# Concentration Inequalities for Euler Schemes

Florent Malrieu[1] and Denis Talay[2]

[1] IRMAR, UMR 6625, Université Rennes 1, France
florent.malrieu@univ-rennes1.fr
[2] INRIA Sophia-Antipolis, 2004 route des Lucioles, B.P. 93, 06902
Sophia-Antipolis Cedex, France Denis.Talay@sophia.inria.fr

**Summary.** We establish a Poincaré inequality for the law at time $t$ of the explicit Euler scheme for a stochastic differential equation. When the diffusion coefficient is constant, we also establish a Logarithmic Sobolev inequality for both the explicit and implicit Euler scheme, with a constant related to the convexity of the drift coefficient. Then we provide exact confidence intervals for the convergence of Monte Carlo methods.

## 1 Poincaré and Logarithmic Sobolev Inequalities

To describe and control the statistical errors of probabilistic numerical methods, one can use better results than limit theorems such as Central Limit Theorems. Indeed, it is worthy having non asymptotic error estimates in order to choose numerical parameters (number of Monte Carlo simulations, or number of particles, or time length of an ergodic simulation) in terms of the desired accuracy and confidence interval. To this end, concentration inequalities are extremely useful and accurate. As reminded in the section 6 below, sufficient conditions for concentration inequalities are Poincaré (or spectral gap) and Logarithmic Sobolev inequalities. Such inequalities consist in bounding from above a variance or an entropy by an energy quantity. We start by defining Poincaré and Logarithmic Sobolev inequalities for measures on $\mathbb{R}^d$.

*Remark 1.* In what follows, we call "smooth" function a $\mathcal{C}^\infty$ function with polynomial growth.

**Definition 1 (Poincaré inequality).** *A probability measure $\mu$ on $\mathbb{R}^d$ satisfies a Poincaré (or spectral gap) inequality with constant $C$ if*

$$\mathrm{Var}_\mu(f) := \mathbb{E}_\mu(f^2) - (\mathbb{E}_\mu f)^2 \le C\,\mathbb{E}_\mu\left(|\nabla f|^2\right) \tag{1}$$

*for all smooth functions $f$.*

**Definition 2 (Logarithmic Sobolev inequality).** *The probability measure $\mu$ on $\mathbb{R}^d$ satisfies a Logarithmic Sobolev inequality with constant $C$ if*

$$\operatorname{Ent}_\mu(f^2) := \int f^2 \log f^2 \, d\mu - \int f^2 \, d\mu \log \int f^2 \, d\mu \leq C \, \mathbb{E}_\mu\Big(|\nabla f|^2\Big) \qquad (2)$$

*for all smooth functions $f$.*

The Logarithmic Sobolev inequality implies the Poincaré inequality and a better concentration inequality (see (19) and (20)) below.

One can easily check that the Gaussian measure $\mathcal{N}(m, S)$ on $\mathbb{R}^d$ satisfies a Poincaré (respectively Logarithmic Sobolev) inequality with constant $\rho$ (respectively $2\rho$), where $\rho$ is the largest eigenvalue of the covariance matrix $S$.

We now consider a much less elementary example and we follow [3]. Let $(X_t)$ be a time continuous Markov process with infinitesimal generator $\mathbf{L}$. Set

$$\alpha(s) := P_s\Big[(P_{t-s}f)^2\Big].$$

As the time derivative of $P_t f$ is $P_t \mathbf{L} f$, one has

$$\alpha'(s) = 2 P_s \boldsymbol{\Gamma} P_{t-s} f, \qquad (3)$$

where

$$\boldsymbol{\Gamma}(f,g) := \frac{1}{2}[\mathbf{L}(fg) - f\mathbf{L}g - g\mathbf{L}f] \quad \text{and} \quad \boldsymbol{\Gamma} f := \boldsymbol{\Gamma}(f,f).$$

Suppose that the semigroup $P_t$ satisfies the commutation relation

$$\boldsymbol{\Gamma} P_t f \leq e^{-2\rho t} P_t \boldsymbol{\Gamma} f. \qquad (4)$$

Then it also satisfies the Poincaré inequality since

$$P_t(f^2) - (P_t f)^2 = \alpha(t) - \alpha(0) = \int_0^t \alpha'(s) \, ds \leq \frac{1 - e^{-2\rho t}}{\rho} P_t \boldsymbol{\Gamma} f.$$

It now remains to get sufficient conditions for (4). Set

$$\boldsymbol{\Gamma_2} f := \frac{1}{2}[\mathbf{L}(\boldsymbol{\Gamma} f) - 2\boldsymbol{\Gamma}(f, \mathbf{L}f)],$$

and notice that $\alpha''(s) = 4 P_s \boldsymbol{\Gamma_2} P_{t-s} f$. Suppose that the Bakry–Émery criterion with curvature $\rho$ holds, that is,

$$\boldsymbol{\Gamma_2} f \geq \rho \, \boldsymbol{\Gamma} f. \qquad (5)$$

Then $\alpha''(s) \geq 2\rho\alpha'(s)$, from which one can deduce (4).

We end this section by considering the special case of diffusion processes. Let $(X_t)_{t\geq 0}$ be the $\mathbb{R}^d$ valued diffusion process solution of the stochastic differential equation

$$X_t = X_0 + \int_0^t b(X_s)\,ds + \int_0^t \sqrt{2}\sigma(X_s)\,dB_s, \tag{6}$$

where $(B_t)_{t\geq 0}$ is a Brownian motion on $\mathbb{R}^d$, $\sigma(x)$ is a $d\times d$ matrix valued function, and $b(x)$ is a $\mathbb{R}^d$ valued function. A straightforward computation provides

$$\boldsymbol{\Gamma}(f) = |\sigma\nabla f|^2.$$

Using the fact that $\mathbf{L}$ is the generator of a diffusion process, one can prove that the logarithmic Sobolev inequality

$$\mathrm{Ent}_{P_t}\left(f^2\right) \leq \frac{2}{\rho}\left(1-e^{-2\rho t}\right)P_t(\boldsymbol{\Gamma} f)$$

is implied by the reinforced commutation relation

$$\sqrt{\boldsymbol{\Gamma} P_t f} \leq e^{-\rho t} P_t\left(\sqrt{\boldsymbol{\Gamma} f}\right),$$

that is,

$$|\sigma\nabla P_t f| \leq e^{-\rho t} P_t(|\sigma\nabla f|). \tag{7}$$

In addition, one can show that this reinforced commutation relation is equivalent to the Bakry–Émery curvature criterion. In the case of one–dimensional diffusions, this criterion is equivalent to the condition

$$\inf_{x\in\mathbb{R}}\left(\sigma(x)\sigma''(x) + \frac{\sigma'(x)}{\sigma(x)}b(x) - b'(x)\right) \geq \rho. \tag{8}$$

We now aim to get Poincaré and Logarithmic Sobolev inequalities for approximation schemes of diffusion processes and particle systems for McKean–Vlasov partial differential equations. Complete proofs will appear in [10].

## 2 Poincaré Inequalities For Multidimensional Euler Schemes

Consider the Euler scheme $(X_n^\gamma)_{n\in\mathbb{N}}$ on $\mathbb{R}^d$ with discretization step $\gamma$:

$$X_{n+1}^\gamma := X_n^\gamma + b(X_n^\gamma)\gamma + \sqrt{2}\sigma(X_n^\gamma)(B_{n+1}-B_n). \tag{9}$$

This scheme discretizes (6) and defines a Markov chain on $\mathbb{R}^d$ with transition kernel

$$K(f)(x) := \mathbb{E}\left[f\left(x + b(x)\gamma + \sqrt{2\gamma}\sigma(x)Y\right)\right],$$

where $Y$ is Gaussian $\mathcal{N}(0, I_d)$. The aim of this section is to establish that the law of $X_n^\gamma$ satisfies a Poincaré inequality and to make the constant explicit.

**Theorem 1.** *If $\sigma$ and $b$ are bounded functions with continuous and bounded derivatives, then, for all $n \in \mathbb{N}$ and all smooth functions $f$,*

$$K^n\big(f^2\big)(x) - (K^n f(x))^2 \leq C_{\gamma,n} K^n\Big(|\nabla f|^2\Big)(x).$$

*The constant $C_{\gamma,n}$ can be chosen as*

$$C_{\gamma,n} = \gamma c \frac{(C_\gamma)^n - 1}{C_\gamma - 1}, \tag{10}$$

*where, denoting by $\rho(AA^*)$ the largest eigenvalue of $AA^*$,*

$$\begin{aligned} c &= \sup_{x\in\mathbb{R}^d} \rho(\sigma(x)\sigma^*(x)) \\ C_\gamma &= \sup_{x\in\mathbb{R}^d} \rho[I_d + \gamma(\text{Jac } B(x) + \text{Jac } B^*(x)) + 2\gamma\sigma(x)\sigma^*(x)]. \end{aligned}$$

*Remark 2.* The constant $C_\gamma$ can be chosen uniformly in $\gamma \leq \gamma_0$. Moreover, $C_\gamma = 1 + c'\gamma + o(\gamma)$, which explains that $\gamma$ appears in the numerator. In addition, if $C_\gamma = 1$, then $((C_\gamma)^n - 1)/(C_\gamma - 1)$ is understood as $c'n$.

*Proof (Theorem 1).* We mimic the continuous time semigroup argument. Observe that

$$\begin{aligned} K^n(f^2) - (K^n f)^2 &= \sum_{i=1}^n \Big\{K^i\big[(K^{n-i}f)^2\big] - K^{i-1}\Big[\big(K^{n-i+1}f\big)^2\Big]\Big\} \\ &= \sum_{i=1}^n K^{i-1}\Big\{K\big[(K^{n-i}f)^2\big] - \Big[K\big(K^{n-i}f\big)^2\Big]\Big\}. \end{aligned}$$

Therefore,

$$\mathbf{Var}_{K^n}(f) = \sum_{i=1}^n K^{i-1}\mathbf{Var}_K\big(K^{n-i}f\big). \tag{11}$$

Notice that the operator $\mathbf{Var}_K()$ is the discrete time version of the operator $\boldsymbol{\Gamma}$. The kernel $K$ is the Gaussian law with mean $x + b(x)\gamma$ and covariance matrix $2\gamma\sigma(x)\sigma^*(x)$. Thus, since $\sigma$ is bounded, it satisfies the Poincaré inequality

$$\mathbf{Var}_K(f)(x) \leq 2\gamma\, K\Big(|\sigma(x)\nabla f|^2\Big)(x) \leq c\gamma\, K\Big(|\nabla f|^2\Big)(x). \tag{12}$$

In addition,

$$\nabla K f(x) = \mathbb{E}[(I_d + \gamma\text{Jac } B(x) + \sqrt{\gamma}\text{Jac } (\sigma(x)Y))\nabla f(x + \gamma b(x) + \sqrt{\gamma}\sigma(x)Y)].$$

For all $d \times d$-matrix $A$ we denote by $\rho(AA^*)$ the largest eigenvalue (or spectral radius) of $AA^*$. For all $v$ in $\mathbb{R}^d$ one has

$$|Av|^2 \leq \rho(AA^*)^2|v|^2.$$

Therefore, Cauchy–Schwarz inequality leads to

$$|\nabla Kf(x)|^2 \leq \mathbb{E}[\rho(A(x)A(x)^*)]K\left(|\nabla f|^2\right)(x) \leq C_\gamma K\left(|\nabla f|^2\right)(x), \qquad (13)$$

with the obvious notation

$$\begin{aligned} A(x) &= I_d + \gamma \text{Jac } B(x) + \sqrt{\gamma}\text{Jac } (\sigma(x)Y), \\ C_\gamma &= \sup_{x\in\mathbb{R}^d} \mathbb{E}[\rho(A(x)A(x)^*)] \\ &= \sup_{x\in\mathbb{R}^d} \rho[I_d + \gamma(\text{Jac } B(x) + \text{Jac } B^*(x)) + 2\gamma\sigma(x)\sigma^*(x)]. \end{aligned}$$

We have thus obtained:

$$\begin{aligned} K^n\left(f^2\right) - \left(K^n f\right)^2 &\leq \sum_{i=1}^n c\gamma K^i\left(\left|\nabla K^{n-i} f\right|^2\right) \\ &\leq c\gamma\left(\sum_{i=1}^n (C_\gamma)^{n-i}\right) K^n\left(|\nabla f|^2\right), \end{aligned}$$

which ends the proof.

## 3 Logarithmic Sobolev Inequalities For One-Dimensional Euler And Milstein Schemes

The aim of this section is to establish Logarithmic Sobolev inequalities for numerical schemes in dimension one and to make the constants explicit in the inequalities in terms of the curvature of the solution of (6).

### The Commutation Relation For The Bernoulli Scheme

Consider the approximation scheme with transition kernel

$$Jf(x) := \mathbb{E}\left[f(x + \gamma b(x) + \sqrt{2\gamma}\sigma(x)Z)\right],$$

where the law of $Z$ is the probability measure $\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$. Then $(\sigma(Jf)')(x)$ is equal to

$$\mathbb{E}\left[\left(\sigma(x)(1 + \gamma b'(x) + \sqrt{2\gamma}\sigma'(x)Z)\right) f'\left(x + \gamma b(x) + \sqrt{2\gamma}\sigma(x)Z\right)\right].$$

Thus

$$\sigma(x)(Jf)'(x) = \mathbb{E}\left[(1 - \alpha_x(\gamma))(\sigma f')\left(x + \gamma b(x) + \sqrt{2\gamma}\sigma(x)Z\right)\right],$$

where

$$\alpha_x(\gamma) := \frac{\sigma(x+\gamma b(x)+\sqrt{2\gamma}\sigma(x)Z) - \sigma(x)(1+\gamma b'(x)+\sqrt{2\gamma}\sigma'(x)Z)}{\sigma(x+\gamma b(x)+\sqrt{2\gamma}\sigma(x)Z)}.$$

In view of the Taylor formula,

$$\begin{aligned}\sigma\Big(x+\gamma b(x)+\sqrt{2\gamma}\sigma(x)Z\Big) &= \sigma(x)+\sigma'(x)\Big(b(x)\gamma+\sqrt{2\gamma}\sigma(x)Z\Big)\\ &\quad + \sigma''(x)\sigma(x)^2 Z^2\gamma + O(\gamma^{3/2}).\end{aligned}$$

Therefore

$$\alpha_x(\gamma) = \left[\sigma(x)\sigma''(x) + \frac{\sigma'(x)b(x)}{\sigma(x)} - b'(x)\right]\gamma + O(\gamma^{3/2}),$$

since $Z^2 = 1$ almost surely. The curvature criterion (8) leads to

$$\alpha_x(\gamma) \geq \rho\gamma + O(\gamma^{3/2}).$$

Consequently, for all $\gamma$ small enough it holds that

$$|\sigma(x)(Jf)'(x)| \leq \Big[1-\rho\gamma+O(\gamma^{3/2})\Big] J(|\sigma f'|)(x).$$

Now, the Bernoulli law satisfies a Logarithmic Sobolev inequality with constant 2 (see [1]). We thus deduce that the iterated kernel $J^n$ of the Bernoulli scheme satisfies a Logarithmic Sobolev inequality with constant

$$\frac{2}{\rho+O(\gamma^{1/2})}\Big(1-(1-\rho\gamma+O(\gamma^{3/2}))^{2n}\Big).$$

### The Milstein Scheme

The previous result seems surprising since we have used that Bernoulli r.v. satisfy $Z^2 = 1$ a.s. Consider the new Markov chain with kernel

$$Jf(x) := \mathbb{E}\Big[f\Big(x+\gamma b(x)+\sqrt{2\gamma}Z+\sigma'(x)\sigma(x)(Z^2-1)\gamma\Big)\Big],$$

where the law of $Z$ is a probability measure with compact support, mean 0 and variance 1. This chain is the one-dimensional Milstein scheme for (6). For a comparison with the Euler scheme, see, e.g. [15]. Similar arguments as above lead to the following result.

**Proposition 1.** *Let $Z$ have a law with compact support, mean 0 and variance 1 which satisfies a Logarithmic Sobolev inequality with constant $c$. Then the iterated kernel $J^n$ of the Milstein scheme satisfies a Logarithmic Sobolev inequality with constant*

$$\frac{c}{\rho+O(\gamma^{1/2})}\Big(1-(1-\rho\gamma+O(\gamma^{3/2}))^{2n}\Big).$$

# 4 Logarithmic Sobolev Inequalities For Multidimensional Euler Schemes With Constant Diffusion Coefficient And Potential Drift Coefficient

In this section, we are given a smooth function $U$ and we consider the equation

$$dX_t = \sqrt{2}dB_t - \nabla U(X_t)\,dt.$$

## 4.1 The Explicit Euler Scheme

Assume in this subsection that $\nabla U$ is a uniformly Lipschitz function on $\mathbb{R}^d$. For $U(x) = |x|^2/2$ one gets the Ornstein–Uhlenbeck process. The transition kernel of the explicit Euler scheme is

$$Kf(x) = \mathbb{E}\Big[f\Big(x - \nabla U(x)\gamma + \sqrt{2\gamma}Y\Big)\Big],$$

where $Y$ is a $d$ dimensional Gaussian vector $\mathcal{N}(0, I_d)$.

Let $\lambda \in \mathbb{R}$ be the largest real number such that

$$\langle \text{Hess } U(x)v, v\rangle \geq \lambda |v|^2 \tag{14}$$

for all $x$ and $v$ in $\mathbb{R}^d$. We now assume that $\lambda\gamma < 1$. This technical assumption is not restrictive since the discretization step $\gamma$ is small.

**Theorem 2.** *For all $n \in \mathbb{N}$, $x \in \mathbb{R}$ and smooth functions $f$ from $\mathbb{R}^d$ to $\mathbb{R}$,*

$$\text{Ent}_{K^n}\left(f^2\right) \leq D_{\gamma,n} K^n\Big(|\nabla f|^2\Big),$$

*where*

$$D_{\gamma,n} := \frac{4}{\lambda(2-\lambda\gamma)}\left(1 - (1-\lambda\gamma)^{2n}\right). \tag{15}$$

*Remark 3.* If $\lambda$ is equal to 0, $D_{\gamma,n}$ needs to be understood as $4n\gamma$.

*Proof.* The kernel $K$ satisfies a Logarithmic Sobolev inequality with constant $4\gamma$. Moreover,

$$\nabla Kf(x) = (I_d - \gamma \text{Hess } U(x))K(\nabla f)(x).$$

Therefore

$$|\nabla Kf(x)| \leq (1 - \gamma\lambda)K(|\nabla f|)(x). \tag{16}$$

Observe that

$$\text{Ent}_{K^n}\left(f^2\right) := K^n(f^2 \log f^2) - K^n(f^2)\log K^n(f^2)$$

is equal to

$$\sum_{i=1}^{n} \left\{ K^i\left[K^{n-i}(f^2)\log K^{n-i}(f^2)\right] - K^{i-1}\left[K^{n-i+1}(f^2)\log K^{n-i+1}(f^2)\right]\right\}.$$

In the sequel, $g_{n-i}$ will stand for $\sqrt{K^{n-i}(f^2)}$. We have

$$\mathrm{Ent}_{K^n}\left(f^2\right) = \sum_{i=1}^{n} K^{i-1}\left[\mathrm{Ent}_K\left(g_{n-i}^2\right)\right] \leq 4\gamma \sum_{i=1}^{n} K^i\left(\left|\nabla g_{n-i}\right|^2\right),$$

since $K$ satisfies a Logarithmic Sobolev inequality with constant $4\gamma$. Now, in view of the commutation relation (16), we get

$$\left|\nabla g_{n-i}\right|^2 = \frac{\left|\nabla K^{n-i}(f^2)\right|^2}{4K^{n-i}(f^2)} \leq (1-\lambda\gamma)^2 \frac{\left[K\left|\nabla K^{n-i-1}(f^2)\right|\right]^2}{4KK^{n-i-1}(f^2)}$$

for all $1 \leq i \leq n$. Therefore, using Cauchy–Schwarz inequality,

$$\frac{(Kf)^2}{K(g)} \leq K\left(\frac{f^2}{g}\right),$$

from which

$$\frac{\left[K\left|\nabla K^{n-i-1}(f^2)\right|\right]^2}{4KK^{n-i-1}(f^2)} \leq K\left[\frac{\left|\nabla K^{n-i-1}(f^2)\right|^2}{4K^{n-i-1}(f^2)}\right] = K\left[\left|\nabla g_{n-i-1}\right|^2\right].$$

A straightforward induction shows that

$$\left|\nabla g_{n-i}\right|^2 \leq (1-\lambda\gamma)^{2(n-i)} K^{n-i}\left[\left|\nabla f\right|^2\right].$$

Consequently,

$$\begin{aligned}\mathrm{Ent}_{K^n}\left(f^2\right) &\leq 4\gamma\left[\sum_{i=0}^{n-1}(1-\lambda\gamma)^{2i}\right]K^n\left[\left|\nabla f\right|^2\right] = 4\gamma\frac{1-(1-\lambda\gamma)^{2n}}{1-(1-\lambda\gamma)^2}K^n\left[\left|\nabla f\right|^2\right]\\ &= \frac{4}{\lambda(2-\lambda\gamma)}\left(1-(1-\lambda\gamma)^{2n}\right)K^n\left[\left|\nabla f\right|^2\right],\end{aligned}$$

which ends the proof.

### 4.2 The Implicit Euler Scheme

In this subsection we assume that $U$ is a uniformly convex function, that is, there exists $\lambda > 0$ such that

$$\langle \mathrm{Hess}\, U(x)v, v\rangle \geq \lambda |v|^2 \text{ for all } x, v \in \mathbb{R}^d.$$

Since the drift coefficient $-\nabla U$ is not necessarily globally Lipschitz, we consider the implicit Euler scheme

$$X_{n+1}^{\gamma} = X_n^{\gamma} - \nabla U\big(X_{n+1}^{\gamma}\big)\gamma + \sqrt{2\gamma}Y,$$

where $Y$ is a standard Gaussian variable on $\mathbb{R}^d$. Setting

$$\varphi(x) := (I + \nabla U(x)\gamma)^{-1}(x),$$

the kernel $\overline{K}$ of the implicit Euler scheme is

$$\overline{K}f(x) = \mathbb{E}\Big[f \circ \varphi\Big(x + \sqrt{2\gamma}Y\Big)\Big].$$

Let $\mathcal{N}(x, 2\gamma I)$ be the Gaussian distribution with mean $x$ and covariance matrix $2\gamma I_d$. We have

$$\mathrm{Ent}_{\overline{K}}\big(f^2\big) = \mathrm{Ent}_{\mathcal{N}(x,2\gamma I)}\big((f \circ \varphi)^2\big) \leq 4\gamma \mathbb{E}_{\mathcal{N}(x,2\gamma I)}\Big[|\nabla(f \circ \varphi)|^2\Big].$$

In view of the definition of $\varphi$ we get

$$\mathrm{Jac}\ \varphi(x) = [I_d + \gamma \mathrm{Hess}\ U(x)]^{-1},$$

and thus

$$\langle \mathrm{Jac}\ \varphi(x)v, v\rangle \leq (1 + \gamma\lambda)^{-1}|v|^2$$

for all $v$ in $\mathbb{R}^d$, from which

$$|\nabla(f \circ \varphi)| = |(\mathrm{Jac}\ \varphi)(\nabla f(\varphi))| \leq \frac{1}{1 + \lambda\gamma}|(\nabla f) \circ \varphi|.$$

Consequently, the kernels $\big(\overline{K}(\cdot)(x)\big)_x$ satisfy a Logarithmic Sobolev inequality with constant $\frac{4\gamma}{1+\lambda\gamma}$.

On the other hand,

$$\nabla \overline{K}f(x) = \mathbb{E}_{\mathcal{N}(x,2\gamma I)}[(\mathrm{Jac}\ \varphi)(\nabla f) \circ \varphi].$$

Then $\overline{K}$ and $\nabla$ satisfy the commutation relation

$$\big|\nabla \overline{K}(f)(x)\big| \leq (1 + \gamma\lambda)^{-1}\overline{K}(|\nabla f|)(x).$$

Obvious adaptations of the proof of Theorem 2 lead to

**Theorem 3.** *For all $n \in \mathbb{N}$, $x \in \mathbb{R}$ and smooth functions $f$ from $\mathbb{R}^d$ to $\mathbb{R}$ one has*

$$\mathrm{Ent}_{\overline{K}^n}\big(f^2\big) \leq \overline{D}_{\gamma,n}\overline{K}^n\Big(|\nabla f|^2\Big),$$

*where*

$$\overline{D}_{\gamma,n} = \frac{4(1 + \lambda\gamma)}{\lambda(2 + \lambda\gamma)}\bigg(1 - \frac{1}{(1 + \lambda\gamma)^{2n}}\bigg). \tag{17}$$

# 5 Uniform Logarithmic Sobolev Inequalities For One–Dimensional Euler Schemes With Constant Diffusion Coefficient And Convex Potential Drift Coefficient

Let $V$ be a smooth functions from $\mathbb{R}$ to $\mathbb{R}$. Let $(X_t)_{t\geq 0}$ be the solution of

$$dX_t = \sqrt{2}\, dB_t - V'(X_t)\, dt.$$

Notice that

$$\nabla P_t f(x) = \mathbb{E}^x\left[\nabla f(X_t) \exp\left(-\int_0^t V''(X_s)\, ds\right)\right]. \tag{18}$$

When $V'' \geq \lambda > 0$ we easily get the commutation relation

$$|\nabla P_t f| \leq e^{-\lambda t} P_t(|\nabla f|).$$

We now consider the less obvious case where $V''$ is supposed nonnegative only.

## 5.1 Poincaré Inequality For The Diffusion Process

**Lemma 1.** *Let*

$$D(t,x) := \mathbb{E}^x\left[\exp\left(-2\int_0^t V''(X_s)\, ds\right)\right].$$

*Then it exist $t_0 > 0$ such that $D(t_0) := \sup_{x\in\mathbb{R}} D(t_0, x) < 1$.*

*Proof.* One has $D(t+s) \leq D(t)D(s)$ for all $t \geq 0$ and $s \geq 0$. Indeed, for all $t \geq 0$, $s \geq 0$ and $x \in \mathbb{R}$, the Markov property ensures that

$$\begin{aligned}
D(t+s,x) &= \mathbb{E}^x\left[\exp\left(-2\int_0^t V''(X_u)\, du\right)\mathbb{E}^{X_t}\left\{\exp\left(-2\int_0^s V''(X_u)\, du\right)\right\}\right] \\
&= \mathbb{E}^x\left[D(s, X_t)\exp\left(-2\int_0^t V''(X_u)\, du\right)\right] \\
&\leq D(s)\mathbb{E}^x\left[\exp\left(-2\int_0^t V''(X_u)\, du\right)\right] = D(s)D(t,x) \leq D(s)D(t).
\end{aligned}$$

For $x \geq a$, set $\tau_a := \inf\{t \geq 0, X_t^x = a\}$. Then,

$$\begin{aligned}
D(t,x) = \mathbb{E}^x&\left[\mathbb{1}_{\{\tau_a < t\}}\exp\left(-2\int_0^t V''(X_s)\, ds\right)\right] \\
&+ \mathbb{E}^x\left[\mathbb{1}_{\{\tau_a \geq t\}}\exp\left(-2\int_0^t V''(X_s)\, ds\right)\right].
\end{aligned}$$

The second term on the r.h.s. is bounded from above by $\exp(-\lambda t)$. The first one can be bounded from above as follows:

$$\begin{aligned}
&\mathbb{E}^x\left[\mathbb{1}_{\{\tau_a<t\}}\exp\left(-2\int_0^t V''(X_s)\,ds\right)\right] \\
&\qquad = \mathbb{E}^x\left[\mathbb{1}_{\{\tau_a<t\}}e^{-\lambda\tau_a}\mathbb{E}\left\{\exp\left(-2\int_{\tau_a}^t V''(X_s)\,ds\right)\middle|\mathcal{F}_{\tau_a}\right\}\right] \\
&\qquad = \mathbb{E}^x\left[\mathbb{1}_{\{\tau_a<t\}}e^{-\lambda\tau_a}D(t-\tau_a,a)\right] \\
&\qquad = \mathbb{E}^x\left[\mathbb{1}_{\{\tau_a<t/2\}}e^{-\lambda\tau_a}D(t-\tau_a,a)\right] \\
&\qquad\quad + \mathbb{E}^x\left[\mathbb{1}_{\{t/2\le\tau_a<t\}}e^{-\lambda\tau_a}D(t-\tau_a,a)\right] \\
&\qquad \le D(t/2,a)+e^{-\lambda t/2}.
\end{aligned}$$

One can easily show that

$$\sup_{x\ge a} D(t,x) \le D(t/2,a)+e^{-\lambda t/2}+e^{-\lambda t}.$$

The right hand side is bounded from above by 1 for all $t$ large enough. By symmetry, one also has

$$\sup_{|x|\ge a} D(t,x)<1.$$

Finally, the continuity of $x\mapsto D(t,x)$ ensures that

$$\sup\{D(t,x),\ x\in[-a,a]\}<1$$

for all $t>0$, which ends the proof.

**Proposition 2.** *Assume that $V$ is convex and it exists $a\ge 0$ and $\lambda>0$ such that* $\sup\{V''(x),\ |x|\ge a\}$. *It then exists $t_0\ge 0$ such that $D(t_0)<1$ and*

$$\mathbf{Var}_{P_t(\cdot)(x)}(f)\le 2t_0\left(1+\frac{1-D(t_0)^{t/t_0-1}}{1-D(t_0)}\right)P_t\left(|\nabla f|^2\right)$$

*for all $t>t_0$. Moreover, the invariant measure $\mu$ of $(X_t)$ satisfies a Poincaré inequality with constant $2t_0(1+1/(1-D(t_0)))$.*

Observe that the proposition implies that, for all $t_0>0$ and $n\in\mathbb{N}$,

$$|\nabla P_{nt_0}f(x)|^2\le D(nt_0)P_{nt_0}\left(|\nabla f|^2\right)(x)\le D(t_0)^nP_{nt_0}\left(|\nabla f|^2\right)(x).$$

*Proof.* Let $t_0>0$ be as in Lemma 1 and set $K(f)(x):=P_{t_0}f(x)$. Arbitrarily choose $t>0$. Let $n$ be the integer part of $t/t_0$. We have

$$\begin{aligned}
\mathbf{Var}_{P_t(\cdot)(x)}(f) &= P_t(f^2)(x)-(P_t(f)(x))^2 \\
&= P_{nt_0}\big(\mathbf{Var}_{P_{t-nt_0}}(f)\big)+\mathbf{Var}_{P_{nt_0}}(P_{t-nt_0}f).
\end{aligned}$$

Since $t - nt_0 < t_0$, $P_{t-nt_0}$ satisfies a Poincaré inequality with a constant bounded by $2t_0$. Therefore,

$$\mathbf{Var}_{P_t(\cdot)(x)}(f) \leq 2t_0 P_t\left(|\nabla f|^2\right) + \mathbf{Var}_{K^n}(P_{t-nt_0}(f)).$$

Moreover, $K^n$ satisfies the Poincaré inequality

$$\mathbf{Var}_{K^n}(f) = \leq \frac{2t_0(1-D(t_0)^n)}{1-D(t_0)} K^n(|\nabla f|^2).$$

In view of the commutation relation $|\nabla P_{t-nt_0}(f)|^2 \leq P_{t-nt_0}(|\nabla f|^2)$, we finally get

$$\mathbf{Var}_{P_t(\cdot)(x)}(f) \leq 2t_0 P_t\left(|\nabla f|^2\right) + \frac{2t_0(1-D(t_0)^n)}{1-D(t_0)} K^n P_{t-nt_0}(|\nabla f|^2).$$

### 5.2 Uniform Poincaré Inequality For The Euler Scheme

We now get an uniform Poincaré inequality for the Euler scheme with kernel

$$K(f)(x) = \mathbb{E}\{f(x - V'(x)\alpha + Y)\},$$

where $Y$ is Gaussian $\mathcal{N}(0, 2\alpha)$. Consider the commutation relation

$$\nabla K^2 f(x) = (1 - \alpha V''(x))\mathbb{E}^x[(1 - \alpha V''(X_1))\nabla f(X_2)].$$

As $1 - \alpha V''(x) \leq 1$, we have

$$\left|\nabla K^2 f(x)\right|^2 \leq \mathbb{E}^x\left[(1 - \alpha V''(X_1))^2\right] K^2(|\nabla f|^2)(x).$$

Since the support of the law of $X_1$ is the entire real line, for all $x \in \mathbb{R}$, $\mathbb{E}^x\left[(1 - \alpha V''(X_1))^2\right] < 1$. Moreover

$$\sup_{x\in\mathbb{R}} \mathbb{E}^x\left[(1 - \alpha V''(X_1))^2\right] < 1$$

since $V''(y) \geq \lambda > 0$ if $|y| \geq a$. This observation leads to Poincaré inequalities for both $K^n$ and the invariant measure of the Euler scheme; in these inequalities the constants are uniform w.r.t. time.

### 5.3 Uniform Logarithmic Sobolev Inequality

In view of (18) we have

$$\begin{aligned}
|\nabla P_t(f)(x)| &\leq \mathbb{E}^x\left[|f'(X_t)| \exp\left(-\int_0^t V''(X_s)\,ds\right)\right] \\
&= \mathbb{E}\left[|f'(X_t)| \mathbb{E}\left\{\exp\left(-\int_0^t V''(X_s)\,ds\right)\middle| X_0 = x, X_t\right\}\right].
\end{aligned}$$

To get the commutation relation $|\nabla P_t(f)| \leq \rho P_t(|f'|)$, one needs to suppose that the following property holds true:

*Property 1.* There exists $t_0 > 0$ such that $D(t_0) := \sup_{x,y\in\mathbb{R}} D(t_0, x, y) < 1$, where

$$D(t_0, x, y) := \mathbb{E}\left[\exp\left(-\int_0^{t_0} V''(X_s)\,ds\right)\middle| X_0 = x, X_{t_0} = y\right]$$

for all $t > 0$ and $x$, $y$ in $\mathbb{R}$.

This property holds true when $V''$ is non negative. We are now trying to relax the convexity condition on $V$, assuming only that $V$ is strictly convex out of a compact set.

# 6 Applications

## 6.1 Monte Carlo Simulations

Poincaré and Logarithmic Sobolev inequalities are important for applications because they provide concentration inequalities for empirical means. The proof of this claim uses a tensorization argument and the Herbst's argument that we now remind.

**Theorem 4.** *Let $\mu$ be a probability measure on $\mathbb{R}^d$. If $\mu$ satisfies a spectral gap (respectively Logarithmic Sobolev) inequality with constant $C$, then the measure $\mu^{\otimes N}$ on $\mathbb{R}^{dN}$ satisfies a spectral gap (respectively Logarithmic Sobolev) inequality with constant $C$.*

**Theorem 5.** *If $\mu$ satisfies a Logarithmic Sobolev inequality with constant $c$, then for all Lipschitz functions $f$ with Lipschitz constant $\varepsilon$ and all $\lambda > 0$,*

$$K\big(e^{\lambda f}\big) \le e^{c\lambda^2\varepsilon^2/4} e^{\lambda K f}.$$

The Herbst's argument ensures that a measure which satisfies a Logarithmic Sobolev inequality has Gaussian tails (see [9]). One then deduces

**Theorem 6.** *Let the measure $\mu$ on $\mathbb{R}^d$ satisfy the Logarithmic Sobolev inequality* (2) *with constant $C$. Let $X_1, \dots, X_N$ be i.i.d. random variables with law $\mu$. Then, for all bounded Lipschitz functions on $\mathbb{R}^d$, it holds*

$$\mathbb{P}\left(\left|\frac{1}{N}\sum_{i=1}^{N} f(X_i) - \mathbb{E}(f(X_1))\right| \ge r\right) \le 2e^{-Nr^2/C}. \tag{19}$$

One can also show

**Theorem 7.** *Assume that the measure $\mu$ on $\mathbb{R}^d$ satisfies the Poincaré inequality* (1) *with constant $c$. Let $X_1, \dots, X_N$ be i.i.d. random variables with law $\mu$. Then, for all bounded Lipschitz functions on $\mathbb{R}^d$ with Lipschitz constant $\alpha$, it holds*

$$\mathbb{P}\left(\left|\frac{1}{N}\sum_{i=1}^{N} f(X_i) - \mathbb{E}(f(X_1))\right| \ge r\right) \le 2\exp\left(-\frac{N}{K}\min\left(\frac{r}{\alpha}, \frac{r^2}{\alpha^2}\right)\right). \tag{20}$$

### 6.2 Ergodic Simulations

Let $(Y_n)_n$ be a Markov chain on $\mathbb{R}^d$ with transition kernel $K$ such that, for all smooth functions $f$,

$$|\nabla K f|(x) \le \alpha K(|\nabla f|)(x), \tag{21}$$

for some $\alpha < 1$. For example, fix $t_0 > 0$ and set $K = P_{t_0}$, where $(P_t)$ is the semi-group of the diffusion

$$dX_t = dB_t - \nabla U(X_t)\,dt$$

with Hess $U(x) \ge \rho I$ and $\rho > 0$. One can then choose $\alpha = e^{-\rho t_0}$. Alternatively, $K$ can be chosen as the transition kernel of the implicit Euler scheme which discretizes $(X_t)$. Using Herbst's argument one can show

**Proposition 3.** *For all 1-Lipschitz functions $f$ on $\mathbb{R}^d$,*

$$\mathbb{P}_x\left(\left|\frac{1}{N}\sum_{i=1}^N f(Y_i) - \int f\,d\mu\right| \ge r + \frac{d_x}{N}\right) \le 2\exp\left(-\frac{N(1-\alpha)^2}{c}r^2\right), \tag{22}$$

*where* $d_x = \dfrac{\alpha}{1-\alpha}\mathbb{E}_x(|x - X_1|)$.

### 6.3 Stochastic Particle Methods For McKean–Vlasov Equations

Consider the McKean–Vlasov equation

$$\frac{\partial}{\partial t}P_t = \frac{1}{2}\sum_{i,j=1}^d \frac{\partial^2}{\partial x_i \partial x_j}(a_{ij}[x, P_t]P_t) - \sum_{i=1}^d \frac{\partial}{\partial x_i}(b_i[x, P_t]P_t), \tag{23}$$

where $P_t$ is a probability measure on $\mathbb{R}^d$ and, for some functions $b$ and $\sigma$,

$$\begin{aligned} b[x,p] &= \int_{\mathbb{R}^d} b(x,y)\,p(dy), \\ \sigma[x,p] &= \int_{\mathbb{R}^d} \sigma(x,y)\,p(dy), \\ a[x,p] &= \sigma[x,p]\sigma[x,p]^* \end{aligned}$$

for all $x$ in $\mathbb{R}^d$ and all probability measures $p$. The functions $b$ and $\sigma$ are the *interaction kernels*. This equation has been introduced by [13] and then widely studied from both probabilistic and analytic points of view (see, e.g., [14] for a review). Under appropriate conditions one can show that $P_t$ is the marginal law at time $t$ of the law of the solution of the nonlinear stochastic differential equation

$$\begin{cases} \overline{X}_t = \overline{X}_0 + \int_0^t \sigma\left[\overline{X}_s, Q_s\right] dB_s + \int_0^t b\left[\overline{X}_s, Q_s\right] ds, \\ \mathcal{L}(\overline{X}_t) = Q_t, \end{cases}$$

where $\mathcal{L}(\overline{X}_t)$ stands for the law of $\overline{X}_t$: one thus has $P_t = Q_t$. This probabilistic interpretation suggests to consider the stochastic particle system in mean field interaction

$$\begin{cases} dX_t^{i,N} = \frac{1}{N} \sum_{j=1}^N \sigma(X_t^{i,N}, X_t^{j,N}) dB_t^i + \frac{1}{N} \sum_{j=1}^N b(X_t^{i,N}, X_t^{j,N})\, dt, \\ X_0^{i,N} = X_0^i, \quad i = 1 \dots, N, \end{cases}$$

where $(B_\cdot^i)_i$ are independent Brownian motions on $\mathbb{R}^d$. One aims to approximate $P_t$ by the empirical measure $\mu_t^N$ of the particle system:

$$\mu_t^N = \frac{1}{N} \sum_{i=1}^N \delta_{X_t^{i,N}}.$$

The convergence of the particle system to the nonlinear process has been deeply studied (see [14]). It can also be shown that the law of the particle system at time $t$ satisfies a Logarithmic Sobolev inequality with a constant which does not depend on the number of particles. However the corresponding confidence intervals are not fully satisfying for numerical purposes since the particle system needs to be discretized in time to be simulated. The convergence rate of the Euler scheme in terms of $N$ and the discretization step are studied in [6, 7, 2, 5]. One can also show that the Euler scheme satisfies a spectral gap inequality with a constant independent of $N$:

**Proposition 4.** *Suppose that the coefficients $b$ and $\sigma$ are bounded Lipschitz functions. Then the Euler scheme for the above particle system satisfies*

$$\mathbb{P}\left( \left| \frac{1}{N} \sum_{i=1}^N f\left(X_t^{\gamma,i,N}(x)\right) - \mathbb{E} f\left(X_t^{\gamma,i,N}(x)\right) \right| \geq r \right) \leq 2 \exp\left( -\frac{N}{C_t^\gamma} \min(r, r^2) \right)$$

*for all Lipschitz functions $f$ with Lipschitz constant equal to 1 and all $r \geq 0$,*

We now consider the granular media equation:

$$\frac{\partial u}{\partial t} = \operatorname{div}\left[\nabla u + u(\nabla V + \nabla W * u)\right],$$

where $*$ stands for the convolution and $V$ and $W$ are convex potentials on $\mathbb{R}^d$. This equation in $\mathbb{R}$ with $V = |x|^2/2$ and $W = |x|^3$ has been introduced by [4] to describe the evolution of media composed of many particles colliding inelastically in a thermal bath. One can show that the solution $u_t$ of the nonlinear partial differential equation converges to an equilibrium distribution $u_\infty$. Indeed, define the generalized relative entropy as

$$\eta(u) = \int u \log u + \int uV + \frac{1}{2} \int\int W(x-y) u(x) u(y).$$

One has

**Theorem 8 ([8]).** *If $V$ is uniformly convex, i.e.* Hess $V \geq \lambda I$ *and $W$ is even and convex then*

$$\eta(u_t) - \eta(u_\infty) \leq K e^{-2\lambda t}$$

*where $u_\infty$ is the unique minimizer of $\eta$ or equivalently the unique solution of*

$$u_\infty = \frac{1}{Z} \exp\left(-V(x) - W * u_\infty(x)\right),$$

*with*

$$Z = \int \exp\left(-V(x) - W * u_\infty(x)\right) dx.$$

The granular media equations can be viewed as McKean–Vlasov equations. The particle system well defined and the propagation of chaos result holds uniformly in time (see [11]):

$$\mathbb{E}\left(\left|X_t^{i,N} - \overline{X}_t^i\right|\right) \leq \frac{c}{\sqrt{N}},$$

where the $\overline{X}^i$'s are independent copies of the solution of the nonlinear equation. As the interaction kernels are not globally Lipschitz, one needs to use the implicit Euler scheme to discretize the particle system. Let $(Y_n^{N,\gamma})_{n\in\mathbb{N}}$ be this implicit Euler scheme with discretization step $\gamma$. We have (see [12]):

**Theorem 9.** *There exists $c > 0$ such that*

$$\mathbb{P}\left(\left|\frac{1}{N}\sum_{i=1}^{N} f(Y_t^{i,N,\gamma}) - \int f\, du_\infty\right| \geq r + c\sqrt{\gamma} + \frac{c}{\sqrt{N}} + ce^{-\lambda t}\right) \leq 2e^{-N\lambda r^2/2}$$

*for all Lipschitz functions $f$ with Lipschitz constant 1.*

## References


1. C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto, and G. Scheffer. *Sur les inégalités de Sobolev logarithmiques*, volume 10 of *Panoramas et Synthèses.* Société Mathématique de France, Paris, 2000.
2. F. Antonelli and A. Kohatsu-Higa. Rate of convergence of a particle method to the solution of the McKean-Vlasov equation. *Ann. Appl. Probab.*, 12:423–476, 2002.
3. D. Bakry. On Sobolev and logarithmic Sobolev inequalities for Markov semi-groups. In *New trends in stochastic analysis (Charingworth, 1994)*, pages 43–75, River Edge, NJ, 1997. Taniguchi symposium, World Sci. Publishing.
4. D. Benedetto, E. Caglioti, and M. Pulvirenti. A kinetic equation for granular media. *RAIRO Modél. Math. Anal. Numér.*, 31(5):615–641, 1997.
5. M. Bossy. Optimal rate of convergence of a stochastic particle method to solutions of 1D viscous scalar conservation laws. *Math. Comp.*, 73(246):777–812, 2004.

6. M. Bossy and D. Talay. Convergence rate for the approximation of the limit law of weakly interacting particles: application to the Burgers equation. *Ann. Appl. Probab.*, 6(3):818–861, 1996.
7. M. Bossy and D. Talay. A stochastic particle method for the McKean-Vlasov and the Burgers equation. *Math. Comp.*, 66(217):157–192, 1997.
8. J. Carrillo, R. McCann, and C. Villani. Kinetic equilibration rates for granular media. *Rev. Matematica Iberoamericana*, 19:1-48, 2003.
9. M. Ledoux. Concentration of measure and logarithmic Sobolev inequalities. In *Séminaire de Probabilités XXXIII. Lectures Notes in Math., vol 1709*, pages 120–216. Springer, Berlin, 1999.
10. F. Malrieu and D. Talay. Poincaré Inequalities for Euler schemes. In preparation.
11. F. Malrieu. Logarithmic Sobolev inequalities for nonlinear PDE's. *Stochastic Process. Appl.*, 95(1):109–132, 2001.
12. F. Malrieu, Convergence to equilibrium for granular media equations and their Euler schemes. *Ann. Appl. Probab.*, 13(2): 540–560, 2003.
13. H. P. McKean, Jr. Propagation of chaos for a class of non-linear parabolic equations. In *Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967)*, pages 41–57. Air Force Office Sci. Res., Arlington, Va., 1967.
14. S. Méléard. Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models. In *Probabilistic models for nonlinear partial differential equations (Montecatini Terme, 1995), D. Talay and L. Tubaro (Eds.)*, pages 42–95. Springer, Berlin, 1996.
15. D. Talay. Probabilistic numerical methods for partial differential equations: elements of analysis. In *Probabilistic models for nonlinear partial differential equations (Montecatini Terme, 1995), D. Talay and L. Tubaro (Eds.)*, pages 148–196. Springer, Berlin, 1996.