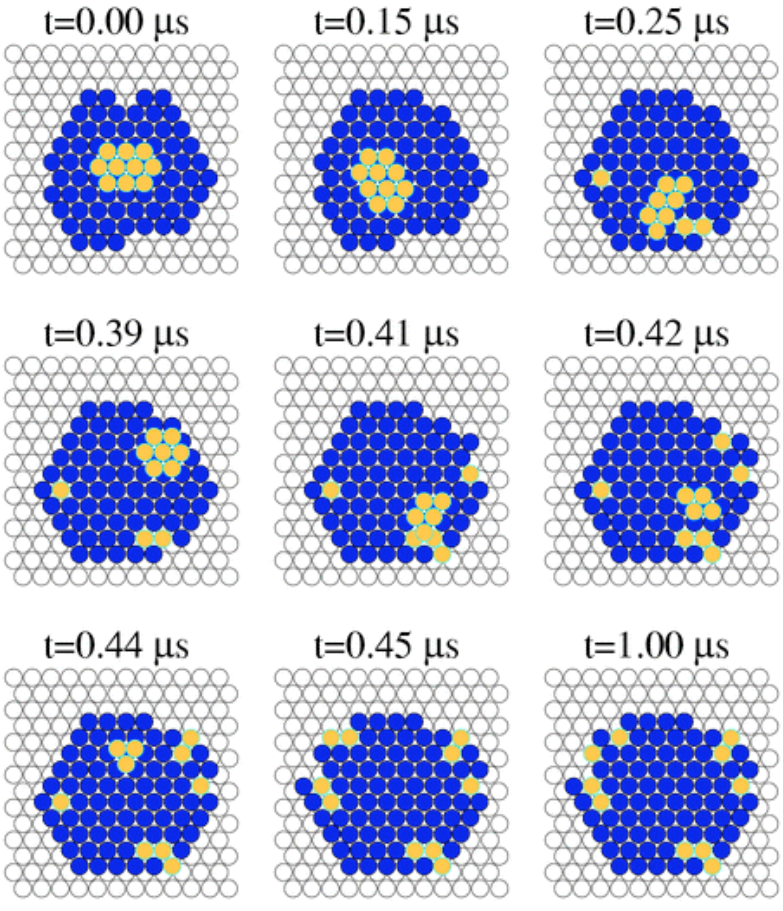
From Langevin dynamics to kinetic Monte Carlo: mathematical foundations of accelerated dynamics algorithms

T. Lelièvre

CERMICS - Ecole des Ponts ParisTech & Equipe Matherials - INRIA

Mathematical and Computational Materials Science,
IMSI workshop, 17th Feb. 2021

Molecular dynamics



Diffusion of adatoms on a surface (Courtesy of A. Voter, Los Alamos National Laboratory)

Molecular dynamics

Molecular dynamics consists in simulating on the computer the evolution of atomistic systems, as a numerical microscope:

- Understand the link bewteen macroscopic properties and microscopic ingredients
- Explore matter at the atomistic scale
- Simulate new materials, new molecules
- Interpret experimental results

Applications: materials science, biology, chemistry

Molecular dynamics comes of age:

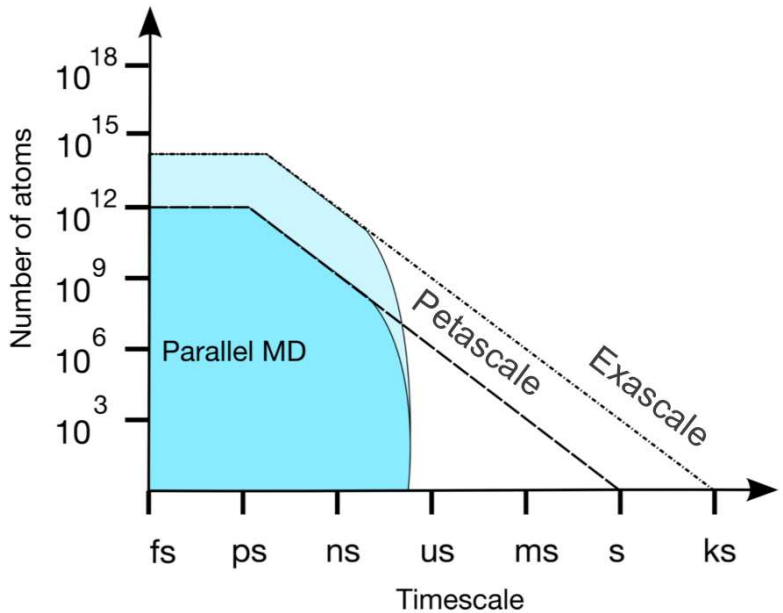
- 1/4 of CPU time worldwide is devoted to computations at the molecular scale
- 2013 Chemistry Nobel prize: Arieh Warshel, Martin Karplus and Michael Levitt. "Today the computer is just as important a tool for chemists as the test tube. Simulations are so realistic that they predict the outcome of traditional experiments."

Challenges

Main challenges:

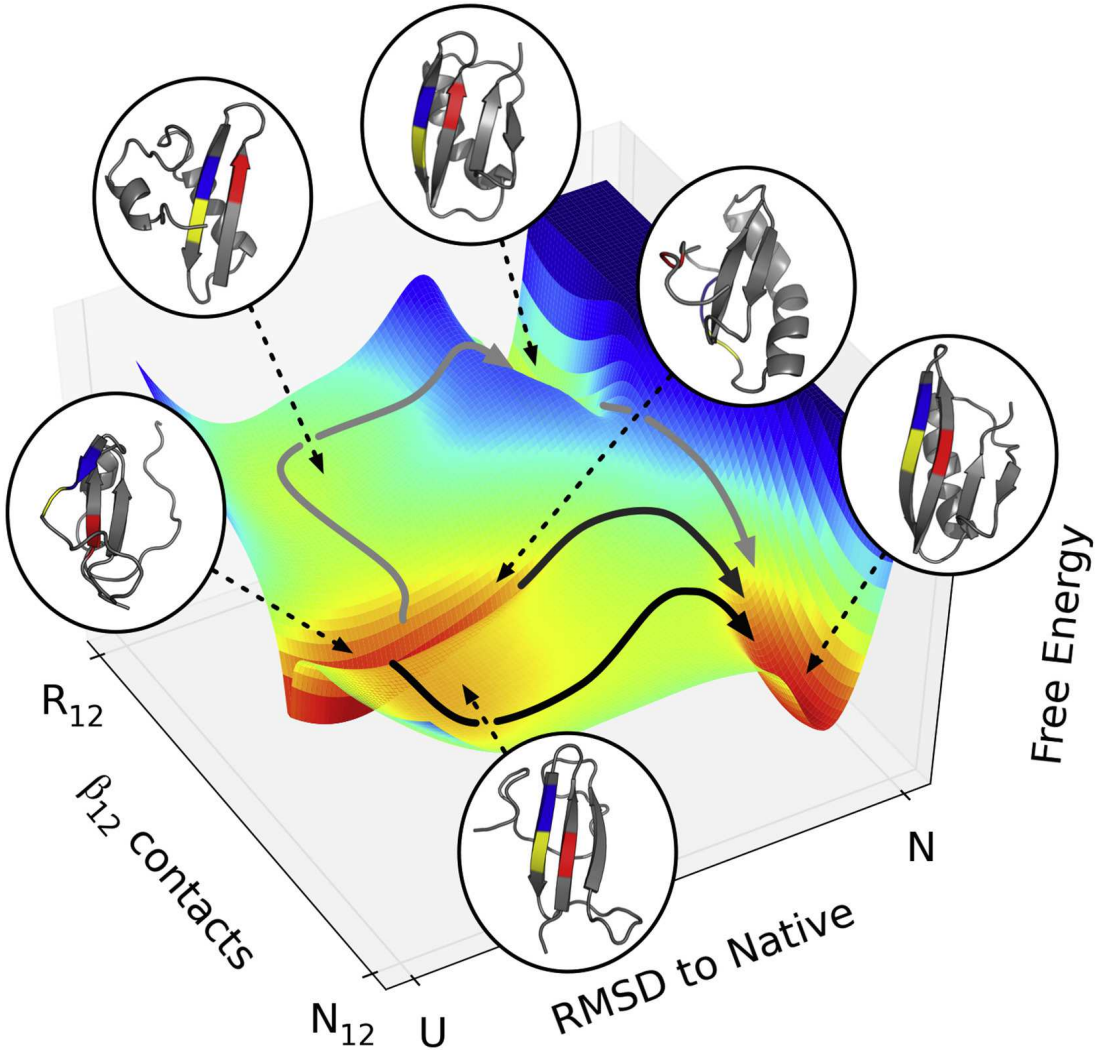
- Improved models (force fields, coarse-grained force fields): polarisability, water, chemical reactions
- Improved sampling methods (access long time scales): thermodynamic quantities, and dynamical properties
- Incorporate data: Bayesian approaches, data sciences

Spatial parallelism is very effective, but temporal reach of heroic brute force MD is limited to $1\mu s$ or less.



Courtesy of Danny Perez (LANL)

From Langevin to kinetic Monte Carlo



C.R. Schwantes, D. Shukla, V.S.Pande, Biophysical Journal, vol. 110, 2016

Two models for dynamics

The basic modeling ingredient in molecular dynamics: a potential function V which associates to a configuration $\boldsymbol{x} = (\boldsymbol{x}_1, ..., \boldsymbol{x}_{N_{atom}}) \in \mathbb{R}^{3N_{atom}}$ an energy $V(\boldsymbol{x}) \in \mathbb{R}$.

From V, two kinds of dynamics are considered:

- Langevin and overdamped Langevin dynamics: Markov processes with values in continuous state space ;
- kinetic Monte Carlo model or Markov state model (first order kinetics): Markov processes with values in discrete state space (jump Markov process).

Question: Can a mathematically rigorous link be made between these two kinds of models ?

Langevin and overdamped Langevin dynamics

Let us introduce the inverse temperature: $\beta^{-1} = k_B T$.

The *Langevin dynamics* write:

$$\begin{cases} d\boldsymbol{Q}_t = M^{-1}\boldsymbol{P}_t\, dt, \\ d\boldsymbol{P}_t = -\nabla V(\boldsymbol{Q}_t)\, dt - \gamma M^{-1}\boldsymbol{P}_t\, dt + \sqrt{2\gamma\beta^{-1}}\, d\boldsymbol{W}_t. \end{cases}$$

We will also consider the *overdamped Langevin dynamics*

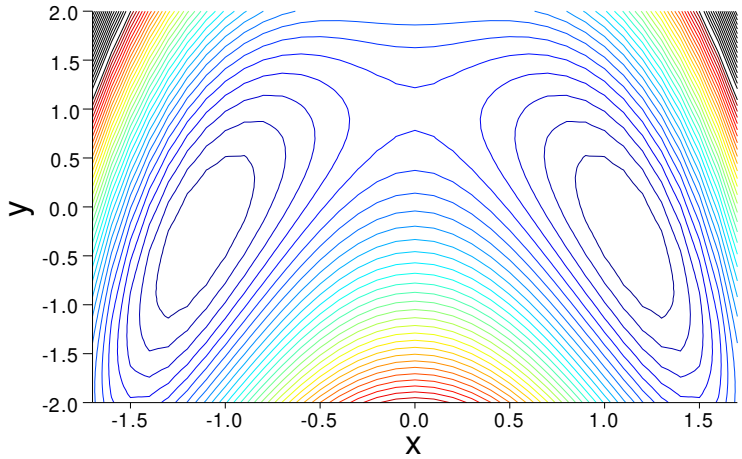
$$d\boldsymbol{Q}_t = -\nabla V(\boldsymbol{Q}_t)\, dt + \sqrt{2\beta^{-1}}\, d\boldsymbol{W}_t.$$

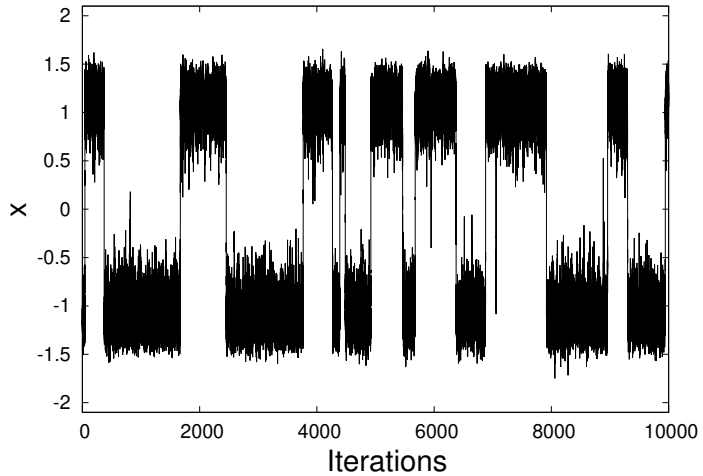
In the following $\boldsymbol{X}_t = (\boldsymbol{Q}_t, \boldsymbol{P}_t)$ or $\boldsymbol{X}_t = \boldsymbol{Q}_t$ denotes the associated Markov process.

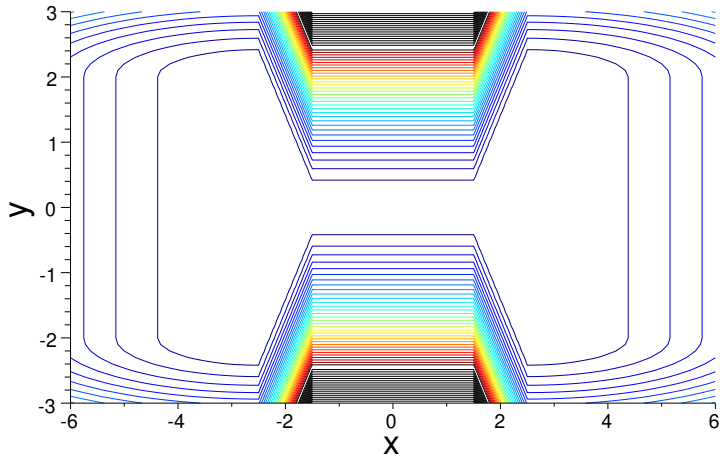
Main practical challenge: these dynamics are metastable.

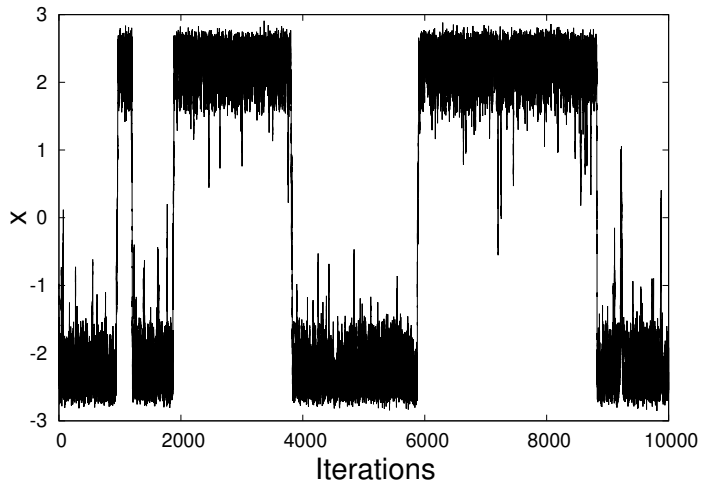
Metastability: energetic and entropic barriers

A two-dimensional schematic picture





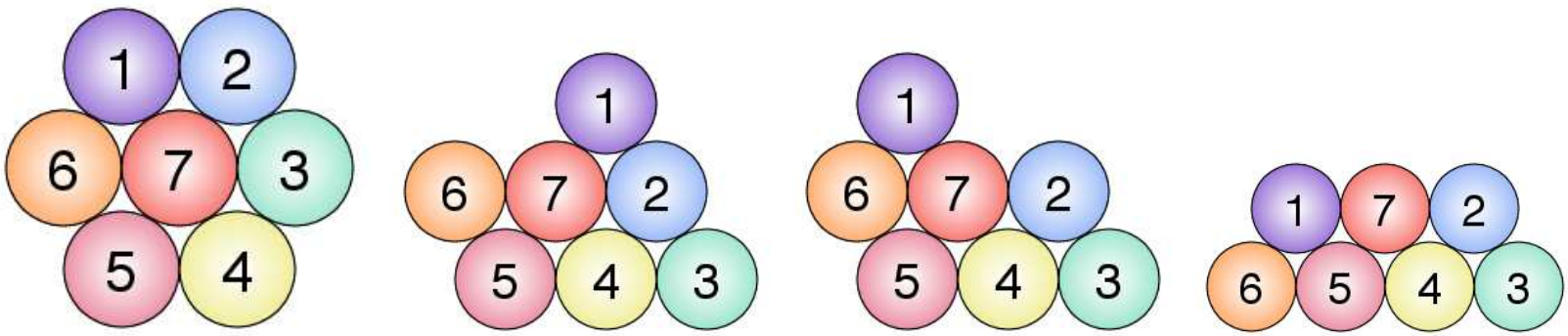




$\longrightarrow$ • Slow convergence of trajectorial averages

• Transitions between metastable states are rare events

Metastability: a toy example



(a) $V = -12.53$ (b) $V = -11.50$ (c) $V = -11.48$ (d) $V = -11.40$

Figure: Low energy conformations of the 7 atoms Lennard-Jones cluster.

$\longrightarrow$ simulation

The exit event

Let us consider a domain $\mathcal{O} \subset \mathbb{R}^d$ defined in position space. The associated state is $\mathcal{S} = \mathcal{O} \times \mathbb{R}^d$ for the Langevin dynamics and $\mathcal{S} = \mathcal{O}$ for the overdamped Langevin dynamics. The exit event from $\mathcal{O}$ is given by

$$(\tau_{\mathcal{O}}, \boldsymbol{X}_{\tau_{\mathcal{O}}})$$

where $\tau_{\mathcal{O}} = \inf\{t > 0,\ \boldsymbol{Q}_t \notin \mathcal{O}\} = \inf\{t > 0,\ \boldsymbol{X}_t \notin \mathcal{S}\}$.

Objective: build a jump Markov model to simulate the exit event $(\tau_{\mathcal{O}}, \boldsymbol{X}_{\tau_{\mathcal{O}}})$.

This is useful theoretically (justification of Markov state models and Eyring-Kramers laws) and numerically (accelerated dynamics *à la* Voter).

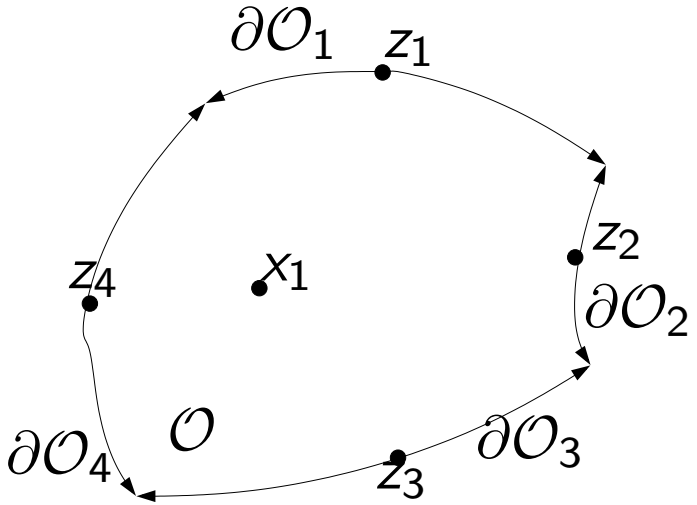
Kinetic Monte Carlo

Kinetic Monte Carlo (or Markov state) models are built as follows:

- define exit regions from $\mathcal{O}$: $\partial\mathcal{O} = \cup_{j=1}^{J} \partial\mathcal{O}_j$
- associate a rate k_j with an exit through $\partial\mathcal{O}_j$

and then (jump Markov model)

- the exit time $\tau_{\mathcal{O}}^{kMC}$ is exponentially distributed with parameter $\sum_{j=1}^{J} k_j$
- the exit region is $I_{\mathcal{O}}^{kMC}$ with law $\mathbb{P}(I_{\mathcal{O}}^{kMC} = i) = \frac{k_i}{\sum_{j=1}^{J} k_j}$
- $I_{\mathcal{O}}^{kMC}$ and $\tau_{\mathcal{O}}^{kMC}$ are independent random variables

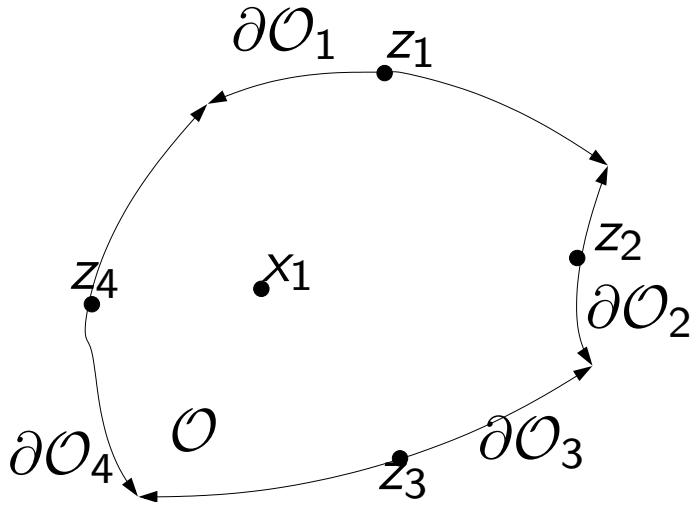


Eyring-Kramers laws

Formulas for transition rates. Let us introduce the local minima $(z_j)_{j=1,\dots,J}$ of V on $\partial\mathcal{O}$, and associated exit regions $\partial\mathcal{O}_j$. The parameters k_j are computed using the Eyring-Kramers formula (Harmonic Transition State Theory):

$$k_j^{HTST} = \nu_j \, \mathrm{e}^{-\beta[V(z_j)-V(x_1)]}$$

where ν_j is an explicit prefactor and $x_1 = \arg\min_{\mathcal{O}} V$.



A theoretical question

Question: can we relate the exit event $(\tau_{\mathcal{O}}, \boldsymbol{X}_{\tau_{\mathcal{O}}})$ for the original dynamics with the exit event $(\tau_{\mathcal{O}}^{kMC}, I_{\mathcal{O}}^{kMC})$ for the jump Markov process?

Two steps:

- Step 1: Introduce the Quasi-Stationary Distribution (both for overdamped Langevin and Langevin)
  $\longrightarrow$ justify the use of a kMC model
- Step 2: Consider the small temperature regime $\beta \to \infty$ (only for the overdamped Langevin)
  $\longrightarrow$ justify the use of the Eyring-Kramers laws

Step 1: The Quasi-Stationary Distribution

Definition of the QSD: A probability measure ν with support $\mathcal{S}$ is a QSD for the Markov process $(\boldsymbol{X}_t)_{t\geq 0}$ iff for all $t > 0$,

$$\boldsymbol{X}_0 \sim \nu \Longrightarrow \mathcal{L}(\boldsymbol{X}_t | \tau_{\mathcal{O}} > t) = \nu$$

Existence, uniqueness, convergence: Assume $\mathcal{O}$ is bounded. For the Langevin and the overdamped Langevin dynamics, there exists a unique QSD ν in $\mathcal{S}$. Moreover, for any $\boldsymbol{X}_0$ in $\mathcal{S}$,

$$\lim_{t\to\infty} \mathcal{L}(\boldsymbol{X}_t | \tau_{\mathcal{O}} > t) = \nu.$$

Remark: Quantitative definition of a metastable exit:

exit time$\gg$ local equilibration time

Quasi-Stationary Distribution for Langevin

Difficulties: (i) $\mathcal{O} \times \mathbb{R}^d$ unbounded and (ii) hypoelliptic and non-reversible infinitesimal generator.

Ingredients of the proof: [TL, Ramil, Reygner, 2021]

- The QSD is the first eigenvector of the transition operator $f \mapsto \mathbb{E}(f(\boldsymbol{Q}_t, \boldsymbol{P}_t)1_{t<\tau_{\mathcal{O}}})$ (and the associated infinitesimal generator) with absorbing boundary conditions
- Probabilistic representation of the solution to the kinetic Fokker-Planck equation on $\mathcal{O} \times \mathbb{R}^d$, with Dirichlet boundary condition on $\Gamma^+ = \{(q,p) \in \partial\mathcal{O} \times \mathbb{R}^d, p \cdot n(q) > 0\}$
- Gaussian upper-bound on the transition density of the absorbed process, using the parametrix method [Konakov, Menozzi, Molchanov, 2010]
- Compactness of the transition operator
- Krein-Rutman theorem

Step 1: The Quasi-Stationary Distribution

Fundamental property of the QSD. Assume $\boldsymbol{X}_0 \sim \nu$, then:

- the first exit time $\tau_{\mathcal{O}}$ is exponentially distributed since:

$$\begin{aligned}\mathbb{P}^{\nu}(\tau_{\mathcal{O}} > s+t) &= \mathbb{P}^{\nu}(\tau_{\mathcal{O}} > s+t | \tau_{\mathcal{O}} > s)\mathbb{P}^{\nu}(\tau_{\mathcal{O}} > s) \\ &= \mathbb{P}^{\nu}(\tau_{\mathcal{O}} > t)\mathbb{P}^{\nu}(\tau_{\mathcal{O}} > s)\end{aligned}$$

- and $\tau_{\mathcal{O}}$ is independent of the first hitting point $\boldsymbol{X}_{\tau_{\mathcal{O}}}$ since:

$$\begin{aligned}\mathbb{P}^{\nu}(\boldsymbol{X}_{\tau_{\mathcal{O}}} \in A, \tau_{\mathcal{O}} \leq t) &= \mathbb{P}^{\nu}(\boldsymbol{X}_{\tau_{\mathcal{O}}} \in A) - \mathbb{P}^{\nu}(\boldsymbol{X}_{\tau_{\mathcal{O}}} \in A | \tau_{\mathcal{O}} > t)\mathbb{P}^{\nu}(\tau_{\mathcal{O}} > t) \\ &= \mathbb{P}^{\nu}(\boldsymbol{X}_{\tau_{\mathcal{O}}} \in A) - \mathbb{P}^{\nu}(\boldsymbol{X}_{\tau_{\mathcal{O}}} \in A)\mathbb{P}^{\nu}(\tau_{\mathcal{O}} > t) \\ &= \mathbb{P}^{\nu}(\boldsymbol{X}_{\tau_{\mathcal{O}}} \in A)\mathbb{P}^{\nu}(\tau_{\mathcal{O}} \leq t)\end{aligned}$$

Consequence: Starting from ν, the exit event from $\mathcal{O}$ can be exactly written as one jump of a kinetic Monte Carlo model with rates

$$k_i = \frac{\mathbb{P}^{\nu}(\boldsymbol{X}_{\tau_{\mathcal{O}}} \in \partial\mathcal{O}_i)}{\mathbb{E}^{\nu}(\tau_{\mathcal{O}})}.$$

Step 2: The small temperature regime

Let us consider the overdamped Langevin dynamics. One has explicit formulas for $\mathbb{E}(\tau_{\mathcal{O}})$ and the distribution of $\boldsymbol{X}_{\tau_{\mathcal{O}}}$. Let us introduce the first eigenstate (λ_1, u_1) of the Fokker-Planck operator associated with the dynamics with Dirichlet boundary conditions on $\partial\mathcal{O}$:

$$\begin{cases} \mathrm{div}\,(\nabla V u_1) + \beta^{-1}\Delta u_1 = -\lambda_1 u_1 \text{ on } \mathcal{O}, \\ u_1 = 0 \text{ on } \partial\mathcal{O}. \end{cases}$$

Then, $\nu = \frac{u_1(\boldsymbol{x})d\boldsymbol{x}}{\int_{\mathcal{O}} u_1}$

$$\mathbb{E}^{\nu}(\tau_{\mathcal{O}}) = \frac{1}{\lambda_1}$$

and

$$\mathbb{P}^{\nu}(\boldsymbol{X}_{\tau_{\mathcal{O}}} \in \partial\mathcal{O}_i) = -\frac{\int_{\partial\mathcal{O}_i} \partial_n u_1 \, d\sigma}{\beta\lambda_1 \int_{\mathcal{O}} u_1(x)\, dx}.$$

Thus, $k_i = -\frac{\int_{\partial\mathcal{O}_i} \partial_n u_1 \, d\sigma}{\beta\lambda_1 \int_{\mathcal{O}} u_1(x)\, dx}$. Can we then show that $k_i \simeq k_i^{HTST}$?

Justifying Eyring-Kramers laws

Theorem [Di Gesu, TL, Le Peutrec, Nectoux, 2019]
For the overdamped Langevin dynamics and under some geometric assumptions, starting from the QSD, the exit rates are

$$k_i = C_i^{OL}\, \mathrm{e}^{-\beta[V(z_i)-V(x_1)]}\,(1+O(\beta^{-1}))$$

in the limit $\beta \to \infty$, where

$$C_i^{OL} = \sqrt{\frac{\beta}{2\pi}}\partial_n V(z_i)\frac{\sqrt{\det(\nabla^2 V)(x_1)}}{\sqrt{\det(\nabla^2 V_{|\partial\mathcal{O}})(z_i)}}.$$

Assumptions (1/2)

- $\mathcal{O}$ is an open bounded smooth domain in $\mathbb{R}^d$.
- $V : \overline{\mathcal{O}} \to \mathbb{R}$ is a Morse function with a single critical point x_1. Moreover, $x_1 \in \mathcal{O}$ and $V(x_1) = \min_{\overline{\mathcal{O}}} V$.
- $\partial_n V > 0$ on $\partial\mathcal{O}$ and $V|_{\partial\mathcal{O}}$ is a Morse function with local minima reached at $z_1, \ldots, z_J$ with $V(z_1) < \ldots < V(z_J)$.
- $V(z_1) - V(x_1) > V(z_J) - V(z_1)$
- $\forall i \in \{1, \ldots, J\}$, consider B_{z_i} the basin of attraction of z_i for the dynamics $\dot{x} = -\nabla_T V(x)$ and assume that

$$\inf_{z \in B_{z_i}^c} d_a(z, z_i) > V(z_J) - V(z_1)$$

Assumptions (2/2)

Here, d_a is the Agmon distance:

$$d_a(x,y) = \inf_{\gamma} \int_0^1 g(\gamma(t))|\gamma'(t)|\,dt$$

where $g = \begin{cases} |\nabla V| \text{ in } \mathcal{O} \\ |\nabla_T V| \text{ in } \partial\mathcal{O} \end{cases}$, and the infimum is over all piecewise $\mathcal{C}^1$ paths $\gamma : [0,1] \to \overline{\mathcal{O}}$ such that $\gamma(0) = x$ and $\gamma(1) = y$.

Numerical tests indicate that the assumption

$$\forall i \in \{1, \dots J\}, \inf_{z \in B_{z_i}^c} d_a(z, z_i) > V(z_I) - V(z_1)$$

seems indeed necessay to get the expected results.

Sketch of the proof (1/3)

The difficult part is to find an approximation for $\int_{\partial\mathcal{O}_i} \partial_n u_1 = \int_{\partial\mathcal{O}_i} \partial_n v_1 e^{-\beta V}$, where $v_1 = u_1 e^{\beta V}$.
We have

$$\begin{cases} L^{(0)} v_1 = -\lambda_1 v_1 \text{ on } \mathcal{O}, \\ v_1 = 0 \text{ on } \partial\mathcal{O}, \end{cases}$$

where $L^{(0)} = \beta^{-1}\Delta - \nabla V \cdot \nabla$ is a self adjoint operator on $L^2(e^{-\beta V})$. We are interested in $\nabla v_1 \cdot n$, and ∇v_1 satisfies

$$\begin{cases} L^{(1)} \nabla v_1 = -\lambda_1 \nabla v_1 \text{ on } \mathcal{O}, \\ \nabla_T v_1 = 0 \text{ on } \partial\mathcal{O}, \\ (\beta^{-1}\mathrm{div} \ - \nabla V \cdot)\nabla v_1 = 0 \text{ on } \partial\mathcal{O}, \end{cases}$$

where

$$L^{(1)} = \beta^{-1}\Delta - \nabla V \cdot \nabla - \mathrm{Hess}(V).$$

Therefore ∇v_1 is an eigenvector (eigen-1-form) of $-L^{(1)}$ associated with the small eigenvalue λ_1.

Sketch of the proof (2/3)

Let $\Pi^{(p)} = 1_{[0,\beta^{-3/2}]}(-L^{(p)})$ be the spectral projection operator on small eigenvalues. We know [Helffer,Sjöstrand] that, for β large, $\dim(\mathrm{Ran}\Pi^{(0)}) = 1$ and $\dim(\mathrm{Ran}\Pi^{(1)}) = J$:

$$\mathrm{Ran}\Pi^{(0)} = \mathrm{Span}(v_1)$$

$$\mathrm{Ran}\Pi^{(1)} = \mathrm{Span}(\psi_1, \ldots, \psi_J).$$

Since $\nabla v_1 \in \mathrm{Ran}\Pi^{(1)}$,

$$\int_{\partial\mathcal{O}_i} \partial_n v_1 e^{-\beta V} = \sum_{j=1}^{J} \langle \nabla v_1, \psi_j \rangle_{L^2(e^{-\beta V})} \int_{\partial\mathcal{O}_i} \psi_j \cdot n e^{-\beta V}.$$

The idea is now to build so-called quasi-modes which approximate the eigenvectors of $L^{(0)}$ and $L^{(1)}$ associated with small eigenvalues in the regime $\beta \to \infty$, in order to approximate the terms in the sum.

Sketch of the proof (3/3)

- $\mathrm{Ran}\Pi^{(0)}$: an approximation of v_1 is given by

$$\tilde{v} = Z^{-1}\chi_{\mathcal{O}'}$$

  where $\mathcal{O}' \subset\subset \mathcal{O}$.
- $\mathrm{Ran}\Pi^{(1)}$: an approximation of $\mathrm{Ran}\Pi^{(1)}$ is $\mathrm{Span}(\tilde{\psi}_1, \ldots, \tilde{\psi}_J)$ where $(\tilde{\psi}_i)_{1\leq i\leq J}$ are solutions to auxiliary eigenvalue problems, attached to the local minima $(z_i)_{1\leq i\leq J}$.

Two tools:

- Agmon estimates (the support of $\tilde{\psi}_i$ is essentially in a neighborhood of z_i):

$$\exists N > 0,\ \|e^{\beta d_a(z_i,\cdot)/2}\tilde{\psi}_i\|_{H^1(e^{-\beta V})} = O(\beta^N).$$

- WKB approximations:

$$\exists N > 0,\ \tilde{\psi}_i \simeq Z_i^{-1} d(e^{\beta V/2} e^{-\beta d_a(z_i,\cdot)/2})\beta^p.$$

Generalizations and perspectives

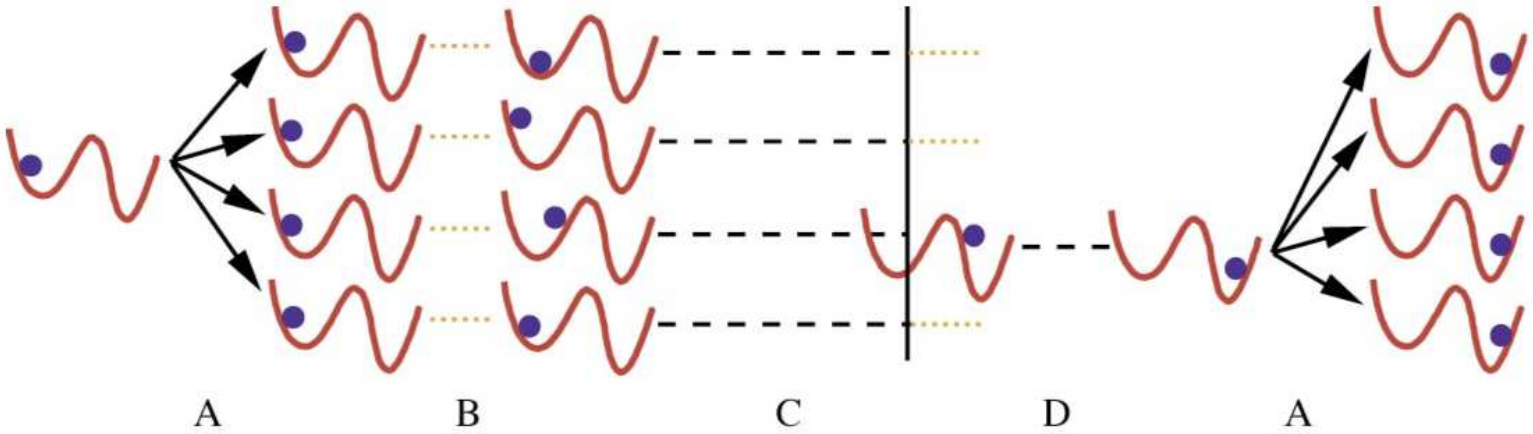
If the state is metastable, the QSD is a good intermediate between continuous-state space dynamics and jump Markov models.

We are working on generalizations:

- Broader geometric setting
- Langevin dynamics
- Non-reversible dynamics

The mathematical analysis gives the proper geometric setting under which the kinetic Monte Carlo model can be built and the Eyring-Kramers formulas can be used to parameterize it.

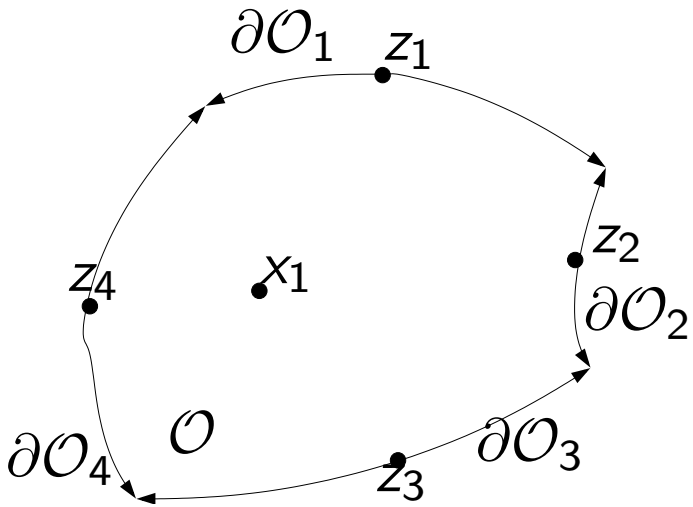
From theory to algorithms



A.F. Voter, Annu. Rev. Mater. Res., vol. 32, 2002.

How to sample efficiently the exit event?

If the process remains sufficiently long in a state, the exit event can be modeled by one jump of a Markov state model. This can be used to simulate efficiently the exit event: accelerated dynamics *à la* A.F. Voter.



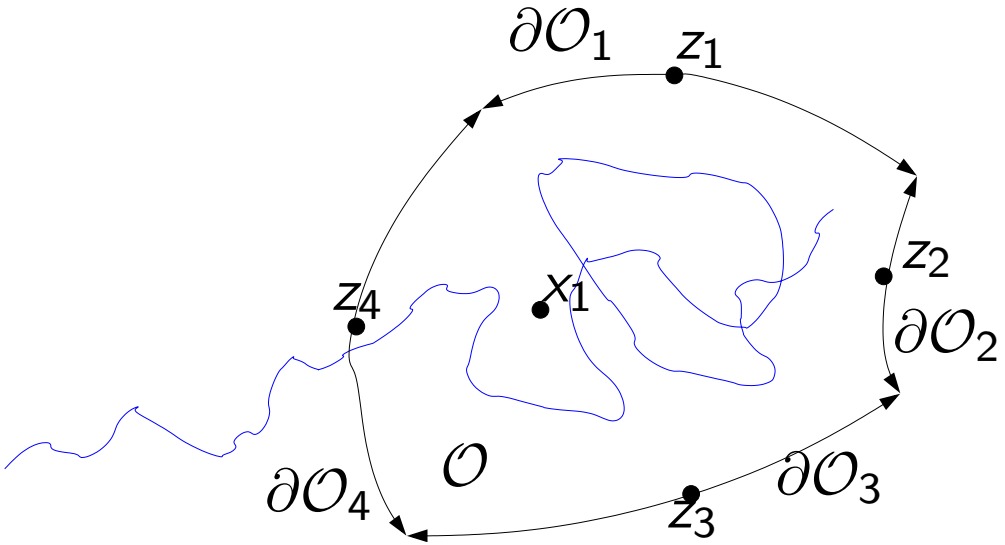
Two steps:

- Estimate the decorrelation time, namely the time to reach the QSD
- Use the underlying jump Markov process to efficiently sample the exit event

Decorrelation time

How long should we wait in practice so that $\mathcal{L}(\boldsymbol{X}_t|\tau_{\mathcal{O}} > t)$ is close to the QSD ν?

- Theoretically: exponential decay
  $\|\mathcal{L}(\boldsymbol{X}_t|\tau_{\mathcal{O}} > t) - \nu\|_{TV} \leq C(\mathcal{L}(\boldsymbol{X}_0)) \exp\left(-(\lambda_2 - \lambda_1)t\right)$;
- Numerically: simulate $\mathcal{L}(\boldsymbol{X}_t|\tau_{\mathcal{O}} > t)$ via an interacting particle system (Fleming-Viot particle system), and test stationarity to estimate the convergence time to the QSD (Gelman-Rubin convergence diagnostic).



The Fleming-Viot particle process

Start N processes i.i.d. from μ_0, and iterate the following steps:

1. Integrate (in parallel) N realizations ($k = 1, \ldots, N$)
$$d\boldsymbol{X}_t^k = -\nabla V(\boldsymbol{X}_t^k)\, dt + \sqrt{2\beta^{-1}} d\boldsymbol{W}_t^k$$
until one of them, say $\boldsymbol{X}_t^1$, exits;
2. Kill the process that exits;
3. With uniform probability $1/(N-1)$, randomly choose one of the survivors, $\boldsymbol{X}_t^2, \ldots, \boldsymbol{X}_t^N$, say $\boldsymbol{X}_t^2$;
4. Branch $\boldsymbol{X}_t^2$, with one copy persisting as $\boldsymbol{X}_t^2$, and the other becoming the new $\boldsymbol{X}_t^1$.

It is known that the empirical distribution [Villemonais]

$$\mu_{t,N} \equiv \frac{1}{N} \sum_{k=1}^{N} \delta_{\boldsymbol{X}_t^k}$$

satisfies:

$$\lim_{N\to\infty} \mu_{t,N} = \mathcal{L}(\boldsymbol{X}_t | t < \tau_{\mathcal{O}}).$$

Accelerated dyamics

Once the QSD has been reached, there are three ideas to efficiently sample $(\tau_{\mathcal{O}}, \boldsymbol{X}_{\tau_{\mathcal{O}}})$:

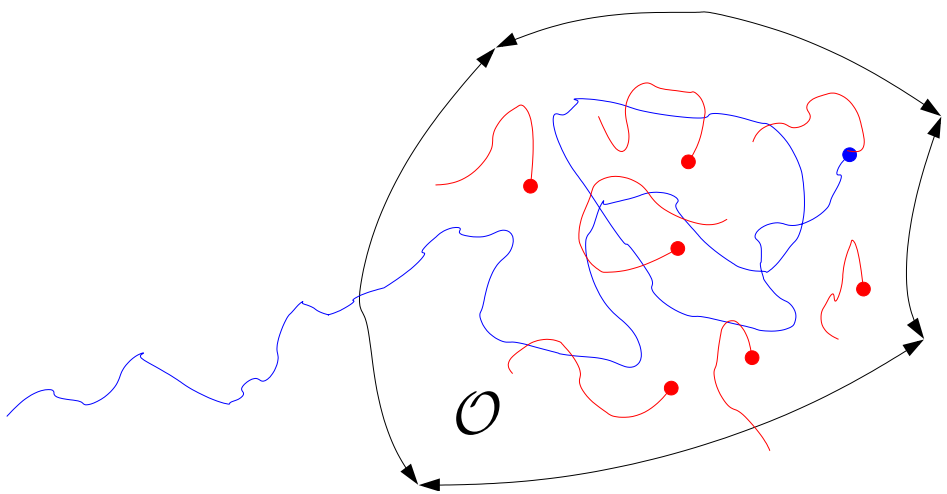
- use parallel architectures to accelerate the sampling: parallel replica, parsplicing
- raise the minimum of the potential inside the state $\mathcal{O}$ (but not on $\partial\mathcal{O}$): hyperdynamics
- raise the temperature: temperature accelerated dynamics

The Parallel Replica Algorithm

Perform many independent exit events in parallel [Voter, 1998]

Two steps:

- Distribute N independent initial conditions in $\mathcal{O}$ according to the QSD ν ;
- Evolve N replicas from these initial conditions, consider the first exiting replica, and multiply the first exit time by the number of replicas.



The Parallel Replica Algorithm

Why is it consistent?

- Exit time is independent of exit point so that

$$\boldsymbol{X}^{I_0}_{\tau^{I_0}_{\mathcal{O}}} \stackrel{\mathcal{L}}{=} \boldsymbol{X}^{1}_{\tau^{1}_{\mathcal{O}}},$$

  where $I_0 = \arg\min_i(\tau^i_{\mathcal{O}})$;
- Exit times are i.i.d. exponentially distributed so that, for all N,

$$N \min(\tau^1_{\mathcal{O}}, \ldots, \tau^N_{\mathcal{O}}) \stackrel{\mathcal{L}}{=} \tau^1_{\mathcal{O}}.$$

Remark: For this algorithm, one just needs two properties: $\tau_{\mathcal{O}}$ is exponentially distributed, and independent of the exit point $\boldsymbol{X}_{\tau_{\mathcal{O}}}$. The Eyring-Kramers formulas are not used.

The generalized Parallel Replica algorithm

[Binder, Hédin, TL, Simpson]

1. Run a reference walker, using standard MD.
2. Each time the reference walker enters a state, start a Fleming-Viot particle process (with N replicas simulated in parallel) with initial condition the entering point.
3. If the reference walker exits before the Fleming Viot particle process reaches stationarity go back to 1. Else go to the parallel step.
4. Parallel step: Starting from the end points of the Fleming-Viot particle process (approximately i.i.d. with law the QSD), run independent MD and consider the first exit event. Multiply the first exit time by N and go back to 1, using the first exit point as initial condition.

The time at which the Fleming-Viot particle process becomes stationary is determined using the Gelman-Rubin statistical test.

The generalized Parallel Replica algorithm

- The algorithm does not require a partition of the state space but only an ensemble of states.
- The time to reach the QSD is estimated each time the process enters a new state (it depends on the state and on the initial condition within the state).

Numerical results

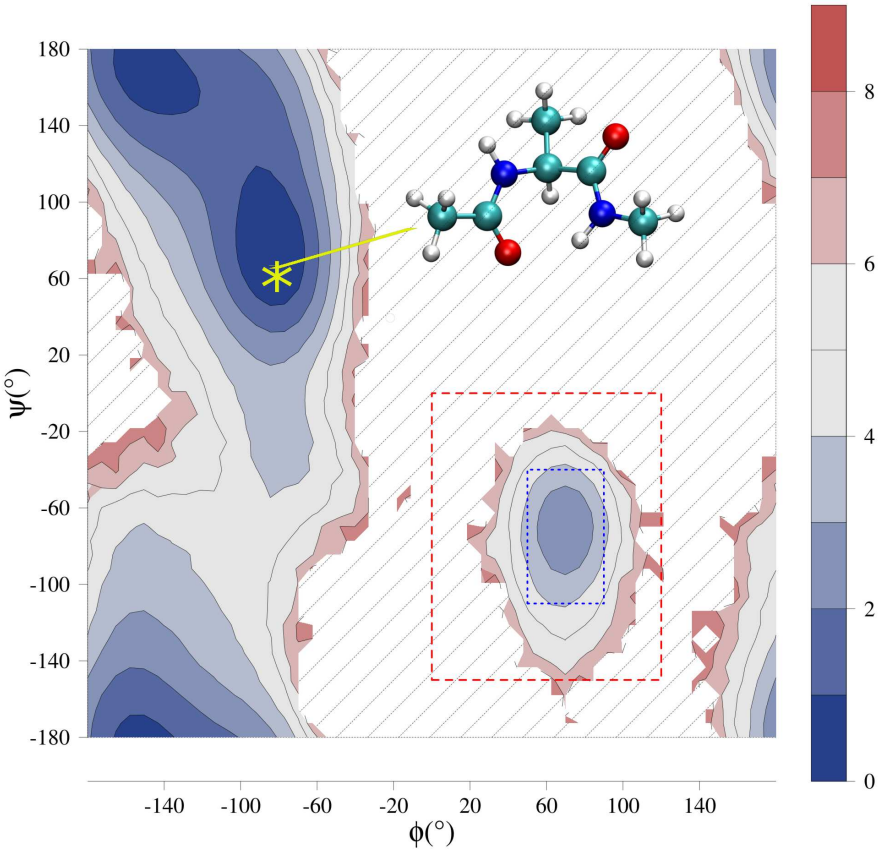
We recently tested the generalized Parallel Replica algorithm applied to biological systems (postdoc Florent Hédin):

- Conformational equilibrium of the alanine dipeptide
- Dissociation of the FKBP-DMSO protein-ligand system

Main differences with materials science: definition of the states using collective variables, the states do not define a partition, much more rugged landscapes.

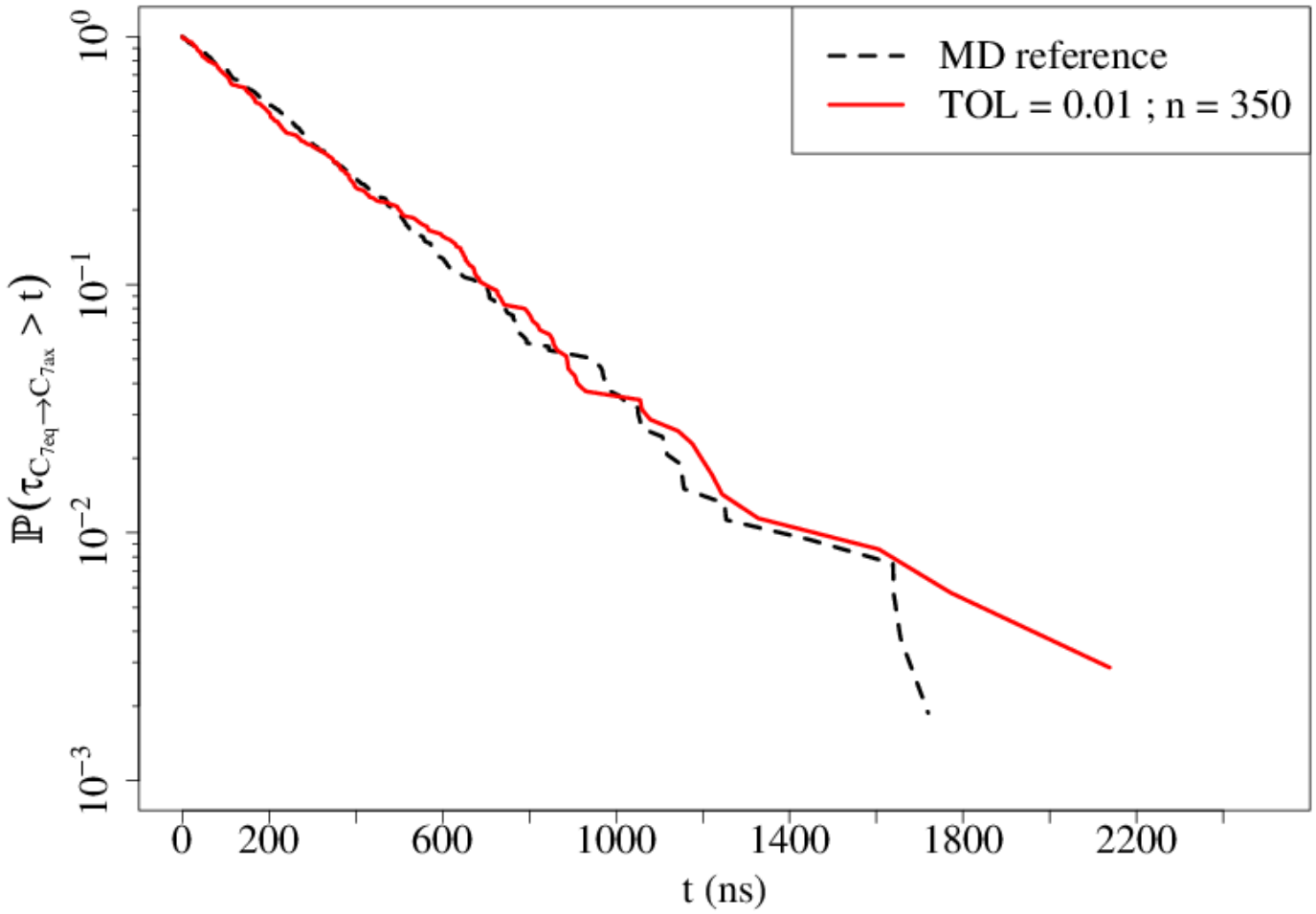
Current implementation within OpenMM, see
`https://gitlab.inria.fr/parallel-replica`

Alanine dipeptide (1/5)



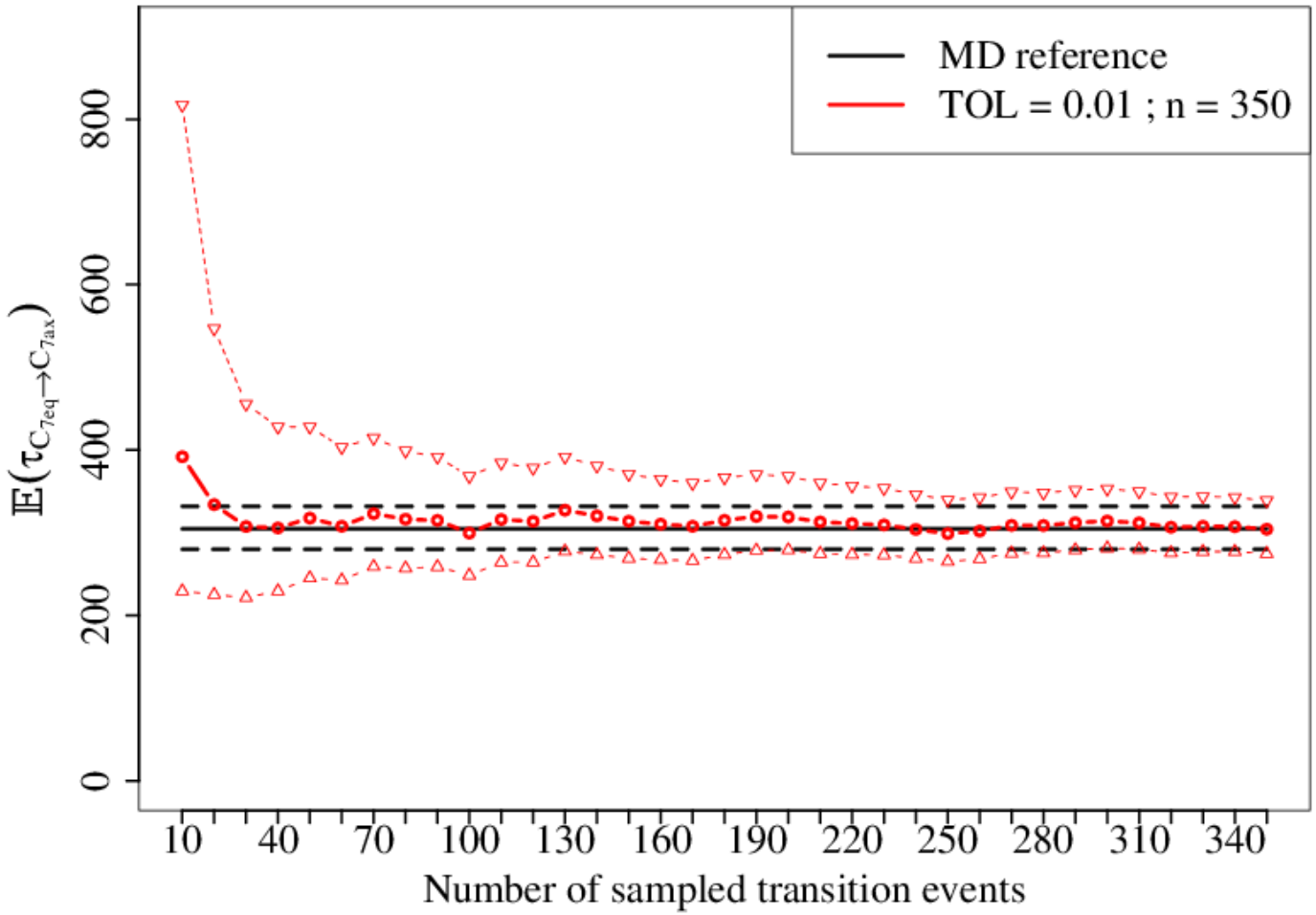
Definition of ParRep domains based on a free energy surface: we study the transition time from C_{7eq} (outside the red rectangle) to C_{7ax} (inside the red rectangle).

Alanine dipeptide (2/5)



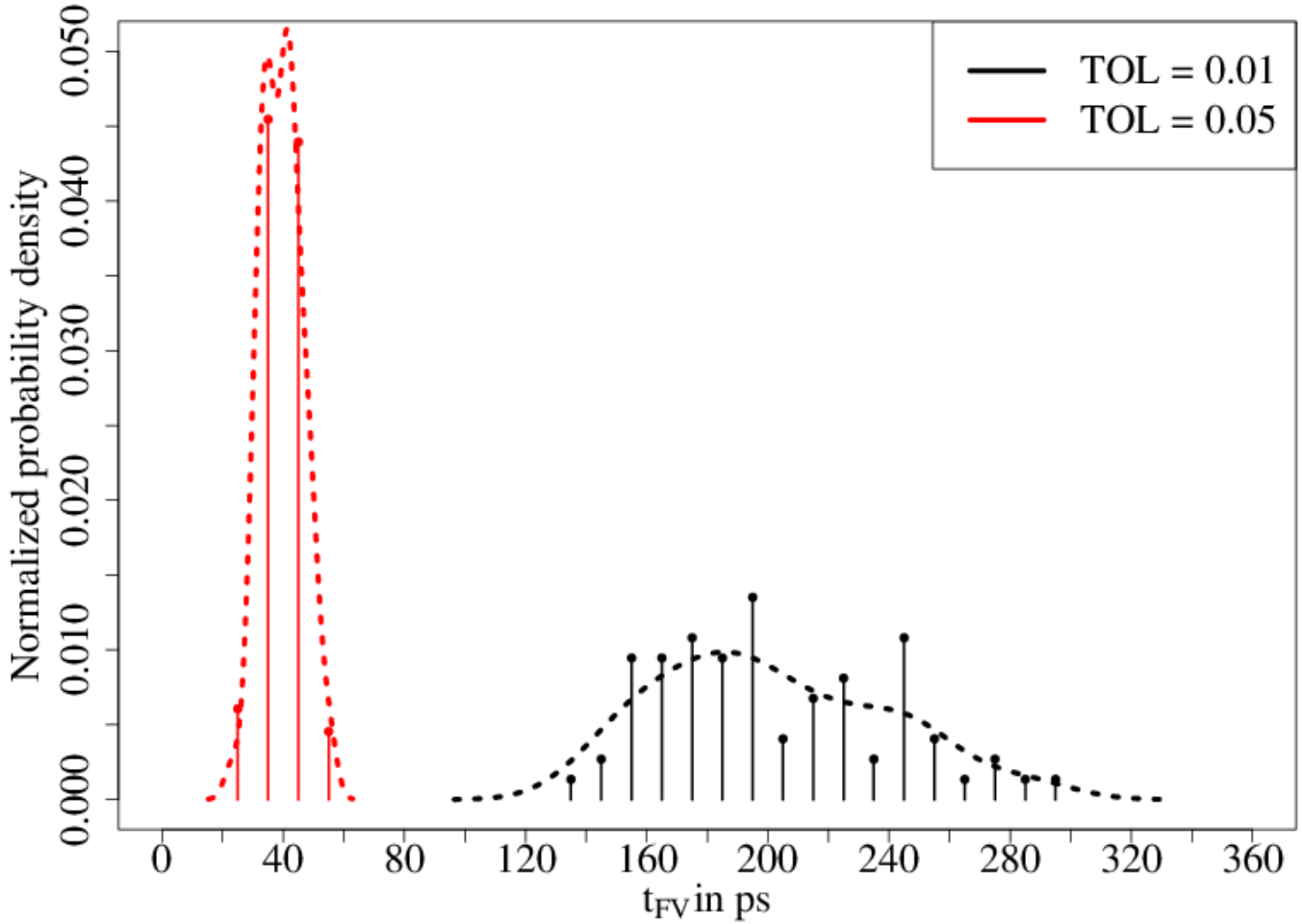
Cumulative distribution function of the transition time.

Alanine dipeptide (3/5)



Convergence of the mean transition time.

Alanine dipeptide (4/5)



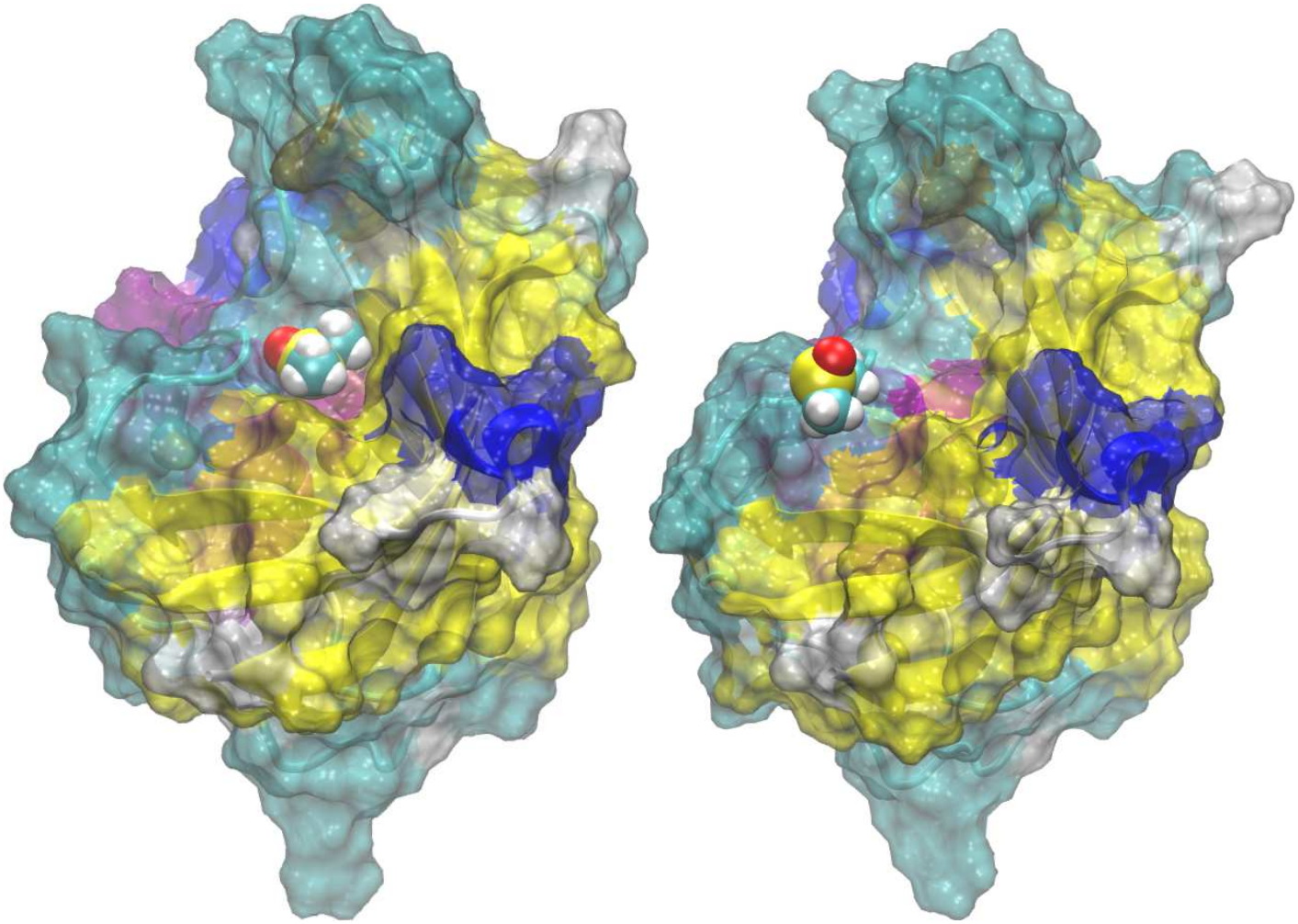
Distribution of the correlation times computed by FV.

Alanine dipeptide (5/5)

tol	WT(s)	t_{sim}(ns)	Speed(ns/day)	Eff. speedup	(Eff./Max)
0.01	6015	10008	143752	156	70%
0.025	5239	10103	166609	181	80%
0.05	4973	10032	174296	189	84%

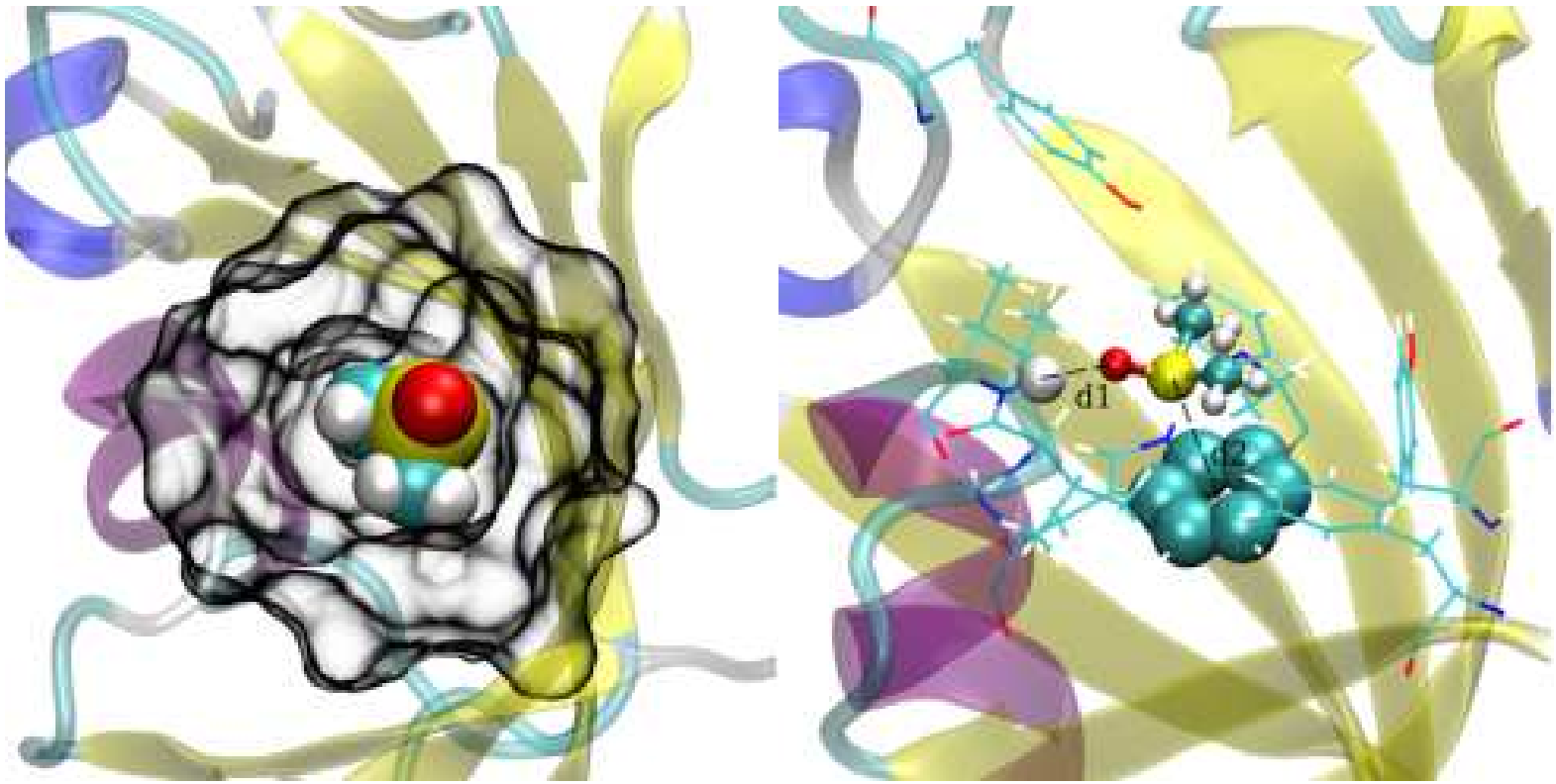
Effective speed-up as a function of the tolerance, for $N = 224$ replicas run in parallel (speed of a reference Langevin dynamics is 921 ns/day).

FKBP-DMSO (1/4)



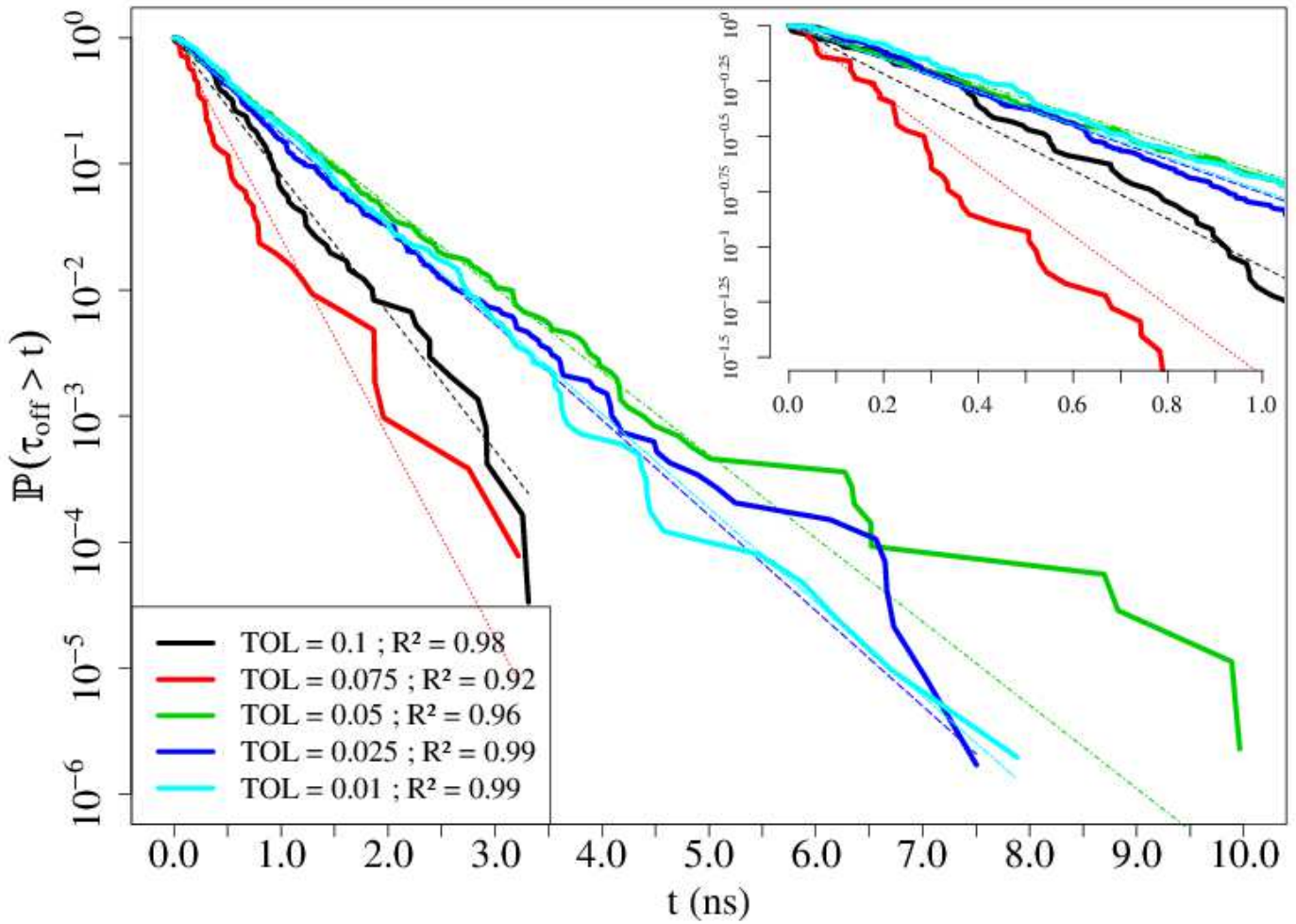
FKBP-DMSO complex,
corresponding to the RCSB-PDB entry “1D7H”.

FKBP-DMSO (2/4)



DMSO in its binding cavity ; distances used to define the cavity.

FKBP-DMSO (3/4)



Cumulative distribution function of the dissociation times.

FKBP-DMSO (4/4)

TOL	WT(s)	t_{sim}(ns)	Speed (ns/day)	Eff. speedup	(Eff./Max)
0.01	85142	403.5	409.4	79.5	56.8%
0.025	79574	457.6	496.8	96.5	68.9%
0.05	84455	482.2	493.4	95.8	68.4%

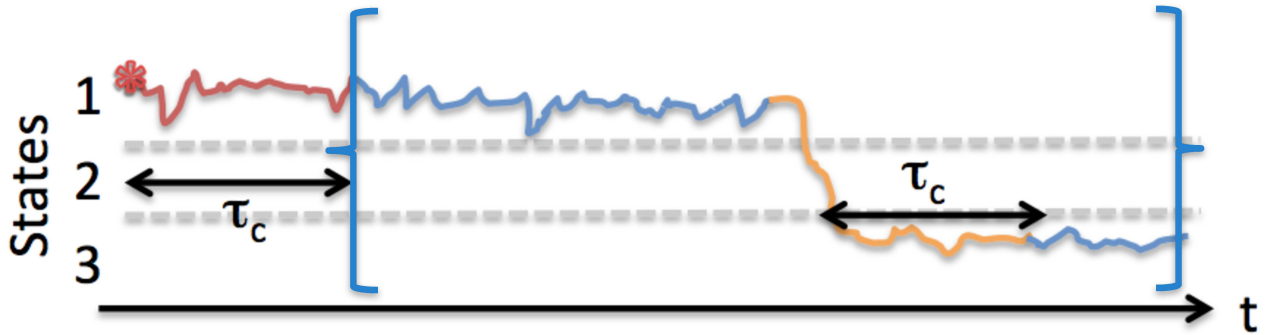
Effective speed-up as a function of the tolerance, for $N = 140$ replicas run in parallel (speed of a reference Langevin dynamics is 5.15 ns/day).

The Parallel Trajectory Splicing algorithm

Precompute the exit events [Perez, Cubuk, Waterland, Kaxiras, Voter, 2015]

Algorithm:

- Simulate in parallel short trajectories which start from the QSD in a state, and end at the QSD in a state.
- Glue together these short trajectories to build the full dynamics.

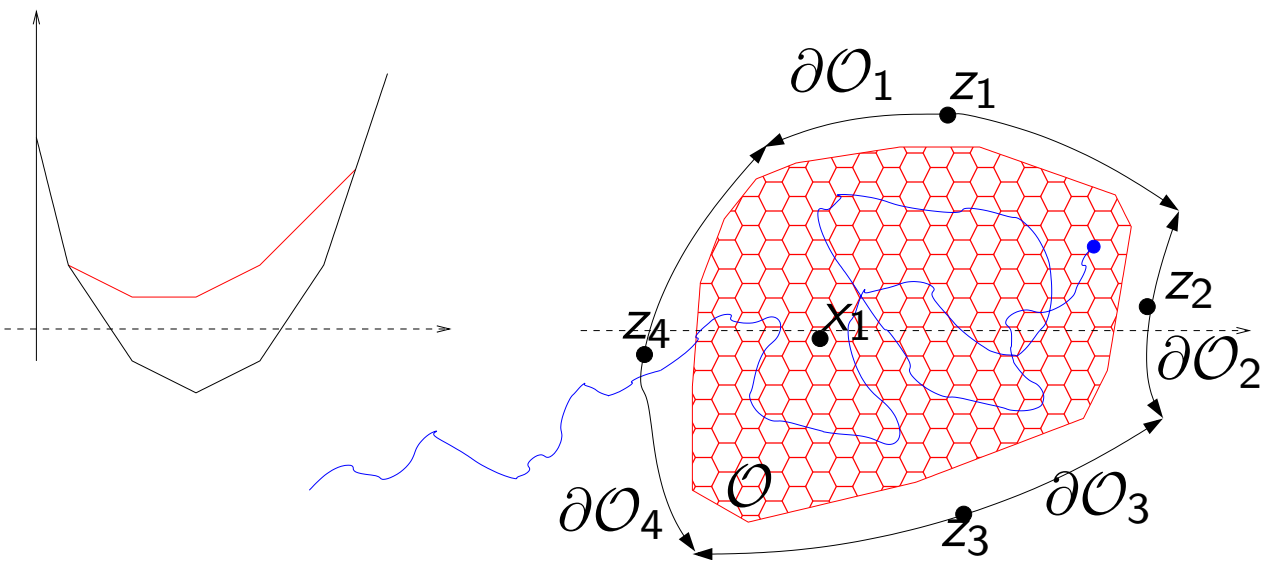


Hyperdynamics (1/2)

Raise the potential in $\mathcal{O}$ to reduce the exit time [Voter, 1997]

Two steps:

- Equilibrate on the biased potential $V + \delta V$;
- Wait for an exit and multiply the exit time $\tau_{\mathcal{O}}^{\delta V}$ by the boost factor $B = \frac{1}{\tau_{\mathcal{O}}^{\delta V}} \int_0^{\tau_{\mathcal{O}}^{\delta V}} \exp(\beta\, \delta V(\boldsymbol{X}_t))\, dt$.



Hyperdynamics (2/2)

Why is it consistent ?

Assumptions on δV: (i) $\delta V = 0$ on ∂S and (ii) δV is sufficiently small so that the Theorem above applies.

Recall the formula for the exit rates:

$$k_i = C_i^{OL}\, \mathrm{e}^{-\beta[V(z_i)-V(x_1)]}\,(1 + O(\beta^{-1}))$$

where $C_i^{OL} = \sqrt{\frac{\beta}{2\pi}}\partial_n V(z_i)\frac{\sqrt{\det(\nabla^2 V)(x_1)}}{\sqrt{\det(\nabla^2 V_{|\partial\mathcal{O}})(z_i)}}$.

One easily check that $k_i / \sum_{j=1}^J k_j$ is independent of δV and

$$\frac{\sum_{j=1}^J k_j(V+\delta V)}{\sum_{j=1}^J k_j(V)} = \sqrt{\frac{\det(\nabla^2(V+\delta V))(x_1)}{\det(\nabla^2(V))(x_1)}}\mathrm{e}^{\beta\delta V(x_1)}(1+O(\beta^{-1}))$$

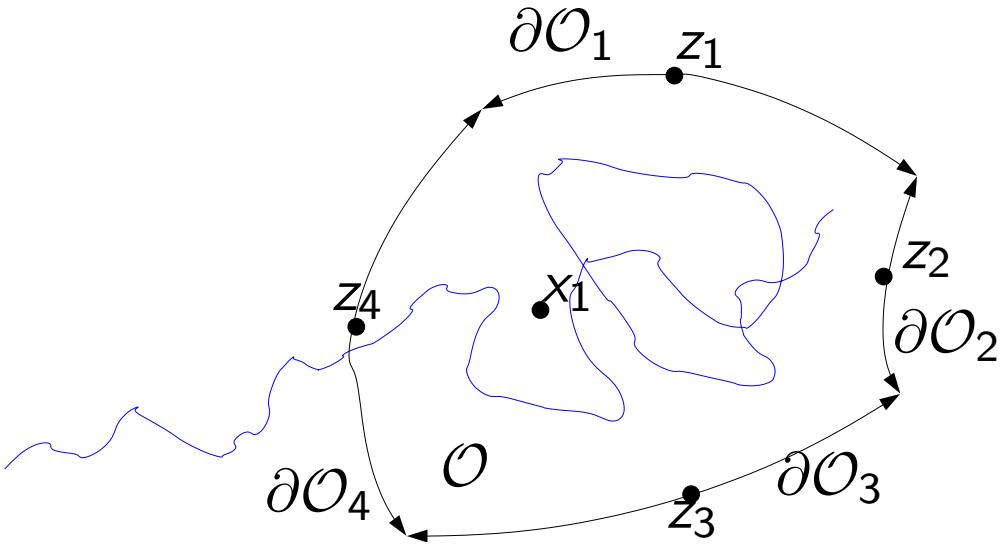
$$= \frac{\int_{\mathcal{O}}\exp(-\beta V)}{\int_{\mathcal{O}}\exp(-\beta(V+\delta V))}(1+O(\beta^{-1})) \simeq B$$

Temperature Accelerated Dynamics (1/2)

Increase the temperature to reduce the exit time [Sorensen, Voter, 2000]

Algorithm:

- Observe the exit events from $\mathcal{O}$ at high temperature ;
- Extrapolate the high temperature exit events to low temperature exit events.



Temperature Accelerated Dynamics (2/2)

Recall that, starting from the QSD, the exit event from a given state $\mathcal{O}$ can exactly be modelled using a kinetic Monte Carlo model with rates

$$k_i = C_i^{OL}\, \mathrm{e}^{-\beta[V(z_i)-V(x_1)]}\,(1+O(\beta^{-1}))$$

where $C_i^{OL} = \sqrt{\frac{\beta}{2\pi}}\partial_n V(z_i)\frac{\sqrt{\det(\nabla^2 V)(x_1)}}{\sqrt{\det(\nabla^2 V_{|\partial\mathcal{O}})(z_i)}}$.

Thus,

$$\frac{k_i^{lo}}{k_i^{hi}} \simeq \sqrt{\frac{\beta^{lo}}{\beta^{hi}}}\exp(-(\beta^{lo}-\beta^{hi})(V(z_i)-V(x_1))).$$

Algorithm: observe exit events at high temperature, extrapolate the rates to low temperature, stop when the extrapolated event will not modify anymore the low temperature exit event.

Remark: TAD can be seen as a smart saddle point search method.

Generalizations and perspectives

- The parallel replica is a very versatile algorithm: it applies *e.g.* to non reversible dynamics, discrete-in-time dynamics, continuous-time Markov Chain [Aristoff, Plechac, Wang]. It does not require estimates of the exit rates.
- Hyper and TAD are more efficient, but require the temperature to be sufficiently small so that estimates of the rates by the Eyring-Kramers formulas hold true.

All these techniques require "good" metastable states:

exit time $>$ convergence time to the QSD.

Conclusion

There are mathematical characterizations of good coarse-graining representations (spectral gaps, convergence times *vs* exit times).

Could we use those characterizations together with advanced learning techniques (auto-encoder, sparse methods) to get better coarse-grained descriptions?

- Identify slow variables
- Sparse representation of the committor function
- Identify metastable states

References

Some papers I mentioned:

- G. Di Gesù, TL, D. Le Peutrec, and B. Nectoux, *Sharp asymptotics of the first exit point density*, Annals of PDE, 5, 2019.
- F. Hédin and TL, *gen.parRep: a first implementation of the Generalized Parallel Replica dynamics for the long time simulation of metastable biochemical systems*, Computer Physics Communications, 239, 2019.
- C. Le Bris, TL, M. Luskin, and D. Perez, *A mathematical formalization of the parallel replica dynamics*, Monte Carlo Methods and Applications, 18(2), 2012.
- TL, M. Ramil, and J. Reygner, *A probabilistic study of the kinetic Fokker-Planck equation in cylindrical domains*, `https://arxiv.org/abs/2010.10157`
- TL, M. Ramil, and J. Reygner, *Quasi-stationary distribution for the Langevin process in cylindrical domains, part I: existence, uniqueness and long-time convergence*, `https://arxiv.org/abs/2101.11999`

References

Review papers:

- TL and Gabriel Stoltz, *Partial differential equations and stochastic methods in molecular dynamics*. Acta Numerica, 25, 2016.
- TL, *Mathematical foundations of Accelerated Molecular Dynamics methods*, In: W. Andreoni and S. Yip (eds) Handbook of Materials Modeling, Springer, 2018.