HAL
open science

Strongly self-interacting processes on the circle

Carl-Erik Gauthier, Pierre Monmarché

To cite this version:

Carl-Erik Gauthier, Pierre Monmarché. Strongly self-interacting processes on the circle. Stochastics: An International Journal of Probability and Stochastic Processes, 2019, pp.1-23. 10.1080/17442508.2019.1594810 . hal-02323671

HAL Id: hal-02323671

https://hal.sorbonne-universite.fr/hal-02323671

Submitted on 21 Oct 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Strongly self-interacting processes on the circle

Carl-Erik Gauthier*and Pierre Monmarché†

January 31, 2019

Abstract

The purpose of this paper is to investigate the long time behaviour for a self-interacting diffusion and a self-interacting velocity jump process. While the diffusion case has already been studied for some particular potential function, the second one, which belongs to the family of piecewise deterministic processes, is new.

Depending on the underlying potential function's shape, we prove either the almost sure convergence or the recurrence for a natural extended process given by a change a variable.

1 Introduction

Our aim is to study the effect of the addition of a self-interaction mechanism to two initially Markovian dynamics. The first one is the classical Fokker-Planck diffusion $X \in \mathbb{R}$ that solves the SDE

$$dX_t = dB_t - V'(X_t)dt, \tag{1}$$

where $(B_t)_{t\geqslant 0}$ is a standard Brownian motion on $\mathbb{R}$. Namely X is the Markov process with generator

$$Lf(x) = \frac{1}{2}f''(x) - V'(x)f'(x).$$

We recall that the generator of a Markov process $(Z_t)_{t\geq 0}$ is formally defined by

$$Lf(z) = (\partial_t)_{|t=0}\, \mathbb{E}\left(f(Z_t) \mid Z_0 = z\right).$$

Equation (1) can be seen as the generalization of the classical Ornstein-Uhlenbeck process, obtained for $V'(x) = \lambda x$, $\lambda > 0$.

*Department of Mathematics, Box 354350, University of Washington, Seattle, WA 98195, United States. Email: carlgaut@uw.edu

†Sorbonne Université, 4 place Jussieu, 75 005 Paris, France. Email: pierre.monmarche@upmc.fr

The second one is the velocity jump process $(X,Y) \in \mathbb{R}\times\{-1,1\}$ which is the piecewise deterministic Markov process (PDMP) introduced in [16] with generator

$$Lf(x,y) \;=\; y\partial_x f(x,y) + \big(\lambda + (yV'(x))_+\big)\,(f(x,-y) - f(x,y))$$

where $\lambda > 0$ is constant and $(\ )_+$ denotes the positive part. A trajectory of the process is defined as follows: starting from an initial state (x,y), the process follows the deterministic flow $(X_t,Y_t) = (x+ty,y)$ up to a random time T with cumulative distribution $\mathbb{P}(T > s) = \exp[-\lambda s + \int_0^s (yV'(x+uy))du]$. At time T, the velocity is reversed, i.e. $Y_T = -y$, while the position is continuous, i.e. $X_T = x+Ty$. By the Markov property, (X_T,Y_T) can then be taken as a new initial state, from which the process again follows free transport up to a new random jump time, etc., and the full trajectory is defined by induction (see [16] and Section 2.2 for details).

In both cases (diffusion or PDMP), if we suppose that the potential V is sufficiently coercive at infinity, X is ergodic and its law converges to the Gibbs measure with density proportional to e^{-V}. Note that when the rate of jump λ goes to infinity and time is correctly accelerated, the velocity jump process (more precisely its first coordinate) converges to the Fokker-Planck diffusion (see [8]).

In both cases we want to replace the potential $V(X_t)$ by a self-interacting potential

$$V_t(X_t) \;=\; \int_0^t W(X_t,X_s)ds$$

where W is a symmetric interaction potential. In other words $V_t(X_t)$ depends both on the current position X_t and the (non-normalized) occupation measure $\int_0^t \delta_{X_s}ds$. This is a strong self-interaction, by contrast with the weak self-interaction such as studied in [2] where the self-interacting potential is a function of X_t and of the normalized occupation measure $\frac{1}{t}\int_0^t \delta_{X_s}ds$.

Self-Interacting processes belong to the family of *path-dependent* processes. The particularity of such processes is their lack of Markov property since the past modifies the environment that drives the particle. New phenomena may arise in their long time behavior, which would be impossible without the path-dependency.

A first example of strong self-interaction is the linear one, that correspond to $W(x,y) = \frac{1}{2}(x-y)^2$. M.Cranston and Y.Le Jan proved in 1995 (see [7]) that the solution of the SDE

$$dX_t = dB_t - \left(\int_0^t (X_t - X_s)ds\right)dt \tag{2}$$

almost surely converges to a Gaussian random variable as t goes to infinity. Later, S.Herrmann and B.Roynette extended this result to a broader class of potentials of the form $W(x,y) = V(x-y)$ with V convex (see [12]). In the case of the circle, the first author obtained the same result (almost sure convergene toward a random variable) for the interaction potential $W(x,y) = -\cos(x-y)$ (see [9]). In all these cases the particle is attracted by its past.

In [1], M.Benaïm and the first author considered the repulsive case, in which the particle is repelled by its past trajectory. More precisely they studied a self-repelling

diffusion on a compact manifold where W can be decomposed as

$$W(x,y) \quad = \quad \sum_{i=1}^{n} a_i e_i(x) e_i(y)$$

with the a_i's being positive numbers and the e_i's being eigenfunctions of the Laplace operator on the manifold. The basic example on the circle would be $W(x,y) = \cos(x-y) = \cos(x)\cos(y) + \sin(x)\sin(y)$. This assumption on the e_i's yields an explicit formula for the invariant measure of the Markov process $\left(X_t, \left(\int_0^t e_i(X_s)ds\right)_{i=1..n}\right)$.

The aim of the present work is to investigate the case where the e_i's are not eigenfunctions of the Laplace operator. On the other hand we restrict the study (in dimension 1) to the case $n = 1$, namely we take a potential of the form

$$W(x,y) = F(x)F(y)$$

with moreover F smooth and 2π-periodic, so that we consider $x \in \mathbb{S}^1 = \mathbb{R}/2\pi\mathbb{Z}$. Following [1], we set

$$U_t \quad = \quad \int_0^t F(X_s)ds, \tag{3}$$

which reduces the study of the non-Markovian process to the study of some Markov process on an extended space. This restriction should be seen as a first step toward the analysis of the more general situation.

As a consequence, in this paper we study the Markov processes (X,U) on $\mathbb{S}^1 \times \mathbb{R}$ and (X,U,Y) on $\mathbb{S}^1 \times \mathbb{R} \times \{-1,1\}$ with respective generators

$$L_1 f(x,u) \quad = \quad \frac{1}{2}\partial_x^2 f(x,u) - uF'(x)\partial_x f(x,u) + F(x)\partial_u f(x,u) \tag{4}$$

and

$$\begin{aligned} &L_2 f(x,u,y) = \\ &\quad y\partial_x f(x,u,y) + F(x)\partial_u f(x,u,y) + \left(\lambda + (yuF'(x))_+\right)\left(f(x,u,-y) - f(x,u,y)\right). \end{aligned} \tag{5}$$

In both cases we call X the position, U the auxiliary variable and, in the case of the velocity jump process, Y the velocity. Remark that (3) would imply that U_0 is always 0, but from now on we consider the general case of the processes with generators (4) and (5) with any initial condition $U_0 \in \mathbb{R}$. The following assumptions are supposed to hold throughout all the paper:

- The function $F : \mathbb{S}^1 \to \mathbb{R}$ is non-constant, $\mathcal{C}^\infty$, changes signs, and $F'(x) = 0$ implies $F(x) \neq 0$. Moreover for all $x \in \mathbb{S}^1$ there exists $k \geq 1$ such that $F^{(k)}(x) \neq 0$. In particular the critical points of F are isolated points.

The assumption that F has no critical point x with $F(x) = 0$ and is nowhere flat is made for simplicity: otherwise, different behaviours may arise and many cases would have to

be distinguished. We concentrate here on the generic case. Throughout this paper, we consider the discrete sets

$$\begin{aligned} M(F,+) &= \{x \in \mathbb{S}^1 \mid x \text{ is a local maximum of } F \text{ and } F(x) > 0\} \\ M(F,-) &= \{x \in \mathbb{S}^1 \mid x \text{ is a local maximum of } F \text{ and } F(x) < 0\} \\ m(F,+) &= \{x \in \mathbb{S}^1 \mid x \text{ is a local minimum of } F \text{ and } F(x) > 0\} \\ m(F,-) &= \{x \in \mathbb{S}^1 \mid x \text{ is a local minimum of } F \text{ and } F(x) < 0\} \end{aligned}$$

and $\mathcal{M} = M(F,-) \cup m(F,+)$. Recall the total variation distance between two probability laws μ and ν is

$$d_{TV}(\mu,\nu) = \inf\{\mathbb{P}(\Xi_1 \neq \Xi_2) \,:\, Law(\Xi_1) = \mu,\ Law(\Xi_2) = \nu\}$$

and a measure μ is said invariant for a Markov process $(Z_t)_{t\geq 0}$ if $\{Law(Z_0) = \mu\}$ implies $\{\forall t \geq 0, Law(Z_t) = \mu\}$. We say that the law of $(Z_t)_{t\geq 0}$ converges exponentially fast to μ in the total variation sense if there exist $C, \rho > 0$, that may depend on the law of Z_0, such that for all $t \geq 0$

$$d_{TV}(Law(Z_t), \mu) \leq Ce^{-\rho t}.$$

Finally, we say that a random variable Z *admits an exponential moment* if there exists $\theta > 0$ such that

$$\mathbb{E}(e^{\theta|Z|}) < \infty.$$

Our main result is the following:

Theorem 1.

1. *If $\mathcal{M} = \emptyset$, then each of the the processes (X,U) with generator (4) and (X,U,Y) with generator (5) admits a unique invariant measure with full support. If the law of U_0 admits an exponential moment then the process converges exponentially fast in the total variation distance sense to this invariant measure.*

2. *If $\mathcal{M} \neq \emptyset$, then, in both cases, the position X_t almost surely converges to a point of $\mathcal{M}$, as t goes to infinity. Any point of $\mathcal{M}$ has a positive probability to be the limit of X.*

Before proceeding to its proof, let us mention why this result may be expected. Suppose that, at some time, $U > 0$. Then, as long as U is large enough, the force $U_t F'(X_t)$ tends to confine X close to the minima of F. If these minima are all negative, while X stays in their neighbourhood, U decreases, up to some point where it becomes negative. From then the effect of the force is reversed, X is attracted by the maxima of F, and the same mechanism comes into play with U and F changed to $-U$ and $-F$. In some sense X and U have then an inhibitory effect one on the other.

On the other hand if X falls in the neighborhood of a positive minimum of F while $U > 0$ (the case of a negative maximum with $U < 0$ being symmetric) then, as long as it stays there, U increases, which make it more and more unlikely for X to escape away from the minimum, so that eventually there is a positive probability that X never leaves and

U goes to infinity. This is reminiscent of the annealing problem (see [18] for the diffusion and [16] for the velocity jump process) where U_t is replaced by a deterministic $(\beta_t)_{t\geq 0}$, called the inverse temperature. It is classical that in this case, if β increases faster than logarithmically then X will eventually stay trapped forever in the cusp of a local minima. Yet, in our present case, as long as X stays close to a positive minimum, U increases linearly in time.

Remarks :

1. The particular form of the interacting potential $W(x,y) = F(x)F(y)$ implies that W is a Mercer Kernel, which means the particle is repulsed by its past (see [1]).

   If furthermore $\int_0^{2\pi} F(y)dy = 0$, it has been shown in [4, Theorem 2.13] that the normalized occupation measure $\frac{1}{t}\int_0^t \delta_{X_s} ds$ converges almost-surely to the uniform distribution on $\mathbb{S}^1$ whether or not $\mathcal{M}$ is empty in the weak self-interaction diffusion case. This is a major difference between strong and weak self-interaction.

   We could also consider the case $W(x,y) = -F(x)F(y)$. Following the proof of Theorem 1, it is not hard to see that in this case X_t almost surely converges as t goes to infinity to a point of $\mathcal{M}' = m(F,-) \cup M(F,+)$ which, as soon as F is not constant and changes signs, is non-empty.

2. If F does not change signs, then, depending on the sign, U_t converges either to ∞ or to $-\infty$ linearly fast. Therefore, Proposition 1 and Proposition 4 imply the almost-sure convergence of X_t respectively either to a local minimum or to a local maximum of F.

We made the choice to write as much as possible notations, results and proofs which are common to both processes, isolating only the few lemmas that deal with the specific technical difficulties of each case. Our arguments are based on bounds for some hitting times of the processes which are established in Section 2. From them we show in Section 3 that, when $\mathcal{M}$ is empty, the time for the processes to return to compact sets is short (i.e. it admits exponential moments). Section 4 is devoted to some uniform bounds of the transition kernel of the processes over compact sets, and Section 5 to the proof of Theorem 1.

Notation: for $s \in \mathbb{R}$, $\lfloor s \rfloor = \max\{k \in \mathbb{Z},\ k \leq s\}$ and $\lceil s \rceil = \min\{k \in \mathbb{Z},\ k \geq s\}$.

2 Hitting times

In this section, for a redaction purpose, we will hide the dependency on U of the evolution of X. More precisely we will consider the (inhomogeneous in time) diffusion

$$dX_t = dB_t - g(t)F'(X_t)dt \tag{6}$$

where g is a Lipschitz-continuous function and, similarly, the inhomogeneous PDMP (X,Y) with generator

$$L_t f(x,y) = y\partial_x f(x,y) + \left(\lambda + (g(t)yF'(x))_+\right)(f(x,-y) - f(x,y)) \tag{7}$$

where the generator of an inhomogeneous Markov process Z is by definition

$$L_t f(z) = (\partial_s)_{|s=0} \mathbb{E}\left(f(Z_{t+s} \mid Z_t = z\right).$$

Note the processes considered in Theorem 1 are particular cases of those defined here.

Let $A = m(F,+) \cup m(F,-)$ be the set of minima of F, and $\delta \leq -\frac{1}{3}\max\{F(x) \,:\, x \in m(F,-)\}$ be positive and small enough so that

- for all $x \in A$, denoting by $I_x^\delta = [z_l, z_r]$ the connected component of $\{F \leq F(x) + 2\delta\}$ containing x, then F decreases on $[z_l, x]$ and increases on $[x, z_r]$.
- there exists $\kappa > 0$ such that for all $x \in A$ and $\eta \in [0, \delta]$,

  $$d(x, B_x^\eta) \geqslant \kappa\sqrt{\eta},$$

  where $B_x^\eta = \{z \in I_x^\delta,\ F(z) = F(x) + \eta\}$.

The existence of κ follows from the fact that $F(x) \leq F(x_0) + \|F''\|_\infty (x - x_0)^2/2$ for all $x \in \mathbb{S}^1$ and $x_0 \in A$. Remark that the definition of δ ensures that for all $x \in m(F,-)$, $\{F \leq F(x) + 2\delta\} \subset \{F \leq -\delta\}$. Finally, for all $\eta \in [0, \delta]$, let

$$B^\eta = \bigcup_{x \in A} B_x^\eta \qquad \text{and} \qquad C^\eta = \left(\bigcup_{x \in A} I_x^\eta\right)^c.$$

In other words C^η is the complementary of a neighbourhood of the minima of F and B^η is a set of intermediary points from A to C^η. These sets (for $\eta = \delta$) are represented in Fig. 1. Note that the choice of δ ensures that if $\mathcal{M} = \emptyset$ then C^δ contains $\{F \geq -\delta\}$.

For $x \in \mathbb{S}^1$ and $D \subset \mathbb{S}^1$ we write

$$\begin{aligned} T_{x \to D} &= \inf\{t \geq 0 \,:\, X_t \in D \mid X_0 = x\} \\ q_{x \to D} &= \mathbb{P}\left(X \text{ reaches } D \text{ before } A \mid X_0 = x\right) \end{aligned}$$

For two real random variables V, W, recall that V is said to be stochastically smaller than W, denoted by $V \overset{sto}{\leq} W$ (or equivalently $W \overset{sto}{\geq} V$), if for all $r \in \mathbb{R}$

$$\mathbb{P}\left(V > r\right) \;\leq\; \mathbb{P}\left(W > r\right).$$

If V and W have same law we write $V \overset{law}{=} W$.

The aim of this section is to prove the following:

Proposition 1. *There exist a constant $K > 1$ and non-negatives random variables S and R with an exponential moment such that, for all $M > 1$ and $\eta \in (0, \delta]$ with $M\eta > 1$, for all Lipschitz function $g \geq M$, if X is defined by (6) or if (X, Y) is defined by the generator (7), the following holds:*

$$\forall x \in B^\eta, \qquad q_{x \to C^\eta} \;\leq\; KMe^{-\eta M} \tag{8}$$

$$\forall x \in A, \qquad T_{x \to B^\eta} \;\overset{sto}{\geq}\; R\eta \tag{9}$$

$$\forall x \in \mathbb{S}^1, \qquad T_{x \to A} \;\overset{sto}{\leq}\; S. \tag{10}$$

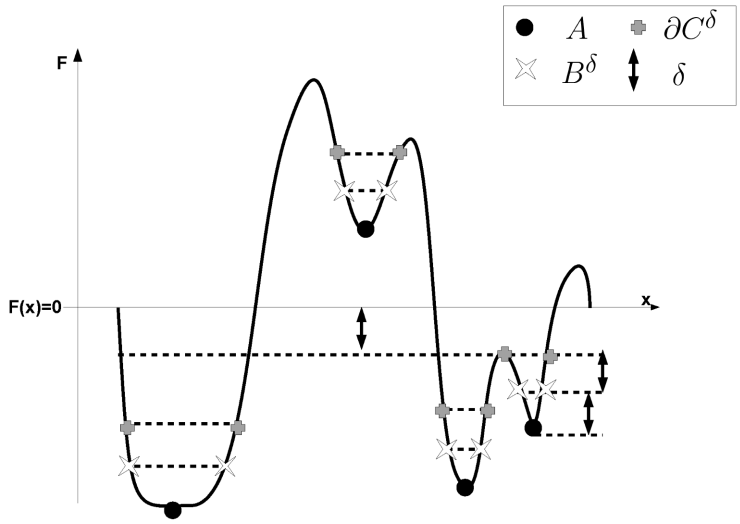

Figure 1: Starting from a minimum in A, the process has to cross an intermediary point of B^δ halfway before reaching C^δ. The energy level difference from A to B^δ, from B^δ to C^δ or (for a negative minimum) from ∂C^δ to $\{F = 0\}$ is always at least δ.

Remark : In the case of the velocity jump process (X, Y), note that these bounds are uniform over the initial velocity Y_0.

The meaning of these bounds is the following. Suppose the auxiliary variable U (whose role here is played by an arbitrary function g) stays for some time above a given level $M > 1$. Then the position X will fall in a local minima of F within a time shorter than S, which does not depend on M (i.e. a high U can only accelerate the hitting time of A). Then to climb back up to an intermediary point of B^η, it takes a time $R\eta$, which is again uniform on $M > 1$. From B^η, the probability to escape from the neighborhood of the minimum in one attempt (namely to reach C^η before having fallen back to A, the bottom) is of order $e^{-\eta M}$, which is a classical metastability result (see [5, 16] for instance) if g is thought as an inverse temperature, since η is the potential barrier to overcome.

The proof of Proposition 1 is split in the two next subsections since the arguments are different for each dynamic. Note that in several proofs we will make assumptions like $x_1 \leq x_2$ where x_1 and x_2 are in $\mathbb{S}^1$, which will make sense since at these times we will only be concerned by the behaviour of the processes on given simply connected intervals of $\mathbb{S}^1$.

2.1 For the diffusion

Proof of Inequality (8) *in the diffusion case.* Let $M > 1$ and $\eta \in [0, \delta]$ with $M\eta > 1$ be fixed and consider the diffusion defined by (6) with $g \geq M$ and $X_0 = x \in B^\eta$. Let $x_0 \in A$ and $x_1 \in \partial C^\eta$ be such that F is monotonous on the interval between x_0 and x_1 that contains x. In particular, $F(x_1) - F(x) = F(x) - F(x_0) = \eta$. Suppose without loss of generality that $x_0 < x < x_1$. Since $g \geq M$, it follows from Ikeda-Watanabe's comparison

result [13, Theorem 1.1, Chapter VI] that

$$q_{x\to C^\eta} \leqslant \mathbb{P}\left(\widetilde{X} \text{ hits } C^\eta \text{ before } A \mid \widetilde{X}_0 = x\right) := \tilde{q}_{x\to C^\eta},$$

where $\widetilde{X}$ solves the SDE

$$d\widetilde{X}_t = dB_t - MF'\left(\widetilde{X}_t\right)dt.$$

Indeed, note that the definitions of $q_{x\to C^\eta}$ and $\tilde{q}_{x\to C^\eta}$ only involve the processes on the interval $[x_0, x_1]$, on which $F' \geq 0$. The scale function of $\widetilde{X}_t)_{t\geqslant 0}$ is defined by

$$\begin{aligned} p(y) &= \int_x^y \exp\left(-2\int_x^z -MF'(s)ds\right)dz \\ &= \int_x^y e^{2M(F(z)-F(x))}dz. \end{aligned}$$

By [14, Proposition 5.22, Chapter 5.5],

$$\tilde{q}_{x\to C^\eta} = \frac{p(x_1)-p(x)}{p(x_1)-p(x_0)} \leq \frac{2\pi e^{2M\eta}}{\int_{x_0}^{x_1} e^{2M(F(z)-F(x))}dz}$$

where we used the local monotony of F. On the other hand,

$$\begin{aligned} \int_{x_0}^{x_1} e^{2M(F(z)-F(x))}dz &\geqslant \frac{1}{2M\|F'\|_\infty}\int_{x_0}^{x_1} 2MF'(z)e^{2M(F(z)-F(x))}dz \\ &= \frac{1}{2M\|F'\|_\infty}\left(e^{4M\eta}-1\right). \end{aligned}$$

Therefore, as $M\eta > 1$,

$$q_{x\to C^\eta} \leqslant 4\pi M\|F'\|_\infty \frac{e^{2M\eta}}{e^{4M\eta}-1} \leqslant 8\pi M\|F'\|_\infty e^{-2M\eta}.$$

□

To prove the two other assertions of Proposition 1, we need the following comparison result:

Lemma 1. *Let x_0 be a local extrema of F and $\varepsilon > 0$ be such that F' is monotonous on $J_\varepsilon := (x_0-\varepsilon, x_0+\varepsilon)$. Consider X the diffusion defined by (6), with $g \geq M > 1$, starting at $X_0 = x \in J_\varepsilon$, and let W be a standard Brownian motion. Denote by*

$$\chi^\varepsilon(x) = \inf\{t > 0 : X_t \notin J_\varepsilon\} \quad \text{and} \quad \iota^\varepsilon = \inf\{t > 0 : |W_t| = \varepsilon\}$$

the respective exit time from J_ε of X and $x_0 + W$. Then:

1. *if x_0 is a local maximum of F, for all $x \in J_\varepsilon$,*

$$\chi^\varepsilon(x) \overset{sto}{\leqslant} \iota^\varepsilon.$$

2. if x_0 is a local minimum of F,

$$\chi^\varepsilon(x_0) \overset{sto}{\geqslant} \iota^\varepsilon.$$

Proof. First, note that by symmetry the exit time from J_ε of $x+W$ has the same law as the exit time of $x+2(x_0-x)+W$, and since the process x_0+W necessarily crosses x or $x+2(x_0-x)$ before leaving J_ε, the exit time of x_0+W is stochastically greater than the one of $x+W$ for any $x \in J_\varepsilon$.

Consider $\Theta_t = (X_t - x_0)^2$, which solves

$$d\Theta_t \;=\; 2\sqrt{\Theta_t}d\widetilde{B}_t + dt - 2g(t)\big((X_t - x_0)F'(X_t)\big)dt,$$

where $\widetilde{B}_t = \int_0^t sign(X_s - x_0)dB_s$ is still a standard Brownian motion. As easily seen, $\left(X_0 - x_0 + \widetilde{B}\right)^2$ is a weak solution of

$$dZ_t \;=\; 2\sqrt{Z_t}d\widetilde{B}_t + dt.$$

Now, we can compare processes $(\Theta_t)_{t\geq 0}$ and $(Z_t)_{t\geq 0}$. When x_0 is a maximum (resp. minimum) of F, $(x-x_0)F'(x)$ is non-positive (resp. non-negative) on J_ε, so that by Ikeda-Watanabe's comparison result ([13, Theorem 1.1, Chapter VI]), $\Theta_t \geqslant Z_t$ (resp. $\Theta_t \leqslant Z_t$) up to the first time where Θ reaches ε^2. As a conclusion, when x_0 is a maximum, Θ reaches ε^2 before Z, and thus in a time stochastically smaller than ι^ε, whereas when x_0 is a minimum, Θ reaches ε^2 after Z and the latter happens at a time with law ι^ε if the starting point is x_0. □

Proof of Inequality (9) *in the diffusion case.* Recall that there exists a constant $\kappa > 0$ such that for all $x \in A$ and $\eta < \delta$, $d(x, B_x^\eta) \geqslant \kappa\sqrt{\eta}$. From Lemma 1 and the Brownian motion's scaling property,

$$T_{x\to B^\eta} \geqslant \chi^{\kappa\sqrt{\eta}}(x) \overset{sto}{\geqslant} \iota^{\kappa\sqrt{\eta}} \overset{law}{=} \eta\iota^\kappa.$$

The fact that ι^κ has an exponential moment is a consequence of [6, Theorem 2]. □

Proof of Inequality (10) *in the diffusion case.* For a given small enough $\varepsilon > 0$, denote by

$$E^\varepsilon = \bigcup_{x\in M(F,+)\cup M(F,-)} (x-\varepsilon, x+\varepsilon)$$

the set of points which are at a distance less than ε from a maximum of F. Let X be the diffusion defined by (6) with $g \geq M$. We apply the following procedure:

1. If, at some time, $X_t \in E^\varepsilon$, wait until it leaves E^ε, which according to the first part of Lemma 1 happens in a time stochastically smaller than ι^ε.

2. If at some time t_0, X leaves E^ε, compare it with $X_{t_0} + B$ where B is the Brownian motion that drives the SDE (6). More precisely by Ikeda-Watanabe's comparison result, $F(X_t) \leqslant F(X_{t_0} + B_t)$ up to the time where either X or $X_{t_0} + B$ reach an extremum of F.

3. Wait until B reaches an extrema of F. If this is a maximum, go back to the first step. If this is a minimum then necessarily, at this time, X has already crossed this minimum, stop the procedure.

Note that, ε being fixed, the probability that $x_0 + B$ reaches a maximum rather than a minimum is bounded above by some $p < 1$ which is uniform over all $x_0 \in \partial E_\varepsilon$. Hence the number of iteration of the procedure is stochastically less than a geometric random variable G with parameter p. Conditionally to whether the Brownian motion reaches a minimum or a maximum in step 3, the law of the duration of the third step is different, but in either cases it is stochastically smaller than $\iota^{2\pi}$. Therefore the total duration of one iteration of the procedure is stochastically smaller than ι^C for some constant $C > 0$, independently from whether this is the last iteration or not. Let $(\iota_k)_{k\geq 0}$ be i.i.d copies of ι^C, independent from G.

We have obtained that for all $x \in \mathbb{S}^1$,

$$T_{x\to A} \overset{sto}{\leqslant} \sum_{k=0}^{G} \iota_k$$

so that

$$\mathbb{E}\left(e^{cT_{x\to A}}\right) \leq \mathbb{E}\left(\left(\mathbb{E}\left(e^{c\iota_0}\right)\right)^G\right)$$

which is finite for c small enough since G admits an exponential moment. □

2.2 For the velocity jump process

This subsection is devoted to the proof of Proposition 1 in the PDMP case, namely for the inhomogeneous Markov process (X, Y) with generator (7). First we construct a trajectory of the process (X, Y) in the following way: consider two independent i.i.d. sequences of standard (with mean 1) exponential random variables $(E_i)_{i\in\mathbb{N}}$ and $(F_i)_{i\in\mathbb{N}}$. Set $T_0 = 0$ and suppose the process has been defined up to some time T_k independently from $(E_i, F_i)_{i\geq k}$. Let

$$\begin{aligned}
\theta_1 &= \inf\left\{t > 0 : \int_0^t g(T_k + s)(Y_{T_k} F'(X_{T_k} + sY_{T_k}))_+ \mathrm{d}s > E_k\right\}, \\
\theta_2 &= \frac{1}{\lambda} F_k,
\end{aligned}$$

and $T_{k+1} = T_k + \theta_1 \wedge \theta_2$, which is the next jump time. If $T_{k+1} = T_k + \theta_1$ we say that the jump is due to the landscape, else we say it is due to the constant rate λ. In either cases, set $X_t = X_{T_k} + (t - T_k)Y_{T_k}$ for all $t \in [T_k, T_{k+1}]$, $Y_t = Y_{T_k}$ for all $t \in [T_k, T_{k+1})$ and $Y_{T_{k+1}} = -Y_{T_k}$. Thus by induction the process is defined up to time T_n for all n. Note that even if, depending on g, the rate of jump may not be bounded, two jumps due to the landscape cannot be arbitrarily close (since at such a jump time, $yF'(x)$ becomes non-positive), so that there cannot be infinitely many jumps in a finite time and $T_n \to \infty$ as $n \to \infty$.

Proof of Inequality (8) *in the PDMP case.* We mainly have to adapt to our inhomogeneous setting the proof of [16, Proposition 4.1]. Without loss of generality, we consider the following configuration: $x_0 \in A$, $x_1 \in B^\eta$ and $x_2 \in \partial C^\eta$ with $x_0 < x_1 < x_2$, and F is increasing on $[x_0, x_2]$.

Let $M \geqslant 1$ and $\mathcal{L}_M$ be the set of Lipschitz functions $g \geqslant M$. For all $x \in [x_1, x_2]$, set

$$\eta_x \;=\; \sup_{g \in \mathcal{L}_M} \mathbb{P}((X,Y) \text{ reaches } (x_2,1) \text{ before } (x,-1) \;\mid (X_0,Y_0) = (x,1)).$$

where the supremum runs over the function g that appears in the generator (7) of the process (X,Y).

Consider a process (X,Y) with generator (7) with some function $g \in \mathcal{L}_M$. For a small $\varepsilon > 0$, suppose that $(X_0,Y_0) = (x-\varepsilon, 1)$.

Then the probability that X goes from $x-\varepsilon$ to x without any jump is less than $1 - \varepsilon\,(MF'(x) + \lambda) + \underset{\varepsilon\to 0}{o}(\varepsilon)$ and the probability it reaches $(x,1)$ before $(x-\varepsilon,-1)$ but with at least one jump is of order ε^2 as $\varepsilon \to 0$ (uniformly over $g \in \mathcal{L}_M$).

If the process has reached $(x,1)$, it has a probability less than η_x to reach $(x_2,1)$ before having fallen back to $(x,-1)$. Nevertheless, if indeed it has fallen back to $(x,-1)$, it has a probability $\varepsilon\lambda + \underset{\varepsilon\to 0}{o}(\varepsilon)$ to jump before reaching $(x-\varepsilon,-1)$, in which case it reaches again $(x,1)$ with probability $1 + \underset{\varepsilon\to 0}{o}(1)$. In this latter case, it reaches $(x_2,1)$ before $(x-\varepsilon,-1)$ with probability less than $\eta_x + \underset{\varepsilon\to 0}{o}(1)$. Thus everything boils down to

$$\begin{aligned}\eta_{x-\varepsilon} &\leqslant (1 - \varepsilon\,(MF'(x)+\lambda))\,\eta_x\,(1+\varepsilon\lambda) + \underset{\varepsilon\to 0}{o}(\varepsilon)\\ &= (1 - \varepsilon MF'(x))\,\eta_x + \underset{\varepsilon\to 0}{o}(\varepsilon).\end{aligned}$$

Together with $\eta_{x_2} = 1$, it yields $\eta_x \leqslant e^{-M(F(x_2)-F(x))}$, and in particular $\eta_{x_1} \leqslant e^{-\eta M}$.

Let

$$r_y \;=\; \sup_{g \in \mathcal{L}_M} \mathbb{P}((X,Y) \text{ reaches } (x_2,1) \text{ before } (x_0,-1) \;\mid (X_0,Y_0) = (x_1,y)).$$

Starting from $(x_1,-1)$ and until the process either jumps or reaches $(x_0,-1)$, we have $YF'(X) < 0$ so that, whatever the function g in (7) is, there cannot be any jump due to the landscape during this time. On the other hand if $\theta_2 > 2\pi$, which happens with probability $e^{-2\lambda\pi}$, there is also no jump due to the constant rate during this time, so that

$$\mathbb{P}\left((X,Y) \text{ reaches } (x_1,1) \text{ before } (x_0,-1) \mid (X_0,Y_0) = (x_1,-1)\right) \;\leqslant\; 1 - e^{-2\lambda\pi}.$$

On the one hand it means that $r_{-1} \leq \left(1 - e^{-2\lambda\pi}\right) r_1$ and on the other hand that

$$r_1 \;\leqslant\; \eta_{x_1} + \left(1 - e^{-2\lambda\pi}\right) r_1$$

and finally that

$$q_{x_1 \to C^\eta} \;\leqslant\; \max(r_1, r_{-1}) \;\leqslant\; e^{2\lambda\pi} e^{-\eta M}.$$

□

Proof of Inequality (9) *in the PDMP case.* Since $|Y| = 1$, the time needed to reach B^η from A is deterministically larger than $d(A, B^\eta) \geqslant \kappa\sqrt{\eta}$. $\square$

Proof of Inequality (10) *in the PDMP case.* Suppose that, at some point in the construction of a trajectory, $\theta_2 > 4\pi$, which happens with probability $e^{-4\lambda\pi}$. If there is also no jump due to the landscape in the meanwhile, X covers the whole circle and in particular reaches A in a time less than 2π. On the other hand if there is a jump due to the landscape before time 2π, the velocity turns to its opposite, and from then and up to the hitting time of A, $YF'(X) < 0$, so that in the meanwhile there cannot be another jump due to the landscape: A is attained in a time less than 4π.

It means that as soon as $\theta_2 > 4\pi$, X reaches A in a time less than 4π, so that starting from any point of $\mathbb{S}^1$, X reaches A in a time stochastically smaller than $4\pi G$ where G is a geometric variable with parameter $e^{-4\lambda\pi}$. $\square$

3 Stability

In this section we consider either $Z = (X, U)$ or $Z = (X, U, Y)$ such as in Theorem 1, and we are interested in the time of return of Z to compact sets. For $M > 1$ we write

$$\tau_M \quad = \quad \inf\{t > 0 \,:\, |U_t| \leqslant M\}$$

and we aim to prove τ_M admits exponential moments when $\mathcal{M} = \emptyset$. For this purpose, we are going to establish that, for some $\kappa > 0$, $V(u) = \exp(\kappa u)$ is a so-called Lyapunov function for both processes, which classically implies the latter.

The notations of Section 2 are kept, in particular the constant δ, and the constant K and the random variables R, S appearing along this section are those given by Proposition 1.

Lemma 2. *Suppose $m(F, +) = \emptyset$. Let $M > 1$ be such that $KMe^{-\delta M} < 1$, and let $(S_i)_{i\in\mathbb{N}}$, $(R_i)_{i\in\mathbb{N}}$ and $(G_i)_{i\in\mathbb{N}}$ be independent i.i.d. sequences where S_0 (resp. R_0) is a copy of S (resp. δR) and G_0 has geometric law with parameter $KMe^{-\delta M}$. For $t \geq 0$ let*

$$N_t \quad = \quad \inf\left\{n \in \mathbb{N} \,:\, \sum_{k=1}^{G_0+\cdots+G_n} R_k \geq t\right\}.$$

Then for all $t > 0$ and for any initial condition Z_0 with $U_0 > M$,

$$\int_0^{t\wedge\tau_M} \mathbb{1}_{\{F(X_s)\geq-\delta\}}\, ds \quad \overset{sto}{\leq} \quad \sum_{k=0}^{N_t} S_k.$$

Proof. While $t \leqslant \tau_M$, the estimates of Proposition 1 hold for X. In particular, independently from its initial condition, the process reaches A in a time stochastically smaller than S_0. Then it takes at least a time R_1 to climb back to B^δ. From there, it reaches C^δ with probability less that $KMe^{-\delta M}$, else it falls back to A. Therefore it remains a time stochastically greater than $\sum_{k=1}^{G_0} R_k$ in $(C^\delta)^c \subset \{F \leqslant -\delta\}$ before reaching C^δ. When this

finally occurs, the process falls again back to A after a time less than S_1 (independently from what occurred before it had reached C). We call this an excursion in C^δ. After n excursions, the process has stayed at least a time $\sum_{k=1}^{G_0+\cdots+G_n} R_k$ in $\{F \leqslant -\delta\}$, which implies in particular that at time t there have been stochastically less than N_t excursions. Thus during a time t, the time spent in C^δ is stochastically less than $\sum_{k=0}^{N_t} S_k$. □

Proposition 2. *If $\mathcal{M} = \emptyset$, there exists $\kappa_0 > 0$ such that, for all $\kappa \in (0, \kappa_0]$, there exists t_0 such that, for all $t \geqslant t_0$, there exists $C_t \geqslant 0$ such that for all initial conditions,*

$$\mathbb{E}\left(e^{\kappa|U_t|}\right) \leqslant \frac{1}{2}e^{\kappa|u_0|} + C_t.$$

Note that, in the sequel, the dependency of C_t with respect to t will not matter.

Proof. Let $\kappa_0 > 0$ be small enough so that $\mathbb{E}\left(e^{\kappa_0(\delta+\max|F|)S}\right) < \infty$ and, for $\kappa \in (0, \kappa_0]$, let t_0 be such that $\mathbb{E}\left(e^{(\kappa(\delta+\max|F|)S}\right) \leq e^{\kappa\delta t_0}/4$. Finally, let $M > 1$ be such that $KMe^{-\delta M} < 1$ and fix $t \geq t_0$.

Let $u_0 > M + t\max|F|$, so that $\tau_M \geqslant t$ almost surely. Hence, Lemma 2 yields

$$\begin{aligned} |U_t| &\leq u_0 - \delta t + (\delta + \max F)\int_0^t \mathbb{1}_{\{F(X_s)\geq -\delta\}}\mathrm{d}s \\ &\overset{sto}{\leq} u_0 - \delta t + (\delta + \max F)\min\left(t, \sum_{k=0}^{N_t} S_k\right). \end{aligned}$$

Hence, distinguishing the cases $N_t = 0$ and $N_t > 0$,

$$\mathbb{E}\left(e^{\kappa(|U_t|-u_0)}\right) \leqslant e^{-\kappa\delta t}\mathbb{E}\left(e^{\kappa(\delta+\max F)S_0}\right) + e^{\kappa t\max F}\mathbb{P}\left(N_t > 0\right).$$

Now, for all $a \in \mathbb{N}^*$,

$$\left\{G_0 \geq a \quad \text{and} \quad \sum_{i=1}^{a} R_i \geq t\right\} \subset \{N_t = 0\},$$

so that, considering the complementary sets,

$$\mathbb{P}\left(N_t > 0\right) \leq \mathbb{P}\left(G_0 < a\right) + \mathbb{P}\left(\sum_{k=1}^{a} R_k < t\right).$$

Applied with $a = \lceil M \rceil$, this reads

$$\mathbb{P}\left(N_t > 0\right) \leq 1 - (1 - KMe^{-\delta M})^M + \mathbb{P}\left(\sum_{k=1}^{\lceil M \rceil} R_k < t\right) \underset{M\to+\infty}{\longrightarrow} 0.$$

In particular, for M large enough, $\mathbb{P}\left(N_t > 0\right) \leq \exp(-\kappa t\max F)/4$. Then we have proved that there exists $M > 0$ such that for all $u_0 > M + t\max|F|$,

$$\mathbb{E}\left(e^{\kappa|U_t|}\right) \leqslant \frac{1}{2}e^{\kappa|u_0|}. \tag{11}$$

The case $u_0 < -M - t\max|F|$ is similar (by changing U and F to their opposites), so that there exists $M > 0$ such that (11) holds for all u_0 with $|u_0| > M + t\max|F|$. Finally, if $|u_0| \leqslant M + t\max|F|$ then

$$\mathbb{E}\left(e^{\kappa|U_t|}\right) \leqslant e^{\kappa(M+2t\max|F|)} := C_t,$$

which concludes. □

In fact, we will also need this similar result, obtained by the same arguments:

Lemma 3. *If $m(F,+) = \emptyset$ then, for M large enough, for all initial condition $u_0 \geq M$, $\inf\{s \geq 0 \ : \ U_s \leq M\}$ is almost surely finite and admits an exponential moment.*

Proof. In the proof of Proposition 2 we have in fact obtained that, under the assumption that $m(F,+) = \emptyset$, there exists $\kappa, t, C_t > 0$ such that for any initial condition $u_0 \geq 0$,

$$\mathbb{E}\left(e^{\kappa U_t}\right) \leq \frac{1}{2}e^{\kappa u_0} + C_t.$$

Denoting $n_0 = \inf\{n \in \mathbb{N},\ \exp(\kappa U_{nt}) \leq 4C_t\}$, then the random sequence $\{(4/3)^{(n\wedge n_0)}\exp(\kappa U_{(n\wedge n_0)t})\}_{n\geq 0}$ is a submartingale. As a consequence, by Fatou's lemma, $\mathbb{E}\left((4/3)^{n_0}\right)$ is finite, which concludes. □

4 Transition kernel bounds

In this section we still consider either $Z = (X, U)$ or $Z = (X, U, Y)$ such as in Theorem 1, and we call E its state space, namely either $\mathbb{S}^1 \times \mathbb{R}_+$ or $\mathbb{S}^1 \times \mathbb{R}_+ \times \{-1, 1\}$. We aim to prove that the following local Doeblin condition holds:

Proposition 3. *Let $\mathcal{K}$ be a compact set of E. There exists $t_0 > 0$ such that, for all $t \geq t_0$, there exist $0 < c < 1$ and a probability measure ν on E such that for all $z \in \mathcal{K}$, for all Borel set D,*

$$\mathbb{P}\left(Z_t \in D \mid Z_0 = z\right) \geq c\nu(D).$$

This condition classically ensures that two processes starting at different states in a compact set $\mathcal{K}$ can be coupled after a time t with some probability $c > 0$. Together with the Lyapunov condition obtained in Proposition 2 which implies that, starting away, the processes reaches $\mathcal{K}$ in a short time, this Doeblin condition ensures exponentially fast mixing for the process (see e.g. [10]).

For the diffusion process, this classically follows from an hypoelliptic argument. By contrast, note that the velocity jump process is not regularizing, in the sense its transition kernel is never absolutely continuous with respect to the Lebesgue measure (at all time there is a positive probability that the process hasn't jumped yet). However the Doeblin condition can still be obtained from some controllability property and a partial regularization.

Since, again, the arguments are different for both processes, we split the proof of Proposition 3 in two paragraphs.

4.1 For the diffusion

In this subsection we consider the process $Z = (X, U)$ induced by the generator (4), namely the solution of the SDE

$$\begin{cases} dX_t &= dB_t - U_t F'(X_t)dt \\ dU_t &= F(X_t)dt. \end{cases} \tag{12}$$

Lemma 4. *For all $z_0 \in \mathbb{S}^1 \times \mathbb{R}$ and $t > 0$, the transition kernel $\mathbb{P}(Z_t \in \cdot \mid Z_0 = z_0)$ admits a smooth density with respect to the Lebesgue measure and its support is $\mathbb{S}^1 \times [u_0 + (\min F)t, u_0 + (\max F)t]$.*

Proof. For $(x, u) \in \mathbb{S}^1 \times \mathbb{R}$, set

$$G_0(x, u) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } G_1(x, u) = \begin{pmatrix} -uF'(x) \\ F(x) \end{pmatrix}.$$

Denoting ∇G_i the Jacobian matrix of G_i for $i = 1, 2$, the Lie-bracket of G_0, G_1 is the vector field $[G_0, G_1]$ defined by

$$[G_0, G_1](x, u) = G_0(x, u)\nabla G_1(x, u) - G_1(x, u)\nabla G_0(x, u)\,,$$

which is here equal to

$$\partial_x G_1(x, u) = \begin{pmatrix} -uF''(x) \\ F'(x) \end{pmatrix}.$$

By induction, replacing F by $F^{(k)}$ for $k \in \mathbb{N}$ in the previous computation, we get

$$[\underbrace{G_0, [G_0, \ldots [G_0}_{k \text{ times}}, G_1] \ldots]](x, u) = \partial_x^{(k)} G_1(x, u) = \begin{pmatrix} -uF^{(k+1)}(x) \\ F^{(k)}(x) \end{pmatrix}.$$

Therefore, by our non-degeneracy assumption on F (namely that for all $x \in \mathbb{S}^1$, $F^k(x) \neq 0$ for some k), the SDE (12) satisfies everywhere the Hörmander condition (see for instance [10]), which gives the first part of the lemma. For the second part, first note that for $z = (x_0, u_0) \in \mathbb{S}^1 \times \mathbb{R}$,

$$\Big((X_t, U_t)\Big)_{t \geqslant 0} \subset \mathbb{S}^1 \times [u_0 + (\min F)t, u_0 + (\max F)t].$$

In order to apply the Stroock-Varadhan support Theorem [19, Theorem 5.2], we are lead to the study of the following deterministic control problem. Denote by $(x_s, u_s)_{s \geqslant 0}$ the solution of the ordinary differential equation

$$\begin{cases} \dot{x} &= v(t) - uF'(x) \\ \dot{u} &= F(x) \end{cases} \tag{13}$$

with initial condition $(x(0), u(0)) = z$ and where $s \mapsto v(s)$ is a piecewise constant function. The proof will be concluded if, given any $z' = (x', u') \in \mathbb{S}^1 \times \Big(u_0 + (\min F)t, u_0 + (\max F)t\Big)$,

we can build a function v such that $(x(t), u(t))$ is arbitrarily close to z'. Let $t_0, t_1 > 0$ be such that $t_0 + t_1 = t$ and $u' - u_0 = t_0(\min F) + t_1(\max F)$. The idea is the following: since we have the choice for v, we can essentially drive $x(t)$ to any position, so we put it first at a minimum of F for a time t_0, then at a maximum of F for a time t_1, and finally we bring it to the end point x'.

More precisely, for any $\varepsilon \in [0, \min(t_0, t_1)/2)$, let $y_0, y_1, y_2 \in \mathbb{R}$ be such that $F(x+y_0) = \min F$, $F(x + y_0 + y_1) = \max F$ and $x + y_1 + y_2 + y_3 = x'$. Set $v(s) = y_0/\varepsilon$ for $t \in [0, \varepsilon]$, $v(s) = 0$ for $t \in (\varepsilon, t_0]$, $v(s) = y_1/\varepsilon$ for $s \in (t_0, t_0 + \varepsilon]$, $v(s) = 0$ for $s \in [t_0 + \varepsilon, t_1 - \varepsilon)$ and finally $v(s) = y_2/\varepsilon$ for $s \in [t_1 - \varepsilon, t_1]$. Let $s \mapsto z_\varepsilon(s) := (x_\varepsilon(s), u_\varepsilon(s))$ be the solution of the associated equation (13) with initial condition z. Remark that for all $s \in [0, t]$ and for all $\varepsilon > 0$, $|u_\varepsilon(s)| \leq |u_0| + t\|F\|_\infty$, and in particular $u_\varepsilon(s)F'(x_\varepsilon(s))$ is bounded uniformly in s and ε. As a consequence, $(x_\varepsilon(\varepsilon), u_\varepsilon(\varepsilon)) \to (x + y_0, u_0)$ as $\varepsilon \to 0$. Since $x + y_0$ is a minimum of F, $s \mapsto z_*(s) := (x + y_0, u_0 + s \min F)$ solves (13) with $v(s) = 0$, so that

$$\sup_{s \in [\varepsilon, t_0]} |z_\varepsilon(s) - z_*(s)| \underset{\varepsilon \to 0}{\longrightarrow} 0\,.$$

Then, similarly, $z_\varepsilon(t_0 + \varepsilon) \to (x + y_0 + y_1, u_0 + t_0 \min F)$ as $\varepsilon \to 0$ and thus

$$\sup_{s \in [t_0 + \varepsilon, t_1 - \varepsilon]} |z_\varepsilon(s) - \tilde{z}_*(s)| \underset{\varepsilon \to 0}{\longrightarrow} 0\,.$$

where $s \mapsto \tilde{z}_*(s) := (x + y_0 + y_1, u_0 + t_0 \min F + (s - t_0) \max F)$ solves again (13) with $v(s) = 0$. Finally $z_\varepsilon(t) \to z'$ as ε vanishes, which concludes.

□

Proof of Proposition 3 in the diffusion case. Denoting by $p_t(\cdot, \cdot)$ the transition density given by Lemma 4, let $z_1, z_2 \in E$ be such that $p_{t_1}(z_1, z_2) > 0$ for some $t_1 > 0$. By continuity, there exist neighbourhood I_1 and I_2 of respectively z_1 and z_2 such that the infimum of p_{t_1} over $I_1 \times I_2$ is $c > 0$.

Let $\mathcal{K}$ be a compact set and let t_0 be large enough so that

$$I_1 \cap \left(\mathbb{S}^1 \times \bigcap_{(x,u) \in \mathcal{K}} [u + (\min F)t_0, u + (\max F)t_0] \right)$$

has a non-empty interior. For $t \geq t_0$, the continuity of p_t and the compactness of $\mathcal{K}$ imply

$$c_t := \inf_{z \in \mathcal{K}} \mathbb{P}(Z_t \in I_1 \mid Z_0 = z) > 0.$$

Let ν be the uniform measure on I_2, namely $\nu(D) = \frac{\lambda(D \cap I_2)}{\lambda(I_2)}$ for any Borel set D of E. Then for all $z \in \mathcal{K}$ and for $t \geq t_0 + t_1$,

$$\begin{aligned} \mathbb{P}\left(Z_t \in D \mid Z_0 = z\right) &\geqslant \mathbb{P}\left(Z_t \in D \mid Z_{t-t_1} \in I_1\right) \mathbb{P}\left(Z_{t_1} \in I_1 \mid Z_0 = z\right) \\ &\geqslant c_{t-t_1} c \lambda(I_2) \nu(D). \end{aligned}$$

□

4.2 For the velocity jump process

In this subsection we consider the process $Z = (X, Y, U)$ with generator (5). The construction of a trajectory is similar to the one exposed in Section 2.2, except from these slight modifications: in the definition of θ_1, $g(T_k + s)$ is replaced by $U_{T_k} + \int_0^s F\left(X_{T_k} + uY_{T_k}\right) du$ and between the two jump times T_k and T_{k+1}, U is defined by $U_t = \int_{T_k}^t F(X_s)ds$.

We start with a controllability result.

Lemma 5. *Let $\mathcal{K}$ and $\mathcal{V}$ respectively be a compact and open set of $\mathbb{S}^1 \times \mathbb{R} \times \{-1, 1\}$. Then there exists $t_0 > 0$ such that, for all $t \geq t_0$,*

$$\inf_{z \in \mathcal{K}} \mathbb{P}\left(Z_t \in \mathcal{V} \mid Z_0 = z\right) \quad > \quad 0.$$

Proof. The boundedness of F implies that for $t > 0$, there exists a compact set $\mathcal{K}_2$ such that for all $s < t$ and for all $z_0 \in \mathcal{K}$, if $Z_0 = z_0$ then $Z_s \in \mathcal{K}_2$. Hence results from [3] apply even if our whole state space is not compact. In particular, the process is Feller, and because $\mathcal{K}$ is compact we only need to prove that, for t large enough,

$$\mathbb{P}\left(Z_t \in \mathcal{V} \mid Z_0 = z\right) > 0$$

for all $z \in \mathcal{K}$. Let $z_0 = (x_0, y_0, u_0) \in \mathcal{K}$ and $z_1 = (x_1, y_1, u_1) \in \mathcal{V}$. We proceed in three steps.

First, suppose that we can choose in a deterministic way a piecewise constant velocity $y(s) \in \{-1, 0, 1\}$ for $s \in [0, t]$, from which $(x(s), u(s))_{s \in [0,t]}$ is defined by an initial condition and by the ODE

$$\begin{pmatrix} \dot{x} \\ \dot{u} \end{pmatrix} = \begin{pmatrix} y \\ F(x) \end{pmatrix}. \tag{14}$$

Let $h_0 < h_1 < h_2 < h_3$ be such that $F(x_0+h_0) = \min F$, $F(x_0+h_1) = \max F$, $F(x_0+h_2) = 0$ and $x_0 + h_3 = x_1$. For $t_0 > h_3$ large enough and any $t \geqslant t_0$, we can build a path of length t between z_0 and z_1 as follows. Given $0 < s_1 < s_2 < t - h_3$, denote

$$\mathcal{I} = [0, h_0) \cup [h_0 + s_1, h_1 + s_1) \cup [h_1 + s_2, h_2 + s_2) \cup [t - (h_3 - h_2), t)$$

set $y(s) = 1$ for $s \in \mathcal{I}$, $y(s) = 0$ for $s \in [0, t) \setminus \mathcal{I}$ and $y(t) = y_1$ and let $s \mapsto (x(s), u(s))$ be the solution of the associated system (14) with initial condition (x_0, u_0). In particular,

$$x(t) = x_0 + h_0 + (h_1 - h_0) + (h_2 - h_1) + (h_3 - h_1) = x_1\,,$$

and

$$u(t) = u_0 + \int_0^{h_3} F(x + s)ds + s_1 \max F + (s_2 - s_1) \min F\,.$$

For t_0 large enough, there exist $s_1 < s_2 < t_0$ such that $u(t) = u_1$. This gives a path from z_0 to z_1 that solves (14) with velocities $y(s) \in \{-1, 0, 1\}$.

In a second instance, we can choose a deterministic $y(s) \in \{-1, 1\}$ such that the solution of the system (14) starting from z_0 is arbitrarily close to z_1 at time t_0. To ensure

this, we simply approximate the case $y(s) = 0$ in the previous step by sufficiently fast and balanced jumps between -1 and 1.

Finally, we consider the PDMP starting from z_0. Since the random jump times have positive density, the PDMP follows arbitrarily closely a trajectory as described in the second step with positive probability. Hence, given any neighbourhood of z_1, the PDMP has positive probability to be in it at time t_0, which concludes.

□

Proof of Proposition 3 in the PDMP case. Consider the following vector fields:

$$G_{-1}(x,u) = \begin{pmatrix} -1 \\ F(x) \end{pmatrix} \text{ and } G_1(x,u) = \begin{pmatrix} 1 \\ F(x) \end{pmatrix}.$$

Then their difference is

$$G_1 - G_{-1} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

so that the Lie bracket $[G_1 - G_{-1}, G_1](x,u)$ is

$$[G_1 - G_{-1}, G_1] = 2\partial_x G_1(x,u) = \begin{pmatrix} 0 \\ 2F'(x) \end{pmatrix}$$

Since F is not constant and smooth, there exists some x such that $F'(x) \neq 0$, at which point the rank of $(G_1 - G_{-1}, [G_1 - G_{-1}, G_1])$ is 2.

According to [3, Theorem 4.4], it implies there exist a non-empty open set $\mathcal{U}$, a probability measure ν and $t_1, c > 0$ such that $\forall z \in \mathcal{U}$,

$$\mathbb{P}\left(Z_{t_1} \in \cdot \mid Z_0 = z\right) \geqslant c\,\nu(\cdot).$$

Considering $t_0 > 0$ given by Lemma 5 with $\mathcal{V} = \mathcal{U}$, we get that for any $z \in \mathcal{K}$, any Borel set D and any $t \geqslant t_0$,

$$\begin{aligned} \mathbb{P}\left(Z_{t+t_1} \in D \mid Z_0 = z\right) &\geq \mathbb{P}\left(Z_t \in \mathcal{U} \mid Z_0 = z\right) \times \inf_{z' \in \mathcal{U}} \mathbb{P}\left(Z_{t+t_1} \in D \mid Z_t = z'\right) \\ &\geq \left(\inf_{z' \in \mathcal{K}} \mathbb{P}\left(Z_t \in \mathcal{U} \mid Z_0 = z'\right)\right) c\,\nu(D) \end{aligned}$$

and Lemma 5 concludes. □

5 Proof of the main theorem

In this section we consider either $Z = (X, U)$ or $Z = (X, U, Y)$ such as in Theorem 1, and we call E the state space, namely either $\mathbb{S}^1 \times \mathbb{R}_+$ or $\mathbb{S}^1 \times \mathbb{R}_+ \times \{-1, 1\}$.

The case $\mathcal{M} = \emptyset$ is a classical consequence of Harris' ergodic theorem:

Proof of point 1 of Theorem 1. Let κ_0 be given by Proposition 2 and $\kappa \leq \kappa_0$ be small enough so that $\mathbb{E}\left(e^{\kappa|U_0|}\right) < \infty$. Let t_0 be large enough for both Propositions 2 and 3 to apply. Let $\mathcal{P}_t$ be the Markov kernel on E defined by $\mathcal{P}_t f(z) = \mathbb{E}\left(f(Z_t) \mid Z_0 = z\right)$. Then

[11, Theorem 1.2], applied to $\mathcal{P}_{t_0}$ with $V(z) = \exp(\kappa|z|)$, implies that $\mathcal{P}_{t_0}$ admits a unique invariant measure μ and that there exists constants $C > 0$ and $\gamma \in (0,1)$ such that for all $n \in \mathbb{N}$,

$$d_{TV}\left(Law(Z_{nt_0}), \mu\right) \;\leq\; C\gamma^n \mathbb{E}\left(e^{\kappa|U_0|}\right).$$

By the semi-group property, for all $t \geq 0$, $\mu\mathcal{P}_t\mathcal{P}_{t_0} = \mu\mathcal{P}_{t_0}\mathcal{P}_t = \mu\mathcal{P}_t$, so that $\mu\mathcal{P}_t$ is invariant for $\mathcal{P}_{t_0}$ and hence, by uniqueness, $\mu\mathcal{P}_t = \mu$. In other words, μ is invariant for $\mathcal{P}_t$ for all $t \geq 0$, and in particular

$$\begin{aligned} d_{TV}\left(Law(Z_t), \mu\right) &= d_{TV}\left(Law(Z_t), \mu\mathcal{P}_{t-\lfloor t/t_0\rfloor t_0}\right) \\ &\leq d_{TV}\left(Law(Z_{\lfloor t/t_0\rfloor t_0}), \mu\right) \\ &\leq C\gamma^{-1}e^{-t|\ln\gamma|/t_0}\mathbb{E}\left(e^{\kappa|U_0|}\right). \end{aligned}$$

Finally, the controllability results of Section 4 (Lemmas 4 and 5) imply that μ has full support. □

The rest of the section is devoted to the proof of the localization of the processes when $\mathcal{M} \neq \emptyset$.

Proposition 4. *Suppose $m(F,+) \neq \emptyset$. Then there exist $p > 0$ and $M > 0$ (which does not depend on Z_0) such that if $X_0 = x_0 \in m(F,+)$ and $U_0 \geq M$, then*

$$\mathbb{P}\left(X_t \underset{t\to\infty}{\longrightarrow} x_0\right) \;\geqslant\; p.$$

Proof. For $j \geqslant 0$, define

$$\eta_j = \frac{4\ln(1+j)}{1+j} \wedge \delta,$$

set $c = \max\{\frac{1}{F(x)},\ x \in m(F,+)\}$ and $S_0 = 0$ and define by induction the following stopping times:

$$\begin{aligned} \tau_{j+1} &= \inf\left\{t > S_j \,:\, X_t \in C^{\eta_{j+1}}\right\}, \\ \tilde{S}_{0,j} &= S_j, \\ \tilde{T}_{k,j} &= \inf\left\{t > \tilde{S}_{k-1,j} \,:\, X_t \in B^{\eta_{j+1}}\right\} \wedge (\tilde{S}_{k-1,j} + c) \wedge \tau_{j+1},\ k \geqslant 1, \\ \tilde{S}_{k,j} &= \inf\left\{t > \tilde{T}_{k,j} \,:\, X_t \in A\right\} \wedge \tau_{j+1},\ k \geqslant 1, \end{aligned}$$

and $S_{j+1} = \tilde{S}_{N_j,j}$ with

$$N_j \;=\; \inf\left\{k \in \mathbb{N} \,:\, \tilde{S}_{k,j} \geqslant S_j + c \text{ or } \tilde{S}_{k,j} = \tau_{j+1}\right\}.$$

Let us give some intuition on these definitions. The connected component of $(C^{\eta_j})^c$ that contains x_0 is a neighbourhood of x_0 whose diameter goes to 0 as j goes to ∞. At time τ_j, the process has escaped from this neighbourhood. For $t \leq \tau_j$, the process makes possibly many oscillations near x_0. When such an oscillation is large enough for the

process to reach B^{η_j} (this is at a time $\widetilde{T}_{k,j}$ for some k), we consider this is the beginning of an attempt to leave $(C^{\eta_j})^c$. If this attempt fails, the process falls back to x_0 (this is $\widetilde{S}_{k,j}$). While X makes those attempts to escape, time goes by, so that U increases: after a time c, U has increased at least by 1. Next time X falls back to x_0 (this is S_{j+1}), we shrink the neighbourhood, namely from then we consider that the process escapes if it reaches $C^{\eta_{j+1}}$. From S_j to S_{j+1}, there have been N_j attempts to leave. The sequence η is scaled so that there is in fact a positive probability that the process never escape from the shrinking neighbourhood that collapses at infinity to $\{x_0\}$.

Let us write these ideas more precisely. Note that as long as $S_{j+1} < \tau_{j+1}$,

$$S_{j+1} - S_j \geqslant c \qquad \text{and} \qquad U_t \geqslant M + j$$

for $t \geqslant S_j$. We take M large enough so that $(M+j)\,\eta_j > 1$ for all $j \in \mathbb{N}$. Therefore, from Proposition 1, for all $k \geq 1$,

$$\mathbb{P}(\tilde{S}_{k,j} = \tau_{j+1} |\, \tilde{T}_{k,j} < \tau_{j+1}) \;\leqslant\; K(j+M)e^{-(j+M)\eta_j}.$$

It implies that $\left(\mathbb{1}_{\tilde{S}_{(i\wedge N_j),j} < \tau_{j+1}} + (i \wedge N_j)K(j+M)e^{-(j+M)\eta_j}\right)_{i \geqslant 0}$ is a submartingale. Thus,

$$\begin{aligned}\mathbb{P}(S_{j+1} < \tau_{j+1} |\, S_j < \tau_j) &= 1 + \mathbb{E}(\mathbb{1}_{S_{j+1} < \tau_{j+1}} - \mathbb{1}_{S_j < \tau_j} |\, S_j < \tau_j) \\ &\geqslant 1 - K(j+M)e^{-(j+M)\eta_{j+1}}\mathbb{E}(N_j |\, S_j < \tau_j). \qquad (15)\end{aligned}$$

From Proposition 1, we have

$$\tilde{S}_{k+1,j} - \tilde{S}_{k,j} \overset{sto}{\geqslant} \eta_j R.$$

Hence, considering a sequence $(R_i)_{i\in\mathbb{N}}$ of i.i.d random variables distributed like R

$$\begin{aligned}N_j &\overset{sto}{\leqslant} \inf\left\{n \geqslant 1 \,:\, \eta_j \sum_{i=1}^n R_i \geqslant c\right\} \\ &\leqslant \left\lceil \frac{2c}{\mathbb{E}(R_1)\eta_j} \right\rceil + \inf\left\{n \geqslant 1 \,:\, \frac{1}{n}\sum_{i=1}^n R_i \geqslant \frac{\mathbb{E}(R_1)}{2}\right\}.\end{aligned}$$

Since R is a positive r.v. with an exponential moment, from Cramer's Theorem (see e.g [17, Chapter 2.4] with the exercise 2.28 in it), there exist $c_1, c_2 > 0$ such that for all $n \geq 0$,

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^n R_i \leqslant \frac{\mathbb{E}(R_1)}{2}\right) \leq c_1 e^{-c_2 n}.$$

Hence, applying the general formula $\mathbb{E}(J) = \sum_{k\in\mathbb{N}} \mathbb{P}(J \geqslant k)$ for a random variable J on $\mathbb{N}$, we get

$$\mathbb{E}(N_j | S_j < \tau_j) \;\leqslant\; \left\lceil \frac{2c}{\mathbb{E}(R_1)\eta_j} \right\rceil + \sum_{n\geqslant 1} \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^n R_i \leqslant \frac{\mathbb{E}(R_1)}{2}\right) \;\leqslant\; \frac{K'}{\eta_j}$$

for some constant K' which does not depend on j, nor M. Thus (15) is now

$$\mathbb{P}(S_{j+1} < \tau_{j+1} | \, S_j < \tau_j) \;\; \geqslant \;\; 1 - \frac{K'K}{\eta_j}(j+M)e^{-(j+M)\eta_{j+1}}.$$

Take M large enough so that the right-hand side is positive for all $j \in \mathbb{N}$. Then by induction

$$\begin{aligned} \mathbb{P}(S_{j+1} < \tau_{j+1}) &= \mathbb{P}(S_{j+1} < \tau_{j+1} \mid S_j < \tau_j)\, \mathbb{P}(S_j < \tau_j) \\ &\geqslant \prod_{i=0}^{j+1} \left(1 - \frac{K'K}{\eta_i}(i+M)e^{-(i+M)\eta_{i+1}}\right). \end{aligned}$$

As $(\{S_j < \tau_j\})_{j \geqslant 1}$ is a decreasing family of events,

$$\begin{aligned} \mathbb{P}(S_j < \tau_j \; \forall j \in \mathbb{N}) &= \lim_{j \to \infty} \mathbb{P}(S_j < \tau_j) \\ &\geqslant \prod_{j=0}^{\infty} \left(1 - \frac{K'K}{\eta_j}(j+M)e^{-(j+M)\eta_{j+1}}\right) \\ &= \exp\left(\sum_{j \geqslant 0} \ln\left(1 - \frac{K'K}{\eta_j}(j+M)e^{-(j+M)\eta_{j+1}}\right)\right). \end{aligned}$$

For j large enough,

$$\frac{K'K}{\eta_j}(j+M)e^{-(j+M)\eta_{j+1}} \;\; \leqslant \;\; \frac{1}{j^2},$$

(where we used that $\eta_j = 4\ln(1+j)/(1+j)$ for j large enough so that the right-hand-side is equivalent to $K'K/(j^2 \ln j)$), and

$$\ln\left(1 - \frac{K'K}{\eta_j}(j+M)e^{-(j+M)\eta_{j+1}}\right) \;\; \geqslant \;\; -\frac{1}{2j^2}$$

which means $\mathbb{P}(S_j < \tau_j \; \forall j \in \mathbb{N}) > p > 0$ where p does not on depend Z_0. Yet,

$$\{S_j < \tau_j \; \forall j \in \mathbb{N}\} = \{\forall j \in \mathbb{N}, \; \forall s \geq S_j, \; X_s \in I_{x_0}^{\eta_j}\}$$

and the S_j's are all a.s. finite, which concludes. □

Remark : The proof even provides an estimation of the speed of convergence. Indeed we can see that $S_{j+1} \overset{sto}{\leqslant} S_j + c + \delta R$, so that the S_j's grow linearly to infinity. From the non-degeneracy assumption on F, there exist $n \in \mathbb{N}$ and $c > 0$ such that the diameter of $I_{x_0}^{\eta_j}$ is less than $c\eta_j^{\frac{1}{n}}$, depending on the first derivative of F at x_0 to be non-zero (if F is a Morse function, $n = 2$). It means when there is convergence, it occurs at least at a speed of order $\left(\frac{\ln t}{t}\right)^{\frac{1}{n}}$.

Proof of point 2 of Theorem 1: First note that by changing U and F to their opposites, Proposition 4 also says that if $M(F,-) \neq \emptyset$ then there exist $p, M > 0$ such that if $U_0 < -M$ and $X_0 \in M(F,-)$ then X_t converges to x_0 with probability at least p.

For $M > 0$ large enough, $\varepsilon > 0$ small enough and $x \in \mathcal{M}$, let

$$\begin{aligned} \mathcal{V}_x^\varepsilon &= \{z' \in E \,:\, |x' - x| < \varepsilon \text{ and } u' \times sign\,(F(x)) > M\} \\ \mathcal{V}^\varepsilon &= \bigcup_{x \in \mathcal{M}} \mathcal{V}_x^\varepsilon. \end{aligned}$$

When ε is fixed, for M large enough, if the process starts in $\mathcal{V}_x^\varepsilon$, from Inequality (8) (which is written for $x \in m(F,+)$ but by symmetry, again, also holds for $x \in M(F,-)$) it has a probability at least $\frac{1}{2}$ to hit $\mathcal{V}_x^0$ before leaving $\mathcal{V}_x^{2\varepsilon}$. Then from Proposition 4, X has a probability at least p to converge to x. By the Markov property, it is then sufficient to prove that the hitting time of $\mathcal{V}^\varepsilon$ is almost surely finite in order to obtain that X will almost surely converge to some point of $\mathcal{M}$.

Denoting by τ_D the first hitting time of a set D and $\mathcal{K} = \{z \in E \,:\, |u| \leqslant M\}$, let us prove that for all $z_0 \in E$

$$\mathbb{P}(\tau_{\mathcal{V}^\varepsilon} \wedge \tau_{\mathcal{K}} < \infty | Z_0 = z_0) = 1.$$

To do so, consider the case $u_0 > M$ (as before, the case $u_0 < -M$ is obtained by symmetry). Consider a smooth 2π-periodic function $\tilde{F}$ such that $\tilde{F}(x) = F(x)$ for all $x \in \{x' \in \mathbb{S}^1 \,:\, |x' - x_0| > \varepsilon \ \forall x_0 \in m(F,+)\}$ and such that all the local minima of $\tilde{F}$ are negative (i.e. $m(\tilde{F},+) = \emptyset$). Let $(\tilde{Z}_t)_{t \geqslant 0}$ be the process constructed like $(Z_t)_{t \geqslant 0}$ but with the function $\tilde{F}$ rather than F. In particular, in the diffusion, we use the same Brownian motion in the SDE (12) for both processes, and in the PDMP case we use the same sequence of i.i.d. exponential variables. That way, $Z_t = \tilde{Z}_t$ up to time $\tau_{\mathcal{V}^\varepsilon} \wedge \tau_{\mathcal{K}}$. By Lemma 3, the hitting time of $\mathcal{K}$ by $\tilde{Z}$, which is greater than $\tau_{\mathcal{V}^\varepsilon} \wedge \tau_{\mathcal{K}}$, is almost surely finite.

On the other hand, by Lemmas 5 (for the PDMP) and 4 (for the diffusion), there exists $t_0 > 0$ such that for all $x \in \mathcal{M}$,

$$\inf_{z \in \mathcal{K}} \mathbb{P}\left(Z_{t_0} \in \mathcal{V}_x^\varepsilon \mid Z_0 = z\right) \;>\; 0.$$

It therefore follows that for any $z \in E$,

$$\mathbb{P}(\tau_{\mathcal{V}^0} < \infty | Z_0 = z) = 1$$

and moreover

$$\mathbb{P}(X_{\tau_{\mathcal{V}^0}} = x) > 0$$

for all $x \in \mathcal{M}$. Proposition 4 concludes. □

Acknowledgement

The first author thanks Olivier Raimond for a discussion that led to this paper. The second author acknowledge financial support from French ANR PIECE. Finally, both authors thank the anonymous referees for their comments, and acknowledge financial support from the Swiss National Science Foundation, Grant 200021_163072. This paper was written when both authors were affiliated to the University of Neuchâtel.

References

[1] Benaïm M. and Gauthier C.E.: *Self-repelling diffusions on a Riemannian manifold*, Probab. Theory Relat. Fields., Volume 169, Issue 1-2, Page 63-104 (2017)

[2] Benaïm M., Ledoux M. and Raimond O., *Self-interacting diffusions*, Probab. Theory Related Fields, volume 122, number 1: 1-41 (2002)

[3] Benaïm M., Le Borgne S., Malrieu F. and Zitt P.-A., *Qualitative properties of certain piecewise deterministic Markov processes*, Annales de l'Institut Henri Poincaré (Probabilités et Statistiques), 51(3): 1040-1075 (2015)

[4] Benaïm M and Raimond O., *Self-interacting diffusions III. Symmetric interactions*, Annals of Probab., volume 33, number 5: 1716-1759 (2005)

[5] Bovier, A., Eckhoff, M., Gayrard, V. and Klein, M.: *Metastability in reversible diffusion processes. I. Sharp asymptotics for capacities and exit times*, J. Eur. Math. Soc., Volume 6, 399–424 (2004)

[6] Ciesielski Z. and Taylor S.J.: *First passage times and sojourn times for Brownian motion in space and the exact Hausdorff measure of the sample path*, Trans. Amer. Math. Soc. 103: 434-450 (1962).

[7] Cranston M. and Le Jan Y.: *Self attracting diffusions : Two case studies*, Math. Ann. 303(1): 87-93 (1995)

[8] Fontbona J., Guerin H. and Malrieu F.: *Long time behavior of telegraph processes under convex potentials*, Stoch. Proc. & Applic. , Vol. 126, no. 10, 3077-3101 (2016)

[9] Gauthier C-E.: *Self attracting diffusions on a sphere and application to a periodic case*, Electronic Communications in Probability, Volume 21, Paper 53, 12pp (2016)

[10] Hairer, M.: *On Malliavin's proof of Hörmander's theorem*, Bull. Sci. Math 135 (6-7): 650-666 (2011)

[11] Hairer, M. and Mattingly, J. C., *Yet another look at Harris' ergodic theorem for Markov chains*, Progr. Probab., Vol. 63, 109–117 (2011)

[12] Herrmann S. and Roynette B.: *Boundedness and convergence of some self-attracting diffusions*, Math. Ann. 325(1): 81-96 (2003)

[13] Ikeda, N., Watanabe, S.: *Stochastic Differential Equations and Diffusion Processes*, North Holland, New-York, 1981

[14] Karatzas, I., Shreve, S.E.: *Brownian motion and stochastic calculus*, Springer-Verlag, New-York, second edition, 1991

[15] Khasminskii, R.: *Stochastic stability of Differential Equations*, Second Edition, Stochastic Modelling and Applied Probability, Springer, volume 66 (2012).

[16] Monmarché, P. *Piecewise deterministic simulated annealing.* ALEA, Lat. Am. J. Probab. Math. Stat. 13 (1), 357-398 (2016)

[17] Rassoul-Agha, F., Seppäläinen, T.: *A Course on Large Deviations with an Introduction to Gibbs Measures*, American Mathematical Society, Graduate Studies in Mathematics, Volume 162 (2015)

[18] Royer G., *A remark on simulated annealing of diffusion processes*, SIAM J. Control Optim., Volume 27, number 6, 1403-1408 (1989)

[19] Stroock, D.W., Varadhan, S.R.S.: *On the support of diffusion processes with applications to the strong maximum principle*, Proc. of Sixth Berkeley Symp. Math. Statist. Prob., Volum 3, 333-359, Univ. California Press, Berkeley, 1972.