Convergence of the kinetic annealing for general potentials
Joint work with Pierre Monmarché

Lucas Journel

Laboratoire Jacques-Louis Lions
Introduction to simulated annealing

\[ U : \mathbb{R}^d \rightarrow \mathbb{R}_+ . \]

Goal :

\[ \min_{\mathbb{R}^d} U . \]
Introduction to simulated annealing

\[ U : \mathbb{R}^d \rightarrow \mathbb{R}_+. \]

Goal:

\[ \min_{\mathbb{R}^d} U. \]

Design stochastic process \((X_t)\) whose law is close to:

\[ \pi_{\beta_t}(dx) \propto e^{-\beta_t U(x)}dx \]

where \(\beta_t \rightarrow \infty\).  

\[ \pi_{\beta_t}(U(x) > \delta) \xrightarrow{t \rightarrow \infty} 0. \]
Examples of process used in simulated annealing

- Overdamped Langevin Process (OLP):
  \[ dX_t = -\nabla U(X_t)dt + \sqrt{2\beta_t^{-1}}dB_t. \]

- Kinetic Langevin Process (KLP):
  \[
  \begin{cases}
  dX_t = Y_t dt \\
  dY_t = -\nabla U(X_t)dt - \gamma_t Y_t dt + \sqrt{2\gamma_t\beta_t^{-1}}dB_t.
  \end{cases}
  \]

- Local equilibrium of KLP:
  \[
  \mu_{\beta_t}(dx dy) \propto e^{-\beta_t H(x, y)} dx dy
  \]
  where \( H(x, y) = U(x) + |y|^2/2. \)
Cooling schedule and energy barrier

Cooling schedule:

\[ \beta_t = \ln\left( e^{\beta_0} + t \right) \frac{c}{c}. \]
Cooling schedule and energy barrier

Cooling schedule:

\[ \beta_t = \frac{\ln(e^{c\beta_0} + t)}{c}. \]

Largest energy barrier:

\[ c^* = \sup_{x_1, x_2} E(x_1, x_2), \]

where

\[ E(x_1, x_2) = \inf_\xi \left\{ \max_{0 \leq t \leq 1} U(\xi(t)) - U(x_1) - U(x_2) \right\}. \]

Figure – Example of potential
Theorem (Holley, Kusuoka, Stroock 89)

Let $M$ be a smooth compact manifold and $U : M \to \mathbb{R}^+$. Consider the process $(X_t)$ obtained by solving an ODE with a cooling schedule $(\beta_t)$.

Then if $c > c^*$, $U(X_t) \to \min U$ in probability. If there exists $p \in M$, bottom of a well of height greater than $c$, then $\mathbb{P}(\inf U(X_t) > U(p)) > 0$. 

---

Figure – Example of potential landscape
Historical review

- \( U : \mathbb{R}^d \rightarrow \mathbb{R}_+ \), \( U(\infty) = \infty \), \( |\nabla U|(\infty) = \infty \),
- \( \inf_{\mathbb{R}^d} |\nabla U|^2 - \Delta U > -\infty \).

(\( X_t \)) Overdamped Langevin Process.

**Theorem (Chiang, Hwang, Sheu 87)**

If \( c > 3/2c^* \), \( U(X_t) \rightarrow \min U \) in probability.

**Theorem (Royer 89, Miclo 92)**

If \( c > c^* \), \( U(X_t) \rightarrow \min U \) in probability.
Historical review

$(X_t)$ Overdamped Langevin Process.

**Theorem (Zitt 08)**

$U : \mathbb{R}^d \rightarrow \mathbb{R}^+$, $U(x) \geq \ln(|x|^m) - C$, $\|\nabla U\|_\infty < \infty$, $\Delta U \leq 0$ outside a compact.

Then if $c > c^*$, $U(X_t) \rightarrow \min U$ in probability.

**Theorem (Fournier, Tardif 21)**

$U : \mathbb{R}^d \rightarrow \mathbb{R}^+$, $U(\infty) = \infty$, $\int_{\mathbb{R}^d} e^{-\alpha_0 U} < \infty$

Then if $c > c^*$, $U(X_t) \rightarrow \min U$ in probability.
Historical review and result

\((X_t, Y_t)\) Kinetic Langevin Process. \(H(x, y) = U(x) + |y|^2/2\).

**Theorem (Monmarché 18)**

\[ U : \mathbb{R}^d \rightarrow \mathbb{R}_+ \text{, } x . \nabla U(x) \geq r|x|^2 - M, \| \nabla^2 U \|_{\infty} < \infty. \]

*Then if \(c > c^*\), \(U(X_t) \rightarrow \min U \text{ in probability.})*

**Theorem (J, Monmarché 21)**

\[ U : \mathbb{R}^d \rightarrow \mathbb{R}_+ \text{, } U(\infty) = \infty, \int_{\mathbb{R}^d} e^{-\alpha_0 U} < \infty. \]

*Then if \(c > c^*\), \(H(X_t, Y_t) \rightarrow \min U \text{ in probability.} \)

*If there exists \(p\), bottom of a well of height greater than \(c^*\), then \(P(\inf U(X_t) > U(p)) > 0.\)*
A basic example

\[ dX_t = -\nabla U(X_t)dt + \sqrt{2\beta^{-1}}dB_t, \]

\(U\) quadratic at infinity.

- \(\pi_{\beta}\) stationary measure.
- \(f_t = \mathcal{L}aw(X_t), \quad h_t = \frac{df_t}{d\pi_{\beta}}.\)

\[ \partial_t h_t = \beta^{-1} \Delta h_t - \nabla U.\nabla h_t =: L^* h_t. \]
A basic example

Definition (Carré du champ)

\[
\Gamma f = \frac{1}{2} \left( L^* f^2 - 2 f L^* f \right).
\]

\[
\Gamma f = \sqrt{2\beta^{-1}} |\nabla f|^2.
\]

\[
H_t = \int_{\mathbb{R}^d} (h_t - 1)^2 d\pi_\beta \geq \|f_t - \pi_\beta\|_{TV}^2.
\]

\[
H'_t = - \int_{\mathbb{R}^d} \Gamma(h_t) d\pi_\beta.
\]
A more basic example

Poincaré inequality

For all \( h : \mathbb{R}^d \to \mathbb{R} \), \( \int_{\mathbb{R}^d} h d\pi_\beta = 1 \):

\[
\int_{\mathbb{R}^d} (h - 1)^2 d\pi_\beta \leq \lambda_\beta \int_{\mathbb{R}^d} \Gamma(h) d\pi_\beta.
\]

\( \lambda_\beta \) satisfies:

\[
\lim_{\beta \to \infty} \beta^{-1} \ln(\lambda_\beta) = c^*.
\]

\[
H_t = \int_{\mathbb{R}^d} (h_t - 1)^2 d\pi_\beta.
\]

\[
H'_t = -\int_{\mathbb{R}^d} \Gamma(h_t) d\pi_\beta \leq -\lambda_\beta^{-1} H_t.
\]

\[
H_t \leq e^{-\lambda_\beta^{-1} t} H_0.
\]
\[ \beta = \beta_t. \]

\[ H'_t = -\int_{\mathbb{R}^d} \Gamma(h_t) d\pi_{\beta_t} + \beta'_t A_t. \]
- $\beta = \beta_t$.

\[ H'_t = -\int_{\mathbb{R}^d} \Gamma(h_t) d\pi_{\beta_t} + \beta'_t A_t. \]

- $U$ is such that $\pi_{\beta}$ does not satisfy Poincaré inequality.
- $\beta = \beta_t$.

\[
H'_t = -\int_{\mathbb{R}^d} \Gamma(h_t) d\pi_{\beta_t} + \beta'_t A_t.
\]

- $U$ is such that $\pi_{\beta}$ does not satisfy Poincaré inequality.

- For kinetic Langevin process:

\[
\Gamma(f) = |\nabla_y f|^2,
\]

hence no Poincaré inequality of the form

\[
\int_{\mathbb{R}^d} (h - 1)^2 d\mu_\beta \leq \lambda_\beta \int_{\mathbb{R}^d} \Gamma(h) d\mu_\beta.
\]
Generator of the process:

$$L_t = y . \nabla_x - \nabla U . \nabla_y - \gamma_t y . \nabla_y + \gamma_t \beta_t^{-1} \Delta_y$$

Let $f_t$ the law of the process and $h_t = \frac{f_t}{\mu_{\beta_t}}$.

$L^*_t$ the dual of $L_t$ in $L^2(\mu_{\beta_t})$:

$$\partial_t h_t = L^*_t h_t - \beta'_t H h_t.$$
Plan of the proof

1. Almost surely, $\sup_t H(X_t, Y_t) < \infty$.
   
a. Almost surely, $\liminf_{t \to \infty} H(X_t, Y_t)$.
   
b. For compact set $C$, there exists $K > 0$ such that
      
      $$\inf_{x \in C} \mathbb{P}(\sup_t H(X_t, Y_t) \leq K) \geq 1/4.$$ 

2. If $X$ lives in a compact set, there is convergence.
Position in a compact set

Fix $K > 1$, $L_K > 1$, $M_K = (\mathbb{R}/2L_K\mathbb{Z})^d$, such that $\{U \leq K\} \subset M_K$. $U^K : M_K \rightarrow \mathbb{R}$, $U^K = U$ on $\{U \leq K\}$

\[
\begin{cases}
    dX^K_t = Y^K_t \, dt \\
    dY^K_t = -\nabla_x U^K(X_t) \, dt - \gamma_t Y^K_t \, dt + \sqrt{2\gamma_t \beta^{-1}_t} \, dB_t.
\end{cases}
\]

(1)

\[
\left\{ \sup_{t \geq 0} H(X_t, Y_t) \leq K \right\} = \left\{ \sup_{t \geq 0} H_K(X^K_t, Y^K_t) \leq K \right\},
\]

where $H_K(x, y) = U^K(x) + |y|^2/2$.

\[
\mu^K_\beta(dx\,dy) \propto e^{-\beta H_K(x, y)} \, dx\,dy.
\]
Hypocoercivity à la Villani:

\[ \phi_t(h) = |(\nabla_x + \nabla_y)h|^2 + \sigma_t h^2 \]

with \( \sigma_t = \frac{1}{2} + 2\sqrt{\gamma_t^{-1}} \beta_t (1 + \|\nabla U^K\|_{\infty} + \gamma_t)^2 \), we introduce

\[ \tilde{N}(t) = \int_{\mathbb{R}^d \times \mathbb{R}^d} \phi_t \left( h^K_t - 1 \right) d\mu^K_{\beta_t} \]

\[ \tilde{I}(t) = \int_{M_K \times \mathbb{R}^d} \left| \nabla h^K_t \right|^2 d\mu^K_{\beta_t} \]

Differentiating \( \tilde{N} \), one can (formally) check that

\[ \tilde{N}'(t) \leq -\frac{1}{2} \tilde{I}(t) + C \beta'_t (1 + \beta_t) \tilde{N}(t) \]
Localisation

Poincaré inequality:

\[ \tilde{N}(t) \leq \lambda_{\beta t} \tilde{I}(t) \]

With:

\[ \frac{1}{\beta} \ln(\lambda_{\beta}) \xrightarrow{\beta \to \infty} c^* \]

Conclusion:

\[ \tilde{N}'(t) \leq \left( -\frac{C'}{(1 + t)^c^*/c} + \frac{C(1 + \ln(1 + t))}{(1 + t)} \right) \tilde{N}(t) \]
Proposition

If $c > c^*$, then for all $K > 1$, all $C^\infty$-probability density $f_0$ with compact support in $M_K \times \mathbb{R}^d$, and all $\delta > 0$,

$$\mathbb{P}_{f_0}\left( H_K(X^K_t, Y^K_t) > \delta \right) \xrightarrow{t \to +\infty} 0.$$ 

$$\mathbb{P}\left( H_K(X^K_t, Y^K_t) > \delta \right) \leq \|h^K_t\|_{L^2(\mu_{\beta_t}(H_K > \delta))}^{1/2}$$
Non-convergence for fast cooling schedule

Goal:

$$\mathbb{P} \left( \sup_{t \geq 0} H(X_t, Y_t) \leq H(x_0, y_0) + c + \delta \right) > 0$$

Figure – Example of $c^*$
THANK YOU FOR YOUR ATTENTION