Parallel Replica algorithm for Langevin dynamics and Adaptative Metadynamics

Mouad Ramil (Seoul National University)

In collaboration with Tony Lelièvre¹, Julien Reygner¹, Caroline Boudier², Alexandra M. Goryaeva³, Mihai-Cosmin Marinica³, and Jean-Bernard Maillet²

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

1 Introduction
   - Metastability in molecular dynamics and the Parallel Replica algorithm
   - Quasi-stationary distribution and Parallel Replica justification

2 Langevin process and kinetic Fokker-Planck equation
   - Transition density of the absorbed semigroup
   - Compactness of the absorbed semigroup

3 Quasi-stationary distributions and overdamped limit of the Langevin process
   - Existence and long-time convergence
   - Overdamped limit of the QSD

4 Metadynamics
   - Description of the algorithm
   - Adaptative Metadynamics
Introduction
- Metastability in molecular dynamics and the Parallel Replica algorithm
- Quasi-stationary distribution and Parallel Replica justification

Langevin process and kinetic Fokker-Planck equation
- Transition density of the absorbed semigroup
- Compactness of the absorbed semigroup

Quasi-stationary distributions and overdamped limit of the Langevin process
- Existence and long-time convergence
- Overdamped limit of the QSD

Metadynamics
- Description of the algorithm
- Adaptative Metadynamics
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** $q_i^t \in \mathbb{R}^3$, and **momentum** $p_i^t \in \mathbb{R}^3$. The process $(X_t = (q_t, p_t))_{t \geq 0} := (q_1^t, \ldots, q_N^t, p_1^t, \ldots, p_N^t)_{t \geq 0}$ is solution of

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

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: $(q_n \Delta t, p_n \Delta t) \sim (b q_n, b p_n)$ such that (Velocity-Verlet integrator)

$$(b q_{n+1}, b p_{n+1}) = \Phi \Delta t (b q_n, b p_n).$$

Problem: The sampling of some physical events takes too many iterations! (Typically $10^{-6}$ s with $\Delta t = 10^{-15}$ s).
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** $q^i_t \in \mathbb{R}^3$, and **momentum** $p^i_t \in \mathbb{R}^3$. The process $(X_t = (q_t, p_t))_{t \geq 0} := (q^1_t, \ldots, q^N_t, p^1_t, \ldots, p^N_t)_{t \geq 0}$ is solution of

$$
\begin{cases}
  dq_t = M^{-1} p_t \, dt, \\
  dp_t = -\nabla V(q_t) \, dt - \gamma M^{-1} p_t \, dt + \sqrt{2\gamma\beta^{-1}} \, dB_t,
\end{cases}
$$

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

$$(\hat{q}_{n+1}, \hat{p}_{n+1}) = \Phi_{\Delta t}(\hat{q}_n, \hat{p}_n).$$
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** $q^i_t \in \mathbb{R}^3$, and **momentum** $p^i_t \in \mathbb{R}^3$. The process $(X_t = (q_t, p_t))_{t \geq 0} := (q^1_t, \ldots, q^N_t, p^1_t, \ldots, p^N_t)_{t \geq 0}$ is solution of

\[
\begin{align*}
\dot{q}_t &= M^{-1} p_t dt, \\
\dot{p}_t &= -\nabla V(q_t) dt - \gamma M^{-1} p_t dt + \sqrt{2\gamma \beta^{-1}} dB_t,
\end{align*}
\]

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

$$(\hat{q}_{n+1}, \hat{p}_{n+1}) = \Phi_{\Delta t}(\hat{q}_n, \hat{p}_n).$$

**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).
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 \((X_t)_{t \geq 0}\).

1. Initialize \( N \) independent replicas \((X_1^t)_{t \geq 0}, \ldots, (X_N^t)_{t \geq 0}\) starting from \( X_{\tau_c} \) and following the same dynamics as \((X_t)_{t \geq 0}\).

2. Make the \( N \) replicas evolve in \( D \) during \( \tau_c \) (rejection sampling).

3. Let \( \tau_i^\partial = \inf \{ t > 0 : X_{i^t + \tau_c} \in D \} \) and \( i^* = \arg \min_{1 \leq i \leq N} \tau_i^\partial \). Define \( \tau^\partial, X_{\tau^\partial} := N \tau_{i^*}^\partial, X_{i^* \tau_{i^*}^\partial} \).

If justified, the last step would ensure a speed-up of \( N \) in wall-clock time.
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 $\tau_c$ ”long enough”. Let $\tau_\partial := \inf\{t > \tau_c : X_t \notin D\}$.

1. Initialize $N$ independent replicas $(X^1_t)_{t \geq 0}, \ldots, (X^N_t)_{t \geq 0}$ starting from $X_{\tau_c}$ and following the same dynamics as $(X_t)_{t \geq 0}$.

2. Make the $N$ replicas evolve in $D$ during $\tau_c$ (*rejection sampling*).

3. Let $\tau^i_\partial = \inf\{t > 0 : X^i_{t+\tau_c} \notin D\}$ and $i^* = \arg\min_{1 \leq i \leq N} \tau^i_\partial$. Define

$$\left(\tau_\partial, X_{\tau_\partial}\right) := \left(N\tau^i_{\tau_\partial}, X^i_{\tau^i_{\tau_\partial}}\right).$$

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).

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

**First exit event starting from the QSD and justification of ParRep**

Assume that \(X_0 L = QSD\), then, see Collet, Martinez, San Martin (2013), \(\tau_{\partial}\) follows the exponential law, \(\tau_{\partial}\) is independent of \(X_{\tau_{\partial}}\).

Let \((X_{N_0}, \ldots, X_{N_0})\) be i.i.d. according to the QSD and \(i^* := \arg \min_{1 \leq i \leq N} \tau_i\), then
\[
(N_{\tau_i}, X_{\tau_i}) \overset{L}{=} (\tau_{\partial}, X_{\tau_{\partial}}).
\]
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 \(\nu\) 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 \(\nu\) 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),

- \(\tau_\partial\) follows the exponential law,
- \(\tau_\partial\) is independent of \(X_{\tau_\partial}\).

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

\[
(N\tau_{i^*}, X_{i^*}^\partial) \overset{\mathcal{L}}{=} (\tau_\partial, X_{\tau_\partial}).
\]

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

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

\[
d\bar{q}_t = -\nabla V(\bar{q}_t)dt + \sqrt{2\beta^{-1}}dB_t,
\]

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

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

\[
d\bar{q}_t = -\nabla V(\bar{q}_t)dt + \sqrt{2\beta^{-1}}dB_t,
\]
on a \(C^2\) bounded connected set \(O\) of \(\mathbb{R}^d\), where \(\beta^{-1} = k_B T > 0\), \(V \in C^\infty(\mathbb{R}^d, \mathbb{R}^d)\) (see Gong, Qian and Zhao (1988), Le Bris, Lelièvre, Luskin, Perez (2012)).

There exists a unique QSD \(\bar{\mu}\) on \(O\). Besides, \(\bar{\mu}\) satisfies

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

\[
\begin{cases}
\bar{L}^* \bar{\psi} = -\lambda \bar{\psi}, & \text{on } O, \\
\bar{\psi} = 0, & \text{on } \partial O,
\end{cases}
\]

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

\[
\|P_\theta(\bar{q}_t \in \cdot | \bar{\tau}_\partial > t) - \bar{\mu}(\cdot)\|_{TV} \leq Ce^{-\alpha t},
\]

where \(\bar{\tau}_\partial = \inf\{t > 0 : \bar{q}_t \notin O\}\).
Overdamped case

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

\[ d\overline{q}_t = -\nabla V(\overline{q}_t)dt + \sqrt{2\beta^{-1}}dB_t, \]

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

There exists a unique QSD \(\overline{\mu}\) on \(O\). Besides, \(\overline{\mu}\) satisfies

1. **(Lebesgue density)** \(\overline{\mu}(dq) = \overline{\psi}(q)dq\),

2. **(Spectral interpretation)** \(\overline{\psi}\) is the unique non-negative, normalized, classical solution in \(C^2(O) \cap C^b(\overline{O})\) of the following eigenvalue problem

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

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

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

\[ ||P_\theta(\overline{q}_t \in \cdot | \overline{\tau}_\partial > t) - \overline{\mu}(\cdot)||_{TV} \leq Ce^{-\alpha t}, \]

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

**Question:** Extension to the Langevin process on \(D := O \times \mathbb{R}^d\)?
1 Introduction
   • Metastability in molecular dynamics and the Parallel Replica algorithm
   • Quasi-stationary distribution and Parallel Replica justification

2 Langevin process and kinetic Fokker-Planck equation
   • Transition density of the absorbed semigroup
   • Compactness of the absorbed semigroup

3 Quasi-stationary distributions and overdamped limit of the Langevin process
   • Existence and long-time convergence
   • Overdamped limit of the QSD

4 Metadynamics
   • Description of the algorithm
   • Adaptative Metadynamics
Langevin process

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

\[
\begin{align*}
    dq_t &= p_t \, dt, \\
    dp_t &= F(q_t) \, dt - \gamma p_t \, dt + \sigma \, dB_t.
\end{align*}
\]

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 $(\tilde{q}_t)_{t \geq 0}$ 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 $\partial D$:

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

where $n(q)$ is the unitary outward normal vector to $O$ at $q \in \partial O$. 
Boundary of $D$

Partition of $\partial D$:

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

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

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

**Theorem (Transition density)**

There exists a smooth function

$$(t, x, y) \mapsto p_t^D(x, y) \in C^\infty(\mathbb{R}_+ \times D \times D) \times C(\mathbb{R}_+ \times \overline{D} \times \overline{D})$$

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

$$P_t^D f(x) := \mathbb{E}_x [1_{\tau_{\partial} > t} f(X_t)] = \int_D p_t^D (x, y)f(y)dy,$$

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

Besides,

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

Mouad Ramil (Seoul National University)
1. Transition density

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

\[
\begin{cases}
    \, d\hat{q}_t^{(\alpha)} = \hat{p}_t^{(\alpha)} \, dt, \\
    \, d\hat{p}_t^{(\alpha)} = -\gamma \hat{p}_t^{(\alpha)} \, dt + \frac{\sigma}{\sqrt{\alpha}} dB_t.
\end{cases}
\]

Let $\hat{p}_t^{(\alpha)}(x, y)$ 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 $x, y \in D$,

\[ p_t^D(x, y) \leq C\hat{p}_t^{(\alpha)}(x, y). \]

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

Proof:

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

Introduction
- Metastability in molecular dynamics and the Parallel Replica algorithm
- Quasi-stationary distribution and Parallel Replica justification

Langevin process and kinetic Fokker-Planck equation
- Transition density of the absorbed semigroup
- Compactness of the absorbed semigroup

Quasi-stationary distributions and overdamped limit of the Langevin process
- Existence and long-time convergence
- Overdamped limit of the QSD

Metadynamics
- Description of the algorithm
- Adaptative Metadynamics
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 \( \mu \) on \( D = \mathcal{O} \times \mathbb{R}^d \) for the Langevin process \((X_t)_{t \geq 0}\). Besides, \( \mu \) satisfies:

1. \( \mu(dx) = \psi(x)dx \) on \( D \),
2. \( \psi \) is the unique, normalized, non-negative classical solution in \( C^2(D) \cap C^b(D \cup \Gamma^-) \) of the following eigenvalue problem
   \[
   \begin{aligned}
   \mathcal{L}^* \psi(x) &= -\lambda \psi(x), & x \in D \\
   \psi(x) &= 0, & x \in \Gamma^-
   \end{aligned}
   \]
3. There exists \( \alpha > 0 \) such that for all \( \theta \) probability on \( D \), there exists \( C_\theta > 0 \) and for all \( t \geq 0 \),
   \[
   \|P_\theta(X_t \in \cdot | \tau_\partial > t) - \mu(\cdot)\|_{TV} \leq C_\theta e^{-\alpha t}.
   \]
Langevin exit point

Overdamped Langevin exit point

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

\[
\overline{q}_{\tau \theta} \sim \frac{\beta^{-1}}{\lambda} |\partial_n \overline{\psi}(q)| \sigma_{\partial \mathcal{O}}(dq),
\]

where \( \sigma_{\partial \mathcal{O}} \) is the surface measure on \( \partial \mathcal{O} \).
Langevin exit point

Overdamped Langevin exit point

If \( \overline{q}_0 \sim \mu(\text{dq}) = \overline{\psi}(q)\text{dq} \) (QSD), then

\[
\overline{q}_{\tau \theta} \sim \frac{\beta^{-1}}{\lambda} |\partial_n \overline{\psi}(q)| \sigma_{\partial \mathcal{O}}(\text{dq}),
\]

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

**Proof:** Integration by parts on

\[
\int_{\mathcal{O}} \mathbb{E}_q \left[ f(\overline{q}_{\tau \theta}) \right] \overline{\psi}(q)\text{dq},
\]

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 $q_0 \sim \mu(dq) = \psi(q)dq$ (QSD), then

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

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

**Proof:** Integration by parts on

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

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

Langevin exit point

If $(q_0, p_0) \sim \mu(dq, dp) = \psi(q, p)dqdp$ (QSD), then

$$(q_{\tau \partial}, p_{\tau \partial}) \sim \frac{1}{\lambda} |p \cdot n(q)| \psi(q, p) \sigma_{\partial \mathcal{O}}(dq) dp,$$

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

Overdamped Langevin exit point

If $\bar{q}_0 \sim \bar{\mu}(dq) = \bar{\psi}(q)dq$ (QSD), then

$$\bar{q}_{\tau_{\partial}} \sim \frac{\beta^{-1}}{\lambda} |\partial_n \bar{\psi}(q)| \sigma_{\partial \mathcal{O}}(dq),$$

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

Proof: Integration by parts on

$$\int_{\mathcal{O}} \mathbb{E}_q \left[ f(\bar{q}_{\tau_{\partial}}) \right] \bar{\psi}(q)dq,$$

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

Langevin exit point

If $(q_0, p_0) \sim \mu(dq, dp) = \psi(q, p)dqdp$ (QSD), then

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

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

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

\[
\begin{align*}
\left\{ \begin{array}{l}
    dq_t^{(\gamma)} = p_t^{(\gamma)} dt, \\
    dp_t^{(\gamma)} = F(q_t^{(\gamma)}) dt - \gamma p_t^{(\gamma)} dt + \sqrt{2\gamma/\beta-1} dB_t.
\end{array} \right.
\end{align*}
\]

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

\[
d\bar{q}_t = F(\bar{q}_t) dt + \sqrt{2\beta^{-1}} dB_t.
\]

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

\[
\text{Law}\left( (q_t^{(\gamma)})_{t \in [0,T]} \right) \xrightarrow{\gamma \to \infty} \text{Law}\left( (\bar{q}_t)_{t \in [0,T]} \right).
\]
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

\[
\begin{align*}
\begin{cases}
  \mathrm{d}q_t^{(\gamma)} &= p_t^{(\gamma)} \, \mathrm{d}t, \\
  \mathrm{d}p_t^{(\gamma)} &= F(q_t^{(\gamma)}) \, \mathrm{d}t - \gamma p_t^{(\gamma)} \, \mathrm{d}t + \sqrt{2/\beta - 1} \, \mathrm{d}B_t.
\end{cases}
\end{align*}
\]

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

\[ \mathrm{d}\bar{q}_t = F(\bar{q}_t) \, \mathrm{d}t + \sqrt{2\beta - 1} \, \mathrm{d}B_t. \]

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

\[ \text{Law}\left( (q_t^{(\gamma)})_{t \in [0,T]} \right) \xrightarrow{\gamma \to \infty} \text{Law}\left( (\bar{q}_t)_{t \in [0,T]} \right). \]

**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 $\bar{\mu}$ be the overdamped Langevin QSD on $O$, then

$$\mu(\gamma)(dqdp) \xrightarrow{\gamma \to \infty} \bar{\mu}(dq) \frac{e^{-\beta |p|^2/2}}{(2\pi \beta^{-1})^{d/2}} dp.$$ 

Besides,

$$\lambda(\gamma) \sim \frac{\bar{\lambda}}{\gamma}.$$ 

Overdamped limit of the Langevin QSD

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

$$
\mu^{(\gamma)}(dqdp) \xrightarrow{\gamma \to \infty} \overline{\mu}(dq) \frac{e^{-\beta |p|^2}}{(2\pi \beta^{-1})^{d/2}} dp.
$$

Besides,

$$
\lambda^{(\gamma)} \xrightarrow{\gamma \to \infty} \overline{\lambda}.
$$


Stationary overdamped limit

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

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

Reference: Monmarché, R. - *Electronic Communications in Probability* (2022)
Objective: extend the QSD formalism to the Underdamped Langevin dynamics.

R. , Lelièvre and Reygner - Mathematical foundations for the Parallel Replica algorithm applied to the underdamped Langevin dynamics - *MRS Communications*, 2022.
1 Introduction
- Metastability in molecular dynamics and the Parallel Replica algorithm
- Quasi-stationary distribution and Parallel Replica justification

2 Langevin process and kinetic Fokker-Planck equation
- Transition density of the absorbed semigroup
- Compactness of the absorbed semigroup

3 Quasi-stationary distributions and overdamped limit of the Langevin process
- Existence and long-time convergence
- Overdamped limit of the QSD

4 Metadynamics
- Description of the algorithm
- Adaptative Metadynamics
Metadynamics for the Mueller potential

Consider the two-dimensional overdamped Langevin dynamics

\[ dX_t = -\nabla V(X_t) dt + \sqrt{2k_B T} dB_t, \]

where \( V \) is the Mueller potential defined by:

\[ V(x_1, x_2) = \sum_{i=1}^{4} K_i e^{a_i(x_1 - \beta_i)^2 + b_i(x_1 - \beta_i)(x_2 - \gamma_i) + c_i(x_2 - \gamma_i)^2}. \]
Metadynamics for the Mueller potential

Consider the two-dimensional overdamped Langevin dynamics

\[ dX_t = -\nabla V(X_t)dt + \sqrt{2k_B T}dB_t, \]

where \( V \) is the Mueller potential defined by:

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

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:

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

Latent equilibrium distribution:

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

Free energy:

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

\[
s(\tilde{X}_\infty) \sim \tilde{p}(s) \propto \int_{\mathbb{R}^2} \delta_{s(x)=s} (dx) e^{-(V(x)+B(s(x)))/k_B T} \\
= e^{-B(s)/k_B T} \int_{\mathbb{R}^2} \delta_{s(x)=s} (dx) e^{-V(x)/k_B T} \\
= e^{-B(s)/k_B T} e^{-F(s)/k_B T} \approx 1.
\]

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:

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

**Figure:** Metadynamics

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

Mouad Ramil (Seoul National University)
Autoencoder

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

\[
\text{Loss} = \|x - y\|^2.
\]
**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 $B$. 

$$
\text{Loss} = \| x - AE(x) \|^2 + \left( \| AE(x_A) - x_A \|^2 + \| AE(x_B) - x_B \|^2 \right) / 2.
$$

The path is defined by:

(LP) $s_A = s(x_A)$,

(D) $s_i = s_A + i (s_B - s_A) / N$, $(1 \leq i \leq N)$.

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 $B$.

\[\text{Loss} = \|x - AE(x)\|^2 + (\|AE(x_A) - x_A\|^2 + \|AE(x_B) - x_B\|^2)/2.\]

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:

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

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.

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

where

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

with \(C > 0\) such that \(V(D(s_i)) + C > 0\).
Adaptative Metadynamics

Penalize high energy configurations happening in the path.

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

where

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

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.

Thank you for your attention!