ANNALES MATHÉMATIQUES

BLAISE PASCAL

Pierre Monmarché

L^2 hypocoercivity, deviation bounds, hitting times and Lyapunov functions

Volume 29, n° 2 (2022), p. 295-306.

https://doi.org/10.5802/ambp.414

Publication éditée par le laboratoire de mathématiques Blaise Pascal de l'université Clermont Auvergne, UMR 6620 du CNRS Clermont-Ferrand — France

CENTRE MERSENNE

Publication membre du Centre Mersenne pour l'édition scientifique ouverte
http://www.centre-mersenne.org/
e-ISSN : 2118-7436

Annales mathématiques Blaise Pascal **29**, 295-306 (2022)

L^2 hypocoercivity, deviation bounds, hitting times and Lyapunov functions

Pierre Monmarché

Abstract

We establish that, for a Markov semi-group, L^2 hypocoercivity, i.e. contractivity for a modified L^2 norm, implies quantitative deviation bounds for additive functionals of the associated Markov process and exponential integrability of the hitting time of sets with positive measure. Moreover, in the case of diffusion processes and under a strong hypoellipticity assumption, we prove that L^2 hypocoercivity implies the existence of a Lyapunov function for the generator. A french version is available [14].

Hypocoercivité L^2, inégalité de concentration, temps d'atteinte et fonctions de Lyapunov

Résumé

On montre que, pour un semi-groupe de Markov, l'hypocoercivité L^2 (c'est-à-dire la contractivité d'une norme L^2 modifiée) implique des inégalités de concentration quantitatives et l'intégrabilité exponentielle des temps d'atteinte des ensembles de mesure positive. D'autre part, pour les diffusions et sous une hypothèse forte d'hypoellipticité, on établit que l'hypocoercivité L^2 implique l'existence d'une fonction de Lyapunov pour le générateur associé. Une version en français est disponible [14].

1. Introduction

This note is primarily motivated by the comparison between two methods for obtaining quantitative rates of convergence for ergodic Markov processes: functional inequalities and entropy dissipation on the one hand, and the classical Meyn–Tweedie approach on the other hand, based on a Foster–Lyapunov criterion together with a local Doeblin or coupling condition. The link between these two kind of arguments is studied by Cattiaux, Guillin and their co-authors in a series of works [2, 9, 10, 11], but mainly in the case of reversible Markov processes, with an invariant measure that typically satisfies a Poincaré (or similar) inequality (with respect to the Dirichlet form associated with the process). Nevertheless, over the last decade, hypocoercivity methods of modified entropies have proven to be able to handle non-reversible, non-elliptic and non-diffusive processes with functional inequality arguments. Here we will focus on the Dolbeault–Mouhot–Schmeiser

This work has been partially supported by the Project EFI ANR-17-CE40-0030 of the French National Research Agency.
Keywords: Hypocoercivité, fonctions de Lyapunov.
2020 *Mathematics Subject Classification*: 60J25, 35F15, 35H10.

(DMS) method [12] for L^2 hypocoercivity. A very appealing feature of the latter is that it provides a general construction for a modified L^2-norm, independently from the process. This is quite different in a Meyn–Tweedie approach where a Lyapunov function has to be constructed based on the dynamics of the process, which can be tricky for degenerate processes. An example is provided by the comparison of the two works [1] and [13] that both study the so-called Bouncy Particle process, which is a kinetic piecewise deterministic Markov process, respectively with the DMS and the Meyn–Tweedie method. With the Meyn–Tweedie approach, the construction of the Lyapunov function is quite intricate and this leads to not-so-nice assumptions on the log-density of the equilibrium, while with the DMS method the construction is standard and the conditions on the log-density are more general and quite simpler. A natural question is thus: provided that (hypocoercive) exponential decay holds in L^2, does it exists a Lyapunov function in the sense of Meyn–Tweedie? A positive answer is provided in [11] in the case of elliptic reversible diffusions for which the exponential decay holds with the usual L^2 norm. The result is based on the exponential integrability of the hitting times of the process, itself obtained from the deviation bounds on additive functionals of the process derived in [8, 16]. These two results are interesting by themselves; in fact, in some cases, getting estimates on the hitting times is the main question and the Lyapunov function is just an intermediate tool to get them. Our main result is that, if the DMS method applies, then these two results also hold. From the exponential integrability of hitting times, in the case of strongly hypoelliptic diffusions, we construct a Lyapunov function for the generator, which answers our initial question. In fact, as we became aware after the redaction of this note, the deviation bounds have been established in the recent preprint [5]. To the best of our knowledge, the other results are new.

2. Results and proofs

Consider $(X_t)_{t\geqslant 0}$ a continuous-time conservative Markov process on a Polish space E, with some invariant probability measure μ. Denote $(P_t)_{t\geqslant 0}$ the associated Markov semi-group on $L^2(\mu)$ and L. We suppose that L is closed and that its domain $D(L)$ is dense in $L^2(\mu)$. Denote $\|\cdot\|_2$ and $\langle\cdot\rangle$ the usual norm and scalar product on $L^2(\mu)$. Our main assumption is the following:

Assumption 2.1. *There exists $\rho > 0$ and a symmetric bounded linear operator S on $L^2(\mu)$ such that for all $f \in D(L)$,*

$$\|Sf\|_2 \leqslant \frac{1}{2}\|f-\mu f\|_2$$
$$\langle f, Lf + SLf\rangle \leqslant -\rho\|f-\mu f\|_2^2\,.$$

The existence of such an S, which will be used in the present note as a black box assumption, is in fact the main tool of the DMS method. More precisely, considering $\varepsilon \in (0,1)$ and the operator A that are used to define the modified entropy H in [12], we set $S = \varepsilon(\mathrm{Id}-\mu)(A+A^*)(\mathrm{Id}-\mu)/2$. It is straightforward to check that, under the assumption (H_1)–(H_4) of [12], this operator S satisfies the conditions of Assumption 2.1. Once such an operator S is obtained, the DMS method to obtain hypocoercive decay in $L^2(\mu)$ is the following: denote $B = \mathrm{Id} + S$ and consider the scalar product and Hilbertian norm

$$\langle f, g\rangle_B = \langle f, Bg\rangle\,, \qquad \|f\|_B = \sqrt{\langle f, f\rangle_B}\,.$$

The latter is equivalent to the usual norm on $L^2(\mu)$, more precisely $1/2\|f\|_2^2 \leqslant \|f\|_B^2 \leqslant 3/2\|f\|_2^2$. Assumption 2.1 implies that for all $f \in D(L)$ with $\mu f = 0$,

$$\langle f, Lf\rangle_B \leqslant -\rho\|f\|_2^2 \leqslant -\frac{2\rho}{3}\|f\|_B^2\,.$$

In other words, $(e^{2\rho t/3}P_t)_{t\geqslant 0}$ is dissipative on $\{f \in L^2(\mu), \mu f = 0\}$ endowed with the scalar product $\langle\cdot\rangle_B$, so that by the Lumer–Philips Theorem [17, Chapter IX, p. 250], for all f with $\mu f = 0$,

$$\|P_t f\|_B \leqslant e^{-2\rho t/3}\|f\|_B\,.$$

This gives an hypocoercive decay in the usual norm of $L^2(\mu)$:

$$\|P_t f - \mu f\|_2 \leqslant \sqrt{2}\|P_t f - \mu f\|_B \leqslant \sqrt{2}e^{-2\rho t/3}\|f - \mu f\|_B \leqslant \sqrt{3}e^{-2\rho t/3}\|f - \mu f\|_2\,.$$

In fact, given a bounded measurable function V on E, the same argument works for the Feynman–Kac semigroup $(P_t^V)_{t\geqslant 0}$ on $L^2(\mu)$ given by

$$P_t^V f(x) = \mathbb{E}_x\left(f(X_t)e^{\int_0^t V(X_s)\mathrm{d}s}\right)\,,$$

where the subscript x denotes the starting point of the process. Indeed, denoting

$$\Lambda(V) = \sup\left\{\langle f, (L+V)f\rangle_B \;:\; f \in D(L), \|f\|_B = 1\right\}\,,$$

we get that $(e^{-t\Lambda(V)}P_t^V)_{t\geqslant 0}$ is dissipative on $(L^2(\mu), \langle\cdot\rangle_B)$ and thus, for all $f \in L^2(\mu)$,

$$\|P_t^V f\|_B \leqslant e^{t\Lambda(V)}\|f\|_B\,. \tag{2.1}$$

From this, following [8] and [16], we can establish the following deviation bounds:

Theorem 2.2. *Under Assumption 2.1, let $\nu \ll \mu$ be a probability measure on E and V be a bounded measurable function on E. Then, for all $t \geqslant 0$ and all $r \geqslant 0$,*

$$\mathbb{P}_\nu\left(\frac{1}{t}\int_0^t V(X_s)\mathrm{d}s - \mu V \geqslant r\right) \leqslant \sqrt{2}\left\|\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right\|_2 e^{-th(r)}\,,$$

where

$$h(r) = \frac{\rho r^2}{25\|V\|_2^2 + 6\|V\|_\infty r}\,.$$

Of course, by changing V to $-V$, this provides a non-asymptotic confidence interval for the empirical mean of bounded functions.

Since the DMS method does not yield sharp estimates, we have prefered a simple expression for $h(r)$ rather than the largest estimate we could obtain.

Although a similar result is already established in [5], since the proof is nice and short, we leave it for completeness.

Proof. Without loss of generality we suppose that $\mu V = 0$. By the Chebyshev and Cauchy–Schwarz inequalities and using (2.1), for all $\nu \ll \mu$ and $\lambda \geqslant 0$,

$$\begin{aligned}
\mathbb{P}_\nu\left(\frac{1}{t}\int_0^t V(X_s)\mathrm{d}s \geqslant r\right) &\leqslant e^{-\lambda t r}\mathbb{E}_\nu\left(e^{\lambda\int_0^t V(X_s)\mathrm{d}s}\right)\\
&= e^{-\lambda t r}\int_E P_t^{\lambda V}\mathbb{1}\mathrm{d}\nu\\
&\leqslant \sqrt{2}e^{-\lambda t r}\left\|\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right\|_2 \left\|P_t^{\lambda V}\mathbb{1}\right\|_B\\
&\leqslant \sqrt{2}\left\|\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right\|_2 e^{-\lambda t r + t\Lambda(\lambda V)}\,. \qquad (2.2)
\end{aligned}$$

We now bound $\Lambda(\lambda V)$ in the spirit of [8]. First, from Assumption 2.1,

$$\Lambda(V) \leqslant 2\sup\left\{-\rho\|f-\mu f\|^2 + \langle f, Vf\rangle_B \ :\ f\in D(L), \|f\|_2 = 1\right\}$$

Fix $f\in D(L)$ with $\|f\|_2 = 1$ and let $\gamma \geqslant 0$ be given by $1+\gamma^2 = 1/(\mu f)^2$, so that

$$g := \frac{1}{\gamma}\left(\sqrt{1+\gamma^2}f - 1\right) = \frac{\sqrt{1+\gamma^2}}{\gamma}(f-\mu f)$$

satisfies $\mu g = 0$ and

$$\mu(g^2) = \frac{1+\gamma^2}{\gamma^2}\left(\mu(f^2) - (\mu f)^2\right) = 1\,.$$

Moreover,

$$\begin{aligned}
\langle f, Vf\rangle_B &= \int Vf^2\mathrm{d}\mu + \langle A(f-\mu f), (Vf - \mu(Vf))\rangle\\
&= \int Vf^2\mathrm{d}\mu + \langle A(f-\mu f), V(f-\mu f)\rangle + \langle A(f-\mu f), (V-\mu V)\rangle\,\mu f\\
&= \frac{1}{1+\gamma^2}\int V\left(\gamma^2 g^2 + 2\gamma g + 1\right)\mathrm{d}\mu + \frac{\gamma^2}{1+\gamma^2}\langle Ag, Vg\rangle + \frac{\gamma}{1+\gamma^2}\langle Ag, V\rangle\\
&\leqslant \frac{3\gamma^2}{2(1+\gamma^2)}\|V\|_\infty + \frac{5\gamma}{2(1+\gamma^2)}\|V\|_2\,,
\end{aligned}$$

where we used the Cauchy–Schwarz inequality and the fact that $\|Ag\| \leqslant 1/2$. As a consequence,

$$\langle f, (L+V)f\rangle_B \leqslant \frac{\gamma}{1+\gamma^2}\left(-\gamma\rho + \frac{3\gamma}{2}\|V\|_\infty + \frac{5}{2}\|V\|_2\right),$$

and for all $\lambda \geqslant 0$,

$$\begin{aligned}
\Lambda(V) &\leqslant 2\sup_{\gamma\geqslant 0}\left\{\frac{\gamma}{1+\gamma^2}\left(-\gamma\rho + \frac{3\gamma}{2}\|V\|_\infty + \frac{5}{2}\|V\|_2\right)\right\} \\
&\leqslant 2\sup_{\gamma\geqslant 0}\left\{\gamma\left(-\gamma\rho + \frac{3\gamma}{2}\|V\|_\infty + \frac{5}{2}\|V\|_2\right)\right\} \\
&= \begin{cases} \dfrac{25\|V\|_2^2}{4\rho - 6\|V\|_\infty} & \text{if } 3\|V\|_\infty < 2\rho\,, \\ +\infty & \text{otherwise.} \end{cases}
\end{aligned}$$

In particular, denoting

$$\lambda_0 := 2\rho/(3\|V\|_\infty)\,, \qquad \beta = \frac{25\|V\|_2^2}{6\|V\|_\infty}\,,$$

we get

$$\begin{aligned}
\sup_{\lambda>0}\{\lambda r - \Lambda(\lambda V)\} &\geqslant \sup_{\lambda\in[0,\lambda_0)}\left\{\lambda r - \frac{\beta\lambda^2}{\lambda_0-\lambda}\right\} \\
&= \frac{\lambda_0 r^2}{\beta\left(1+\sqrt{1+r/\beta}\right)^2} \geqslant \frac{\lambda_0 r^2}{4(\beta+r)} = h(r)\,.
\end{aligned}$$

Taking the supremum over $\lambda \in [0,\lambda_0)$ in (2.2) concludes. □

Let U be a measurable subset of E and $T_U = \inf\{t \geqslant 0, X_t \in U\}$ be the first hitting time of U of the process $(X_t)_{t\geqslant 0}$.

Theorem 2.3. *Under Assumption 2.1, if $\mu(U) > 0$, then for all $\theta < h(\mu(U))$ and all probability measure $\nu \ll \mu$,*

$$\mathbb{E}_\nu\left(e^{\theta T_U}\right) \leqslant 1 + \sqrt{2}\left\|\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right\|_2 \frac{\theta}{h(\mu(U)) - \theta}\,.$$

Proof. Following [11] we remark that for all $t \geqslant 0$,

$$\{T_U \geqslant t\} \subset \left\{\frac{1}{t}\int_0^t \mathbb{1}_U(X_s)\mathrm{d}s = 0\right\}.$$

Theorem 2.2 applied with $V = -\mathbb{1}_U$ and $r = \mu(U)$ yields

$$\mathbb{P}_\nu\left(-\frac{1}{t}\int_0^t \mathbb{1}_U(X_s)\mathrm{d}s + \mu(U) \geqslant \mu(U)\right) \leqslant \sqrt{2}\left\|\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right\|_2 e^{-th(\mu(U))}\,.$$

Then, conclusion follows from

$$\mathbb{E}_\nu\left(e^{\theta T_U}\right) = 1 + \int_0^\infty \theta e^{\theta t}\mathbb{P}_\nu\left(T_U \geqslant t\right)\mathrm{d}t \leqslant 1 + \sqrt{2}\left\|\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right\|_2 \int_0^\infty \theta e^{t(\theta - h(\mu(U)))}\mathrm{d}t\,. \quad \square$$

In particular, Theorem 2.3 implies that, for all U with $\mu(U) > 0$, $W(x) := \mathbb{E}_x\left(e^{\theta T_U}\right)$ is finite for μ-almost every $x \in E$, and $W \in L^1(\mu)$. Moreover, for $\theta < h(\mu(U))/2$, $W \in L^2(\mu)$, and one can try to make sense of LW in $L^2(\mu)$ based on the theory of Dirichlet form, see e.g. [3, Section 1.7] and references within. We won't detail this: instead, we will see that, in fact, we can have better under some regularity assumption.

Assumption 2.4. *The transition kernel $p_t(x, \mathrm{d}y)$ of the process admits for all $t > 0$ and $x \in E$ a density r_t with respect to μ, i.e. $p_t(x, \mathrm{d}y) = r_t(x, y)\mu(\mathrm{d}y)$, such that $y \mapsto r_t(x, y)$ is in $L^2(\mu)$ for all $x \in E$ with $x \mapsto \|r_t(x, \cdot)\|_2$ locally bounded.*

Under Assumptions 2.1 and 2.4, denoting $T_U' = \inf\{t \geqslant 1, X_t \in U\}$, then $T_U' \geqslant T_U$ and

$$\mathbb{E}_x\left(e^{\theta T_U}\right) \leqslant \mathbb{E}_x\left(e^{\theta T_U'}\right) = \mathbb{E}_{p_1(x,\cdot)}\left(e^{\theta T_U}\right)\,. \tag{2.3}$$

From Theorem 2.3 and the assumption on $x \mapsto \|r_t(x, \cdot)\|_2$, $W(x)$ is then finite for all $x \in E$ and W is locally bounded.

Assumption 2.5. *E is a d-dimensional C^∞ manifold and $L = Y_0 + \sum_{i=1}^m Y_i^2$ for some $m \geqslant 1$ where $Y_0, \dots, Y_m$ are bounded C^∞ vector fields with all their derivative bounded, and such that for some $\alpha > 0$ and $N \in \mathbb{N}$,*

$$\forall\, x \in E\,,\ y \in \mathbb{R}^d\,, \qquad \sum_{j=1}^m \left\langle Y_j(x), y\right\rangle^2 + \sum_{Z \in L_N} \langle Z(x), y\rangle^2 \geqslant \alpha |y|^2\,, \tag{2.4}$$

where L_N denotes the set of Lie brackets of $Y_0, \dots, Y_m$ with length in $[\![1, N]\!]$.

Corollary 2.6. *Under Assumptions 2.1, 2.4 and 2.5, let U be a compact measurable subset of E with $\mu(U) > 0$ and $\theta \in (0, h(\mu(U)))$. Set $W(x) = \mathbb{E}_x(e^{\theta T_U})$. Then W is in $\in C^\infty(\mathbb{R}^d) \cap L^1(\mu)$ and solves the Dirichlet boundary problem*

$$LW + \theta W = 0 \quad \text{on } U^c\,, \qquad\qquad W = 1 \quad \text{on } U\,.$$

In particular, in that case, W is a Lyapunov function for L in the sense that

$$LW \leqslant -\theta W + C\,, \qquad C := \sup\{LW(x) + \theta W(x)\ :\ x \in U\}\,.$$

Proof. The proof is based on [7, Theorem 5.14] and similar to the proof of (H2) $\Rightarrow$ (H1) in [11]. $\square$

The strong hypo-ellipticity required by Assumption 2.5, which is already enforced in the reversible case in [11], is a quite restrictive condition, especially if E is not compact (which is the case where Corollary 2.6 is interesting). It should be possible to prove that W is a Lyapunov function for L under weaker conditions, but this question exceeds the scope of the present elementary note.

3. A few examples

The classical benchmark for hypocoercivity is the Langevin (or kinetic Fokker–Planck) diffusion on $\mathbb{R}^d \times \mathbb{R}^d$ whose generator is

$$Lf = v \cdot \nabla_x f - (\nabla U(x) + v) \cdot \nabla_v f + \Delta_v f$$

for some $U \in C^2(\mathbb{R}^d)$. Assume that $\int_{\mathbb{R}^d} e^{-U(x)} \mathrm{d}x < +\infty$ and let μ be the probability measure with Lebesgue density proportional to $e^{-U(x)-|v|^2/2}$. Assume that

$$\liminf_{|x| \to +\infty} \left(|\nabla U(x)|^2 - 2\Delta U(x) \right) > 0$$

and that there exists $c_1 > 0$, $c_2 \in [0,1)$ and $c_3 > 0$ such that, on $\mathbb{R}^d$,

$$\Delta U \leqslant c_1 + \frac{c_2}{2} |\nabla U|^2 , \qquad |\nabla^2 U| \leqslant c_3 \left(1 + |\nabla U|\right) .$$

Then, following the proof of [12, Theorem 10] (more precisely the construction of the bounded operator A), Assumption 2.1 is satisfied, so that Theorem 2.2 and 2.3 hold. As a comparison, the Meyn–Tweedie approach, that also yields the exponential integrability of the hitting times, has been applied to the Langevin diffusion under various conditions on U. For instance, due to the difficulty of constructing a suitable Lyapunov function, seven technical conditions are required in [15, Hypothesis 1.1] on U, that involve a function R whose existence is then checked on various examples. The conditions of [12] are arguably more general and easier to check.

As mentioned in the introduction, the comparison between [1] and [13] for the Bouncy Particle process is similarly at the advantage of the DMS method. Other examples where the DMS method is successfully applied (and thus where Assumption 2.1 holds) can be found in [5] and references within.

Next, consider the strongly self-interacting diffusion studied in [4], which is the process $(X_t)_{t \geqslant 0}$ on the torus $\mathbb{T}^d$ (with $\mathbb{T} = \mathbb{R}/(2\pi\mathbb{Z})$) that solves

$$\mathrm{d}X_t = \mathrm{d}B_t - \int_0^t \nabla_{x_1} V(X_t, X_s) \mathrm{d}s$$

where $(x_1, x_2) \in \mathbb{T}^d \times \mathbb{T}^d \mapsto V(x_1, x_2) \in \mathbb{R}$ is a C^∞ potential and $(B_t)_{t\geqslant 0}$ is a standard d-dimensional Brownian motion. Under the additional assumption that V can be decomposed as

$$V(x_1, x_2) = \sum_{j=1}^{n} a_j e_j(x_1) e_j(x_2)$$

where $n \in \mathbb{N}_*$ and, for all $j \in [\![1, n]\!]$, e_j is an eigenfunction of the Laplacian (with $\langle e_j, e_k \rangle = 0$ if $j \neq k$) and $a_j \in \mathbb{R}$, then the process can be extended to a finite-dimensional Markov process by setting $U_{j,t} = \int_0^t e_j(X_s)\mathrm{d}s$. Indeed, $(X, U_1, \dots, U_j) \in \mathbb{T}^d \times \mathbb{R}^n$ then solves the system of stochastic equations

$$\begin{aligned} \mathrm{d}X_t &= \mathrm{d}B_t - \sum_{j=1}^{n} a_j \nabla e_j(X_t) U_{j,t} \mathrm{d}t \\ \forall\, j \in [\![1, n]\!]\,, \qquad \mathrm{d}U_{j,t} &= e_j(X_t)\mathrm{d}t\,. \end{aligned}$$

The basic example is $V(x_1, x_2) = \cos(x_1 - x_2) = \cos(x_1)\cos(x_2) + \sin(x_1)\sin(x_2)$ when $d = 1$, in which case the system reads

$$\begin{aligned} \mathrm{d}X_t &= \mathrm{d}B_t + \sin(X_t)U_{1,t}\mathrm{d}t - \cos(X_t)U_{2,t}\mathrm{d}t \\ \mathrm{d}U_{1,t} &= \cos(X_t)\mathrm{d}t \\ \mathrm{d}U_{2,t} &= \sin(X_t)\mathrm{d}t\,. \end{aligned}$$

In the general case, provided $a_j > 0$ for all $j \in [\![1, n]\!]$, the process admits a unique invariant probability measure

$$\mu(\mathrm{d}x, \mathrm{d}u_1, \dots, \mathrm{d}u_n) \propto \exp\left(-\frac{1}{2}\sum_{j=1}^{n} a_j |\lambda_j|^2 u_j^2\right) \mathrm{d}x\mathrm{d}u_1 \dots \mathrm{d}u_n$$

where λ_j is the eigenvalue of the Laplacian associated to e_j. It is proven in [4, Section 5] that the DMS method, hence Assumption 2.1, holds. As a comparison, we are aware of unpublished attempts to construct a Lyapunov function for the system $(X, U_1, \dots, U_n)$, that were succesful but intricate in the particular case $V(x_1, x_2) = \cos(x_1 - x_2)$ and couldn't be extended to the general case (which highlight again the fact that constructing a Lyapunov function is a very *ad hoc* task that should be repeated for every new system).

Finally, remark that Assumption 2.5 is not satisfied in the examples above (in particular, for the Langevin and the self-interacting diffusions, Y_0 is not bounded). We now give an example that ensures that the scope of Corollary 2.6 is not empty. Consider the diffusion $(X_t, U_t)_{t\geqslant 0}$ on $E = \mathbb{R} \times \mathbb{T}$ that solves

$$\begin{cases} \mathrm{d}X_t = \cos(U_t)\mathrm{d}t \\ \mathrm{d}U_t = -V'(X_t)\sin(U_t)\mathrm{d}t + \mathrm{d}B_t\,, \end{cases}$$

with $V(x) = x^2/\sqrt{1+x^2}$. This is a non-elliptic, hypoelliptic diffusion on a non-compact space. The associated generator is $L = Y_0 + Y_1^2$ with

$$Y_1 = \partial_u\,, \qquad Y_0 = \cos(u)\partial_x - V'(x)\sin(u)\partial_u\,,$$

which are bounded C^∞ vector fields with all their derivative bounded. Consider $Z_1 = \partial_u$,

$$\begin{aligned} Z_2 = [Y_1, Y_0] &= -\sin(u)\partial_x - V'(x)\cos(u)\partial_u \\ Z_3 = [Y_1, Z_2] &= -\cos(u)\partial_x + V'(x)\sin(u)\partial_u\,. \end{aligned}$$

Then for all $f \in C^\infty(E)$ and all $(x,u) \in E$,

$$(Z_2 f(x,u))^2 \geqslant \frac{1}{2}\sin^2(u)\,(\partial_x f(x,u))^2 - \|V'\|_\infty^2 \cos^2(u)\,(\partial_u f(x,u))^2\,,$$

and similarly for Z_3. Setting $a_1 = \|V'\|_\infty^2 + 1/2$ and $a_2 = a_3 = 1$, we get that

$$\sum_{j=1}^{3} a_j\left(Z_j f(x,u)\right)^2 \geqslant \frac{1}{2}|\nabla f(x,u)|^2$$

for all $f \in C^\infty(E)$ and all $(x,u) \in E$, and thus Assumption 2.5 holds.

Second, let us check that the DMS method applies here, i.e. that the conditions (H_1)–(H_4) of [12] are satisfied. Let μ be the probability measure on E with Lebesgue density proportional to $\exp(-V(x))$. Denote $Q = \partial_u^2$ and $T = Y_0$. Then by integration by part, we see that $Q^* = Q$ and $T^* = -T$, where G^* denotes the dual in $L^2(\mu)$ of an operator G. In particular, for all $f \in C_c^\infty(E)$,

$$\int_E Lf\mathrm{d}\mu = \int_E f(Q^* + T^*)\mathbb{1}\mathrm{d}\mu = \int_E f(Q-T)\mathbb{1}\mathrm{d}\mu = 0\,,$$

which proves μ is invariant for L. The Poincaré inequality on $\mathbb{T}$ states that

$$\int_E (\partial_u f(x,u))^2 e^{-V(x)}\mathrm{d}x\mathrm{d}u \geqslant 4\pi^2 \int_E \left(f(x,u) - \int_{\mathbb{T}} f(x,v)\mathrm{d}v\right)^2 e^{-V(x)}\mathrm{d}x\mathrm{d}u$$

for all $f \in C_c^\infty(E)$ (and thus, by density, for all $f \in D(Q)$), which is exactly the microscopic coercivity Assumption (H_1) of [12]. Denote $\Pi f(x,u) = \int_{\mathbb{T}} f(x,v)\mathrm{d}v$. Then

$$T\Pi f(x,u) = \cos(u)\partial_x \int_{\mathbb{T}} f(x,v)\mathrm{d}v\,,$$

and in particular, since $\int_T \cos(u)\mathrm{d}u = 0$, $T\Pi T = 0$, which is (H_3) of [12]. Moreover, integrating by part again,

$$(T\Pi)^* f(x,u) = \int_{\mathbb{T}} \cos(v)\left(V'(x)f(x,v) - \partial_x f(x,v)\right)\mathrm{d}v\,,$$

so that

$$\begin{aligned}(T\Pi)^*T\Pi f(x,u) &= \int_{\mathbb{T}} \cos(v)^2 \left(V'(x)\partial_x - \partial_x^2\right) \int f(x,w)\mathrm{d}w\mathrm{d}v \\ &= \frac{1}{2}\left(V'(x)\partial_x - \partial_x^2\right)\Pi f(x,u)\,.\end{aligned}$$

As a consequence,

$$\int_E (T\Pi f)^2\,\mathrm{d}\mu = \frac{1}{2}\int_{\mathbb{R}} (\partial_x \Pi f)^2\, e^{-V} \geqslant \frac{1}{2} C_P \int (\Pi f - \mu f)^2\,\mathrm{d}\mu\,,$$

where C_P is the Poincaré constant of the measure e^{-V}. This is the macroscopic coercivity condition (H_2) of [12]. In view of the expression of $(T\Pi)^*T\Pi$, the last condition (H_4) of [12] follows from [12, Lemma 4]. As a consequence, Assumption 2.1 holds.

Finally, let us prove that Assumption 2.4 holds. Denoting $\widetilde{U}_t = U_t + \pi$, remark that

$$\begin{cases} \mathrm{d}X_t = -\cos(\widetilde{U}_t)\mathrm{d}t \\ \mathrm{d}\widetilde{U}_t = V'(X_t)\sin(\widetilde{U}_t)\mathrm{d}t + \mathrm{d}B_t\,. \end{cases}$$

In other words $(X_t, U_t + \pi)_{t\geqslant 0}$ is a Markov process with generator L^* (this is reminiscent of the change of variable $w = -v$ for the Langevin process). Denoting p_t and p_t^* the transition kernels of L and L^*, we thus have, for all $(x,u),(y,w)\in E$,

$$p_t\left((y,w),(x,u)\right) = p_t^*\left((x,u),(y,w)\right) = p_t\left((x,u-\pi),(y,w-\pi)\right)\,.$$

Since μ is invariant by the transformation $(x,u)\mapsto(x,u-\pi)$, the densities $r_t = p_t/\mu$ and $r_t^* = p_t/\mu$ satisfy the same relation. From [6, Theorem 1.5], Assumption 2.5 implies that $p_t((x,u),(y,w)) \leqslant C(1\wedge t)^{-M}$ for some $C, M > 0$ uniformly in $(x,u),(y,w)\in E$. Then we bound

$$\begin{aligned}\|r_t\left((x,u),\cdot\right)\|_2^2 &= \int_E r_t\left((x,u),(y,w)\right) r_t\left((y,w+\pi),(x,u+\pi)\right) e^{-V(y)}\mathrm{d}y\mathrm{d}w \\ &\leqslant \frac{Ce^{V(x)}}{(1\wedge t)^M}\int_E r_t\left((x,u),(y,w)\right) e^{-V(y)}\mathrm{d}y\mathrm{d}w = \frac{Ce^{V(x)}}{(1\wedge t)^M}\,,\end{aligned}$$

which establishes Assumption 2.4.

As a conclusion, Corollary 2.6 holds for the process $(X_t, U_t)_{t\geqslant 0}$. Moreover, Theorem 2.3, Inequality (2.3) and the quantitative bound on $\|r_t\left((x,u),\cdot\right)\|_2$ yield

$$\mathbb{E}_{(x,u)}\left(e^{\theta T_U}\right) \leqslant Ce^{\frac{1}{2}V(x)}$$

for some $C > 0$ for all $x\in E$ all set U with $\mu(U) > 0$ and all $\theta < h\left(\mu(U)\right)$.

References

[1] Christophe Andrieu, Alain Durmus, Nikolas Nüsken, and Julien Roussel. Hypocoercivity of Piecewise Deterministic Markov Process-Monte Carlo. `https://arxiv.org/abs/1808.08592`, 2018.

[2] Dominique Bakry, Patrick Cattiaux, and Arnaud Guillin. Rate of convergence for ergodic continuous Markov processes: Lyapunov versus Poincaré. *J. Funct. Anal.*, 254(3):727–759, 2008.

[3] Dominique Bakry, Ivan Gentil, and Michel Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren der Mathematischen Wissenschaften*. Springer, 2014.

[4] Michel Benaïm and Carl-Erik Gauthier. Self-repelling diffusions on a Riemannian manifold. *Probab. Theory Relat. Fields*, 169(1-2):63–104, 2017.

[5] Jeremiah Birrell and Luc Rey-Bellet. Concentration inequalities and performance guarantees for hypocoercive MCMC samplers. `https://arxiv.org/abs/1907.11973`, 2019.

[6] Patrick Cattiaux. Calcul stochastique et opérateurs dégénérés du second ordre. I. Résolvantes, théorème de Hörmander et applications. *Bull. Sci. Math.*, 114(4):421–462, 1990.

[7] Patrick Cattiaux. Calcul stochastique et opérateurs dégénérés du second ordre. II. Problème de Dirichlet. *Bull. Sci. Math.*, 115(1):81–122, 1991.

[8] Patrick Cattiaux and Arnaud Guillin. Deviation bounds for additive functionals of Markov processes. *ESAIM, Probab. Stat.*, 12:12–29, 2008.

[9] Patrick Cattiaux and Arnaud Guillin. Functional inequalities via Lyapunov conditions. In *Optimal transportation*, volume 413 of *London Mathematical Society Lecture Note Series*, pages 274–287. Cambridge University Press, 2014.

[10] Patrick Cattiaux and Arnaud Guillin. Hitting times, functional inequalities, Lyapunov conditions and uniform ergodicity. *J. Funct. Anal.*, 272(6):2361–2391, 2017.

[11] Patrick Cattiaux, Arnaud Guillin, and Pierre-André Zitt. Poincaré inequalities and hitting times. *Ann. Inst. Henri Poincaré, Probab. Stat.*, 49(1):95–118, 2013.

[12] Jean Dolbeault, Clément Mouhot, and Christian Schmeiser. Hypocoercivity for linear kinetic equations conserving mass. *Trans. Am. Math. Soc.*, 367(6):3807–3828, 2015.

[13] Alain Durmus, Arnaud Guillin, and Pierre Monmarché. Geometric ergodicity of the Bouncy Particle Sampler. *Ann. Appl. Probab.*, 30(5):2069–2098, 2020.

[14] Pierre Monmarché. Hypocoercivité L^2, inégalité de concentration, temps d'atteinte et fonctions de Lyapunov. `https://arxiv.org/abs/1911.01748`, 2019.

[15] Denis Talay. Stochastic Hamiltonian systems: exponential convergence to the invariant measure, and discretization by the implicit Euler scheme. *Markov Process. Relat. Fields*, 8(2):163–198, 2002.

[16] Liming Wu. A deviation inequality for non-reversible Markov processes. *Ann. Inst. Henri Poincaré, Probab. Stat.*, 36(4):435–445, 2000.

[17] Kosaku Yosida. *Functional analysis*. Classics in Mathematics. Springer, 1995. Reprint of the sixth (1980) edition.

Pierre Monmarché
Sorbonne Université Laboratoire
Jacques-Louis Lions
4 place Jussieu
75011 Paris,
France
pierre.monmarche@sorbonne-universite.fr