Parallel Replica algorithm for Langevin dynamics and Adaptative Metadynamics

Mouad Ramil (Seoul National University)

In collaboration with Tony Lelièvre[1], Julien Reygner[1], Caroline Boudier[2], Alexandra M. Goryaeva[3], Mihai-Cosmin Marinica[3], and Jean-Bernard Maillet[2]

[1] CERMICS
École Nationale des Ponts et Chaussées
[2] CEA - DAM, DIF
[3] Service de Recherches de Métallurgie Physique
Université Paris-Saclay

Workshop I: Increasing the Length, Time, and Accuracy of Materials Modeling Using Exascale Computing, IPAM, March 2023.

Motivation

- **Molecular Dynamics** methods are used in **Biology, Material Science, Nuclear Physics** (protein folding, nuclear fuels propagation inside the nuclear reactor).

Motivation

- **Molecular Dynamics** methods are used in **Biology, Material Science, Nuclear Physics** (protein folding, nuclear fuels propagation inside the nuclear reactor).
- The **Underdamped Langevin dynamics** model the evolution of thermostated molecular systems.

Let N particles described by **position** $q_t^i \in \mathbb{R}^3$, and **momentum** $p_t^i \in \mathbb{R}^3$. The process $(X_t = (q_t, p_t))_{t\geq 0} := (q_t^1, \ldots, q_t^N, p_t^1, \ldots, p_t^N)_{t\geq 0}$ is solution of

$$\begin{cases} \mathrm{d}q_t = M^{-1} p_t \mathrm{d}t, \\ \mathrm{d}p_t = -\nabla V(q_t)\mathrm{d}t - \gamma M^{-1} p_t \mathrm{d}t + \sqrt{2\gamma\beta^{-1}}\mathrm{d}B_t, \end{cases}$$

with $V : \mathbb{R}^{3N} \mapsto \mathbb{R}$ the interaction potential, $\gamma > 0$ the friction parameter, M the mass matrix and $\beta^{-1} = k_B T$.

Motivation

- **Molecular Dynamics** methods are used in **Biology, Material Science, Nuclear Physics** (protein folding, nuclear fuels propagation inside the nuclear reactor).
- The **Underdamped Langevin dynamics** model the evolution of thermostated molecular systems.

Let N particles described by **position** $q_t^i \in \mathbb{R}^3$, and **momentum** $p_t^i \in \mathbb{R}^3$. The process $(X_t = (q_t, p_t))_{t \geq 0} := (q_t^1, \ldots, q_t^N, p_t^1, \ldots, p_t^N)_{t \geq 0}$ is solution of

$$\begin{cases} \mathrm{d}q_t = M^{-1} p_t \mathrm{d}t, \\ \mathrm{d}p_t = -\nabla V(q_t)\mathrm{d}t - \gamma M^{-1} p_t \mathrm{d}t + \sqrt{2\gamma\beta^{-1}}\mathrm{d}B_t, \end{cases}$$

with $V : \mathbb{R}^{3N} \mapsto \mathbb{R}$ the interaction potential, $\gamma > 0$ the friction parameter, M the mass matrix and $\beta^{-1} = k_B T$.

Numerical discretization: $(q_{n\Delta t}, p_{n\Delta t}) \sim (\widehat{q}_n, \widehat{p}_n)$ such that (Velocity-Verlet integrator)

$$(\widehat{q}_{n+1}, \widehat{p}_{n+1}) = \Phi_{\Delta_t}(\widehat{q}_n, \widehat{p}_n).$$

Motivation

- **Molecular Dynamics** methods are used in **Biology, Material Science, Nuclear Physics** (protein folding, nuclear fuels propagation inside the nuclear reactor).
- The **Underdamped Langevin dynamics** model the evolution of thermostated molecular systems.

Let N particles described by **position** $q_t^i \in \mathbb{R}^3$, and **momentum** $p_t^i \in \mathbb{R}^3$. The process $(X_t = (q_t, p_t))_{t\geq 0} := (q_t^1, \ldots, q_t^N, p_t^1, \ldots, p_t^N)_{t\geq 0}$ is solution of

$$\begin{cases} \mathrm{d}q_t = M^{-1} p_t \mathrm{d}t, \\ \mathrm{d}p_t = -\nabla V(q_t)\mathrm{d}t - \gamma M^{-1} p_t \mathrm{d}t + \sqrt{2\gamma\beta^{-1}}\mathrm{d}B_t, \end{cases}$$

with $V : \mathbb{R}^{3N} \mapsto \mathbb{R}$ the interaction potential, $\gamma > 0$ the friction parameter, M the mass matrix and $\beta^{-1} = k_B T$.

Numerical discretization: $(q_{n\Delta t}, p_{n\Delta t}) \sim (\widehat{q}_n, \widehat{p}_n)$ such that (Velocity-Verlet integrator)

$$(\widehat{q}_{n+1}, \widehat{p}_{n+1}) = \Phi_{\Delta_t}(\widehat{q}_n, \widehat{p}_n).$$

Problem: The sampling of some physical events takes too many iterations! (Typically 10^{-6} s with $\Delta t = 10^{-15}$ s).

An example in dimension 2

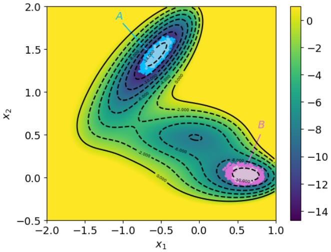


Figure: Sampling in a double well potential. 100 000 iterations

An example in dimension 2

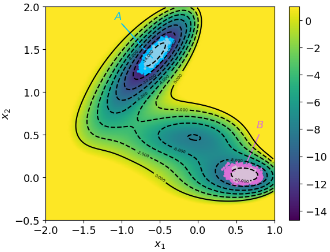


Figure: Sampling in a double well potential. 100 000 iterations

- Oscillation inside basins of attraction of the potential.
- Transition events take a very long time: metastability because the system needs to overcome an energetic gap.
- Problem in large dimension where metastability correspond to entropic effects (narrow escapes).

An example in dimension 2

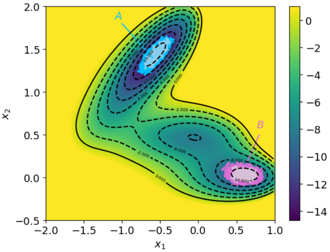


Figure: Sampling in a double well potential. 100 000 iterations

- Oscillation inside basins of attraction of the potential.
- Transition events take a very long time: metastability because the system needs to overcome an energetic gap.
- Problem in large dimension where metastability correspond to entropic effects (narrow escapes).

How to sample **precisely** these transition events?

Parallel Replica algorithm

Conceived by Arthur Voter (Los Alamos National Laboratory) in 1998.

Objective: Parallelize the sampling of the first exit event (first exit time, exit point) from a domain D for the process $(X_t)_{t\geq 0}$.

Parallel Replica algorithm

Conceived by Arthur Voter (Los Alamos National Laboratory) in 1998.

Objective: Parallelize the sampling of the first exit event (first exit time, exit point) from a domain D for the process $(X_t)_{t\geq 0}$.

Parallel Replica: Assume that $(X_t)_{t\geq 0}$ stayed in D during τ_c "long enough". Let $\tau_\partial := \inf\{t > \tau_c : X_t \notin D\}$.

1. Initialize N independent replicas $(X_t^1)_{t\geq 0}, \dots, (X_t^N)_{t\geq 0}$ starting from X_{τ_c} and following the same dynamics as $(X_t)_{t\geq 0}$.
2. Make the N replicas evolve in D during τ_c (*rejection sampling*).
3. Let $\tau_\partial^i = \inf\{t > 0 : X_{t+\tau_c}^i \notin D\}$ and $i^* = \arg\min_{1\leq i\leq N} \tau_\partial^i$. Define

$$\left(\tau_\partial, X_{\tau_\partial}\right) := \left(N\tau_\partial^{i^*}, X_{\tau_\partial^{i^*}}^{i^*}\right).$$

If justified, the last step would ensure a **speed-up** of N in wall-clock time.

Quasi-stationary distribution

If $(X_t)_{t\geq 0}$ stays long enough in a state D it reaches a "local equilibrium", called **quasi-stationary distribution** (QSD).

Quasi-stationary distribution

If $(X_t)_{t\geq 0}$ stays long enough in a state D it reaches a "local equilibrium", called **quasi-stationary distribution** (QSD).

Definition: Let $\tau_\partial := \inf\{t > 0 : X_t \notin D\}$. A probability measure ν on D is said to be a **QSD** on D of the process $(X_t)_{t\geq 0}$, if for all $A \subset D$,

$$\frac{\mathbb{P}_\nu(X_t \in A, \tau_\partial > t)}{\mathbb{P}_\nu(\tau_\partial > t)} = \nu(A).$$

Quasi-stationary distribution

If $(X_t)_{t\geq 0}$ stays long enough in a state D it reaches a "local equilibrium", called **quasi-stationary distribution** (QSD).

Definition: Let $\tau_\partial := \inf\{t > 0 : X_t \notin D\}$. A probability measure ν on D is said to be a **QSD** on D of the process $(X_t)_{t\geq 0}$, if for all $A \subset D$,

$$\frac{\mathbb{P}_\nu(X_t \in A, \tau_\partial > t)}{\mathbb{P}_\nu(\tau_\partial > t)} = \nu(A).$$

First exit event starting from the QSD and justification of ParRep

Assume that $X_0 \overset{\mathcal{L}}{=} QSD$, then, see Collet, Martinez, San Martin (2013),

- τ_∂ follows the exponential law,
- τ_∂ is independent of X_{τ_∂}.

Let $(X_0^N, \ldots, X_0^N)$ be i.i.d. according to the QSD and $i^* := \arg\min_{1\leq i\leq N} \tau_\partial^i$, then

$$(N\tau_\partial^{i^*}, X^{i^*}_{\tau_\partial^{i^*}}) \overset{\mathcal{L}}{=} (\tau_\partial, X_{\tau_\partial}).$$

Parallel Replica: Existence of a QSD and long time convergence to the QSD ?

Overdamped case

QSDs have been investigated for the overdamped Langevin dynamics $(\overline{q}_t)_{t\geq 0}$

$$\mathrm{d}\overline{q}_t = -\nabla V(\overline{q}_t)\mathrm{d}t + \sqrt{2\beta^{-1}}\mathrm{d}B_t,$$

on a $\mathcal{C}^2$ **bounded connected** set $\mathcal{O}$ of $\mathbb{R}^d$, where $\beta^{-1} = k_B T > 0$, $V \in \mathcal{C}^\infty(\mathbb{R}^d, \mathbb{R}^d)$ (see Gong, Qian and Zhao (1988), Le Bris, Lelièvre, Luskin, Perez (2012)).

Overdamped case

QSDs have been investigated for the overdamped Langevin dynamics $(\overline{q}_t)_{t\geq 0}$

$$\mathrm{d}\overline{q}_t = -\nabla V(\overline{q}_t)\mathrm{d}t + \sqrt{2\beta^{-1}}\mathrm{d}B_t,$$

on a $\mathcal{C}^2$ **bounded connected** set $\mathcal{O}$ of $\mathbb{R}^d$, where $\beta^{-1} = k_B T > 0$, $V \in \mathcal{C}^\infty(\mathbb{R}^d, \mathbb{R}^d)$ (see Gong, Qian and Zhao (1988), Le Bris, Lelièvre, Luskin, Perez (2012)).
There exists a **unique** QSD $\overline{\mu}$ on $\mathcal{O}$. Besides, $\overline{\mu}$ satisfies

1. (Lebesgue density) $\overline{\mu}(\mathrm{d}q) = \overline{\psi}(q)\mathrm{d}q$,
2. (Spectral interpretation) $\overline{\psi}$ is the unique non-negative, normalized, classical solution in $\mathcal{C}^2(\mathcal{O}) \cap \mathcal{C}^b(\overline{\mathcal{O}})$ of the following eigenvalue problem

$$\begin{cases} \overline{\mathcal{L}}^*\overline{\psi} = -\overline{\lambda}\,\overline{\psi}, & \text{on } \mathcal{O}, \\ \overline{\psi} = 0, & \text{on } \partial\mathcal{O}, \end{cases}$$

where $\overline{\mathcal{L}}^* = \mathrm{div}(\nabla V \cdot) + \beta^{-1}\Delta$,

3. (Convergence) $\exists C > 0$, $\exists \alpha > 0$ s.t. $\forall t \geq 0$, $\forall \theta$ probability on $\mathcal{O}$,

$$\left\| \mathbb{P}_\theta(\overline{q}_t \in \cdot | \overline{\tau}_\partial > t) - \overline{\mu}(\cdot) \right\|_{TV} \leq C\mathrm{e}^{-\alpha t},$$

where $\overline{\tau}_\partial = \inf\{t > 0 : \overline{q}_t \notin \mathcal{O}\}$.

Overdamped case

QSDs have been investigated for the overdamped Langevin dynamics $(\overline{q}_t)_{t\geq 0}$

$$\mathrm{d}\overline{q}_t = -\nabla V(\overline{q}_t)\mathrm{d}t + \sqrt{2\beta^{-1}}\mathrm{d}B_t,$$

on a $\mathcal{C}^2$ **bounded connected** set $\mathcal{O}$ of $\mathbb{R}^d$, where $\beta^{-1} = k_B T > 0$, $V \in \mathcal{C}^\infty(\mathbb{R}^d, \mathbb{R}^d)$ (see Gong, Qian and Zhao (1988), Le Bris, Lelièvre, Luskin, Perez (2012)).
There exists a **unique** QSD $\overline{\mu}$ on $\mathcal{O}$. Besides, $\overline{\mu}$ satisfies

1. (Lebesgue density) $\overline{\mu}(\mathrm{d}q) = \overline{\psi}(q)\mathrm{d}q$,
2. (Spectral interpretation) $\overline{\psi}$ is the unique non-negative, normalized, classical solution in $\mathcal{C}^2(\mathcal{O}) \cap \mathcal{C}^b(\overline{\mathcal{O}})$ of the following eigenvalue problem

$$\begin{cases} \overline{\mathcal{L}}^*\overline{\psi} = -\overline{\lambda}\,\overline{\psi}, & \text{on } \mathcal{O}, \\ \overline{\psi} = 0, & \text{on } \partial\mathcal{O}, \end{cases}$$

   where $\overline{\mathcal{L}}^* = \mathrm{div}(\nabla V \cdot) + \beta^{-1}\Delta$,
3. (Convergence) $\exists C > 0$, $\exists \alpha > 0$ s.t. $\forall t \geq 0$, $\forall \theta$ probability on $\mathcal{O}$,

$$\left\| \mathbb{P}_\theta(\overline{q}_t \in \cdot | \overline{\tau}_\partial > t) - \overline{\mu}(\cdot) \right\|_{TV} \leq C\mathrm{e}^{-\alpha t},$$

   where $\overline{\tau}_\partial = \inf\{t > 0 : \overline{q}_t \notin \mathcal{O}\}$.

Question: Extension to the Langevin process on $D := \mathcal{O} \times \mathbb{R}^d$?

Langevin process

Let $F \in \mathcal{C}^\infty(\mathbb{R}^d, \mathbb{R}^d)$, $\gamma \in \mathbb{R}$, $\sigma > 0$. Consider the **Langevin** process $(X_t = (q_t, p_t))_{t\geq 0}$ on $\mathbb{R}^{2d}$

$$\begin{cases} \mathrm{d}q_t = p_t \mathrm{d}t, \\ \mathrm{d}p_t = F(q_t)\mathrm{d}t - \gamma p_t \mathrm{d}t + \sigma \, \mathrm{d}B_t. \end{cases}$$

The **infinitesimal generator** $\mathcal{L}$ of $(q_t, p_t)_{t\geq 0}$ on $\mathbb{R}^{2d}$ (kinetic Fokker-Planck operator) is given by

$$\mathcal{L} = p \cdot \nabla_q + F(q) \cdot \nabla_p - \gamma p \cdot \nabla_p + \frac{\sigma^2}{2} \Delta_p.$$

Differences between the study of $(\overline{q}_t)_{t\geq 0}$ in $\mathcal{O}$ and $(q_t, p_t)_{t\geq 0}$ in $D = \mathcal{O} \times \mathbb{R}^d$:

- $\mathcal{L}$ is only **hypoelliptic** on $\mathbb{R}^{2d}$ but not elliptic,
- $\mathcal{O}$ is **bounded** but $D = \mathcal{O} \times \mathbb{R}^d$ is not.

Boundary of D

Partition of ∂D:

- $\Gamma^+ := \{(q,p) \in \partial\mathcal{O} \times \mathbb{R}^d : \langle p, n(q) \rangle > 0\}$ (*exiting velocities*),
- $\Gamma^- := \{(q,p) \in \partial\mathcal{O} \times \mathbb{R}^d : \langle p, n(q) \rangle < 0\}$ (*entering velocities*),
- $\Gamma^0 := \{(q,p) \in \partial\mathcal{O} \times \mathbb{R}^d : \langle p, n(q) \rangle = 0\}$ (*tangential velocities*),

where $n(q)$ is the unitary outward normal vector to $\mathcal{O}$ at $q \in \partial\mathcal{O}$.

Boundary of D

Partition of ∂D:

- $\Gamma^+ := \{(q,p) \in \partial\mathcal{O} \times \mathbb{R}^d : \langle p, n(q)\rangle > 0\}$ (*exiting velocities*),
- $\Gamma^- := \{(q,p) \in \partial\mathcal{O} \times \mathbb{R}^d : \langle p, n(q)\rangle < 0\}$ (*entering velocities*),
- $\Gamma^0 := \{(q,p) \in \partial\mathcal{O} \times \mathbb{R}^d : \langle p, n(q)\rangle = 0\}$ (*tangential velocities*),

where $n(q)$ is the unitary outward normal vector to $\mathcal{O}$ at $q \in \partial\mathcal{O}$.

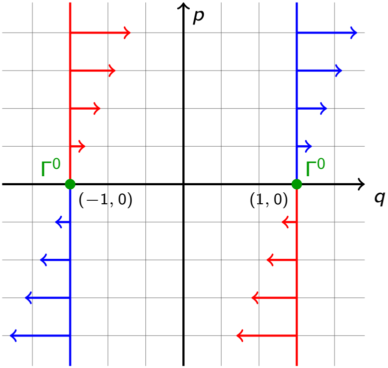


Figure: Boundary of $D = (-1,1) \times \mathbb{R}$

1. Transition density

Theorem (Transition density)

There exists a smooth function

$$(t, x, y) \mapsto \mathrm{p}_t^D(x, y) \in \mathcal{C}^\infty(\mathbb{R}_+^* \times D \times D) \times \mathcal{C}(\mathbb{R}_+^* \times \overline{D} \times \overline{D})$$

such that for all $t > 0$, $x \in \overline{D}$ and $f \in \mathrm{L}^\infty(D)$,

$$P_t^D f(x) := \mathbb{E}_x \left[1_{\tau_\partial > t} f(X_t) \right] = \int_D \mathrm{p}_t^D(x, y) f(y) \mathrm{d}y,$$

where $\tau_\partial = \inf\{t > 0 : X_t \notin D\}$.

Besides,

- $\forall t > 0, x \in D$, $\partial_t \mathrm{p}_t^D(x, y) = \mathcal{L}_x \mathrm{p}_t^D(x, y) = \mathcal{L}_y^* \mathrm{p}_t^D(x, y)$,
- $\mathrm{p}_t^D(x, y) = 0$ if $x \in \Gamma^+ \cup \Gamma^0$ or $y \in \Gamma^- \cup \Gamma^0$,
- $\mathrm{p}_t^D(x, y) > 0$ if $x \notin \Gamma^+ \cup \Gamma^0$ and $y \notin \Gamma^- \cup \Gamma^0$.

1. Transition density

For any $\alpha \in (0,1]$, let $(\widehat{X}_t^{(\alpha)} = (\widehat{q}_t^{(\alpha)}, \widehat{p}_t^{(\alpha)}))_{t\geq 0}$ be the solution on $\mathbb{R}^{2d}$ of

$$\begin{cases} \mathrm{d}\widehat{q}_t^{(\alpha)} = \widehat{p}_t^{(\alpha)}\mathrm{d}t, \\ \mathrm{d}\widehat{p}_t^{(\alpha)} = -\gamma\widehat{p}_t^{(\alpha)}\mathrm{d}t + \frac{\sigma}{\sqrt{\alpha}}\mathrm{d}B_t. \end{cases}$$

Let $\widehat{\mathrm{p}}_t^{(\alpha)}(x,y)$ be its transition density.

Theorem (Gaussian upper-bound)

For any $\alpha \in (0,1)$, $T > 0$, there exists $C > 0$ such that for all $t \in (0,T]$, for all $x, y \in D$,

$$\mathrm{p}_t^D(x,y) \leq C\widehat{\mathrm{p}}_t^{(\alpha)}(x,y).$$

This result is inspired from work on the parametrix method by Konakov, Menozzi, Molchanov (2010). One can show that $\widehat{\mathrm{p}}_t^{(\alpha)} \in \mathrm{L}^\infty(D\times D)\cap\mathrm{L}^1(D\times D)$ (thus in any $\mathrm{L}^p(D\times D)$ for $p\geq 1$).

2. Compactness

Compactness of the semigroup

For any $p \in [1, +\infty]$ and $t > 0$, the operator P_t^D is compact from $\mathcal{C}^b(\overline{D})$ to $\mathcal{C}^b(\overline{D})$.

Proof:

- $\widehat{\mathrm{p}}_t^{(\alpha)} \in \mathrm{L}^2(D \times D) \Rightarrow P_t^D$ is a Hilbert-Schmidt integral operator, hence compact in $\mathrm{L}^2(D)$.
- Propagate to $\mathcal{C}^b(\overline{D})$ using the Gaussian upper-bound.

Reference: Lelievre, R., Reygner - *Journal of Evolution Equations* (2022).

Krein-Rutman theorem

The compactness provides key spectral properties leading to the existence of a QSD, using in particular Krein-Rutman theorem.

Krein-Rutman theorem

The compactness provides key spectral properties leading to the existence of a QSD, using in particular Krein-Rutman theorem.

Theorem (QSD existence/uniqueness and convergence properties)

There exists a unique QSD μ on $D = \mathcal{O} \times \mathbb{R}^d$ for the Langevin process $(X_t)_{t\geq 0}$. Besides, μ satisfies:

1. $\mu(\mathrm{d}x) = \psi(x)\mathrm{d}x$ on D,
2. ψ is the unique, normalized, non-negative classical solution in $\mathcal{C}^2(D) \cap \mathcal{C}^b(D \cup \Gamma^-)$ of the following eigenvalue problem
$$\begin{cases} \mathcal{L}^*\psi(x) = -\lambda\psi(x), & x \in D \\ \psi(x) = 0, & x \in \Gamma^- \end{cases}$$
3. There exists $\alpha > 0$ such that for all θ probability on D, there exists $C_\theta > 0$ and for all $t \geq 0$,
$$\left\|\mathbb{P}_\theta(X_t \in \cdot | \tau_\partial > t) - \mu(\cdot)\right\|_{TV} \leq C_\theta \mathrm{e}^{-\alpha t}.$$

Langevin exit point

Overdamped Langevin exit point

If $\overline{q}_0 \sim \overline{\mu}(\mathrm{d}q) = \overline{\psi}(q)\mathrm{d}q$ (QSD), then

$$\overline{q}_{\overline{\tau}_\partial} \sim \frac{\beta^{-1}}{\overline{\lambda}} |\partial_n \overline{\psi}(q)| \sigma_{\partial\mathcal{O}}(\mathrm{d}q),$$

where $\sigma_{\partial\mathcal{O}}$ is the surface measure on $\partial\mathcal{O}$.

Langevin exit point

Overdamped Langevin exit point

If $\overline{q}_0 \sim \overline{\mu}(\mathrm{d}q) = \overline{\psi}(q)\mathrm{d}q$ (QSD), then

$$\overline{q}_{\overline{\tau}_\partial} \sim \frac{\beta^{-1}}{\overline{\lambda}} |\partial_n \overline{\psi}(q)| \sigma_{\partial\mathcal{O}}(\mathrm{d}q),$$

where $\sigma_{\partial\mathcal{O}}$ is the surface measure on $\partial\mathcal{O}$.

Proof: Integration by parts on

$$\int_{\mathcal{O}} \mathbb{E}_q \left[f(\overline{q}_{\overline{\tau}_\partial}) \right] \overline{\psi}(q)\mathrm{d}q,$$

using that $\mathcal{L}^* \overline{\psi} = -\overline{\lambda}\,\overline{\psi}$ (see Le Bris, Lelievre, Luskin, Perez (2010)).

Langevin exit point

Overdamped Langevin exit point

If $\overline{q}_0 \sim \overline{\mu}(\mathrm{d}q) = \overline{\psi}(q)\mathrm{d}q$ (QSD), then

$$\overline{q}_{\overline{\tau}_\partial} \sim \frac{\beta^{-1}}{\overline{\lambda}} |\partial_n \overline{\psi}(q)| \sigma_{\partial\mathcal{O}}(\mathrm{d}q),$$

where $\sigma_{\partial\mathcal{O}}$ is the surface measure on $\partial\mathcal{O}$.

Proof: Integration by parts on

$$\int_{\mathcal{O}} \mathbb{E}_q \left[f(\overline{q}_{\overline{\tau}_\partial}) \right] \overline{\psi}(q)\mathrm{d}q,$$

using that $\mathcal{L}^* \overline{\psi} = -\overline{\lambda}\,\overline{\psi}$ (see Le Bris, Lelievre, Luskin, Perez (2010)).

Langevin exit point

If $(q_0, p_0) \sim \mu(\mathrm{d}q, \mathrm{d}p) = \psi(q,p)\mathrm{d}q\mathrm{d}p$ (QSD), then

$$(q_{\tau_\partial}, p_{\tau_\partial}) \sim \frac{1}{\lambda} |p \cdot n(q)| \, \psi(q,p) \sigma_{\partial\mathcal{O}}(\mathrm{d}q)\mathrm{d}p,$$

where $\sigma_{\partial\mathcal{O}}$ is the surface measure on $\partial\mathcal{O}$.

Langevin exit point

Overdamped Langevin exit point

If $\overline{q}_0 \sim \overline{\mu}(\mathrm{d}q) = \overline{\psi}(q)\mathrm{d}q$ (QSD), then

$$\overline{q}_{\overline{\tau}_\partial} \sim \frac{\beta^{-1}}{\overline{\lambda}} |\partial_n \overline{\psi}(q)| \sigma_{\partial\mathcal{O}}(\mathrm{d}q),$$

where $\sigma_{\partial\mathcal{O}}$ is the surface measure on $\partial\mathcal{O}$.

Proof: Integration by parts on

$$\int_{\mathcal{O}} \mathbb{E}_q \left[f(\overline{q}_{\overline{\tau}_\partial}) \right] \overline{\psi}(q)\mathrm{d}q,$$

using that $\mathcal{L}^* \overline{\psi} = -\overline{\lambda}\,\overline{\psi}$ (see Le Bris, Lelievre, Luskin, Perez (2010)).

Langevin exit point

If $(q_0, p_0) \sim \mu(\mathrm{d}q, \mathrm{d}p) = \psi(q,p)\mathrm{d}q\mathrm{d}p$ (QSD), then

$$(q_{\tau_\partial}, p_{\tau_\partial}) \sim \frac{1}{\lambda} |p \cdot n(q)| \, \psi(q,p) \sigma_{\partial\mathcal{O}}(\mathrm{d}q)\mathrm{d}p,$$

where $\sigma_{\partial\mathcal{O}}$ is the surface measure on $\partial\mathcal{O}$.

Reference: Lelievre, R., Reygner - *Stochastic Process and its Applications* (2022)

Overdamped limit

Consider here $\gamma > 0$ and $\sigma = \sqrt{2\gamma\beta^{-1}}$. Let $\left(X_t^{(\gamma)} = (q_t^{(\gamma)}, p_t^{(\gamma)})\right)_{t\geq 0}$ be the Langevin process

$$\begin{cases} \mathrm{d}q_t^{(\gamma)} = p_t^{(\gamma)}\mathrm{d}t, \\ \mathrm{d}p_t^{(\gamma)} = F(q_t^{(\gamma)})\mathrm{d}t - \gamma p_t^{(\gamma)}\mathrm{d}t + \sqrt{2\gamma\beta^{-1}}\,\mathrm{d}B_t. \end{cases}$$

Let $(\overline{q}_t)_{t\geq 0}$ be the overdamped Langevin process

$$\mathrm{d}\overline{q}_t = F(\overline{q}_t)\mathrm{d}t + \sqrt{2\beta^{-1}}\mathrm{d}B_t.$$

Assume that F is globally Lipschitz, then for $T > 0$,

$$Law((q_{\gamma t}^{(\gamma)})_{t\in[0,T]}) \underset{\gamma\to\infty}{\longrightarrow} Law((\overline{q}_t)_{t\in[0,T]}).$$

Overdamped limit

Consider here $\gamma > 0$ and $\sigma = \sqrt{2\gamma\beta^{-1}}$. Let $\left(X_t^{(\gamma)} = (q_t^{(\gamma)}, p_t^{(\gamma)})\right)_{t\geq 0}$ be the Langevin process

$$\begin{cases} \mathrm{d}q_t^{(\gamma)} = p_t^{(\gamma)}\mathrm{d}t, \\ \mathrm{d}p_t^{(\gamma)} = F(q_t^{(\gamma)})\mathrm{d}t - \gamma p_t^{(\gamma)}\mathrm{d}t + \sqrt{2\gamma\beta^{-1}}\,\mathrm{d}B_t. \end{cases}$$

Let $(\overline{q}_t)_{t\geq 0}$ be the overdamped Langevin process

$$\mathrm{d}\overline{q}_t = F(\overline{q}_t)\mathrm{d}t + \sqrt{2\beta^{-1}}\mathrm{d}B_t.$$

Assume that F is globally Lipschitz, then for $T > 0$,

$$Law((q_{\gamma t}^{(\gamma)})_{t\in[0,T]}) \underset{\gamma\to\infty}{\longrightarrow} Law((\overline{q}_t)_{t\in[0,T]}).$$

Question: Overdamped limit of the Langevin QSD ?

Overdamped limit

Overdamped limit of the Langevin QSD

Let $\mu^{(\gamma)}$ be the Langevin QSD on D and $\overline{\mu}$ be the overdamped Langevin QSD on $\mathcal{O}$, then

$$\mu^{(\gamma)}(\mathrm{d}\boldsymbol{q}\mathrm{d}\boldsymbol{p}) \underset{\gamma\to\infty}{\longrightarrow} \overline{\mu}(\mathrm{d}\boldsymbol{q})\frac{\mathrm{e}^{-\beta\frac{|\boldsymbol{p}|^2}{2}}}{(2\pi\beta^{-1})^{d/2}}\mathrm{d}\boldsymbol{p}.$$

Besides,

$$\lambda^{(\gamma)} \underset{\gamma\to\infty}{\sim} \frac{\overline{\lambda}}{\gamma}.$$

Reference: R. - *Electronic Journal of Probability* (2022).

Overdamped limit

Overdamped limit of the Langevin QSD

Let $\mu^{(\gamma)}$ be the Langevin QSD on D and $\overline{\mu}$ be the overdamped Langevin QSD on $\mathcal{O}$, then

$$\mu^{(\gamma)}(\mathrm{d}\boldsymbol{q}\mathrm{d}\boldsymbol{p}) \underset{\gamma\to\infty}{\longrightarrow} \overline{\mu}(\mathrm{d}\boldsymbol{q}) \frac{\mathrm{e}^{-\beta\frac{|p|^2}{2}}}{(2\pi\beta^{-1})^{d/2}}\mathrm{d}\boldsymbol{p}.$$

Besides,

$$\lambda^{(\gamma)} \underset{\gamma\to\infty}{\sim} \frac{\overline{\lambda}}{\gamma}.$$

Reference: R. - *Electronic Journal of Probability* (2022).

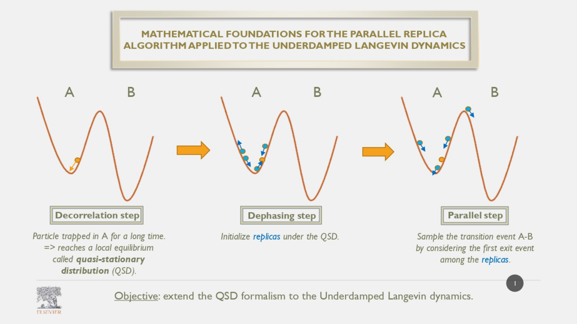
Stationary overdamped limit

Let $\mu_\infty^{(\gamma)}$ be the Langevin stationary distribution and $\overline{\mu_\infty}$ be the overdamped Langevin stationary distribution, then there exists $C > 0$ such that for all $\gamma \geq 2$,

$$\mathcal{W}\left(\mu_\infty^{(\gamma)}, \overline{\mu}_\infty \otimes \mathcal{L}(Z)\right) \leq C\frac{\sqrt{\log(\gamma)}}{\gamma}.$$

Reference: Monmarché, R. - *Electronic Communications in Probability* (2022)

Research letter



R. , Lelièvre and Reygner - Mathematical foundations for the Parallel Replica algorithm applied to the underdamped Langevin dynamics - *MRS Communications*, 2022.

Metadynamics for the Mueller potential

Consider the two-dimensional overdamped Langevin dynamics

$$\mathrm{d}X_t = -\nabla V(X_t)\mathrm{d}t + \sqrt{2k_B T}\mathrm{d}B_t,$$

where V is the Mueller potential defined by:

$$V(x_1, x_2) = \sum_{i=1}^{4} K_i \mathrm{e}^{a_i(x_1-\beta_i)^2 + b_i(x_1-\beta_i)(x_2-\gamma_i) + c_i(x_2-\gamma_i)^2}.$$

Metadynamics for the Mueller potential

Consider the two-dimensional overdamped Langevin dynamics

$$\mathrm{d}X_t = -\nabla V(X_t)\mathrm{d}t + \sqrt{2k_B T}\mathrm{d}B_t,$$

where V is the Mueller potential defined by:

$$V(x_1, x_2) = \sum_{i=1}^{4} K_i \mathrm{e}^{a_i(x_1-\beta_i)^2 + b_i(x_1-\beta_i)(x_2-\gamma_i) + c_i(x_2-\gamma_i)^2}.$$

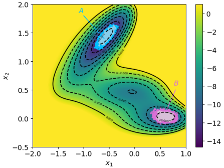


Figure: Sampling in the Mueller potential. 100 000 iterations

We want to sample a transition path between A and B using Metadynamics.

Metadynamics

Definition: A collective variable $s : \mathbb{R}^2 \mapsto \mathbb{R}$ captures the low-dimensional information of the system. A good collective variable takes different values in relevant *metastable states* and *transition states*.

Equilibrium distribution:

$$X_\infty \sim \mu(x) = \frac{\mathrm{e}^{-V(x)/k_B T}}{\alpha}.$$

Latent equilibrium distribution:

$$s(X_\infty) \sim p(s) = \int_{\mathbb{R}^2} \mu(x) \delta_{s(x)=s}(\mathrm{d}x).$$

Free energy:

$$F(s) = -k_B T \log(p(s)).$$

Metadynamics

Assuming one has an estimate $B \approx -F$.If we perturb the potential into $\tilde{V}(x) = V(x) + B(s(x))$.

$$
\begin{aligned}
s(\tilde{X}_\infty) \sim \tilde{p}(s) &\propto \int_{\mathbb{R}^2} \delta_{s(x)=s}(\mathrm{d}x)\mathrm{e}^{-(V(x)+B(s(x)))/k_B T} \\
&= \mathrm{e}^{-B(s)/k_B T} \int_{\mathbb{R}^2} \delta_{s(x)=s}(\mathrm{d}x)\mathrm{e}^{-V(x)/k_B T} \\
&= \mathrm{e}^{-B(s)/k_B T}\mathrm{e}^{-F(s)/k_B T} \approx 1.
\end{aligned}
$$

As a result,

$$
s(\tilde{X}_\infty) \sim \text{ Uniform}
$$

Problem: We need a free energy estimate first ...

Metadynamics

Build an approximation of F iteratively:

$$B_{t+1}(s) = B_t(s) + w\mathrm{e}^{-\frac{(s_t - s)^2}{2\sigma^2}}.$$

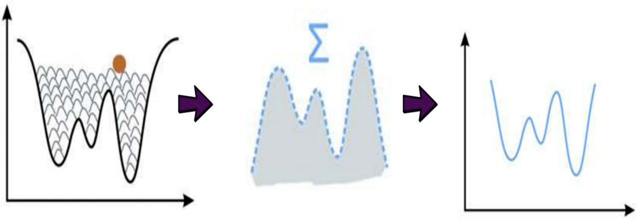
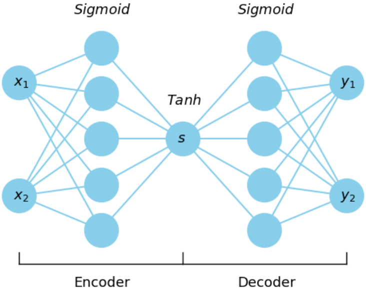


Figure: Metadynamics

Generates uniform samples on the path $A \leftrightarrow B$.

Autoencoder

Train the sampled data on an **autoencoder** to obtain the collective variable.



$$Loss = \|x - y\|^2.$$

Metadynamics trajectories

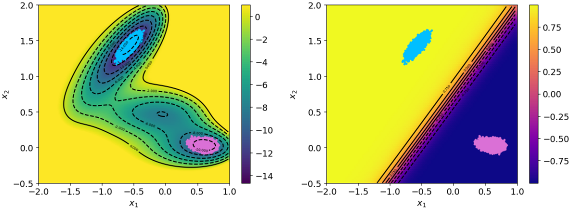


Figure: 50 000 Metadynamics iterations with autoencoder trained on the database of A and B

Metadynamics trajectories

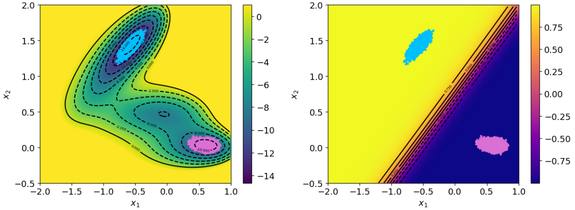


Figure: 50 000 Metadynamics iterations with autoencoder trained on the database of A and B

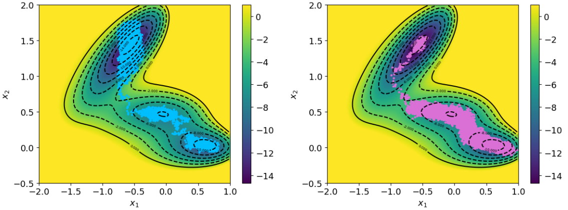


Figure: 25 000 Metadynamics iterations with CV as the orthogonal projection on $[AB]$

Adaptative Metadynamics

Train the autoencoder adaptatively on the previous trajectory after every 1000 Metadynamics iterations.

Adaptative Metadynamics

Train the autoencoder adaptatively on the previous trajectory after every 1000 Metadynamics iterations.

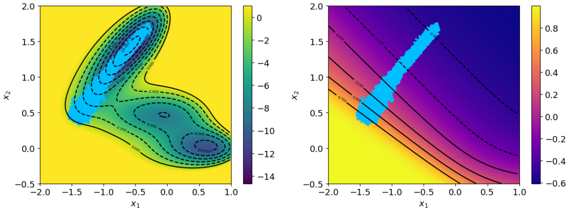


Figure: 40 000 Metadynamics iterations with autoencoder trained iteratively on the trajectory

Adaptative Metadynamics

Impose conditions on the path such that it visits A and B.

Adaptative Metadynamics

Impose conditions on the path such that it visits A and B.

$$Loss = \|x - AE(x)\|^2 + (\|AE(x_A) - x_A\|^2 + \|AE(x_B) - x_B\|^2)/2.$$

The path is defined by:

- (Latent projection) $s_A = s(x_A)$, $s_B = s(x_B)$
- (Discretization) $s_i = s_A + i(s_B - s_A)/N$, $(1 \leq i \leq N)$.
- Path is given by $(D(s_i))_{1 \leq i \leq N}$,

where D is the decoder of the autoencoder.

Adaptative Metadynamics

Impose conditions on the path such that it visits A and B.

$$Loss = \|x - AE(x)\|^2 + (\|AE(x_A) - x_A\|^2 + \|AE(x_B) - x_B\|^2)/2.$$

The path is defined by:

- (Latent projection) $s_A = s(x_A)$, $s_B = s(x_B)$
- (Discretization) $s_i = s_A + i(s_B - s_A)/N$, $(1 \leq i \leq N)$.
- Path is given by $(D(s_i))_{1 \leq i \leq N}$,

where D is the decoder of the autoencoder.

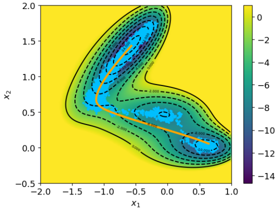


Figure: 38 000 Metadynamics iterations with autoencoder trained iteratively on the trajectory. Path is plotted in orange.

Adaptative Metadynamics

Consider a modified Mueller potential $\tilde{V}$ defined by:

$$\tilde{V}(x) = V(x) + (-100 + \|x - \eta\|^2)\mathrm{e}^{-2\|x-\eta\|^2},$$

with $\eta = [-1.7, 0.2]$.

Adaptative Metadynamics

Consider a modified Mueller potential $\tilde{V}$ defined by:

$$\tilde{V}(x) = V(x) + (-100 + \|x - \eta\|^2)e^{-2\|x-\eta\|^2},$$

with $\eta = [-1.7, 0.2]$.

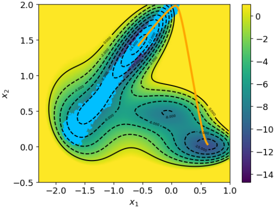


Figure: 100 000 Metadynamics iterations with autoencoder trained iteratively on the trajectory. Path in orange.

Adaptative Metadynamics

Penalize high energy configurations happening in the path.

$$Loss = \|x - y\|^2 + (\|AE(x_A) - x_A\|^2 + \|AE(x_B) - x_B\|^2)/2 + E_{path},$$

where

$$E_{path} = \sum_{i=1}^{N} \|D(s_{i+1}) - D(s_i)\|(V(D(s_i)) + C),$$

with $C > 0$ such that $V(D(s_i)) + C > 0$.

Adaptative Metadynamics

Penalize high energy configurations happening in the path.

$$Loss = \|x - y\|^2 + (\|AE(x_A) - x_A\|^2 + \|AE(x_B) - x_B\|^2)/2 + E_{path},$$

where

$$E_{path} = \sum_{i=1}^{N} \|D(s_{i+1}) - D(s_i)\|(V(D(s_i)) + C),$$

with $C > 0$ such that $V(D(s_i)) + C > 0$.

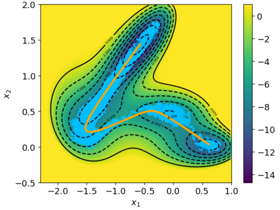


Figure: 100 000 Metadynamics iterations with autoencoder trained iteratively on the trajectory. Path in orange.

Reference: R., Boudier, Goryaeva, Marinica and Maillet - *Journal of Chemical Theory and Computation*, (2022).

Thank you for your attention!