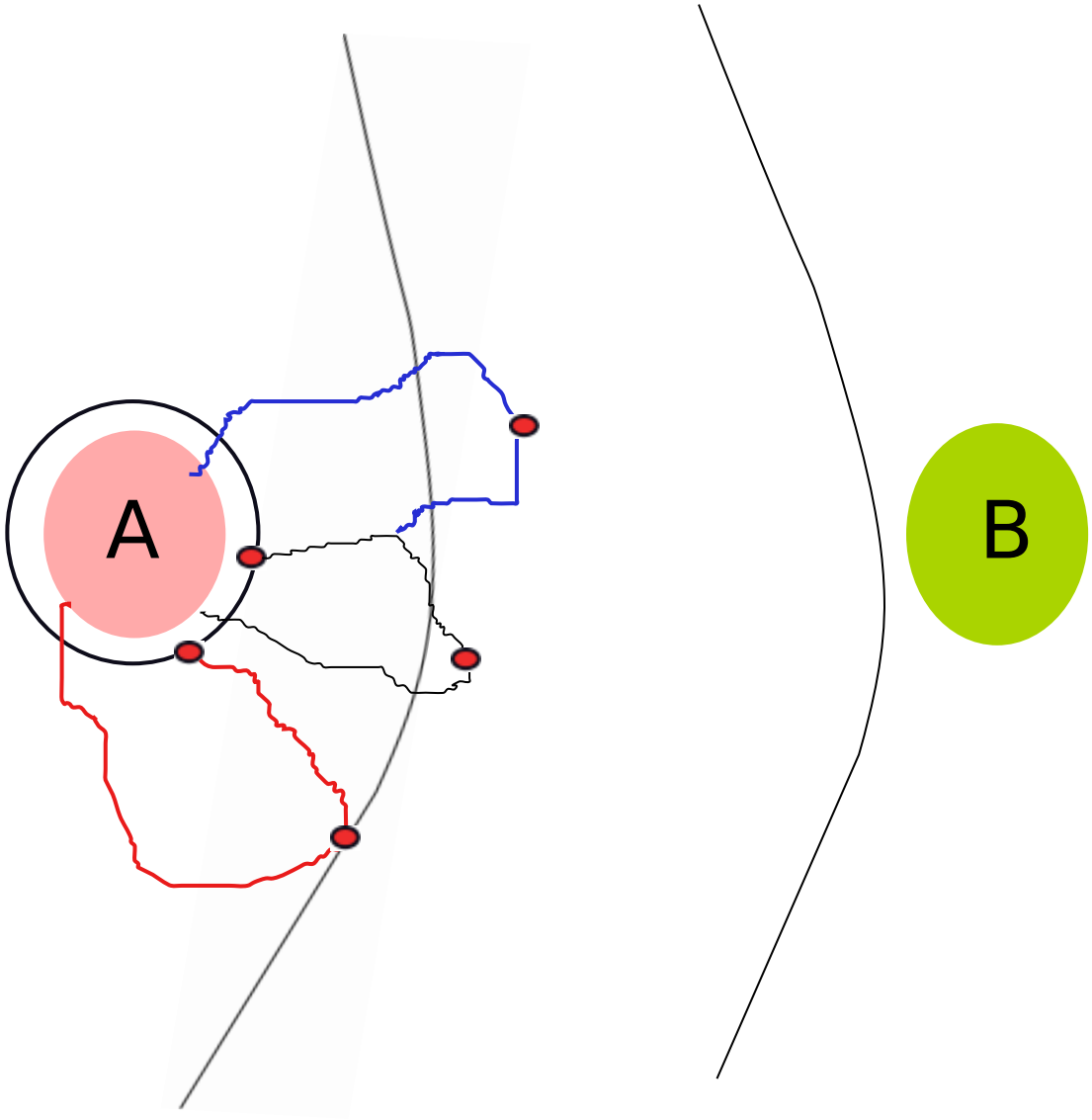
For computing thermodynamics quantities, there is a clear classification of available methods, and the difficulties are now well understood (in particular for free energy computations, see for example [TL, Rousset, Stoltz, 2010]). On the opposite, computing efficiently dynamical quantities remains a challenge.

In this talk, we would like to discuss two methods to sample metastable dynamics and compute dynamical quantities:

1. Adaptive multilevel splitting method: A new algorithm related to the forward flux sampling or the interface sampling method.
2. The Parallel Replica dynamics: An algorithm proposed by A. Voter in 1998 for which we propose a mathematical analysis.

There are many other techniques: hyperdynamics and temperature accelerated dynamics [Voter, Fichthorn], the string method [E, Ren, Vanden-Eijnden], transition path sampling methods [Chandler, Bolhuis, Dellago], milestoning techniques [Elber, Schuette, Vanden-Eijnden], etc...

Splitting strategies



Multilevel splitting

We would like to sample trajectories between two given metastable states A and B. The main assumption in this section is that we are given a smooth one dimensional function $\xi : \mathbb{R}^d \to \mathbb{R}$ (s.t. $|\nabla \xi| \neq 0$) which "indexes" the transition from A to B in the following sense:

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

where $z_{\min} < z_{\max}$, and $\Sigma_{z_{\min}}$ (resp. $\Sigma_{z_{\max}}$) is "close" to ∂A (resp. ∂B).

Example: $\xi(x) = \|x - x_A\|$ where $x_A \in A$ is a reference configuration in A.

Multilevel splitting

Question: How to compute dynamical quantities using ξ ? More precisely, we consider: (a) Reactive trajectories and (b) Transition times between the two metastable states A and B.

We propose a multilevel splitting approach [Kahn, Harris, 1951] [Rosenbluth, 1955] to discard failed trajectories and branch trajectories approaching the rare set. We focus on an adaptive variant [Cérou, Guyader, 2007] [Cérou, Guyader, TL, Pommier, 2010]: the Adaptive Multilevel Splitting (AMS) algorithm.

Reactive trajectory

A reactive trajectory between two metastable sets A and B is a piece of equilibrium trajectory that leaves A and goes to B without going back to A in the meantime [Hummer,2004] [Metzner, Schütte, Vanden-Eijnden, 2006].

Difficulty: A trajectory leaving A is more likely to go back to A than to reach B.

AMS Algorithm

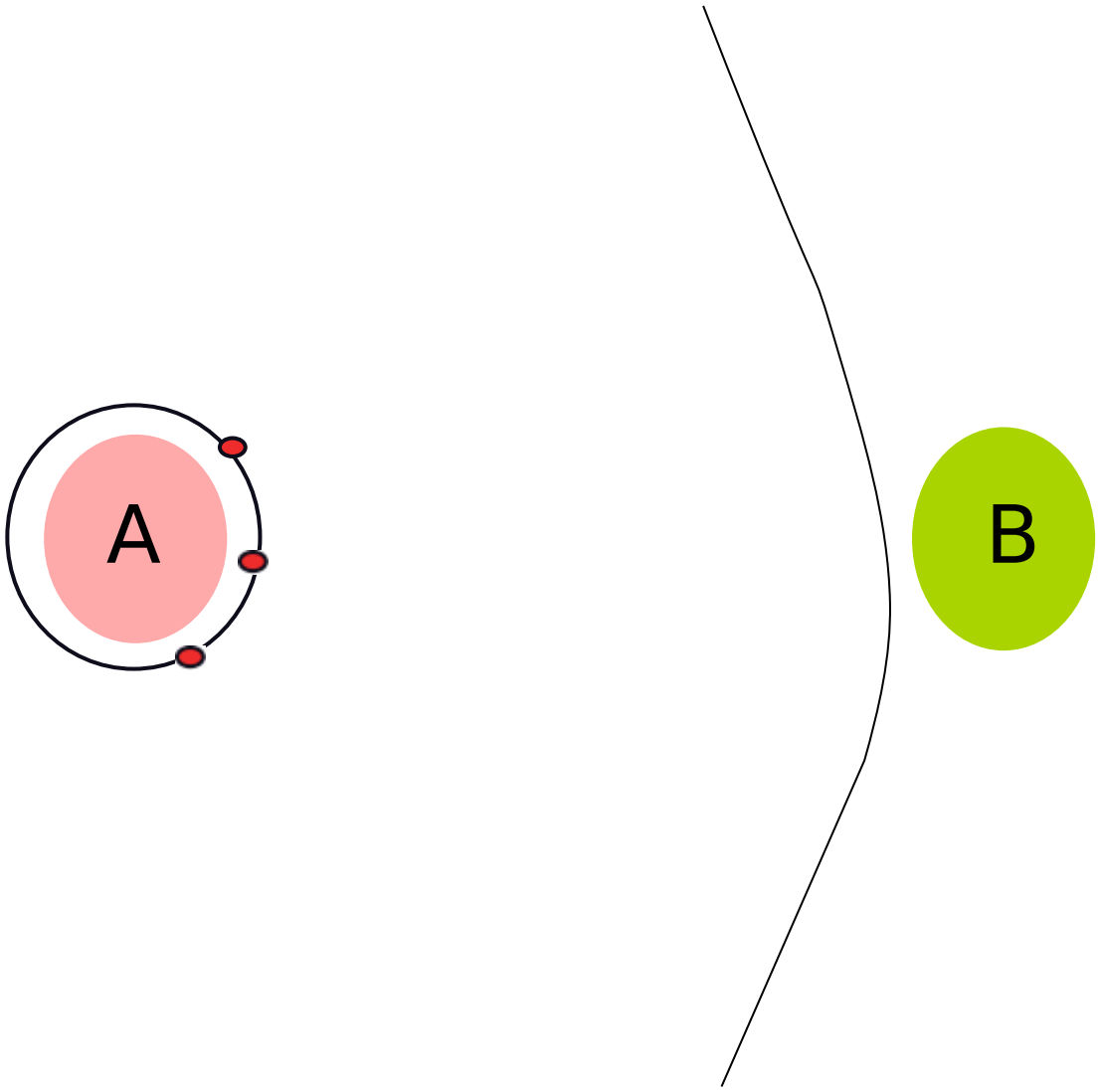
Initialisation: Generate an ensemble of N equilibrium trajectories starting from A, up to the time it reaches A or B, conditionnally to the fact that $\sup_{t\in(0,\tau^n)} \xi(X_t^n) > z_{\min}$. This is easily done by DNS.

Algorithm: (i) Order the $z^n = \sup_{t\in(0,\tau^n)} \xi(X_t^n)$; (ii) Kill the trajectory with the smallest supremum (say z^{n_0}) ; (iii) Create a new trajectory by branching another trajectory from the first time it crosses $\Sigma_{z^{n_0}}$; Go back to (i) until z^{n_0} is larger than $z_{\max}$.
This generates an ensemble of N equilibrium trajectories starting from A, up to the time it reaches A or B, conditionnally to the fact that $\sup_{t\in(0,\tau^n)} \xi(X_t^n) > z_{\max}$.

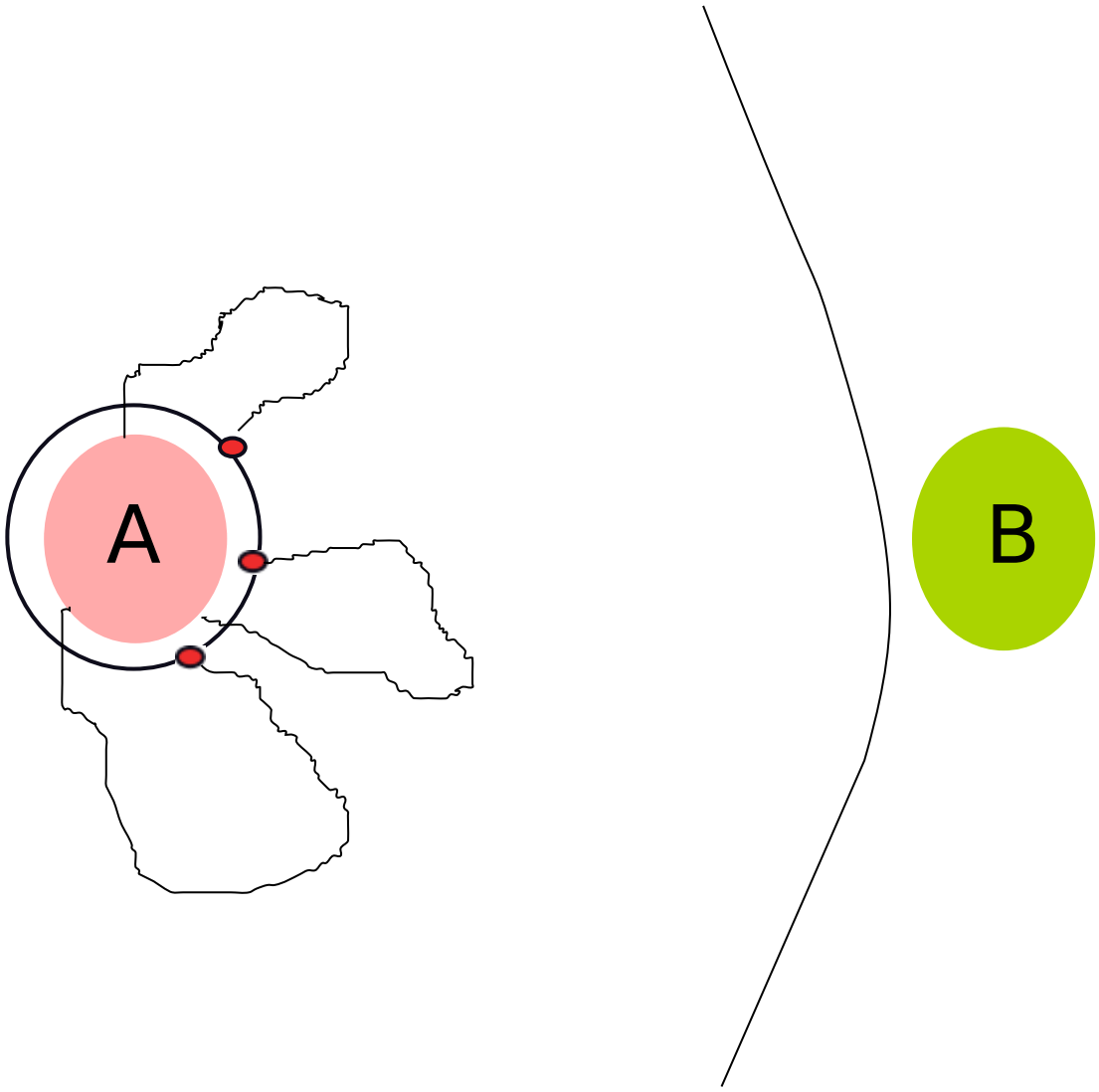
Final step: To get reactive trajectories, one only retains paths which indeed end in B, and the part of the trajectory between A and B.

Remark: The algorithm can be seen as a kind of adaptive Forward Flux Sampling [Allen, Valeriani, Ten Wolde, 2009]. It is also related to the Interface Sampling Method [Bolhuis, van Erp, Moroni 2003] and the Milestoning method [Elber, Faradjian 2004]. See the review paper [Bolhuis, Dellago, 2009]

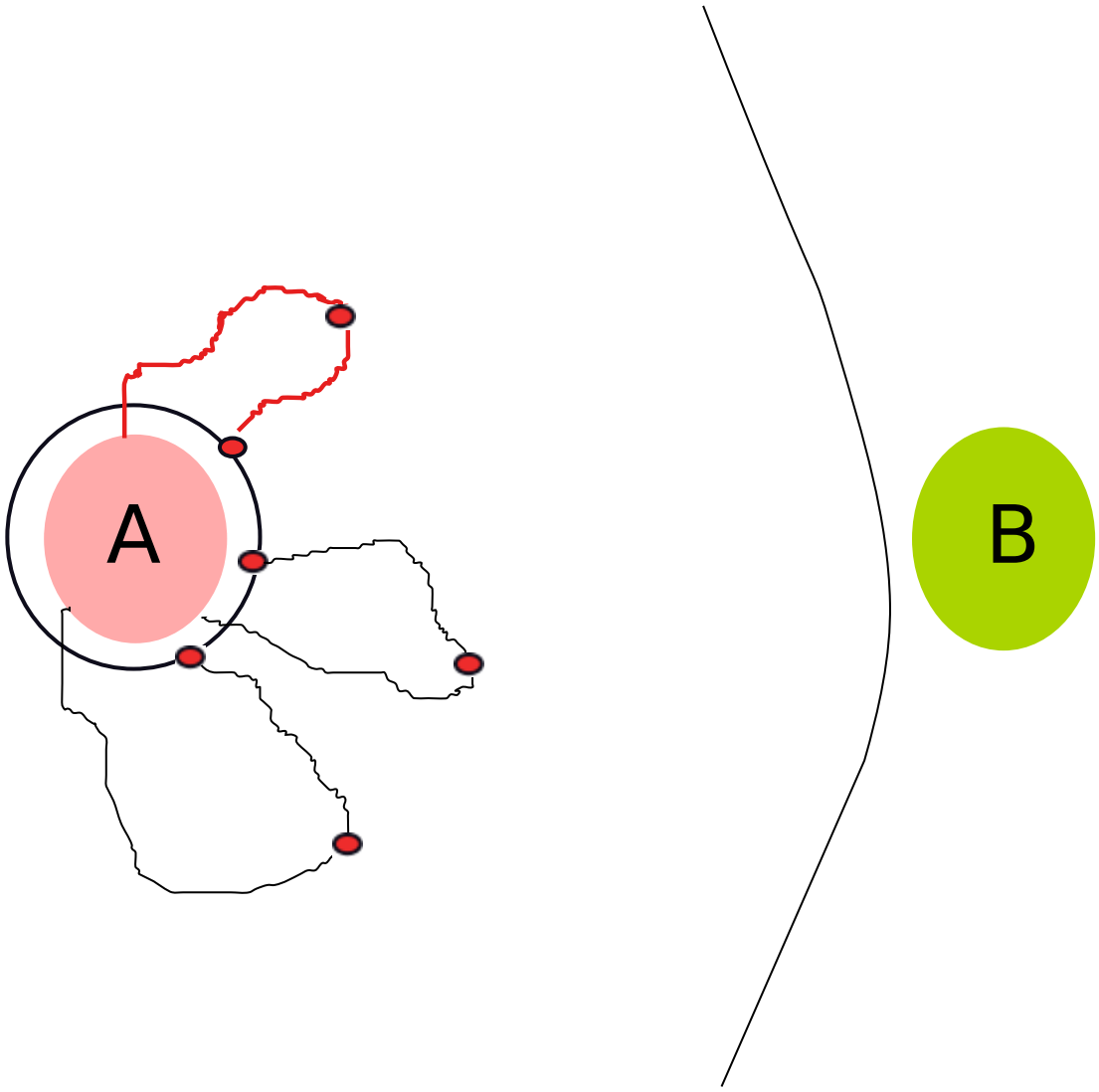
AMS Algorithm



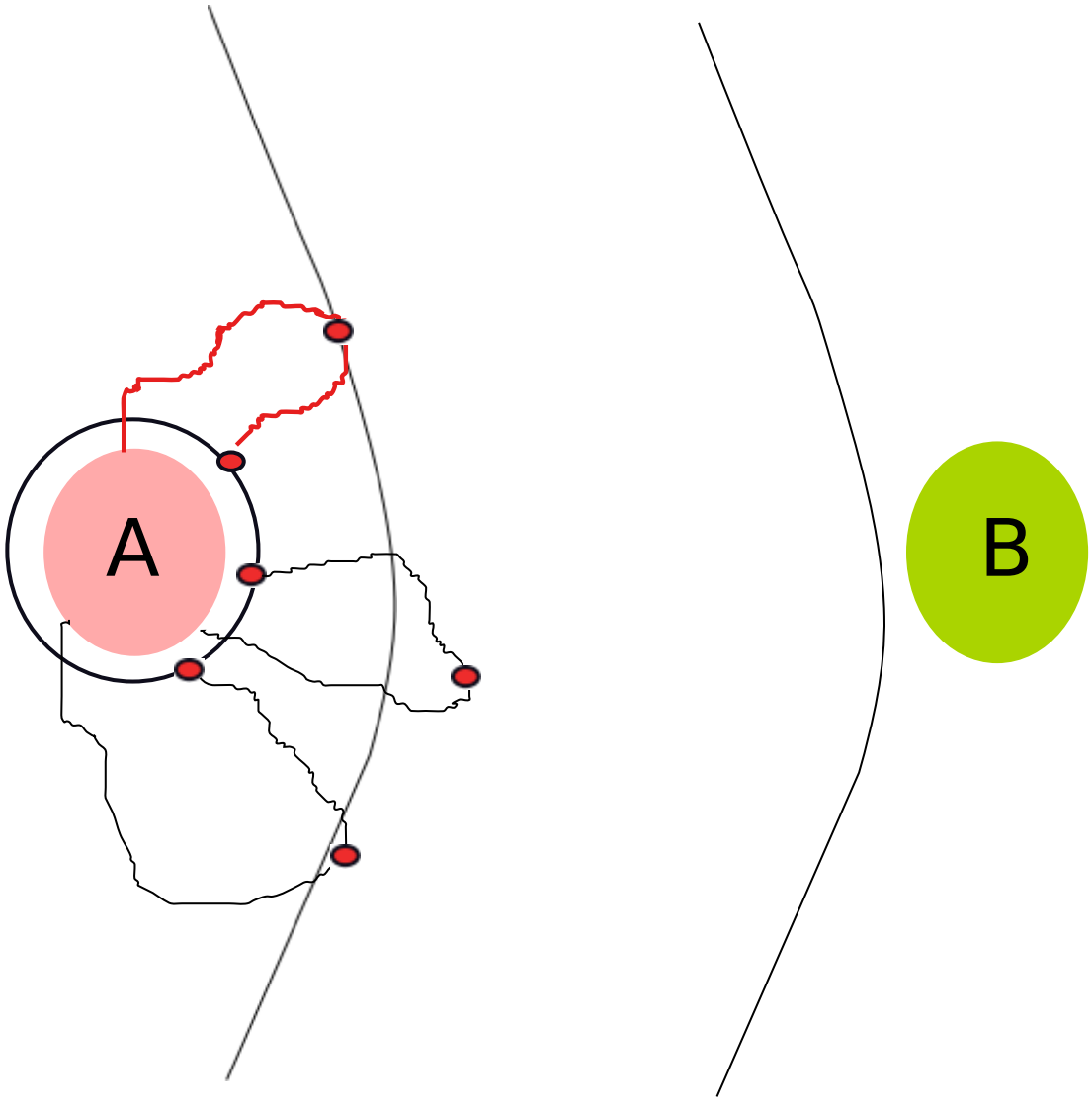
AMS Algorithm



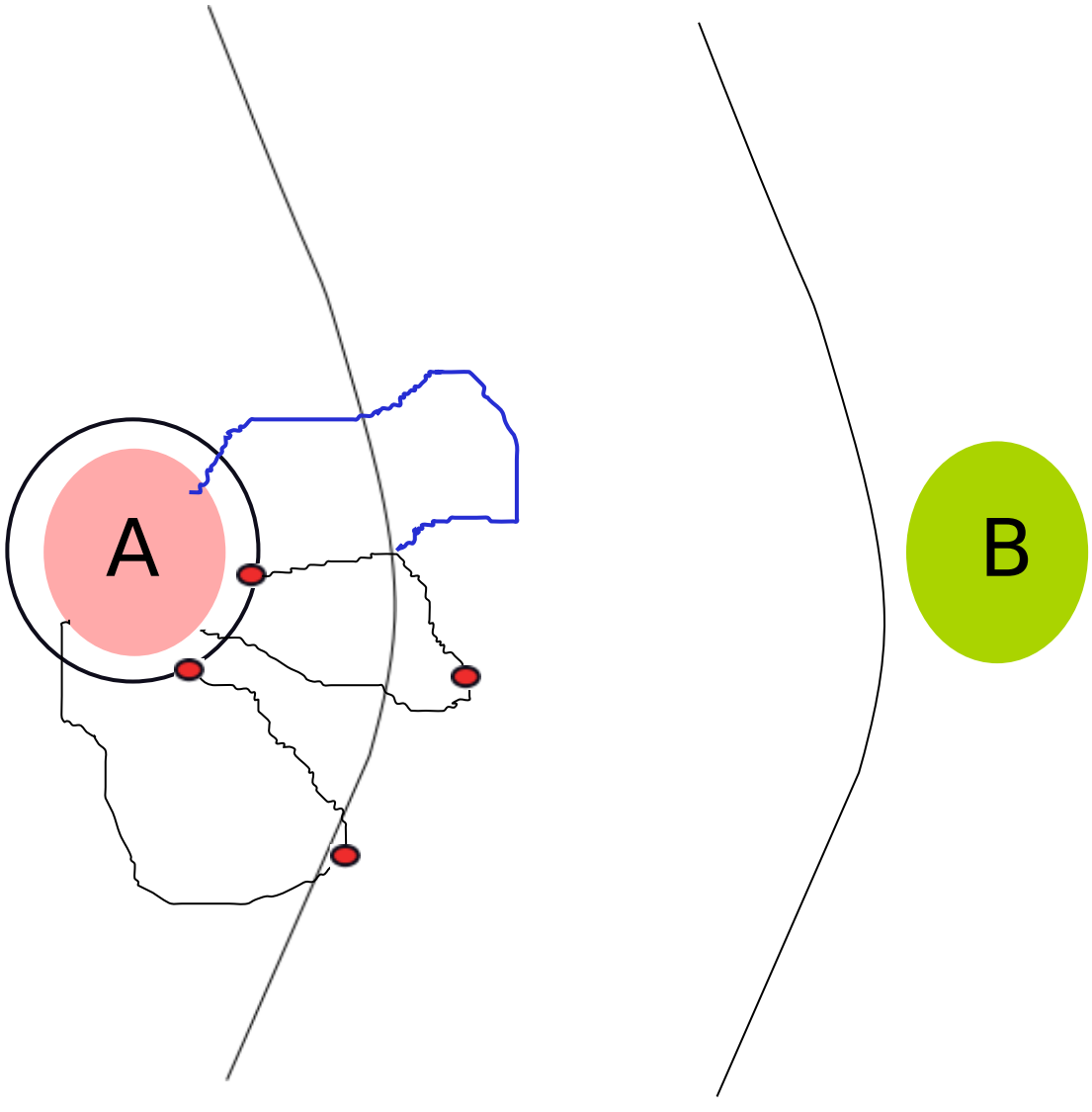
AMS Algorithm



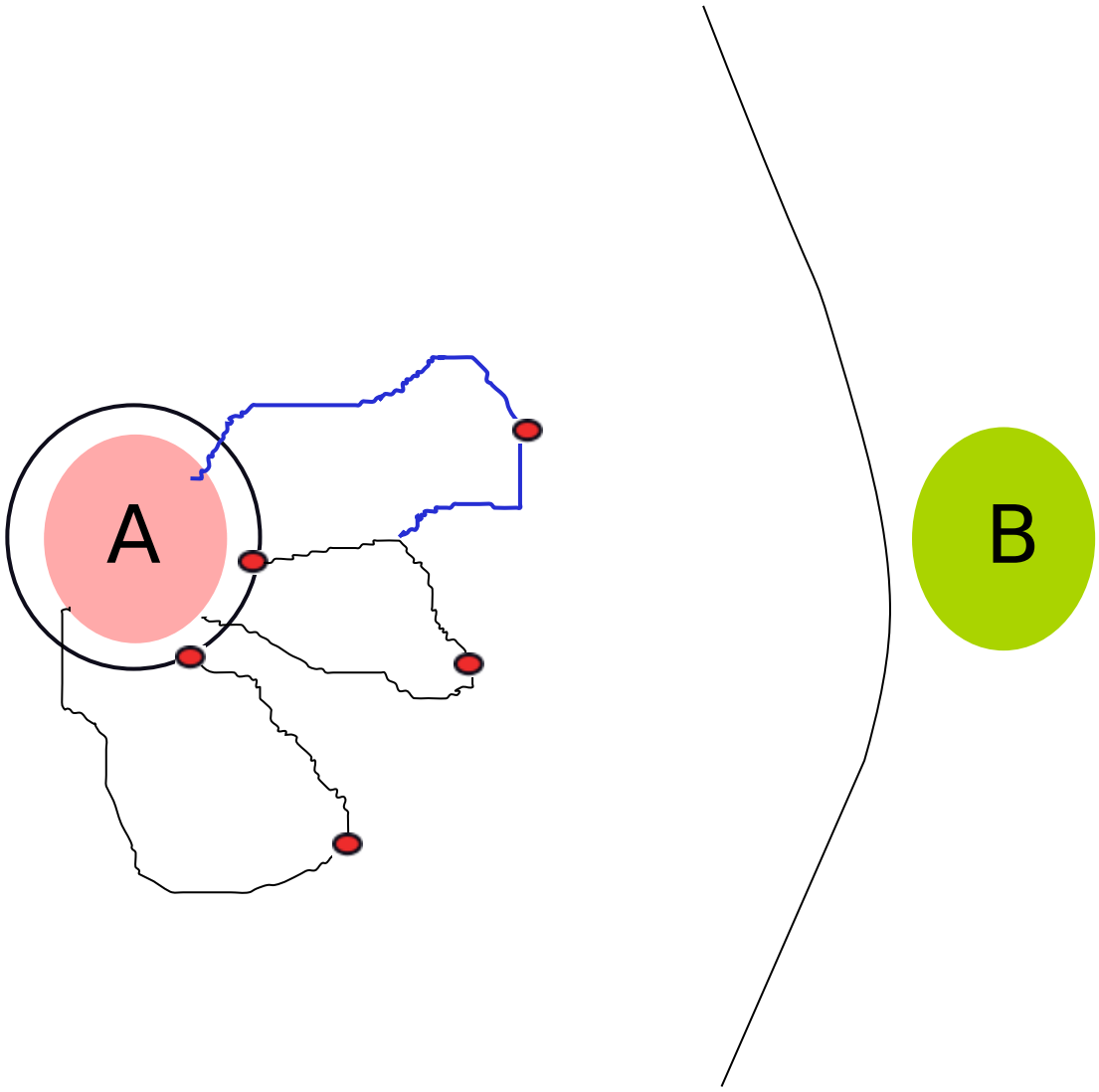
AMS Algorithm



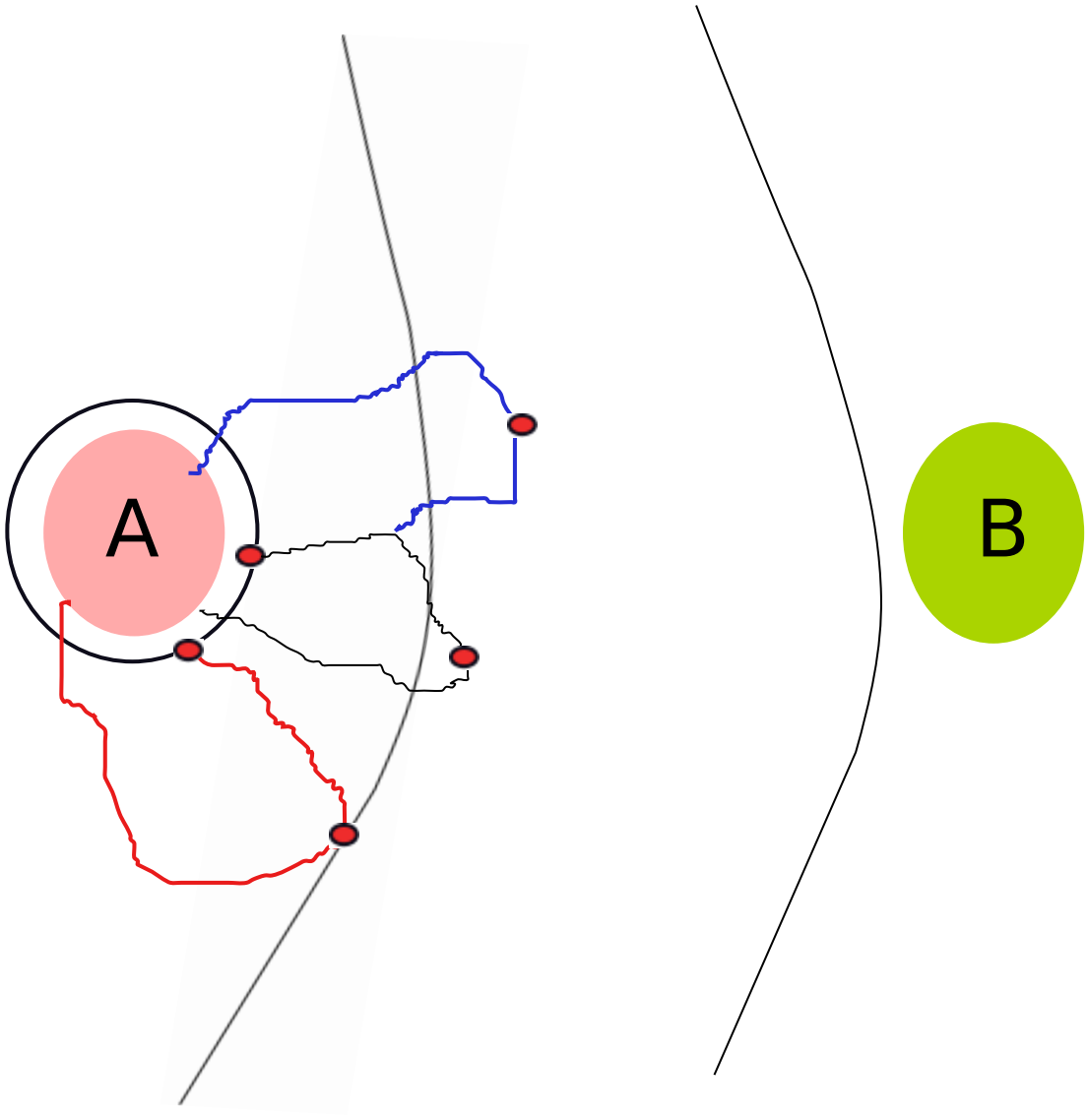
AMS Algorithm



AMS Algorithm



AMS Algorithm



AMS Algorithm

An important output of the algorithm is

$$\hat{\alpha}_N = (1 - 1/N)^{k_{\max}}\, r$$

where N is the number of trajectories, $k_{\max}$ the number of iterations to reach the quantile $z_{\max}$, and r the proportion of trajectories that do finally end in B and not in A (at the final step).

The probability $\hat{\alpha}_N$ is an estimator of the probability for an equilibrium trajectory leaving A and reaching $\Sigma_{z_{\min}}$ to reach B before A. It may be interprated as a "probability of observing a reactive trajectory".

Values for $\hat{\alpha}_N$ in the numerical results below vary between 10^{-18} and 10^{-4} depending on the test cases: these are very rare events.

Numerical results

In all the numerical tests, we use overdamped dynamics, but the algorithm applies to any stochastic Markovian dynamics, is straightforward to parallelize, and requires only a small modification of an original MD code.

A 1D example: We consider the double-well potential:

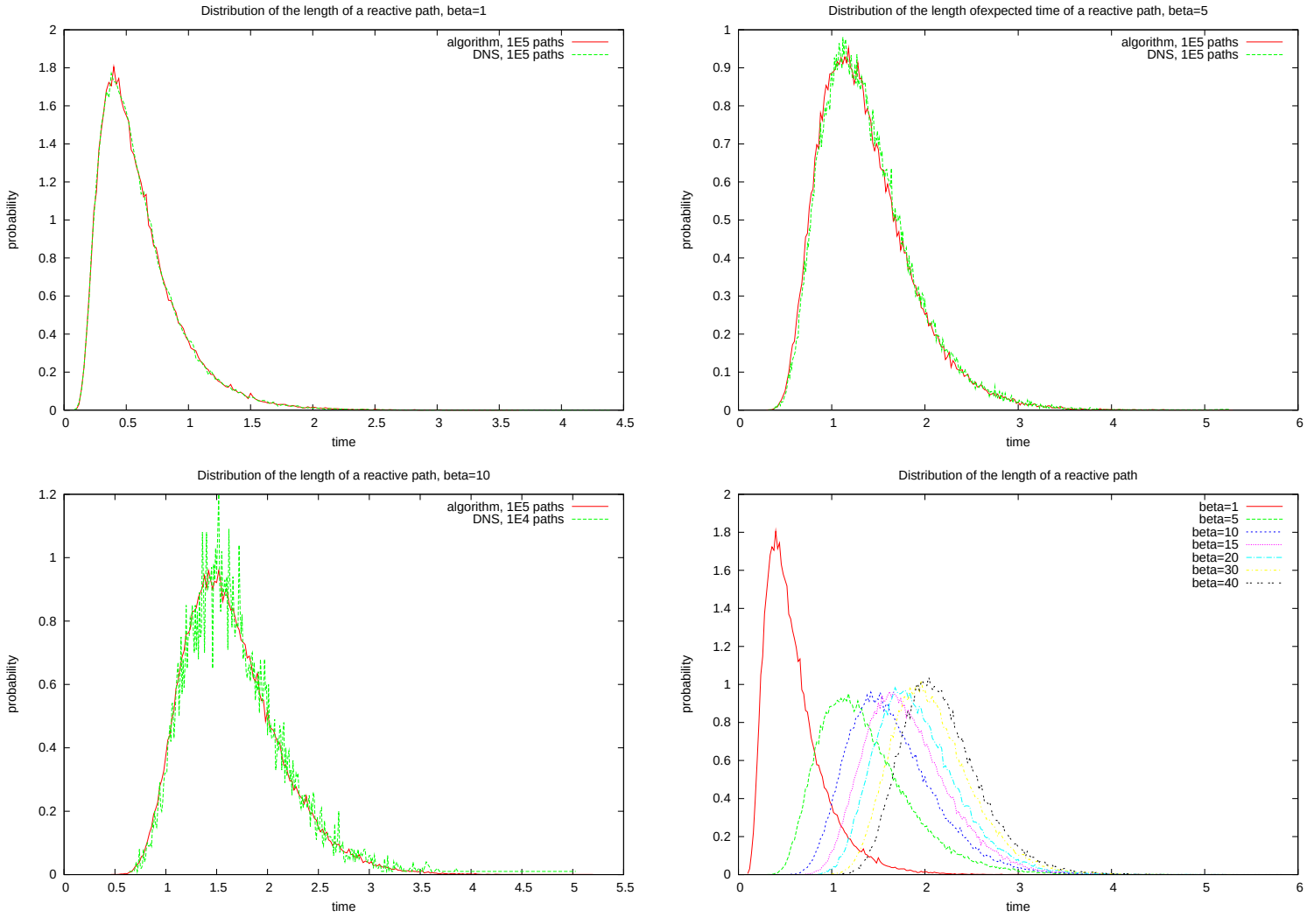
$$V(x) = x^4 - 2x^2,$$

which has two minima at ± 1 and one saddle point at 0.
In this simple one dimensional setting, we set as metastable states $A = \{-1\}$ and $B = \{+1\}$, and the reaction coordinate is taken to be simply

$$\xi(x) = x.$$

To test the consistency of the algorithm, we compute the distribution of the time-lengths of the reactive paths and compare to DNS (when possible).

A 1D example



Distributions of time-lengths of reactive paths. Comparison with DNS for $\beta = 1, 5, 10$ and distributions for $\beta = 1, 5, 10, 15, 20, 30, 40$.

A 1D example

β	N	$\hat{\alpha}_N$	DNS CPU	AMS CPU
1	10^4	$1.03\,10^{-1}$	2s	2s
1	10^5	$1.01\,10^{-1}$	21s	1 min 19 s
10	10^4	$2.04\,10^{-5}$	140 min 05 s	5 s
10	10^5	$1.98\,10^{-5}$	1400 min $\star$	5 min 22 s
15	10^5	$1.78\,10^{-7}$	92000 min $\star$	7 min 52 s
20	10^5	$1.33\,10^{-9}$		8 min 36 s
40	10^5	$5.82\,10^{-18}$		10 min 09 s

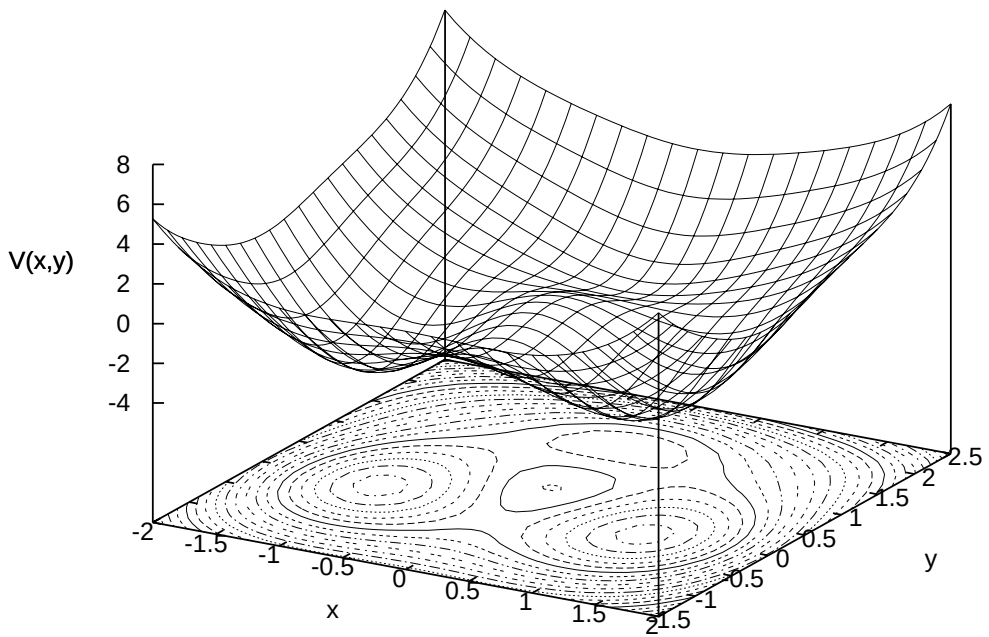
Probability $\hat{\alpha}_N$, and computational time for different values of β and numbers of paths N. DNS CPU time with $\star$ is an extrapolated time deduced from a small number of generated reactive trajectories.

The algorithm makes possible the generation of reactive trajectories for some parameter values for which DNS would be impractible.

A 2D example

Let us consider the potential ([Park, Sener, Lu, Schulten, 2003] [Metzner, Schütte and Vanden-Eijnden, 2006]):

$$V(x,y) = 3\mathrm{e}^{-x^2-(y-\frac{1}{3})^2} - 3\mathrm{e}^{-x^2-(y-\frac{5}{3})^2} - 5\mathrm{e}^{-(x-1)^2-y^2} - 5\mathrm{e}^{-(x+1)^2-y^2} + 0.2x^4 + 0.2\left(y-\frac{1}{3}\right)^4.$$



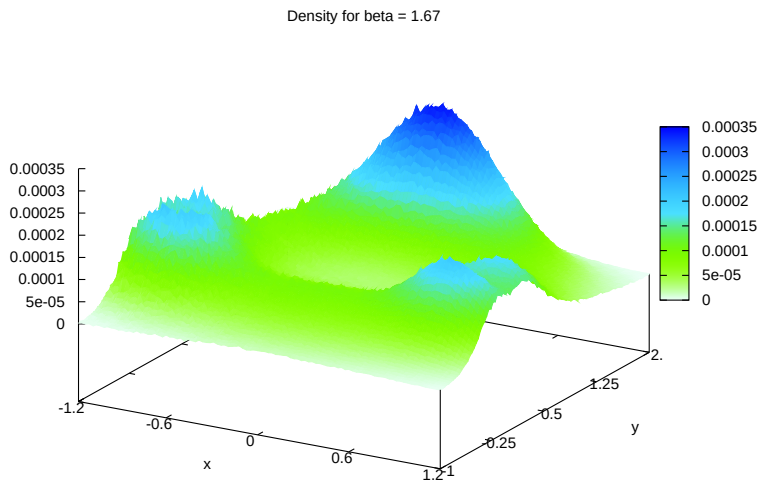
A 2D example

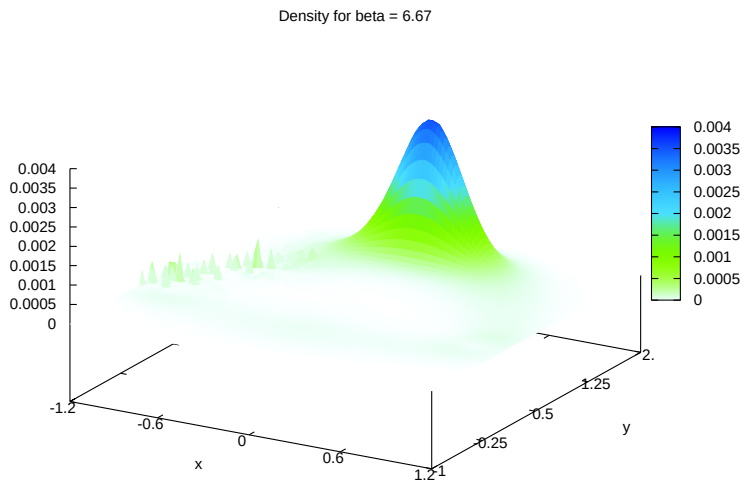
The interest of this "bi-channel" potential is that, depending on the temperature, one or the other channel is prefered to go from A (around $H_- = (-1, 0)$) to B (around $H_+ = (1, 0)$).

We will look at two quantities ([Hummer,2004] [Metzner, Schütte, Vanden-Eijnden, 2006]): the density of configurations along reactive paths, and the flux of reactive trajectories, for two values of the inverse temperature $\beta = 1.67$ and $\beta = 6.67$.

Two reaction coordinates are tested: $\xi_1(x, y) = x$ or $\xi_2(x, y) = \|(x, y) - H_-\|$. Both give reliable results, even though: (i) They are not the commitor functions ; (ii) The system exhibits metastabilities at fixed values of ξ_i.

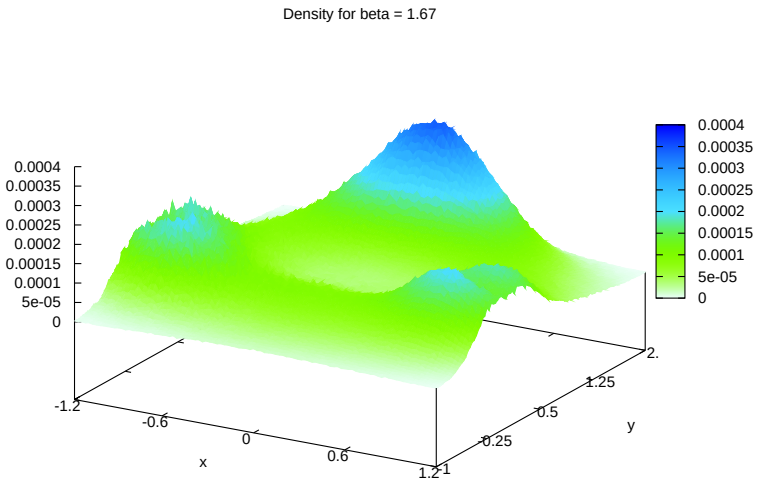
Density on reactive paths

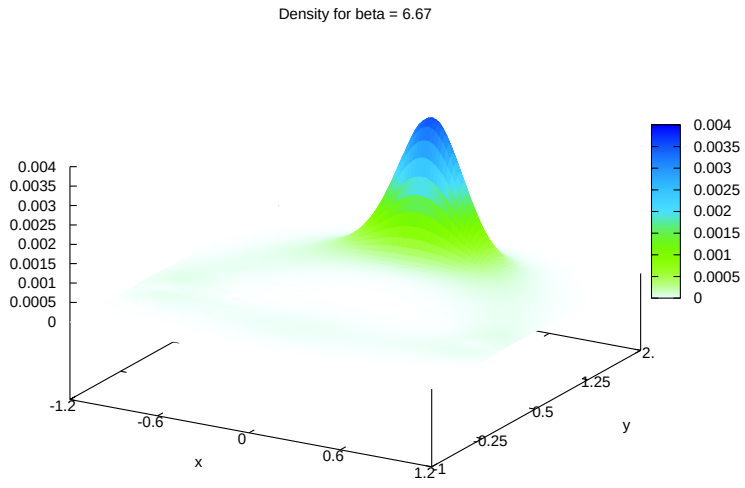




Results obtained for $\xi_1(x, y) = x$ and $N = 10^5$. Left: $\beta = 1.67$; Right: $\beta = 6.67$.

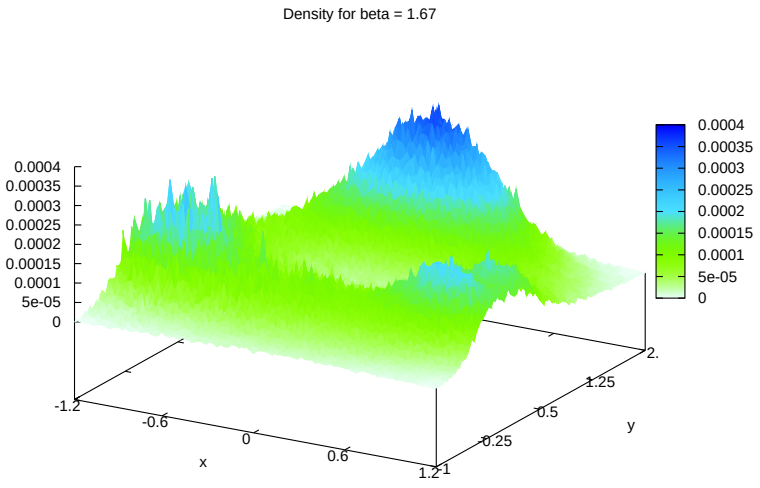
Density on reactive paths

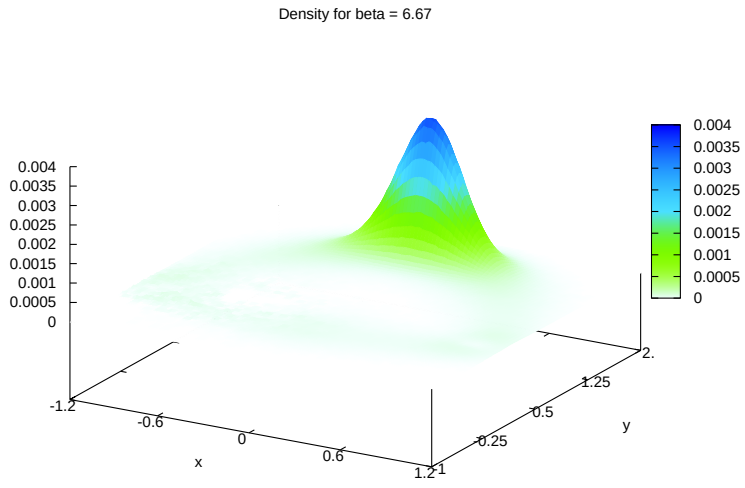




Results obtained for $\xi_2(x,y) = \|(x,y) - H_-\|$ and $N = 10^5$. Left: $\beta = 1.67$; Right: $\beta = 6.67$.

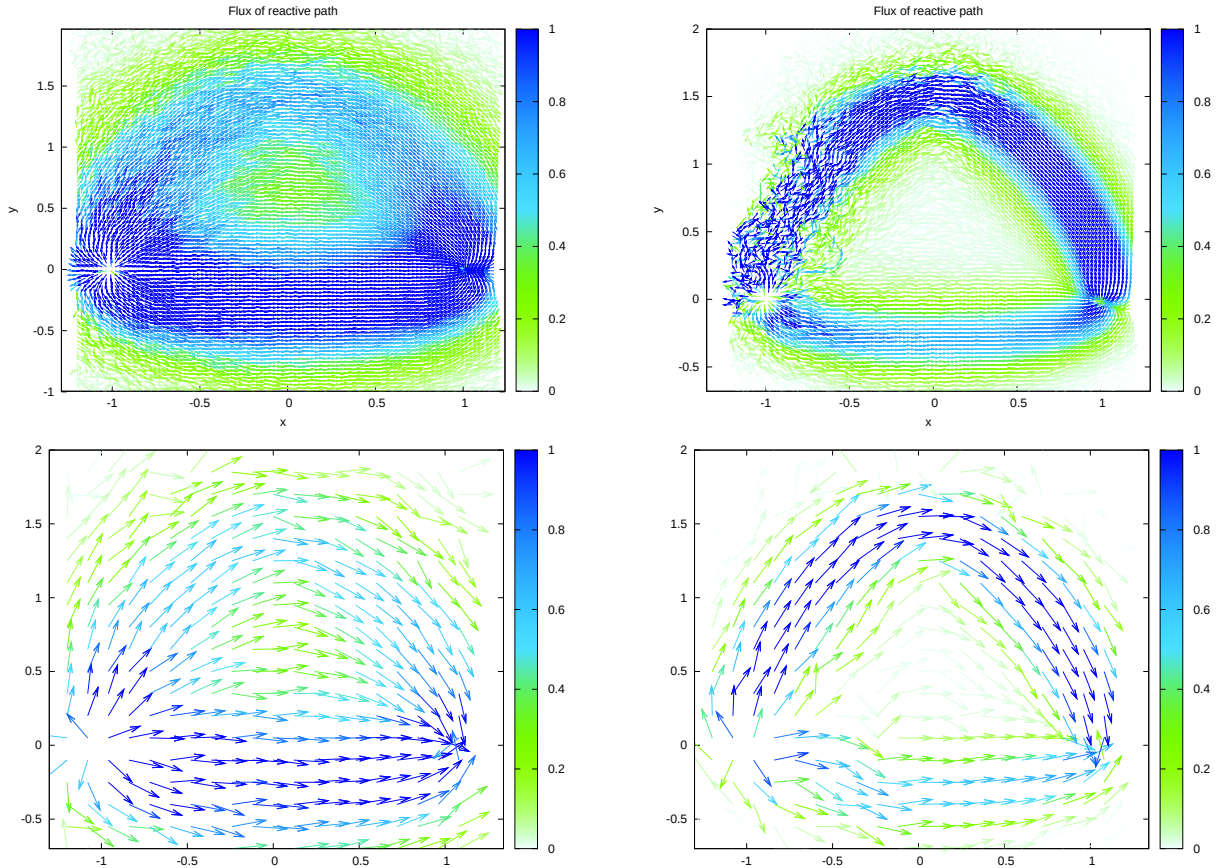
Density on reactive paths





Results obtained for $\xi_2(x, y) = \|(x, y) - H_-\|$ and $N = 10^4$. Left: $\beta = 1.67$; Right: $\beta = 6.67$.

Flux of reactive trajectories



Flux of reactive trajectories, at $\beta = 1.67$ on the left, and $\beta = 6.67$ on the right.

Computing transition times

To use the algorithm to compute transition times, we split a transition path from A to B into: excursions from ∂A to $\Sigma_{z_{\min}}$ and then back to ∂A, and finally an excursion from ∂A to $\Sigma_{z_{\min}}$ and then to B. Assuming $\Sigma_{z_{\min}}$ is close to an iso-commitor (Markov property), one obtains that the mean transition time is:

$$\mathbb{E}(T) = \left(\frac{1}{p} - 1\right) \mathbb{E}(T_1 + T_2) + \mathbb{E}(T_1 + T_3)$$

where:

- p is the probability, once $\Sigma_{z_{\min}}$ has been reached, to go to B rather than A (approximated by $\hat{\alpha}_N$) ;
- $\mathbb{E}(T_1 + T_2)$ is the mean time for an excursion from ∂A to $\Sigma_{z_{\min}}$ and then back to ∂A (approximated by DNS) ;
- $\mathbb{E}(T_1 + T_3)$ is the mean time for an excursion from ∂A to $\Sigma_{z_{\min}}$ and then to B (approximated by the AMS algorithm).

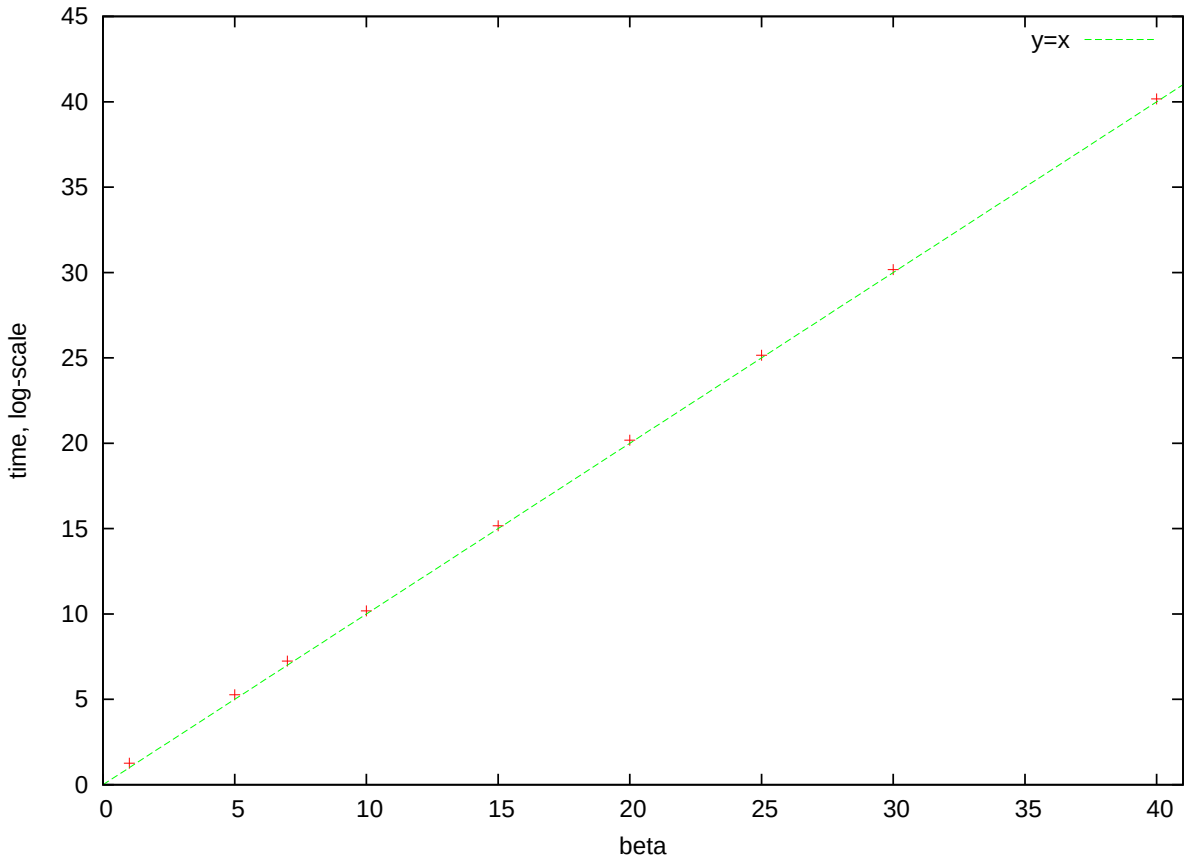
Numerical results

Numerical results on the 1D case

β	Δt	$z_{\min}$	$\mathbb{E}(T)$ (DNS)	$\mathbb{E}(T)$ (AMS)	C.I. in $\mathbb{E}(T)$ (AMS)	Error (%)
5	0.010	-0.9	185	208.3	$[199.6, 217.7]$	12.591
5	0.010	-0.6	185	221.2	$[214.3, 228.4]$	19.577
5	0.001	-0.9	185	187.4	$[180.5, 194.8]$	1.292
5	0.001	-0.6	185	193.2	$[188.3, 198.3]$	4.459
7	0.010	-0.9	1405	1515	$[1468, 1565]$	7.832
7	0.010	-0.6	1405	1642	$[1567, 1722]$	16.847
7	0.001	-0.9	1405	1380	$[1316, 1449]$	1.808
7	0.001	-0.6	1405	1400	$[1358, 1444]$	0.370

Transition times for small values of β, with various time steps Δt and $z_{\min}$. Reference values are computed by DNS.

Numerical results



Comparison of the estimated mean transition times as a function of β with the asymptotic law from large deviation theory:

$$\mathbb{E}(T) \propto \exp(\beta \Delta V)$$

where $\Delta V = 1$ is the height of the energy barrier.

Numerical results

Numerical results on the 2D case

N $\times 10^3$	β	k_{AB} (AMS)	C.I. on k_{AB} (AMS)
2	1.67	$2.03\ 10^{-2}$	$[1.83; 2.22]\ 10^{-2}$
10	1.67	$1.84\ 10^{-2}$	$[1.82; 1.86]\ 10^{-2}$
50	1.67	$1.88\ 10^{-2}$	$[1.87; 1.88]\ 10^{-2}$
100	1.67	$1.89\ 10^{-2}$	$[1.89; 1.90]\ 10^{-2}$
2	6.67	$9.97\ 10^{-8}$	$[7.74; 12.2]\ 10^{-8}$
10	6.67	$9.20\ 10^{-8}$	$[7.71; 10.7]\ 10^{-8}$
50	6.67	$8.88\ 10^{-8}$	$[8.42; 9.34]\ 10^{-8}$
100	6.67	$9.32\ 10^{-8}$	$[9.08; 9.57]\ 10^{-8}$

Estimates of the reaction rate $k_{AB} = 2/\mathbb{E}(T)$, with$\xi = \xi_2$. Values from [Metzner, Schütte, Vanden-Eijnden, 2006] are $k_{AB} = 1.912\ 10^{-2}$ for $\beta = 1.67$ and $k_{AB} = 9.47\ 10^{-8}$ for $\beta = 6.67$.

Adaptive multilevel splitting: conclusions

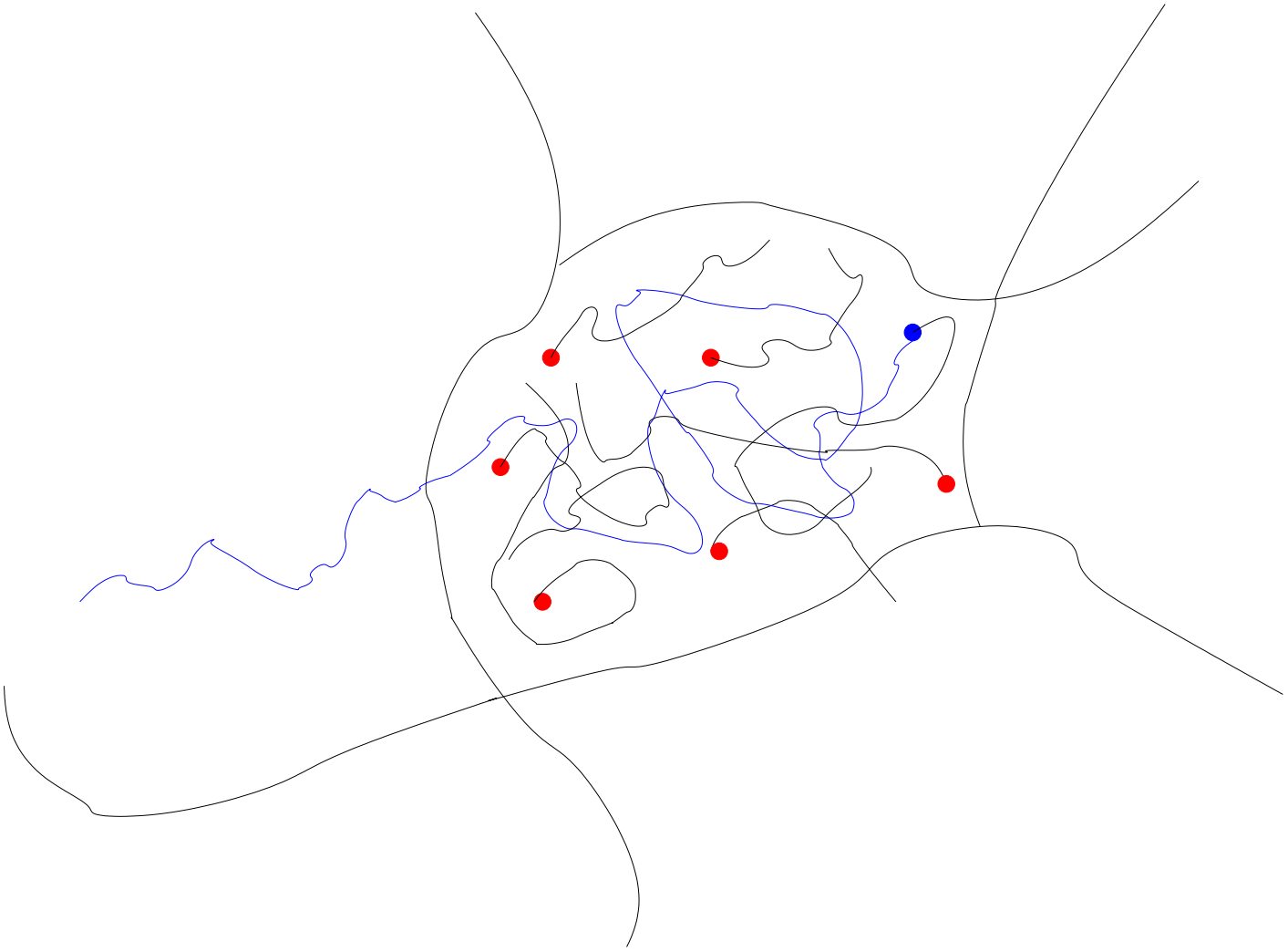
Possible extensions:

- Generate equilibrium trajectories which leave a metastable state A without knowing the neighborhood.
- Adaptively compute a better and better ξ (proportional to the committor function).

The mathematical analysis of the AMS algorithm remains essentially to be done:

- Analysis of the time discretization error,
- Asymptotic variance of the estimator and optimization of ξ,
- Precise estimate of the bias.

The Parallel Replica Algorithm



The Parallel Replica Algorithm

The Parallel Replica Algorithm, proposed by A.F. Voter in 1998, is a method to get efficiently a "coarse-grained projection" of a dynamics.
Let us consider again the overdamped Langevin dyanmics:

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

and let assume that we are given a smooth mapping

$$\mathcal{S} : \mathbb{R}^d \to \mathbb{N}$$

which to a configuration in $\mathbb{R}^d$ associates a state number. Think of a numbering of the wells of the potential V.

The aim of the parallel replica dynamics is to generate very efficiently a trajectory $(S_t)_{t\geq 0}$ which has (almost) the same law as $(\mathcal{S}(\boldsymbol{X}_t))_{t\geq 0}$.

The Parallel Replica Algorithm

Initialization: Consider an initial condition $\boldsymbol{X}_0^{ref}$ for a reference walker, the associated initial condition $S_0 = \mathcal{S}(\boldsymbol{X}_0^{ref})$, and a simulation time counter $T_{simu} = 0$.

Then, one iteration of the algorithm goes through three steps.

- The decorrelation step: Let the reference walker $(\boldsymbol{X}_{T_{simu}+t}^{ref})_{t\geq 0}$ evolve over a time interval $t \in [0, \tau_{corr}]$. Then,
  - If the process leaves the well during the time interval (*i.e.* $\exists t \leq \tau_{corr}$ such that $\mathcal{S}\left(\boldsymbol{X}_{T_{simu}+t}^{ref}\right) \neq \mathcal{S}\left(\boldsymbol{X}_{T_{simu}}^{ref}\right)$) advance the simulation clock by τ_{corr} and restart the decorrelation step ;
  - otherwise, advance the simulation clock by τ_{corr} and proceed to the dephasing step.

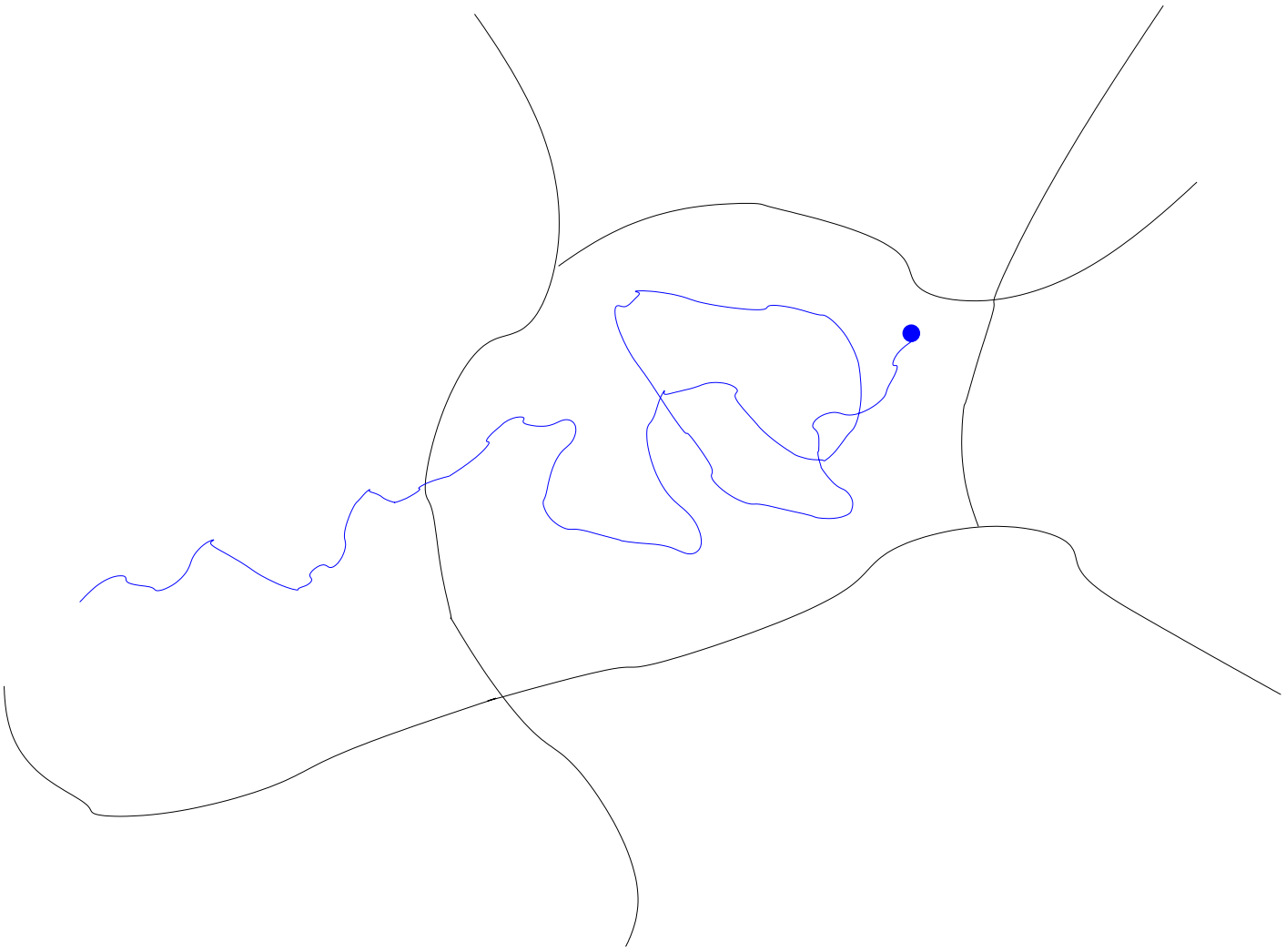
  During all this step, $S_{T_{simu}+t} := \mathcal{S}\left(\boldsymbol{X}_{T_{simu}+t}^{ref}\right)$.

The Parallel Replica Algorithm

The reference walker enters a new state

The Parallel Replica Algorithm

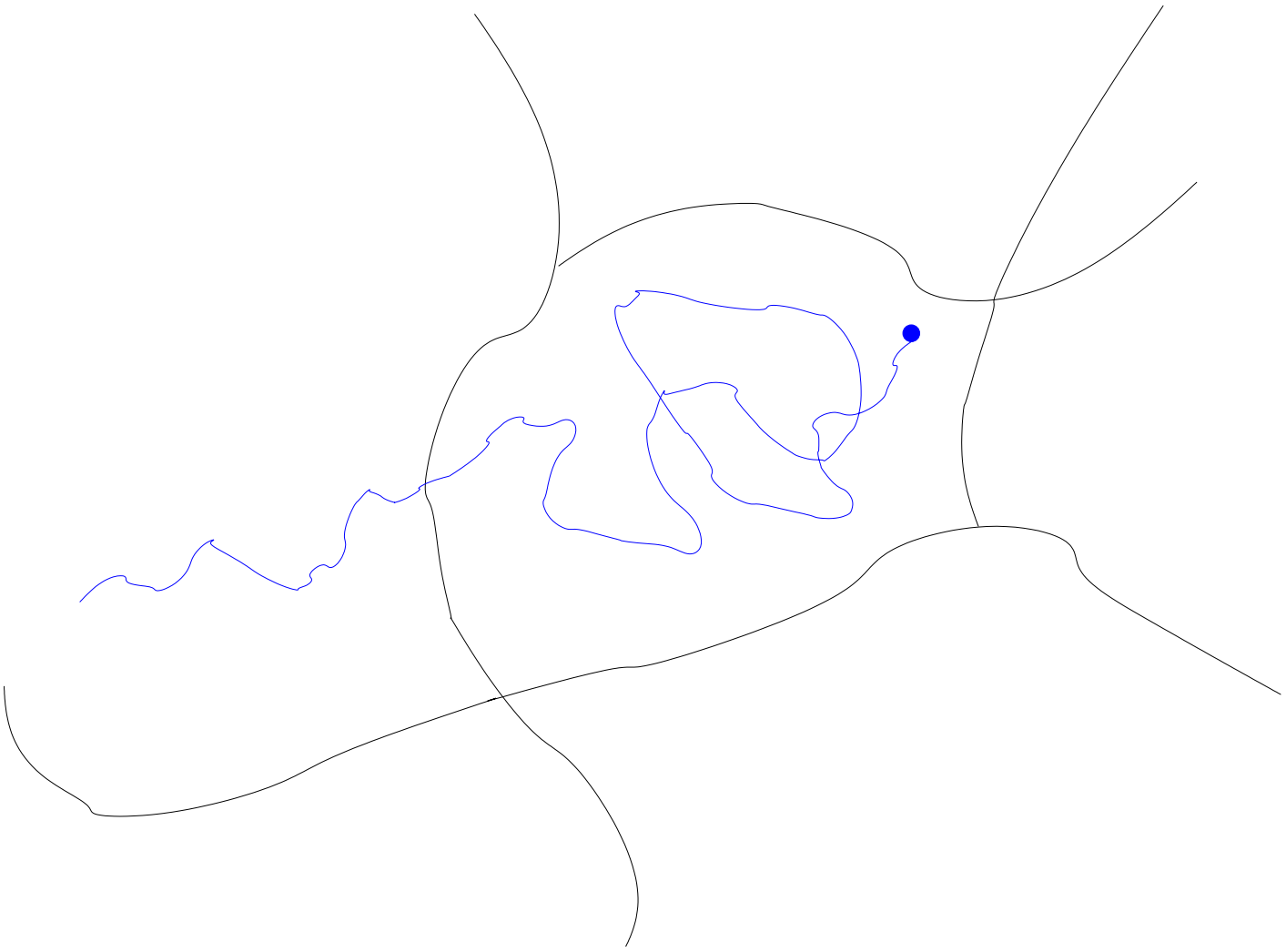
Decorrelation step: wait for a time τ_{corr}.



- The dephasing step: Duplicate the walker $\boldsymbol{X}_{T_{simu}}^{ref}$ into N replicas. Let these replicas evolve independently and in parallel over a time interval of length $\tau_{dephase}$. If a replica leaves the well during this time interval, restart the dephasing step for this replica. Throughout this step, the simulation counter is stopped.

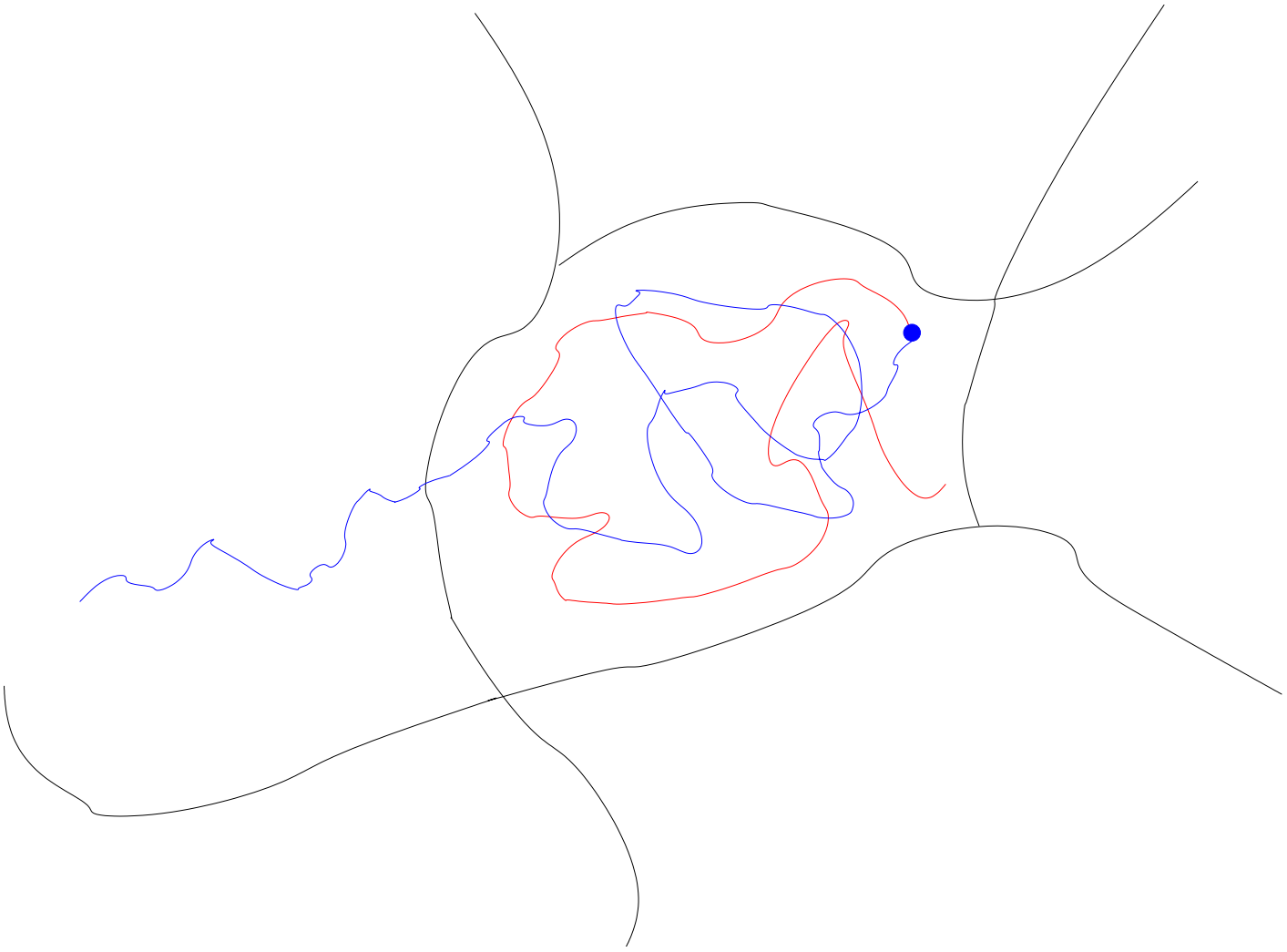
The Parallel Replica Algorithm

Dephasing step.



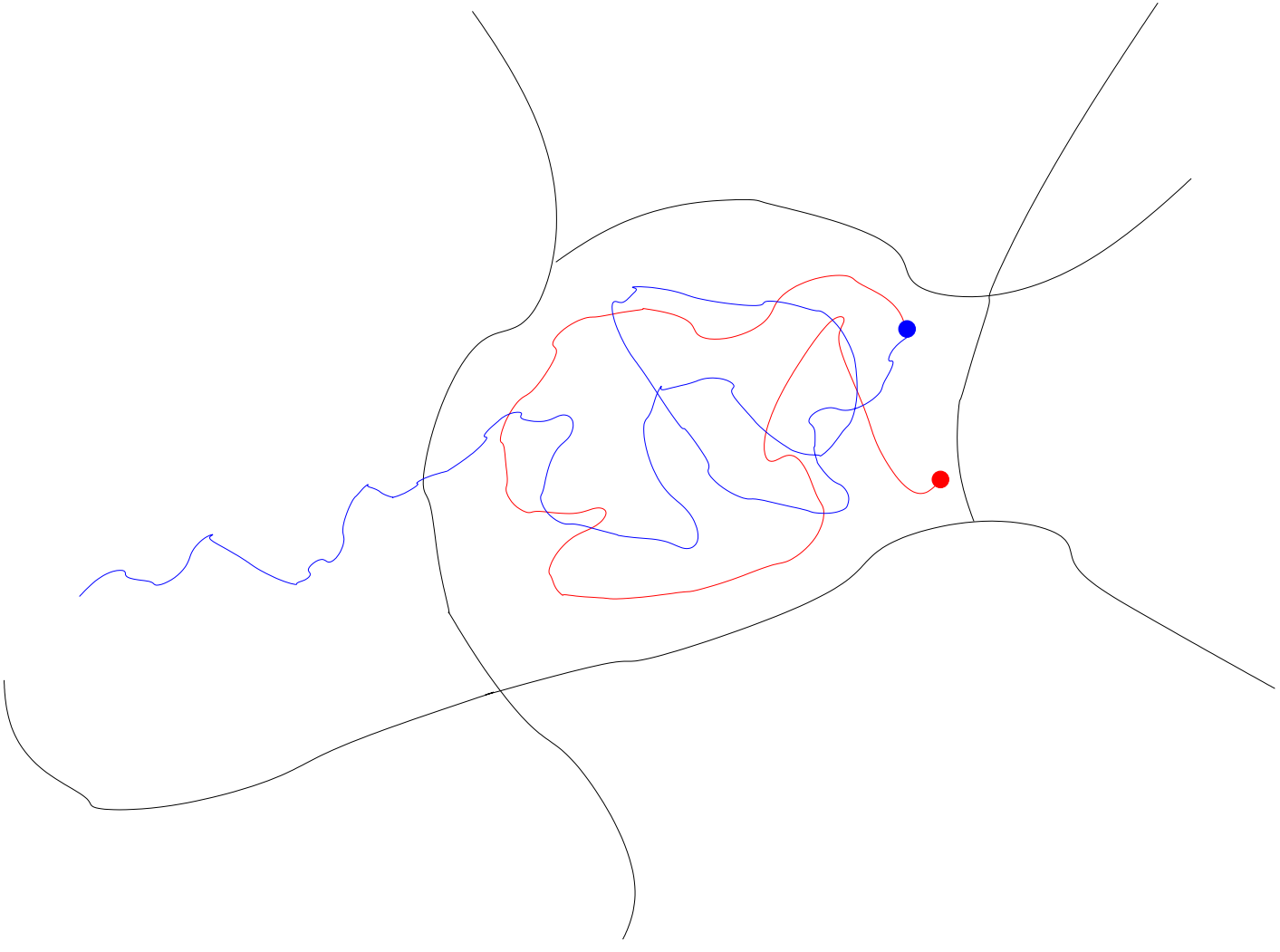
The Parallel Replica Algorithm

Dephasing step: generate new initial conditions in the state.



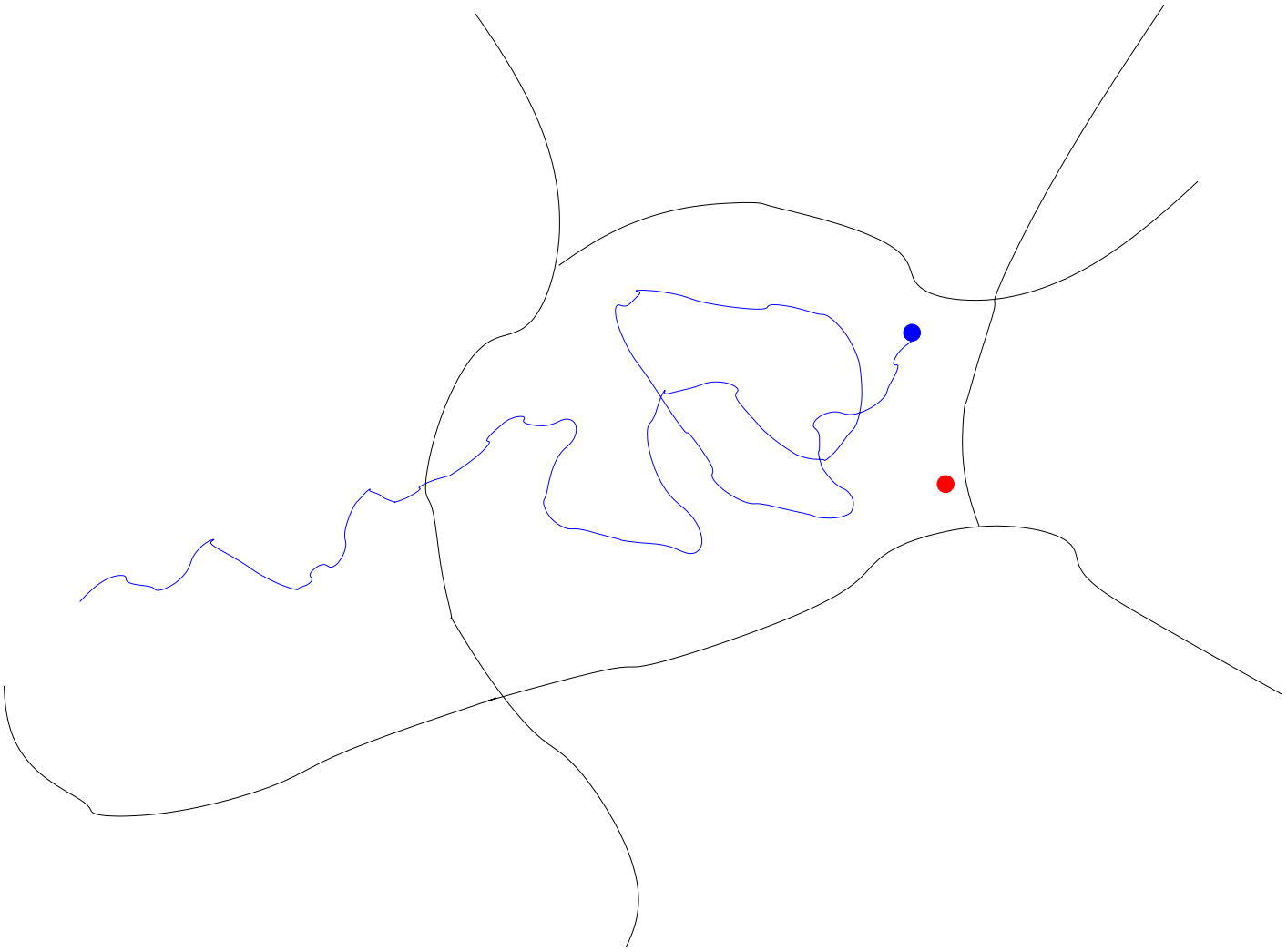
The Parallel Replica Algorithm

Dephasing step: generate new initial conditions in the state.



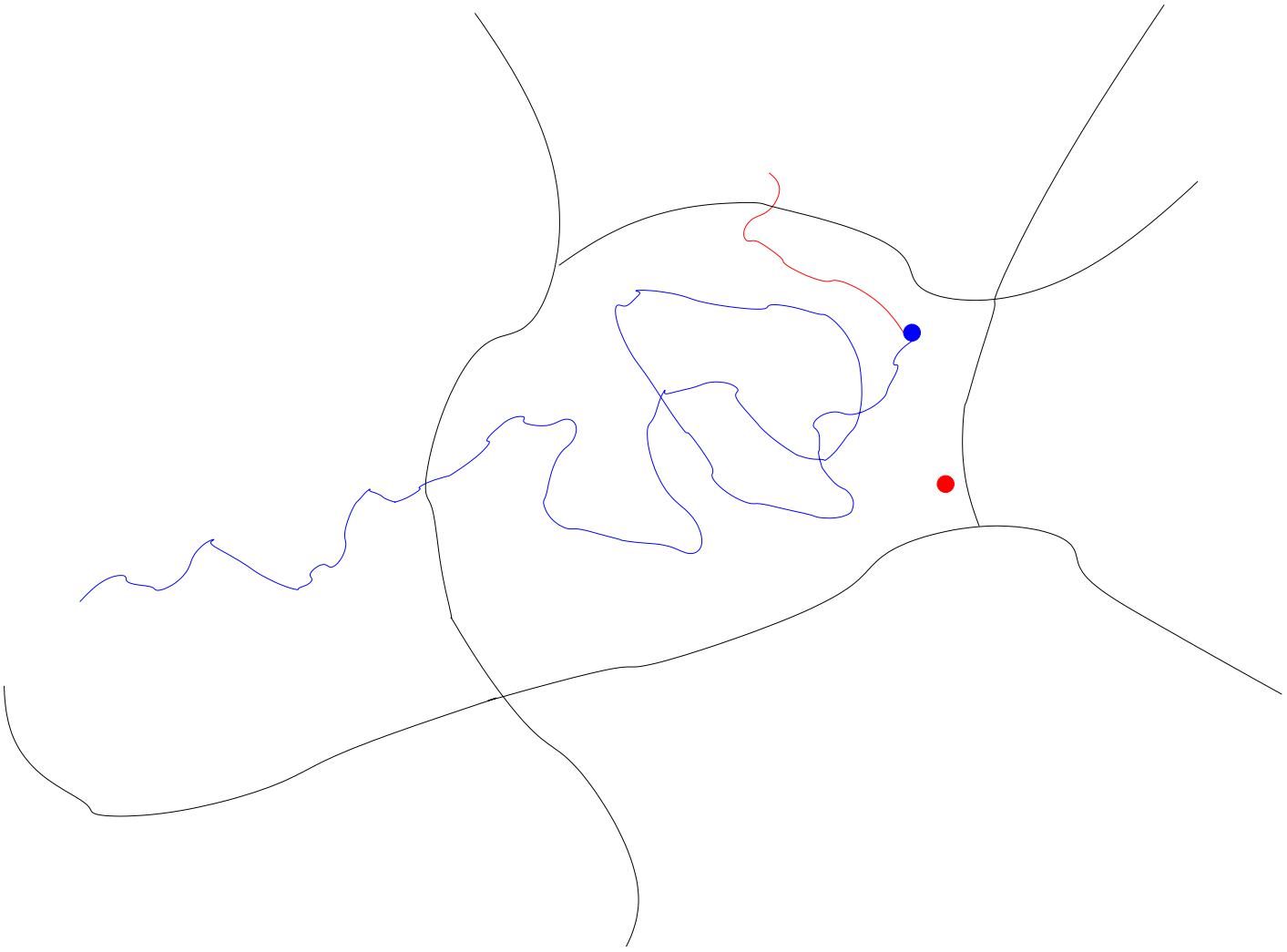
The Parallel Replica Algorithm

Dephasing step: generate new initial conditions in the state.



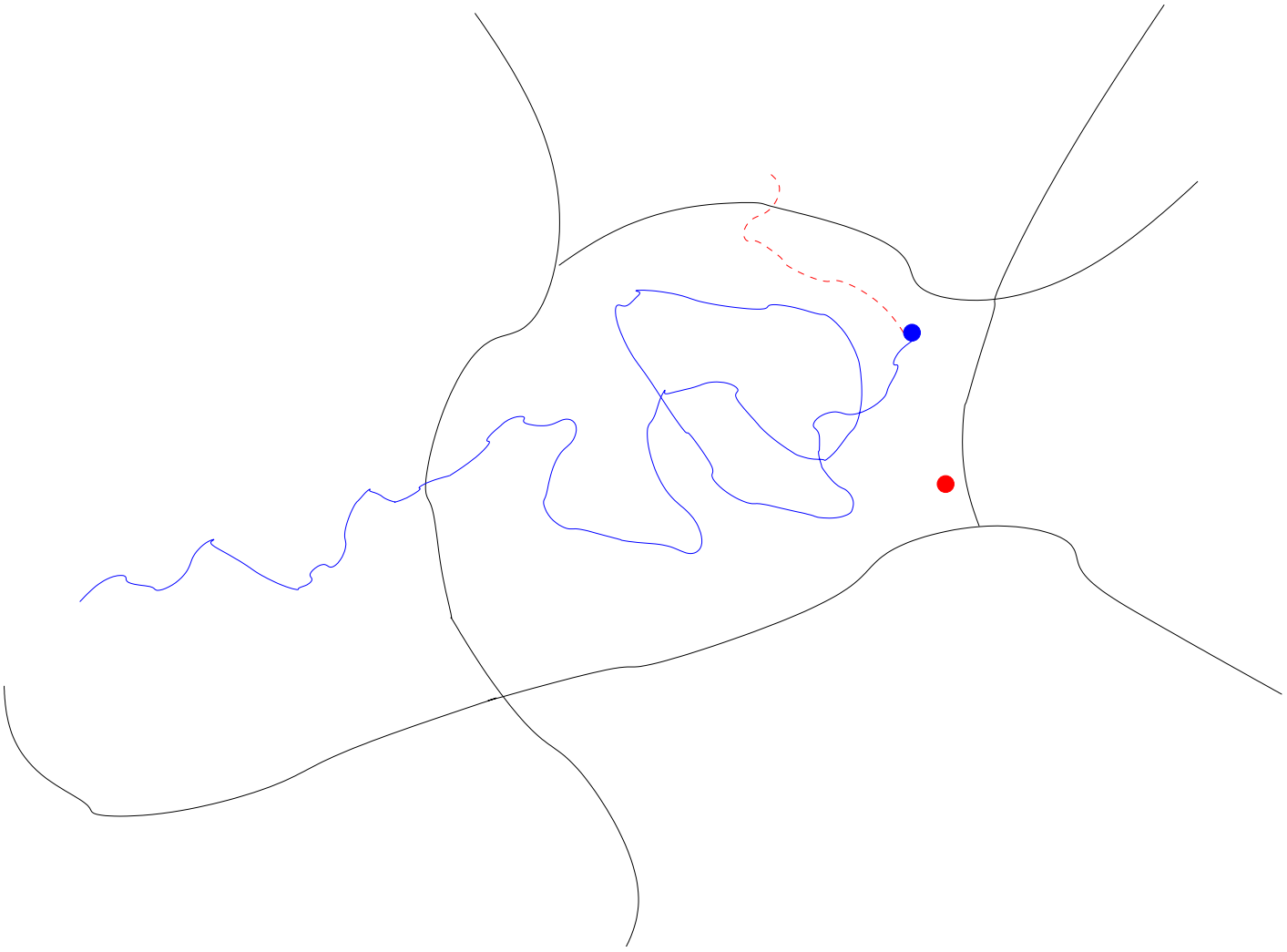
The Parallel Replica Algorithm

Dephasing step: generate new initial conditions in the state.



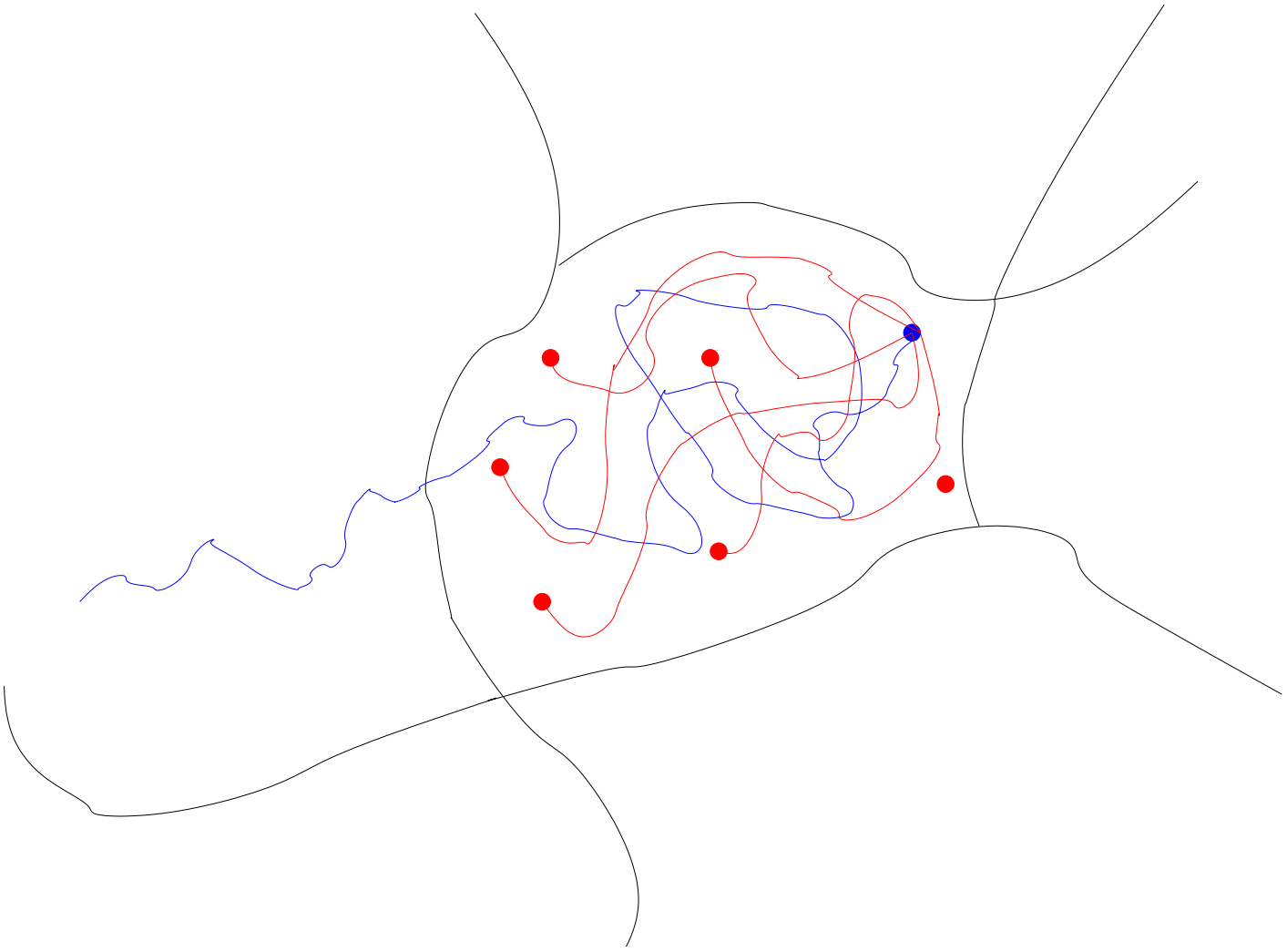
The Parallel Replica Algorithm

Dephasing step: generate new initial conditions in the state.



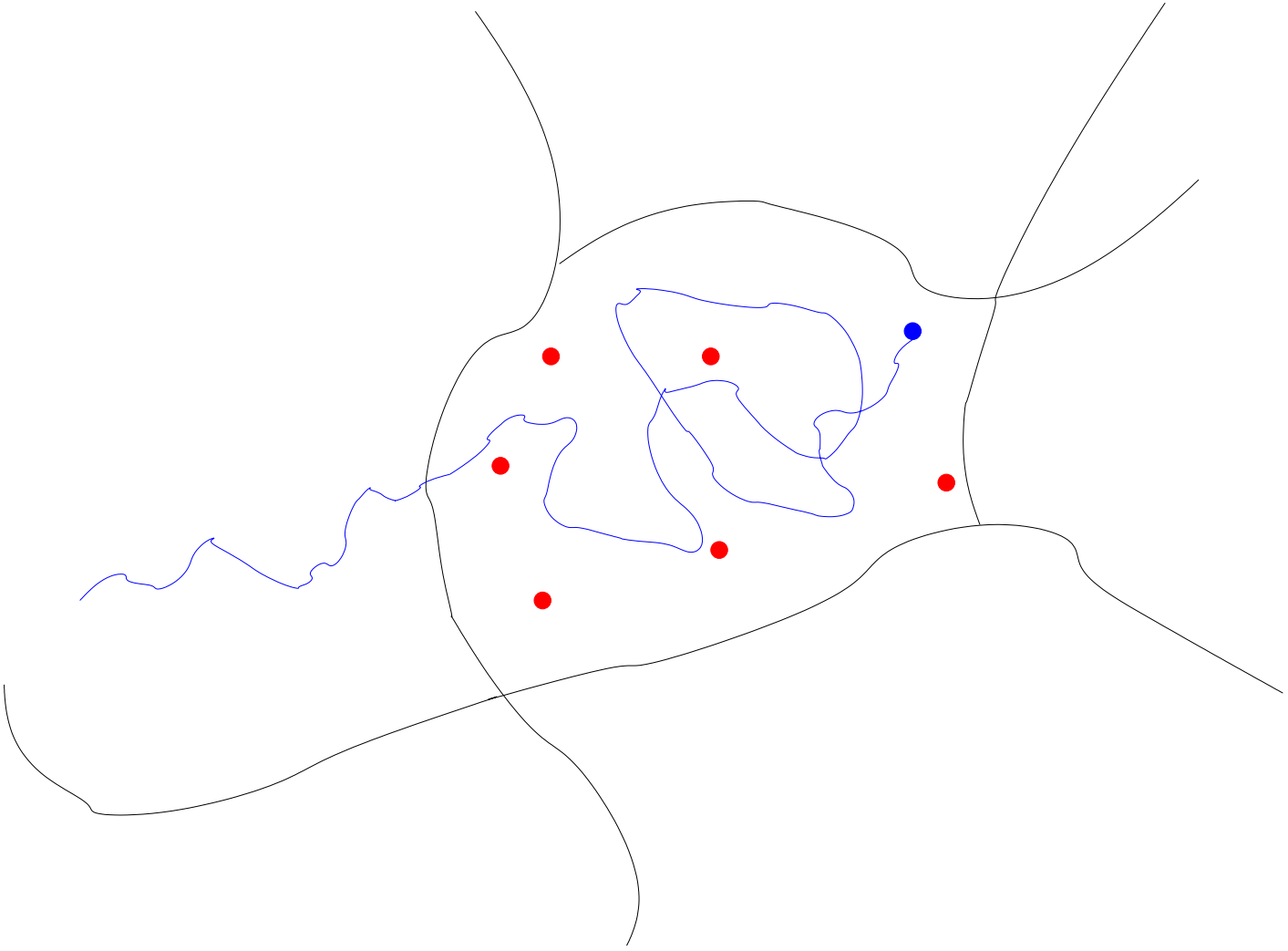
The Parallel Replica Algorithm

Dephasing step: generate new initial conditions in the state.



The Parallel Replica Algorithm

Dephasing step: generate new initial conditions in the state.



- The parallel step: Let all the replicas evolve independently and track the first escape event:

$$T = \inf_k T_W^k = T_W^{K_0}$$

where $K_0 = \arg\inf_k T_W^k$ and

$$T_W^k = \inf\{t \geq 0,\ \mathcal{S}(\boldsymbol{X}^k_{T_{simu}+t}) \neq \mathcal{S}(\boldsymbol{X}^k_{T_{simu}})\}$$

is the first time the k-th replica leaves the well. Then:

$$T_{simu} = T_{simu} + NT \text{ and } \boldsymbol{X}^{ref}_{T_{simu}+NT} = \boldsymbol{X}^{K_0}_{T_{simu}+T}.$$

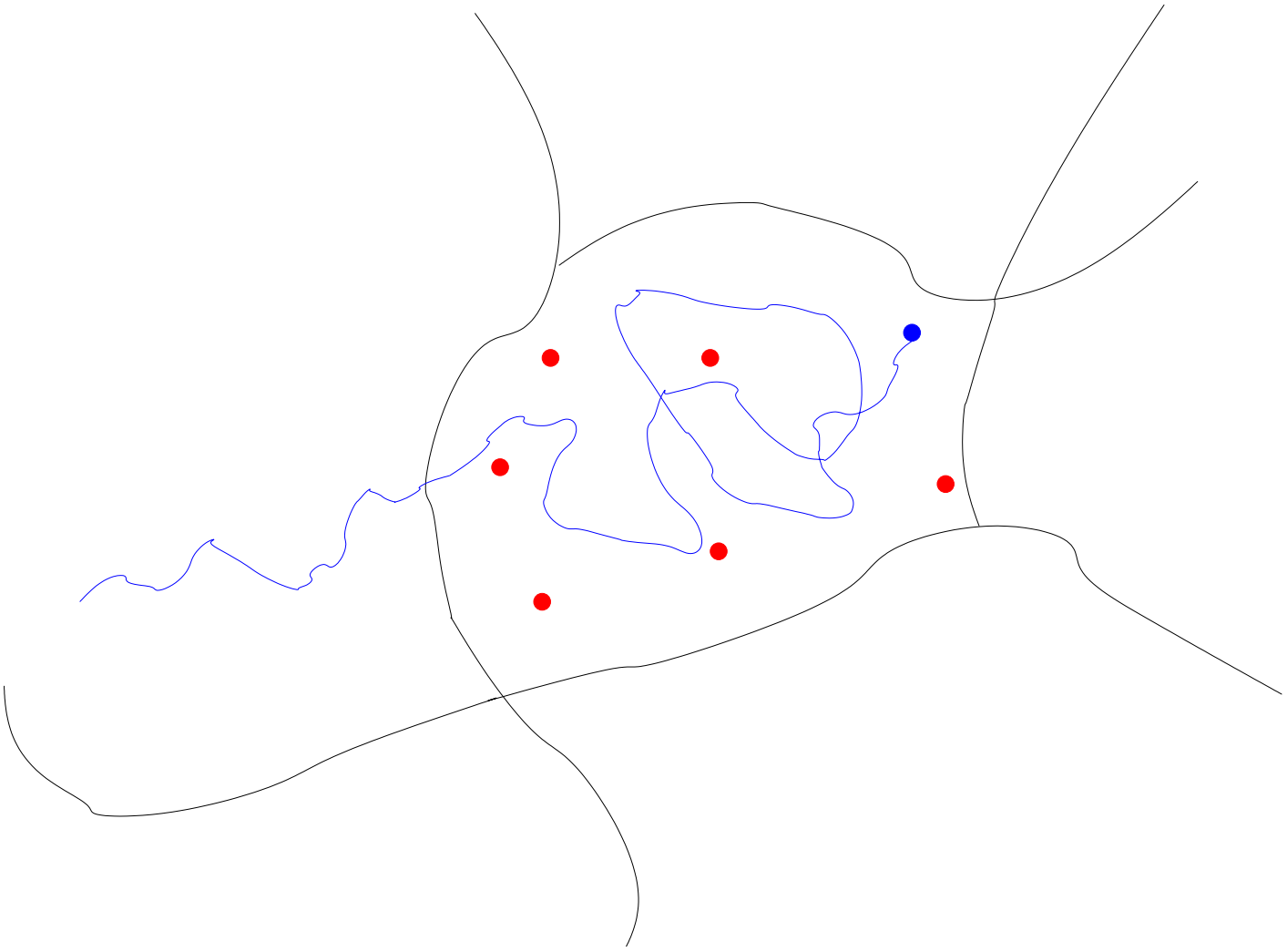
Moreover, over $[T_{simu}, T_{simu} + NT]$, the state dynamics S_t is constant and defined as:

$$S_t = \mathcal{S}(\boldsymbol{X}^1_{T_{simu}}).$$

Then, go back to the decorrelation step...

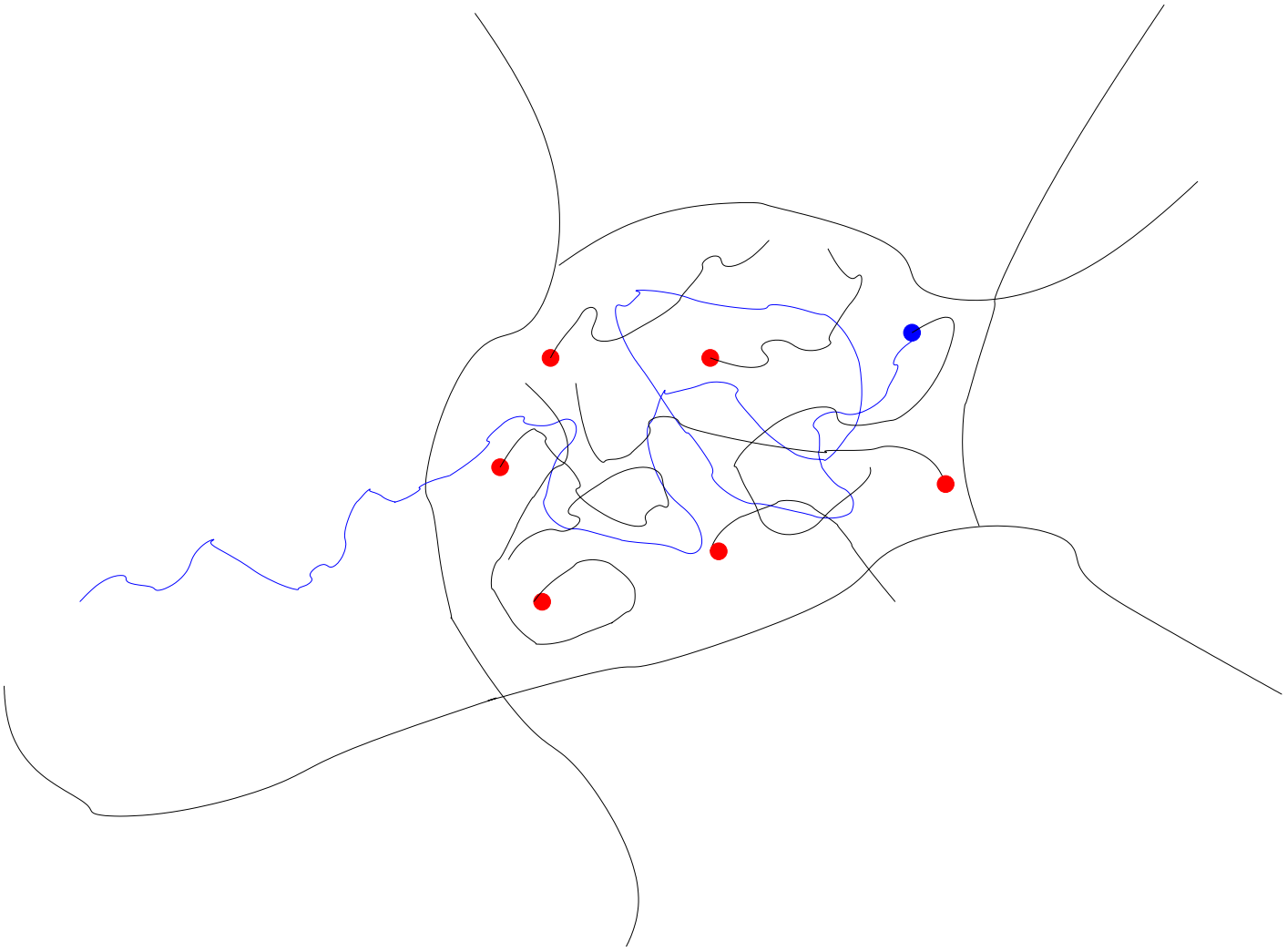
The Parallel Replica Algorithm

Parallel step.



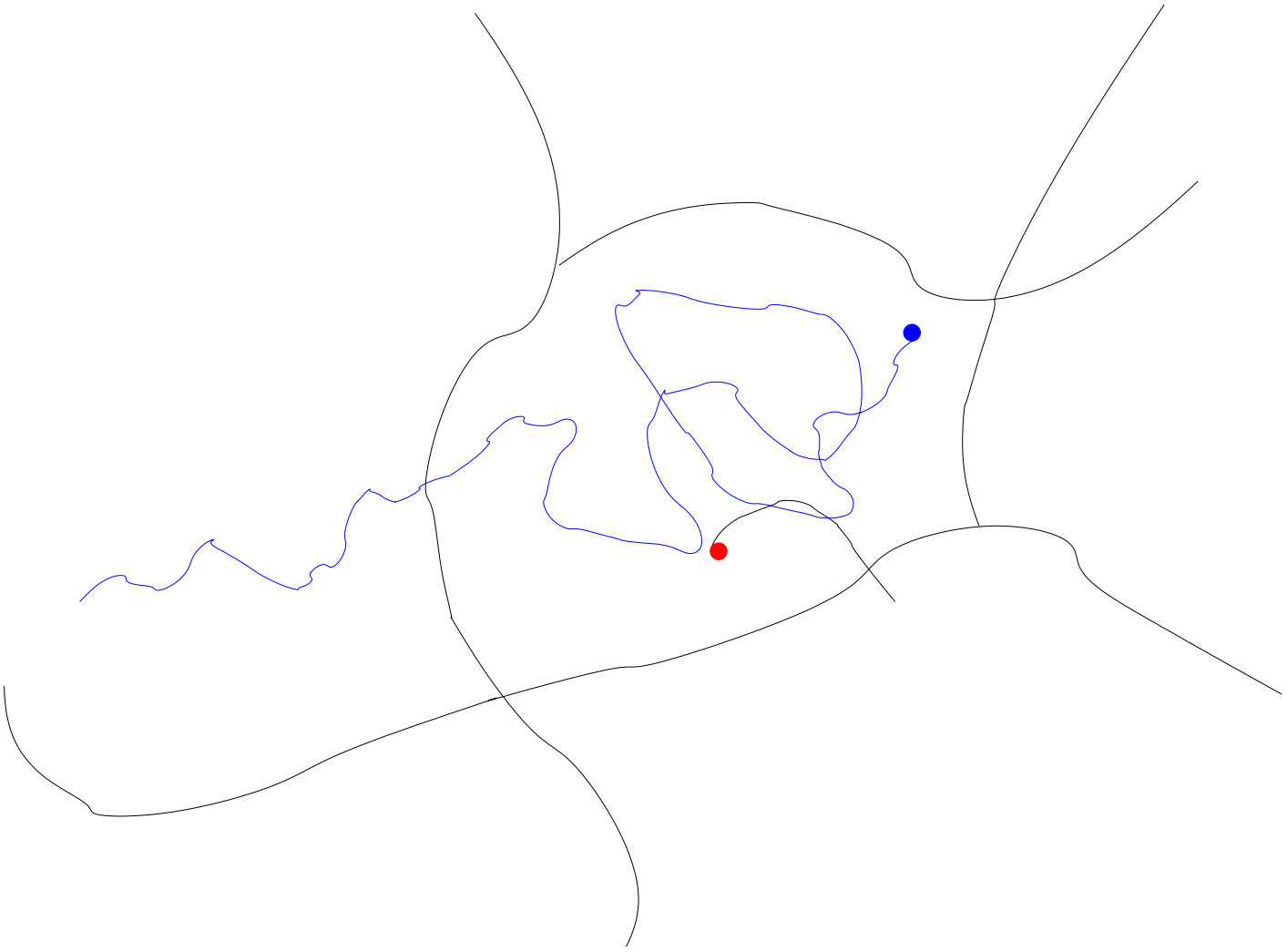
The Parallel Replica Algorithm

Parallel step: run independent trajectories in parallel...



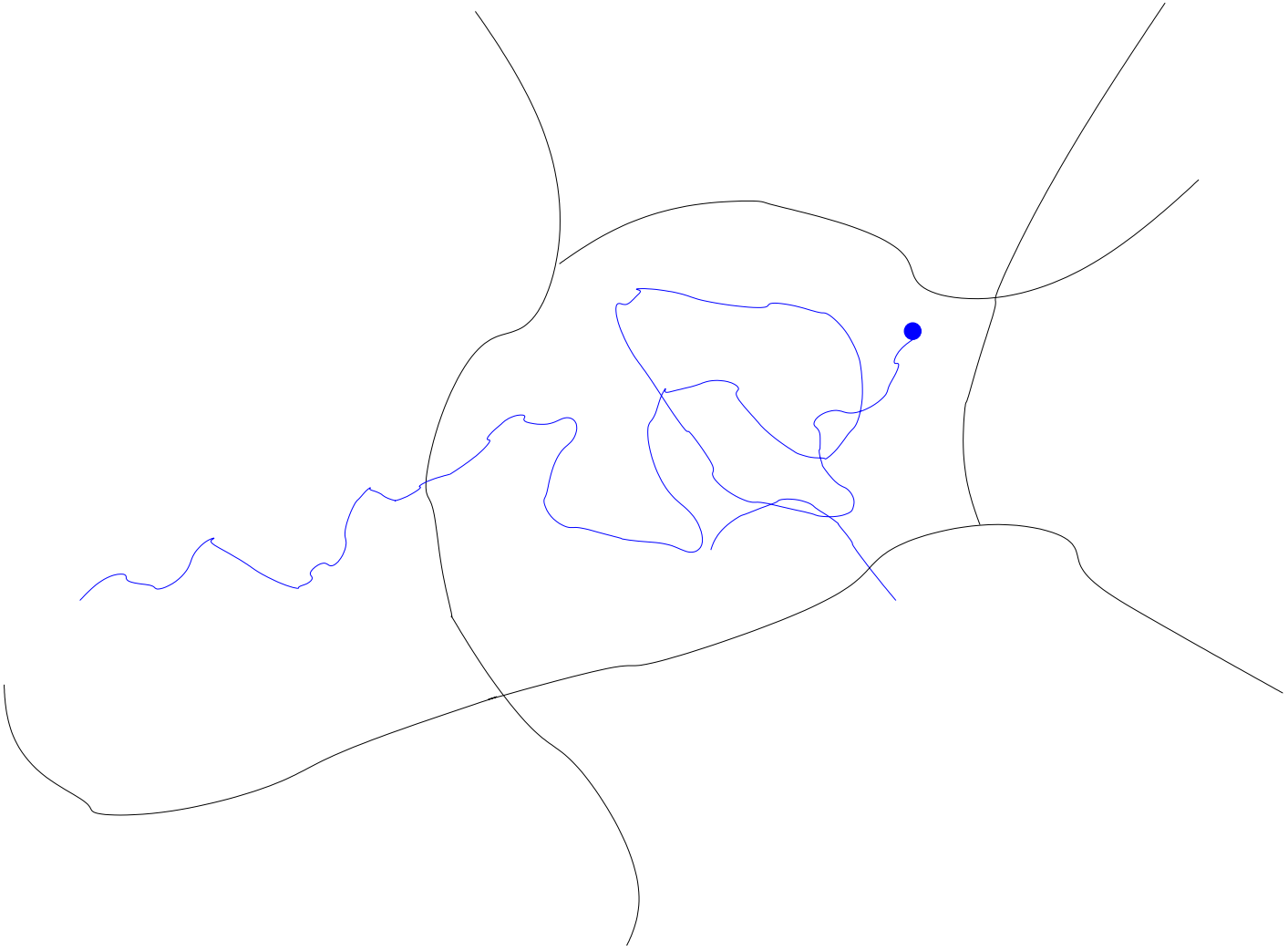
The Parallel Replica Algorithm

Parallel step: ... and detect the first transition event.



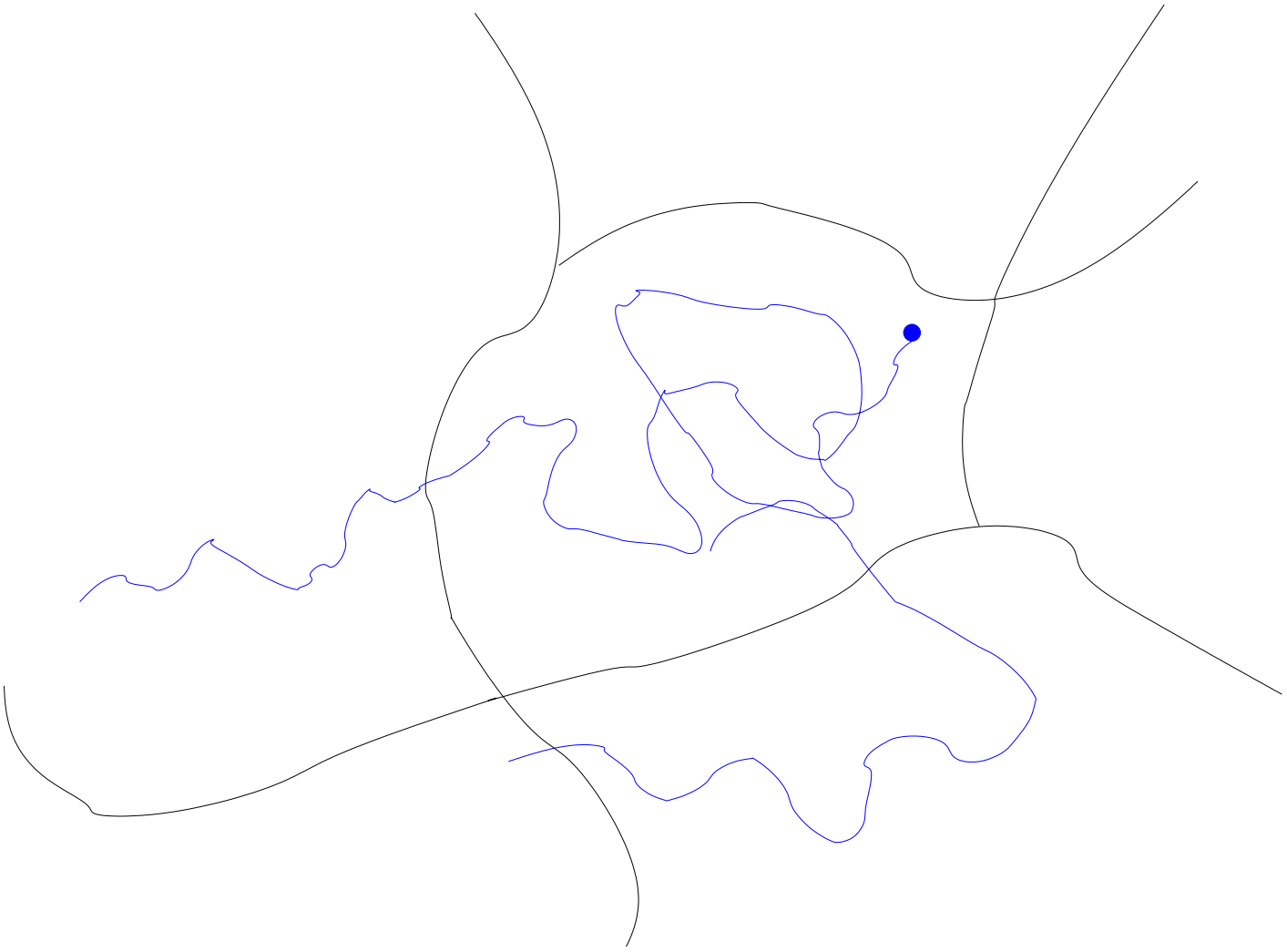
The Parallel Replica Algorithm

Parallel step: update the time clock: $T_{simu} = T_{simu} + NT$.



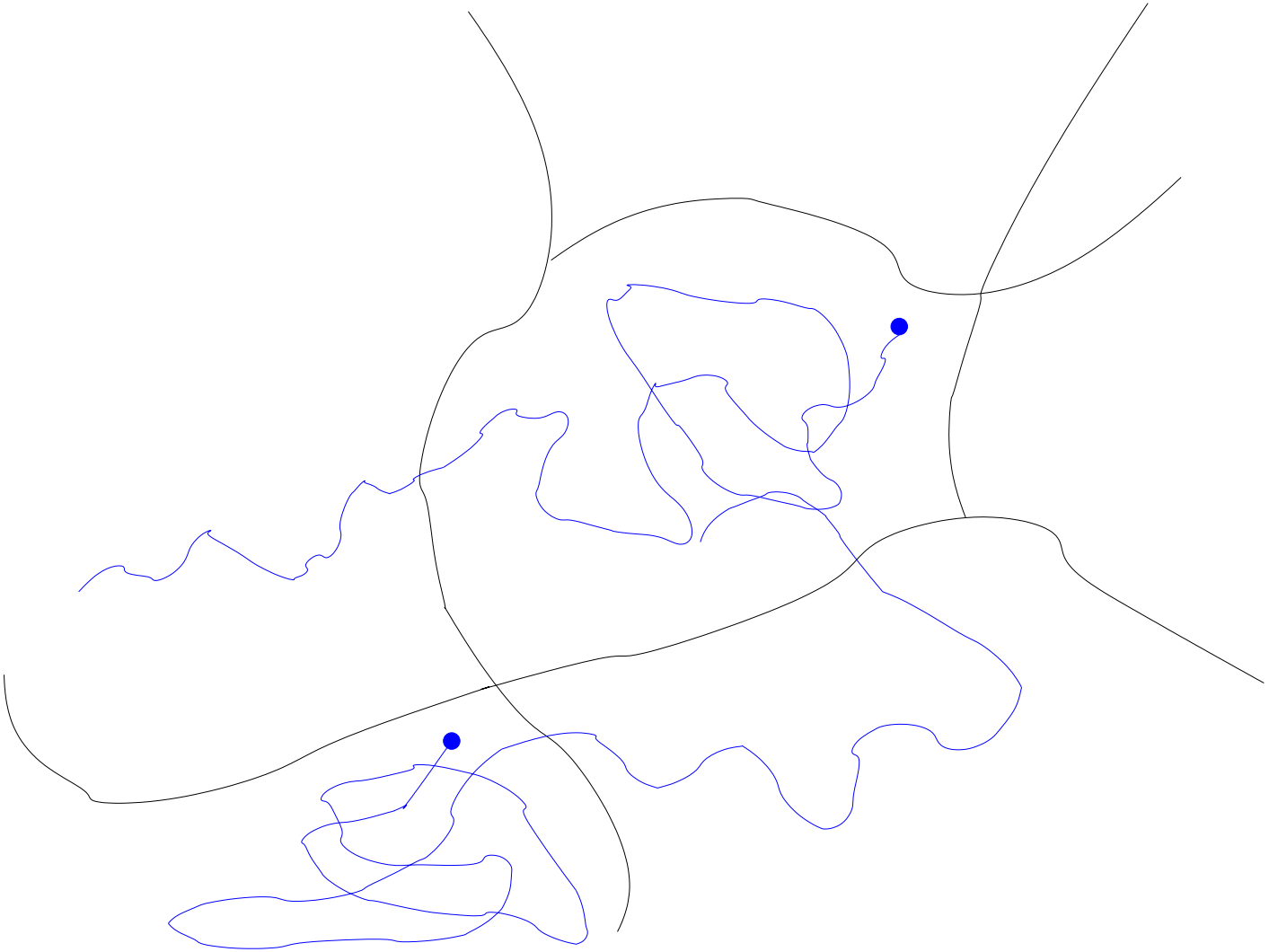
The Parallel Replica Algorithm

A new decorrelation step starts...



The Parallel Replica Algorithm

New decorrelation step



The Parallel Replica Algorithm

Analysis of the algorithm: the parallel step would introduce no error if

- the escape time T_W^1 was exponentially distributed
- and independent of the next visited state.

This essentially amounts to assuming that $\mathcal{S}(\boldsymbol{X}_t)$ is a Markov chain...

How to analyze the error introduced by the algorithm ?

This is related to the general question: how to relate a continuous state space Markov dynamics to a discrete state space Markov dynamics ? Pitfalls: (i) the temperature is not necessarily small (ii) the partition of the state space may be anything (iii) no thermodynamic limit in general (non-homogeneous systems).

The quasi-stationary distribution

The quasi-stationary distribution (QSD) ν for $\boldsymbol{X}_t$ and associated to the actual well W is a probability measure which is (i) supported by W and such that (ii): $\forall t > 0$, $\forall A \subset W$,

$$\nu(A) = \frac{\displaystyle\int_W \mathbb{P}(\boldsymbol{X}_t^{\boldsymbol{x}} \in A,\, t < T_W^{\boldsymbol{x}})\, \nu(d\boldsymbol{x})}{\displaystyle\int_W \mathbb{P}(t < T_W^{\boldsymbol{x}})\, \nu(d\boldsymbol{x})}.$$

If $\boldsymbol{X}_0 \sim \nu$ and if $(\boldsymbol{X}_s)_{0 \leq s \leq t}$ has not left the well, then $\boldsymbol{X}_t \sim \nu$.

Let $L = -\nabla V \cdot \nabla + \beta^{-1}\Delta$ be the infinitesimal generator of $(\boldsymbol{X}_t)$. Then the density u of ν ($d\nu = u(\boldsymbol{x})d\boldsymbol{x}$) is the first eigenfunction of $L^* = \operatorname{div}(\nabla V + \beta^{-1}\nabla)$ with absorbing boundary conditions:

$$\begin{cases} L^* u = -\lambda_1 u \text{ on } W, \\ \quad u = 0 \text{ on } \partial W. \end{cases}$$

The quasi-stationary distribution and the dephasing step

Property of the QSD: If $\boldsymbol{X}_0 \sim \nu$ then, the first exit time T_W from W is exponentially distributed with parameter λ_1 and is a random variable independent of the first hitting point $\boldsymbol{X}_{T_W}$ on ∂W.

The dephasing step is very much related to the so-called Fleming-Viot process and may be seen as a way to get N i.i.d. random variables distributed according to the QSD.

Remark: In general, T_W exponentially distribtued is *not* sufficient for $\boldsymbol{X}_0$ to be distributed according to ν.

The parallel step

As announced above, starting from the QSD, the parallel step is exact. This is stated precisely here.

Let us start from N initial conditions $\boldsymbol{X}_0^k$ i.i.d. in the well W and let the processes evolve independently. Let us denote

$$T_W^k = \inf\{t > 0, \boldsymbol{X}_t^k \notin W\}$$

the escape time for the k-th replica, and

$$T = T_W^{K_0} \text{ where } K_0 = \arg \min_{k \in \{1,\dots,N\}} T_W^k$$

the *first* escape time over all processes.

- Assume that T_W^1 is exponentially distributed [OK starting from QSD.] Then NT has the same law as T_W^1.
- Assume that T_W^1 is independent of $\boldsymbol{X}_{T_W^1}^1$ [OK starting from QSD.] Then $\boldsymbol{X}_{T_W^{K_0}}^{K_0}$ has the same distribution as $\boldsymbol{X}_{T_W^1}^1$ and is independent of $T_W^{K_0}$.

The decorrelation step

We would like to quantify the error introduced by the dephasing and parallel steps, when the decorrelation step is successful.

As shown above, when the decorrelation step is successful, it is assumed that the reference walker is distributed according to the QSD. If it was indeed the case, the algorithm would be exact. The decorrelation step can be seen as a way to probe this assumption. What is the error introduced there ?

The decorrelation step

We have the following error estimate in total variation norm: for $t \geq \frac{C}{\lambda_2 - \lambda_1}$,

$$\sup_{f, \|f\|_{L^\infty} \leq 1} \left| \mathbb{E}(f(T_W - t, \boldsymbol{X}_{T_W}) | T_W \geq t) - \mathbb{E}^\nu(f(T_W, \boldsymbol{X}_{T_W})) \right| \leq C \exp(-(\lambda_2 - \lambda_1)t),$$

where $-\lambda_2 < -\lambda_1 < 0$ are the two first eigenvalues of L^* with absorbing boundary conditions on ∂W.

This shows that τ_{corr} should be chosen such that:

$$\tau_{corr} \geq \frac{\overline{C}}{\lambda_2 - \lambda_1}.$$

On the other hand, it should be smaller than the typical time to leave the well, $\mathbb{E}(T_W)$. Since $\mathbb{E}^\nu(T_W) = 1/\lambda_1$, this typically implies the spectral gap requirement,

$$\frac{\overline{C}}{\lambda_2 - \lambda_1} \leq \frac{1}{\lambda_1}.$$

The Parallel Replica Algorithm: conclusions

This can be generalized to other dynamics (coarse-graining of kMC).

Main results:

- The QSD is a good intermediate between continuous state dynamics and kMC-like approximations.
- The error analysis holds whatever the partition. But the method requires metastability between the states to be computationally efficient.
- The parameter τ_{corr} should be adjusted in terms of the two first eigenvalues of the Fokker-Planck operator with absorbing boundary conditions.

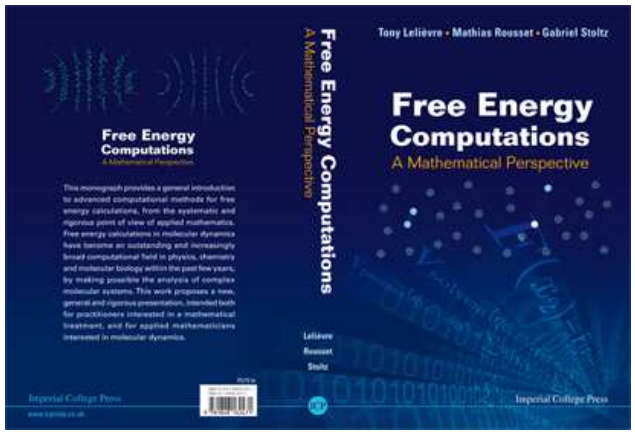
Conclusion

Notice that the numerical methods we have presented are based on some dimension reduction and coarse-graining techniques (through the functions ξ or $\mathcal{S}$). They are easy to parallelize.

Conclusion

A few references:

- F. Cérou, A. Guyader, T. Lelièvre and D. Pommier, *A multiple replica approach to simulate reactive trajectories*, Journal of Chemical Physics 134, 054108, 2011.
- T. Lelièvre, M. Rousset and G. Stoltz, *Free energy computations, a mathematical perspective*, Imperial College Press, 2010.



- C. Le Bris, T. Lelièvre, M. Luskin and D. Perez, *A mathematical formalization of the parallel replica dynamics*, `http://arxiv.org/abs/1105.4636`, to appear in Monte Carlo Methods and Applications.