Probab. Theory Relat. Fields
DOI 10.1007/s00440-016-0717-1

Self-repelling diffusions on a Riemannian manifold

Michel Benaïm[1] · **Carl-Erik Gauthier**[1]

Received: 7 December 2015 / Revised: 25 April 2016

Abstract Let M be a compact connected oriented Riemannian manifold. The purpose of this paper is to investigate the long time behavior of a degenerate stochastic differential equation on the state space $M \times \mathbb{R}^n$; which is obtained via a natural change of variable from a self-repelling diffusion taking the form

$$dX_t = \sigma dB_t(X_t) - \int_0^t \nabla V_{X_s}(X_t) ds dt, \qquad X_0 = x$$

where $\{B_t\}$ is a Brownian vector field on M, $\sigma > 0$ and $V_x(y) = V(x, y)$ is a diagonal Mercer kernel. We prove that the induced semi-group enjoys the strong Feller property and has a unique invariant probability μ given as the product of the normalized Riemannian measure on M and a Gaussian measure on $\mathbb{R}^n$. We then prove an exponential decay to this invariant probability in $L^2(\mu)$ and in total variation.

Keywords Self-interacting diffusions · Strong Feller property · Degenerate diffusions · Hypocoercivity · Invariant probability measure

Mathematics Subject Classification Primary 58J65 · 60K35 · 60H10 · 60J60; Secondary 37A25 · 37A30

We acknowledge financial support from the Swiss National Science Foundation Grant 200020-149871/1. We thank B. Colbois and H. Donnelly for useful discussions on eigenfunctions of the Laplace operator and P. Monmarché for useful discussion about hypocoercivity.

✉ Carl-Erik Gauthier
carl-erik.gauthier@unine.ch

[1] Institut de Mathématiques, Université de Neuchâtel, Rue Émile Argand 11, 2000 Neuchâtel, Switzerland

Published online: 06 May 2016

1 Introduction

Let M be a smooth (i.e $\mathscr{C}^\infty$) Riemannian manifold, $V : M \times M \to \mathbb{R}$ a smooth function and $w : [0, \infty[\to [0, \infty[$ a continuous function. Adopting the terminology now coined in the literature we define a *Self Interacting Diffusion with potential V and weight function w* to be a continuous time stochastic process $(X_t)_{t\geq 0}$ living on M defined by the stochastic differential equation

$$dX_t = \sigma dB_t(X_t) - \nabla V_t(X_t)dt, \tag{1}$$

where $\sigma > 0$, $\{B_t\}$ is a Brownian vector field on M and

$$V_t(x) = w_t \int_0^t V(X_s, x)ds, \tag{2}$$

The case M compact and $w_t = t^{-1}$ has been thoroughly analyzed in a series of papers by the first named author in collaboration with Raimond [4–6] and Ledoux [4]. In particular, it was shown that long term behavior of the normalized occupation measure $\mu_t = \frac{1}{t}\int_0^t \delta_{X_s} ds$ can be precisely related to the long term behavior of a deterministic semi-flow defined on the space of probability measures over M. Pemantle's survey paper [26] contains a comprehensive discussion of these results among others and further references. Some extensions to noncompact spaces have been considered by Kurtzmann in [23,24] and other weight functions decreasing to zero by Raimond in [30].

When w doesn't converge to zero, say $w_t = 1$, the literature on the subject mainly consists of case studies under the assumption that $M = \mathbb{R}$ (or $\mathbb{R}^d$) and $V(x, y) = v(y - x)$. Self attracting processes, that is $xv'(x) \geq 0$ (or $\langle x, v'(x)\rangle \geq 0$ in $\mathbb{R}^d$), have been considered by Cranston and Le Jan [7], Raimond [29], Herrmann and Roynette [17], Herrmann and Scheutzow [18] and typically converge almost surely. For self repelling processes, that is $xv'(x) \leq 0$, the process tends to be "transient" and strong law of large numbers and rate of escapes have been obtained under various assumptions by Cranston and Mountford [8], Durrett and Rogers [13], Mountford and Tarrès [25]. In [32], Tarrès, Tóth and Valkó consider the situation when v is a sufficiently smooth function having a nonnegative Fourier transform. Under this condition and other technical assumptions, they show that the *environment seen from* X_t, that is the mapping $x \mapsto \int_0^t v'(x + X_t - X_s)ds$, admits an ergodic invariant Gaussian measure.

In this paper we will pursue this line of research and investigate the long term behavior of (1) under the assumptions that:

(i) (**Strong interaction**) $w_t = 1$.
(ii) (**Compactness**) M is smooth, finite dimensional, compact, oriented, connected and without boundary.
(iii) (**Self repulsion**) V is a *Mercer kernel*. That is, $V(x, y) = V(y, x)$ and

$$\int_M \int_M V(x, y) f(x) f(y) dx dy \geqslant 0$$

for all $f \in L^2(dx)$, where dx stands for the Riemannian measure.

By Mercer Theorem, V can be written as

$$V(x, y) = \sum_i a_i e_i(x) e_i(y) \tag{3}$$

where $a_i \geq 0$ and $\{e_i\}$ is an orthonormal (in $L^2(dx)$) family of eigenfunctions of the operator $f \mapsto Vf$, where $Vf(x) = \int V(x, y) f(y) dy$.

Thus, if one interpret the sequence

$$\Psi(x) = (\sqrt{a_i} e_i(x))_i$$

as a *feature vector* representing x in l^2,

$$V(x, y) = \langle \Psi(x), \Psi(y) \rangle_{l^2}$$

can be thought of as a similarity between the feature vectors $\Psi(x)$ and $\Psi(y)$. The process is therefore *self-repelling* in the sense that the drift term $-\nabla V_t(X_t)$ in equation (1) tends to minimize the similarity between the current feature vector $\Psi(X_t)$ and the cumulative feature $\int_0^t \Psi(X_s) ds$.

Here we will focus on the particular situation where

(iii') (**Diagonal decomposition**) The sum in (3) is finite and the $\{e_i\}$ are eigenfunctions of the Laplace operator.

Our motivation for such a restriction is twofold. First, for a suitable choice of n and (a_i), the feature map

$$\Psi : M \mapsto \mathbb{R}^n,$$
$$x \mapsto (\sqrt{a_1} e_1(x), \ldots, \sqrt{a_n} e_n(x))$$

is a quasi-isometric embedding of M in $\mathbb{R}^n$. We refer the reader to the recent paper (Portegies [28]) for a precise statement (Theorem 5.1), and further interesting discussions and references on embedding by eigenfunctions. In particular, for some $\varepsilon > 0$

$$-V(x, y) \leq \frac{1}{2} \|\Psi(x) - \Psi(y)\|^2 \leq (1 + \varepsilon) \frac{d(x, y)^2}{2},$$

where d stands for the Riemannian distance on M. Hence, with this choice of (a_i), the smaller is $V_t(X_t)$ the larger is the cumulative quadratic distance $\int_0^t d^2(X_t, X_s) ds$.

Secondly, under hypothesis (iii)$'$, an invariant probability measure of the process $(X_t, V_t(x))$ can be explicitly computed. It turns out that this will be of fundamental importance for our analysis.

A motivating example: the periodic case

Let $M = \mathbb{S}^1 = \mathbb{R}/2\pi\mathbb{Z}$ denote the unit circle and let $V : M \times M \to \mathbb{R}$ be the map defined by

$$V(x, y) = \cos(y - x) = 1 - \frac{1}{2}d^2(y, x),$$

where $d(y, x) = |e^{iy} - e^{ix}|$.

Noting that $\nabla V_x(y) = -\sin(y - x)$, (1) can be rewritten as

$$dX_t = \sigma dB_t + \int_0^t \sin(X_t)\cos(X_s) - \cos(X_t)\sin(X_s)dsdt. \tag{4}$$

Setting $U_t = \int_0^t \cos(X_s)ds$ and $V_t = \int_0^t \sin(X_s)ds$ we get the following SDE on $\mathbb{S}^1 \times \mathbb{R}^2$:

$$\begin{cases} dX_t &= \sigma dB_t + (\sin(X_t)U_t - \cos(X_t)V_t)dt \\ dU_t &= \cos(X_t)dt. \\ dV_t &= \sin(X_t)dt \end{cases} \tag{5}$$

Some motions are shown in (Figs. 1, 2). This system enjoys the following properties, summarized by the next Theorem, which proof follows from Theorems 5, 6, 7 and Proposition 1. Given $y = (x, u, v) \in \mathbb{S}^1 \times \mathbb{R}^2$, we let $(Y_t^y)_{t\geq 0} = ((X_t^y, U_t^y, V_t^y))_{t\geq 0}$ denote the solution to (5) with initial condition $Y_0^y = y$. Here $\mathbb{S}^1$ is identified with $\mathbb{R}/2\pi\mathbb{Z}$.

Theorem 1 *The Markov process induced by* (5) *is a positive Harris process and admits a unique invariant probability given as*

$$\mu(dxdudv) = \frac{dx}{2\pi} \otimes \frac{\exp(-u^2/2)}{\sqrt{2\pi}}du \otimes \frac{\exp(-v^2/2)}{\sqrt{2\pi}}dv.$$

Furthermore, the law of Y_t^y converges exponentially fast to μ in $L^2(\mu)$ and in total variation.

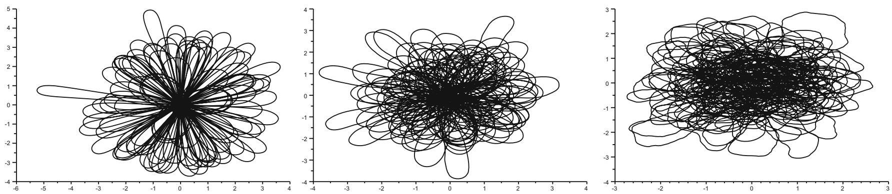

Fig. 1 Evolution of the coordinate (U_t, V_t) after time $T = 750$, where σ is respectively 0.1, 1 and 4

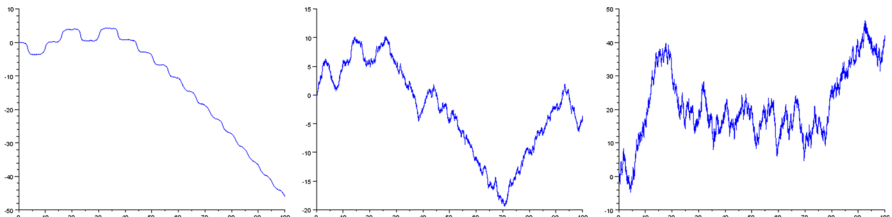

Fig. 2 Evolution of the angle X_t after time $T = 100$, where σ is respectively 0.1, 1 and 4

Remark 1 A similar result holds for the decoupled SDE when $V(x, y) = \sum_{j=1}^{n} a_j \cos(j(y - x))$ and $a_j > 0$ for all $j = 1, \ldots, n$, by setting $U_j(t) = \int_0^t \cos(jX_s)ds$ and $V_j(t) = \int_0^t \sin(jX_s)ds$.

Theorem 2 *Almost surely, the solution of* (4) *with initial condition* $(X_0, U_0, V_0) = (0, 0, 0)$ *does not converge on* $\mathbb{S}^1$ *and a fortiori on* $\mathbb{R}$*. However, on* $\mathbb{R}$,

$$\frac{X_t}{t} \to 0 \quad a.s. \quad as\ t \to \infty.$$

Proof Let $\varepsilon > 0$ and set $R_j^{\varepsilon} = \bigcup_{k\in\mathbb{Z}}((2k + j)\pi - \varepsilon, (2k + j)\pi + \varepsilon) \times \mathbb{R}^2$, $j = 0, 1$. Then by positive Harris recurrence of $(X_t, U_t, V_t)_t$, we have that

$$X_t \in \bigcup_{k\in\mathbb{Z}}((2k + j)\pi - \varepsilon, (2k + j)\pi + \varepsilon),$$

infinitely often for $j = 0, 1$. This proves the first assertion.

Applying now Corollary 1 in Sect. 3 to the function $f(x, u, v) = \sin(x)u - \cos(x)v$ gives us

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t f(X_s, U_s, V_s)ds = \int_{\mathbb{S}^1\times\mathbb{R}^2} f(x, u, v)\mu(dx, du, dv) = 0 \quad \mathbb{P}_{(0,0,0)}a.s.$$

Consequently,

$$\frac{X_t}{t} = \sigma\frac{B_t}{t} + \frac{1}{t}\int_0^t f(X_s, U_s, V_s)ds$$

converges $\mathbb{P}_{(0,0,0)}$ almost surely to 0. □

The zero noise limit

We point out that (5) is -for $\sigma \ll 1$- a random perturbation of the following ordinary differential equation (ODE)

$$\begin{cases} \dot{X}_t = \sin(X_t)U_t - \cos(X_t)V_t \\ \dot{U}_t = \cos(X_t) \\ \dot{V}_t = \sin(X_t) \end{cases} \tag{6}$$

Below (Fig. 3) is a typical motion for (6). The dynamics of (6) can be fully described as follows:

Let $\mathscr{R} : \mathbb{S}^1 \times \mathbb{R}^2 \mapsto \mathbb{S}^1 \times \mathbb{R}^2$ be the map defined by $\mathscr{R}(x, z) = (x, \mathscr{R}_x z)$ where $\mathscr{R}_x$ is the rotation of angle x. Let $H : \mathbb{R}^2 \mapsto [1, \infty]$ be the map defined by

$$H(u, v) = \begin{cases} \frac{1}{2}(u^2 + v^2 - \log(v^2)), \text{ if } v \neq 0, \\ \infty, \text{ if } v = 0. \end{cases}$$

Some level sets are shown in (Fig. 4). Set $H_c = H^{-1}(c)$. Then H_∞ is the line $v = 0$, while for $c < \infty$, H_c has two components H_c^+ and H_c^- obtained from each other by

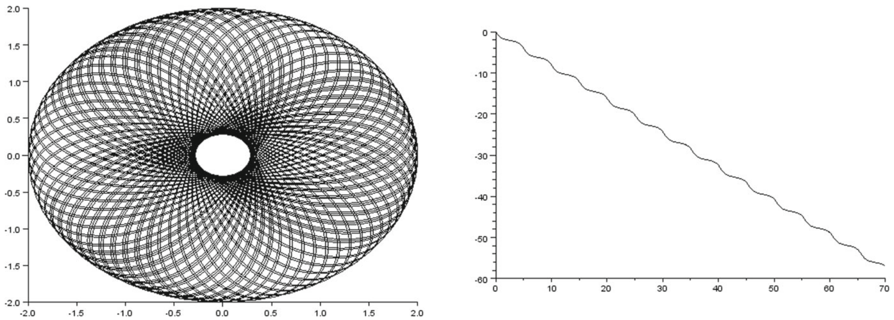

Fig. 3 Evolution of (U_t, V_t) after time $T = 1000$ (*left*) and evolution of X_t until time $T = 70$ (*right*). Both simulations started with initial condition $(x, u, v) = (0, 0, 2)$

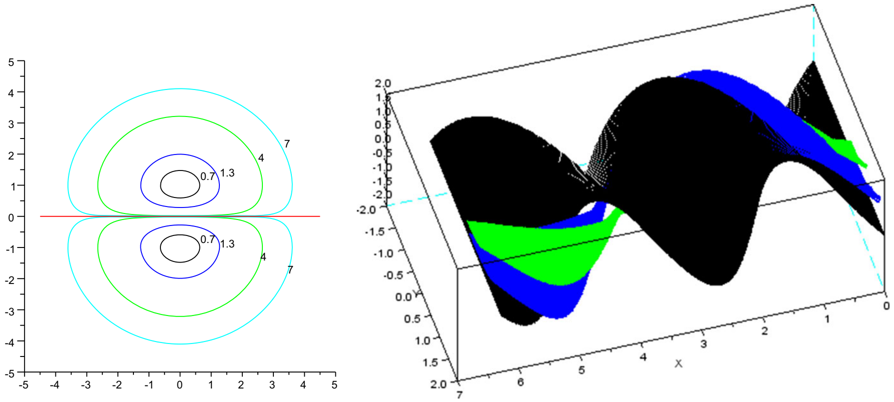

Fig. 4 The *left* picture shows level sets of the function H whereas the *right* picture shows the full twisted strip (in *black*) and two torus T_c^+, with $c = \sqrt{2}$ (in *green*) and $c = 2$ (in *blue*) (color figure online)

reflection along the line $v = 0$. For $c > 1/2$, H_c^+ is a closed curve around $(0, 1)$, and $H_{1/2}^+ = \{(0, 1)\}$.

Given $\alpha \in \{-, +\}$ and $c \in [1, \infty[$ set $\mathbb{T}_c^\alpha = \mathscr{R}(\mathbb{S}^1 \times H_c^\alpha)$ and $\mathbb{T}_\infty = \mathscr{R}(\mathbb{S}^1 \times H_\infty)$. Then $\mathbb{T}_{1/2}^\alpha$ is a closed curve, $\mathbb{T}_c^\alpha$ is, for $c > 1$, a torus and $\mathbb{T}_\infty$ is a full twisted strip. Furthermore

$$\mathbb{S}^1 \times \mathbb{R}^2 = \bigcup_{c \geq 1, \alpha \in \{-,+\}} \mathbb{T}_c^\alpha \cup \mathbb{T}_\infty \tag{7}$$

Theorem 3 *The foliation* (7) *is invariant under the dynamics* (6). *More precisely,*

(i) $\mathbb{T}_{1/2}^\alpha$ *consists of a periodic orbit having period* 2π;
(ii) *For* $c > 1/2$ *the orbits on* $\mathbb{T}_c^\alpha$ *are either all periodic or all dense in* $\mathbb{T}_c^\alpha$. *Furthermore, the set of* c *such that the orbits on* $\mathbb{T}_c^\alpha$ *are periodic is a countable and dense subset of* $]1/2, \infty[$;
(iii) *On* $\mathbb{T}_\infty$ *the solution to* (6) *with initial condition* (x_0, u_0, v_0) *is given by*

$$(X_t, U_t, V_t) = (x_0, u_0 + t\cos(x_0), v_0 + t\sin(x_0)).$$

The proof is the purpose of the appendix.

Remark 2 To determine whether or not the orbits on $\mathbb{T}_c^\alpha$ are periodic, we introduce (see appendix) some function $T_{.,2} :]1/2, \infty[\to]2\sqrt{2}, \sqrt{2}\pi[$ which is continuous and decreasing and prove that the orbits on $\mathbb{T}_c^\alpha$ are periodic if and only if $\frac{T_{c,2}}{2\pi} \in \mathbb{Q}$. Details are given in the appendix.

2 Description of the model

Let us start by fixing some notation. Throughout all the paper, we let ∇ denote the gradient on M, Δ_M the Laplacian on M and for some vector field $\mathscr{X}$ on a manifold $\mathscr{N}$, we denote by $\mathscr{X}(f)$ the Lie derivative of f along $\mathscr{X}$; f being a smooth function.

For a smooth function $V : M \times M \to \mathbb{R}$ and for a Borel measure μ, we let $V\mu : M \to \mathbb{R}$ denotes the function defined by

$$V\mu(x) = \int_M V(u,x)\mu(du).$$

We then consider the model

$$dX_t = \sigma \sum_{j=1}^N F_j(X_t) \circ dB_t^{(j)} - \nabla V\mu_t(X_t)dt, \qquad X_0 = x, \tag{8}$$

where $\sigma > 0$, $(B^{(1)}, \ldots, B^{(N)})$ is a standard Brownian motion on $\mathbb{R}^N$, $\circ$ denotes the Stratonovitch integral, $\{F_i\}$ is a family of smooth vectors fields on M such that

$$\sum_{i=1}^N F_i(F_i f) = \Delta_M f, \quad f \in \mathscr{C}^\infty$$

and μ_t is the random occupation measure defined by

$$\mu_t = \int_0^t \delta_{X_s} ds.$$

Note that there exists at least one such family $\{F_i\}$ since by Nash's embedding Theorem, there exists $N \in \mathbb{N}$ large enough such that M is isometrically embedded in $\mathbb{R}^N$ with the standard metric (see Theorem 3.1.4 in [20] or Proposition 2.5 in [4]).

In this paper, we suppose that the function V has the following form

$$V(x,y) = \sum_{j=1}^n a_j e_j(x) e_j(y), \tag{9}$$

where $(e_j)_{j=1,\dots,n}$ are eigenfunctions for the Laplacian associated to non zero eigenvalues $\lambda_1, \dots, \lambda_n < 0$ such that

$$\int_M e_j(x)e_k(x)dx = \delta_{k,j},$$

where $\delta_{k,j}$ is the Kronecker symbol and dx stands for the Riemannian measure on M. We also assume that $a_j > 0$ for all $j = 1, \dots, n$.

Due to the particular form for V, we can obtain a "true" stochastic differential equation by introducing the new variables $U_{k,t} = \int_0^t e_k(X_s)ds$. Therefore we get the following system on $\mathbb{M} := M \times \mathbb{R}^n$

$$\begin{cases} dX_t = \sigma \sum_{j=1}^N F_j(X_t) \circ dB_t^{(j)} - \sum_{j=1}^n a_j \nabla e_j(X_t)U_{j,t}dt \\ dU_{k,t} = e_k(X_t)dt, \qquad k = 1, \dots, n \end{cases} \tag{10}$$

with initial condition $(x, 0, \dots, 0)$.

In the rest of the paper, we will work with the system (10) and prove that:

1. There exists a unique global strong solution for the system (10);
2. Strong Feller property holds;
3. The system admits a unique invariant measure which is given explicitly as the product of the uniform probability on M and a Gaussian probability on $\mathbb{R}^n$;
4. The law of the solution converges to μ exponentially fast.

The paper is organized as follows. In the next section, we present the main results and the proof of point 1.

In Sect. 4, we provide the proofs of points 2 and 3. To this end, we introduce a property, called *condition* (E') and prove that it implies the Strong Feller property.

In Sect. 5 is given the proof of an exponential decay in $L^2(\mathbb{M}, \mu)$, where μ is the unique invariant probability whereas a proof for an exponential decay in the Total Variation norm is presented in Sect. 6.

3 Presentation of the results

Recall that $\mathbb{M} = M \times \mathbb{R}^n$. Throughout, we denote by $\mathscr{C}_0(\mathbb{M})$ the set of function $f : \mathbb{M} \to \mathbb{R} : (x, u) \mapsto f(x, u)$ which are continuous and such that $f(x, u) \to 0$ when $\|u\| \to \infty$, and by $\mathscr{C}_c^k(\mathbb{M})$ the set of function which are k times continuously differentiable with compact support.

We equip $\mathscr{C}_0(\mathbb{M})$ with the supremum norm

$$\|f\|_\infty := sup_{y\in\mathbb{M}}|f(y)|.$$

Let $G_0, G_1, \ldots, G_N$ be the vector fields on $\mathbb{M}$ defined by

$$G_0(x,u) = \begin{bmatrix} -\sum_{j=1}^n a_j \nabla e_j(x) u_j \\ e_1(x) \\ \vdots \\ e_n(x) \end{bmatrix},$$

and for $j = 1, \ldots, N$,

$$G_j(x,u) = \begin{bmatrix} \sigma F_j(x) \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

with $x \in M$ and $u \in \mathbb{R}^n$. So (10) can be rewritten as:

$$dY_t = \sum_{j=1}^N G_j(Y_t) \circ dB_t^j + G_0(Y_t)dt. \tag{11}$$

Proposition 1 *For all $y = (x,u) \in \mathbb{M}$ there exists a unique global strong solution $(Y_t^y)_{t \geqslant 0}$ to (11) with initial condition $Y_0^y = y = (x,u)$. Moreover, we have*

$$Y_t^y = (X_t^y, U_t^y) \in M \times \bar{B}(u, Kt), \tag{12}$$

where $K = (\max_{y \in M} \sum_{j=1}^n e_j(y)^2)^{1/2}$ and $\bar{B}(u,R) = \{v \in \mathbb{R}^n : \|v-u\| \leqslant R\}$.

Proof Existence and uniqueness is standard since G_0 is locally Lipschitz and sublinear (see for example [31, page 383]). Concerning (12), note that we have

$$\begin{aligned} \sum_{j=1}^n (U_{j,t} - u_j)^2 &\leqslant t \int_0^t \sum_{j=1}^n e_j(X_s)^2 ds \\ &\leqslant t^2 \max_{y \in M} \sum_{j=1}^n e_j(y)^2 < \infty; \end{aligned}$$

which proves (12). □

Throughout, we let $(P_t)_{t \geqslant 0}$ denote the semi-group induced by (11). Recall that for any bounded or nonnegative measurable function $f : \mathbb{M} \to \mathbb{R}$, $P_t f$ is the function defined by

$$P_t f(y) = \mathbb{E}(f(Y_t^y)) \quad \text{for all} \quad y \in \mathbb{M}. \tag{13}$$

Lemma 1 *The semi-group $(P_t)_{t \geqslant 0}$ is Feller, meaning that*

1. *For all $t \geqslant 0$, $P_t(\mathscr{C}_0(\mathbb{M})) \subset \mathscr{C}_0(\mathbb{M})$.*

2. *For all* $f \in \mathscr{C}_0(\mathbb{M})$, $\lim_{t\to 0} \|P_t f - f\|_\infty = 0$.

Proof By Proposition 1, for all $T > 0$, $(Y_t^y)_{t\in[0,T]}$ lies on a deterministic compact set depending only y and T. Hence, by standard results (see eg Theorem IX.2.4 in [31]) , $y \mapsto Y_t^y$ is continuous. Thus, by dominated convergence, $y \mapsto P_t f(y)$ lies in $\mathscr{C}_0(\mathbb{M})$ for all $f \in \mathscr{C}_0(\mathbb{M})$.

In order to prove the second point, it suffices to show that $\lim_{t\downarrow 0} P_t f(y) = f(y)$ (see Proposition III.2.4 in [31]). This follows again from continuity of $t \mapsto Y_t^y$ and dominated convergence. □

The next result gives further informations on the semi-group.

Proposition 2 *The set* $\mathscr{C}_c^2(\mathbb{M})$ *is stable for* P_t, $t \geqslant 0$, *ie for all* $t \geqslant 0$, $P_t(\mathscr{C}_c^2(\mathbb{M})) \subset \mathscr{C}_c^2(\mathbb{M})$.

Proof Let $f \in \mathscr{C}_c^2(\mathbb{M})$. The fact that $P_t f$ has a compact support is a consequence of Eq. (12). Let us now prove that $P_t f$ is twice continuously differentiable.

Let $y = (x_0, u) \in \mathbb{M}$ and $R > 0$. For $\tilde{y} \in M \times B(u, R)$, we have, by Proposition 1,

$$(Y_s^{\tilde{y}})_{0\leqslant s\leqslant t} \subset M \times \bar{B}(u, Kt + R). \tag{14}$$

Pick a smooth function $\psi : \mathbb{R}^n \to \mathbb{R}_+$ which is 1 on the ball $B(u, Kt + R)$, 0 outside the ball $\bar{B}(u, Kt + R + 1)$ and $\psi(v) \leqslant 1$ for all v.

Consider now the SDE defined by

$$d\tilde{Y}_t = \sum_{j=1}^{N} G_j(\tilde{Y}_t) \circ dB_t^j + \tilde{G}_0(\tilde{Y}_t)dt, \tag{15}$$

where $\tilde{G}_0(x, v) = G_0(x, u + \psi(v)(v - u))$. Let us denote by $\tilde{P}_t$ its associated semi-group. The fact that G_0 is smooth and locally Lipschitz implies that $\tilde{G}_0$ is smooth and Lipschitz. By Nash's embedding Theorem and proceeding in the same way as in Proposition 2.5 in [4], we can extend (15) to a SDE on $\mathbb{R}^N \times \mathbb{R}^n$ and f to a function in $\mathscr{C}^2(\mathbb{R}^N \times \mathbb{R}^n)$. Therefore, in view of subsection 3.2.1 in [9] and of Proposition 2.5 in [10], it follows that $\tilde{P}_s f$ is a function of class $\mathscr{C}^2$ for all $s \geqslant 0$. Since

$$P_s f(\tilde{y}) = \tilde{P}_s f(\tilde{y}) \quad \text{for all} \quad 0 \leqslant s \leqslant t \quad \text{and all} \quad \tilde{y} \in M \times B(u, R), \tag{16}$$

it follows that $P_t f$ is of class $\mathscr{C}^2$ on $M \times B(u, R)$.

Consequently, $P_t f \in \mathscr{C}_c^2(\mathbb{M})$. □

The infinitesimal generator of $(P_t)_{t\geqslant 0}$ is the operator

$$\mathscr{L} : D(\mathscr{L}) \to \mathscr{C}_0(\mathbb{M}) : f \mapsto \lim_{t\downarrow 0} \frac{P_t f - f}{t}, \tag{17}$$

where $D(\mathscr{L}) := \{f \in \mathscr{C}_0(\mathbb{M}) : \frac{P_t f - f}{t}$ converges in $\mathscr{C}_0(E)$ when $t \downarrow 0\}$. Then (see for example Theorem 17.6 in [22]) for all $f \in D(\mathscr{L})$,

$$P_t f - f = \int_0^t \mathscr{L}(P_s f)ds = \int_0^t P_s(\mathscr{L} f)ds \tag{18}$$

We briefly recall the following result which characterize the elements of $D(\mathscr{L})$:

Theorem 4 (Propositions VII.1.6 and VII.1.7 in [31]) *For $g, h \in \mathscr{C}_0(\mathbb{M})$, the following assertions are equivalent:*

1. *$h \in D(\mathscr{L})$ and $\mathscr{L}h = g$.*
2. *For all $y \in E$, the process*

$$h(Y_t^y) - \int_0^t g(Y_s^y)ds$$

is a martingale with respect to the filtration $\mathscr{F}_t = \sigma(Y_s^y : 0 \leqslant s \leqslant t)$.

Since the definition of the infinitesimal generator is implicit, it is convenient to introduce a more tractable operator: the *Kolmogorov operator*.

Definition 1 The Kolmogorov operator associated to (10) is the operator defined on $\mathscr{C}^2$ bounded functions having first and second bounded derivatives by

$$L = \frac{\sigma^2}{2}\Delta_M - \sum_{k=1}^n a_k u_k (\nabla e_k(x), \nabla_x \cdot)_{TM} + \sum_{k=1}^n e_k(x)\partial_{u_k},$$

with the convention $(\Delta_M f)(x, u) = (\Delta_M f(., u))(x)$ and $(., .)_{TM}$ stands for the inner product on the tangent bundle of M.

The link between the infinitesimal and the Kolmogorov operator is given by the next proposition.

Proposition 3 *Let f be a $\mathscr{C}^2$ bounded function having first and second bounded derivatives, then $f \in D(\mathscr{L})$ and*

$$Lf = \mathscr{L}f.$$

Proof It follows from Itô's formula and Theorem 4. □

Definition 2 Let $\Phi : \mathbb{R}^n \to \mathbb{R} : u = (u_1, \ldots, u_n) \mapsto \ln(C(\Phi)) + \frac{1}{2}\sum_{k=1}^n a_k|\lambda_k|u_k^2$ with

$$C(\Phi) = \int_{\mathbb{R}^n} \exp\left(-\frac{1}{2}\sum_{k=1}^n a_k|\lambda_k|u_k^2\right)du = \prod_{i=1}^n \sqrt{\frac{2\pi}{|\lambda_i|a_i}} < \infty.$$

Recall that $\lambda_i < 0$ is the eigenvalue associated to the eigenfunction e_i of Δ_M. On $\mathbb{M}$, we define the probability measure

$$\mu(dx \otimes du) = \nu(dx) \otimes e^{-\Phi(u)}du =: \varphi(y)dy, \tag{19}$$

with $y = (x, u)$ and $\nu(dx) = \frac{dx}{\int_M dz}$ is the uniform probability measure on M.

Remark 3 Note that $\mu(dy)$ does not depend on the noise term σ.

We can now state our first main result.

Theorem 5 *Let $(P_t)_{t\geqslant 0}$ be the semi-group associated to the system (10) and $P_t(y_0, dy)$ its transition probability. Then*

(1) *The semi-group $(P_t)_{t\geqslant 0}$ is strongly Feller (meaning that $P_t f$ is a bounded continuous function for whatever bounded measurable function f) and there exists a $\mathscr{C}^\infty((0,\infty),\mathbb{M},\mathbb{M})$ function $p_t(y_0, y)$ such that $P_t(y_0, dy) = p_t(y_0, y)dy$ for all $y_0 \in \mathbb{M}$ and $(L_z^* - \partial_t)p_t(y, z) = 0$,*
(2) *The probability $\mu(dy) = \varphi(y)dy$, where φ is given in Definition 2, is the unique invariant probability. Moreover for all $y \in \mathbb{M}$ and for all bounded measurable function f, we have*

$$\lim_{t\to\infty} P_t f(y) = \int_{\mathbb{M}} f(z)\mu(dz).$$

Furthermore, the process $(Y_t)_t$ is positive Harris recurrent, ie for all Borelian set R such that $\mu(R) > 0$, then

$$\int_0^\infty \mathbf{1}_R(Y_t^y)dt = \infty \ a.s$$

for all $y \in \mathbb{M}$.
(3) $\lim_{t\to\infty}\int_{\mathbb{M}} |p_t(z, y) - \varphi(y)|dy = 0$ *for all $z \in \mathbb{M}$.*

Remark 4 The fact that μ is independent of the parameter σ implies that it is also an invariant probability of the deterministic system obtained with $\sigma = 0$. However, in that case it is not necessarily unique (compare with Theorem 3, where there exists infinitely many compact disjoint invariant sets, thus infinitely many ergodic probabilities.)

As an immediate consequence of the Harris positive recurrence property, we have

Corollary 1 *For all $f \in L^1(\mu)$,*

$$\frac{1}{t}\int_0^t f(Y_s^y)ds \to \int_{\mathbb{M}} f(y)\mu(dy)$$

almost surely for any $y \in \mathbb{M}$.

Proof Apply Theorem 3.1 in [3] to the positive and negative part of f. □

The next results establish exponential rate of convergence of $(P_t)_{t\geqslant 0}$ to μ.

Theorem 6 *For every $\eta > 0$ and $g \in L^2(\mu)$*

$$\left\| P_t g - \int_{\mathbb{M}} g(y)\mu(dy)\right\|_{L^2(\mu)} \leqslant \sqrt{1+2\eta}\|g - \int_{\mathbb{M}} g(y)\mu(dy)\|_{L^2(\mu)} e^{-\lambda t},$$

where

$$\lambda = \frac{\eta}{1+\eta} \frac{K_1 \sigma^2}{1 + K_2 \sigma^2 + K_3 \sigma^4},$$

with

$$K_1 = \frac{1}{4(2 + (1+N_2)^2)} \left(\frac{\Lambda}{1+\Lambda} \right)^2,$$

$$K_2 = \frac{(1+N_2) \sum_{j=1}^n |\lambda_j|}{2 + (1+N_2)^2},$$

$$K_3 = \frac{(\sum_{j=1}^n |\lambda_j|)^2}{4(2 + (1+N_2)^2)},$$

$$\Lambda = \min_{i=1,\ldots,n} |\lambda_i| a_i$$

and

$$N_2 = 2 \frac{n}{\min\{|\lambda_j|, \, j = 1, \ldots, n\}} \sup_{i=1,\ldots,n} \|\nabla e_i\|_\infty^2 \sqrt{4 + \sum_{i=1}^n |\lambda_i| a_i + 4 \| \sum_i e_i^2 \|_\infty}.$$

Remark 5 Note that if $g \in L^2(\mu)$, then it is not clear at first glance that $P_t g$ is meaningful. However it is. In order to prove it, set $h_t(y, z) = p_t(y, z)/\varphi(z)$. Due to the properties of $p_t(y, .)$ and φ for all $t > 0$ and $x \in \mathbb{M}$ (see Theorem 5, Proposition 1 and Definition 2), then $h_t(y, .)$ has compact support. Thus, by the Cauchy–Schwarz inequality, we obtain

$$\mathbb{E}(|g|(Y_t^y)) = \int_{\mathbb{M}} |g|(z) p_t(y, z) dz = \int_{\mathbb{M}} |g|(z) h_t(y, z) \mu(dz) \leqslant \|g\|_{L^2(\mu)} \|h_t(y, .)\|_{L^2(\mu)}. \tag{20}$$

Furthermore, we have $P_t g \in L^2(\mu)$. Indeed by Jensen inequality and invariance of μ, we have $\int_{\mathbb{M}} (P_t g)^2(y) \mu(dy) \leqslant \int_{\mathbb{M}} P_t(g^2)(y) \mu(dy) = \int_{\mathbb{M}} g^2(y) \mu(dy) < \infty$.

Since both $\mu(dy)$ and $P_t(y_0, dy)$ have smooth densities with respect to the Lebesgue measure for all $y_0 \in \mathbb{M}$ and in view of the third point of Theorem 5, we would hope to get a convergence speed for the total variation norm. Once again the answer is positive as shown by the following theorem.

Theorem 7 *For all $z_0 \in \mathbb{M}$ and $t \geqslant 1$,*

$$\|P_t(z_0, dz) - \mu(dz)\|_{TV} \leqslant \sqrt{1 + 2\eta} \|h(1, z_0, z) - 1\|_{L^2(\mu)} e^{-\lambda(t-1)},$$

where $h(1, z_0, z) = \frac{p_1(z_0,z)}{\varphi(z)}$ *and*

$$\lambda = \frac{\eta}{1+\eta} \frac{K_1\sigma^2}{1 + K_2\sigma^2 + K_3\sigma^4} ,$$

μ *is the probability given in Theorem 5 and the constants* $K_j < \infty$, $j = 1, 2, 3$, *are the same as in Theorem 6.*

The proofs of Theorem 6 and 7 are postponed to Sects. 5 and 6.

4 Proof of Theorem 5

We emphasize, from Eq. (11), that the Kolmogorov operator L can be expressed in Hörmander's form (as a sum of squares):

$$L = \frac{1}{2} \sum_{j=1}^{N} G_j^2 + G_0, \tag{21}$$

where $G_j^2(f) = G_j(G_j f)$. The proof mainly relies on classical results by Kanji Ichihara and Hiroshi Kunita in [21] dealing with this type of operator.

Proof of assertion (1): the Strong Feller Property

Throughout, we use the following notation. If $\mathcal{N}$ is a smooth manifold (such as M, $\mathbb{M}$ or $\mathbb{R}^m$), $W : \mathcal{C}^\infty(\mathcal{N}) \to \mathcal{C}^\infty(\mathcal{N})$ a linear map (typically a differential operator) and $f : \mathcal{N} \to \mathbb{R}^n : x \mapsto (f_1(x), \ldots, f_n(x))$ a smooth map, we let $W(f) : \mathcal{N} \to \mathbb{R}^n$ denote the map defined by

$$W(f)(x) = (W(f_1)(x), \ldots, W(f_n)(x)).$$

Given two smooth vector fields A and B on $\mathcal{N}$ recall that the *Lie-bracket* of A and B is the vector field on $\mathcal{N}$ characterized by

$$[A, B](f) = A(B(f)) - B(A(f))$$

for all $f \in \mathcal{C}^\infty(\mathcal{N})$. In case $\mathcal{N} = \mathbb{R}^m$ then for all $x \in \mathbb{R}^m$

$$[A, B](x) = DB(x)A(x) - DA(x)B(x)$$

where $DA(x)$ (resp. $DB(x)$) stands for the derivative of A (resp. B) at x.

Let $\mathcal{G}_0 = \{G_1, \ldots, G_N\}$. Define then recursively $\mathcal{G}_k$, $k \geqslant 1$, by

$$\mathcal{G}_k = \mathcal{G}_{k-1} \cup \{[B, G_j], \quad B \in \mathcal{G}_{k-1} \quad \text{and} \quad j = 0, \ldots, N.\}$$

Let then $\mathcal{G}_\infty = \bigcup_{k \geqslant 0} \mathcal{G}_k$ and for all $(x, u) \in \mathbb{M}$

$$\mathcal{G}_\infty(x, u) = \{V(x, u) : V \in \mathcal{G}_\infty\}.$$

Using the terminology of [21], we say that

Definition 3 The dynamics (11) satisfies the ellipticity condition (E) if for all $(x, u) \in \mathbb{M}$, $\mathscr{G}_\infty(x, u)$ spans $T_{(x,u)}\mathbb{M} = T_x M \times \mathbb{R}^n$.

The next result rephrases Lemma 5.1 (ii) and Theorem 3 (i) and (iii) of [21].

Lemma 2 *If* (10) *satisfies* (E) *then the induced semi-group* (P_t) *is strongly Feller and there exists a* $\mathscr{C}^\infty((0,\infty), \mathbb{M}, \mathbb{M})$ *function* $p_t(y_0, y)$ *such that* $P_t(y_0, dy) = p_t(y_0, y)dy$ *for all* $y_0 \in \mathbb{M}$ *and* $(L_z^* - \partial_t)p_t(y, z) = 0$.

Remark 6 Note that when $\sigma = 0$, the condition (E) is never satisfied since $\mathscr{G}_0$ is reduced to $\{0\}$; hence $\mathscr{G}_\infty = \{0\}$.

Let $\mathscr{A}_0 = \{F_1, \dots, F_N\}$ and for all $k \geq 1$

$$\mathscr{A}_k = \mathscr{A}_{k-1} \cup \{F_j B, \quad B \in \mathscr{A}_{k-1} \quad \text{and} \quad j = 1, \dots, N\}, \tag{22}$$

where $F_j B$ is the operator on $\mathscr{C}^\infty(M)$ defined by $(F_j B)(f) = F_j(B(f))$.
Let then $\mathscr{A}_\infty = \bigcup_{k \geqslant 0} \mathscr{A}_k$ and for all $x \in M$

$$\mathscr{A}_\infty(x) = \{W(e)(x) : W \in \mathscr{A}_\infty\}$$

where $e : M \to \mathbb{R}^n$ is the map defined by $e(x) = (e_1(x), \dots, e_n(x))$. Note that while $\mathscr{G}_\infty$ is a set of vector fields on $\mathbb{M}$, $\mathscr{A}_\infty$ is a set of differential operators of all orders on $\mathscr{C}^\infty(M)$.

Definition 4 We say that the condition (E') is fulfilled if and only if for all $x \in M$, $\mathscr{A}_\infty(x)$ spans $\mathbb{R}^n$.

Lemma 3 *Suppose* $\sigma > 0$. *Then, condition* (E') *implies condition* (E).

The proof relies on the following lemma.

Lemma 4 *Let* $\mathfrak{e} : \mathbb{R}^m \to \mathbb{R}^n$ *be a smooth function and let* $F(x, u) = \begin{bmatrix} A(x) \\ 0 \end{bmatrix}$ *and* $G(x, u) = \begin{bmatrix} B(x, u) \\ \mathfrak{e}(x) \end{bmatrix}$ *be two vector fields on* $\mathbb{R}^{m+n}$, *where* $A : \mathbb{R}^m \to \mathbb{R}^m$ *and* $B : \mathbb{R}^{m+n} \to \mathbb{R}^m$ *are smooth functions. Then*

$$[F, G](x, u) = \begin{bmatrix} [A, B(., u)](x) \\ A(\mathfrak{e})(x) \end{bmatrix},$$

with $B(., u) : \mathbb{R}^m \to \mathbb{R}^m : x \mapsto B(x, u)$

Proof Let $(x, u) \in \mathbb{R}^m \times \mathbb{R}^n$. We then get that

$$DF(x, u) = \begin{bmatrix} DA & 0 \\ 0 & 0 \end{bmatrix}(x, u)$$

and

$$DG(x,u) = \begin{bmatrix} D_x B & D_u B \\ D\mathfrak{e} & 0 \end{bmatrix}(x,u).$$

Hence

$$\begin{aligned} [F,G](x,u) &= DG(x,u)F(x,u) - DF(x,u)G(x,u) \\ &= \begin{bmatrix} D_x B(x,u)A(x) - DA(x)B(x,u) \\ D\mathfrak{e}(x)A(x) \end{bmatrix} \\ &= \begin{bmatrix} [A, B(.,u)](x) \\ A(\mathfrak{e})(x) \end{bmatrix} \end{aligned} \tag{23}$$

as stated. □

Proof of Lemma 3 Let

$$W = \prod_{j=1}^{l} F_{i_j}, \quad (i_1,\ldots,i_l) \in \{1,\ldots,N\}^l. \tag{24}$$

By definition of G_0, and Lemma 4 (used in a local chart) it follows that

$$G_W(x,u) := [G_{i_1}, [\ldots,[G_{i_l}, G_0]\ldots]] = \sigma^l \begin{bmatrix} \divideontimes \\ W(e)(x) \end{bmatrix} \tag{25}$$

Thus, by hypothesis and the definition of G_j for $j = 1,\ldots,N$,

$$\{G_1(x,u),\ldots,G_N(x,u)\} \cup \{G_W(x,u) \,:\, W \in \mathscr{A}_\infty\}$$

spans $T_{(x,u)}\mathbb{M}$. This set being a subset of $\mathscr{G}_\infty(x,u)$, this proves the lemma. □

Lemma 5 *Suppose that $\{e_1,\ldots,e_n\}$ are eigenfunctions associated to the same nonzero eigenvalue of Δ_M. Then condition (E') holds true.*

Proof Let $(U,(x_1,\ldots,x_m))$ be a local chart with U an open set in M. Let $D_1,\ldots,D_m$ be the vector fields defined on U by $D_i(f) = \frac{\partial}{\partial x_i} f$. Define $\mathscr{A}_\infty^D$ like $\mathscr{A}_\infty$ by replacing $F_1,\ldots,F_N$ by $D_1,\ldots,D_m$, and set $\mathscr{A}_\infty^D(x) = \{W(e)(x) \,:\, W \in \mathscr{A}_\infty^D\}$ for all $x \in U$. We claim that $\mathscr{A}_\infty^D(x)$ spans $\mathbb{R}^n$. Suppose to the contrary that there exists some $x^* \in U$ and some vector $t \in \mathbb{R}^n\backslash\{0\}$ such that $\mathscr{A}_\infty^D(x^*) \subset t^\perp$. Let $f(x) = \sum_i t_i e_i(x)$. Then f is an eigenfunction of Δ_M and for all $W \in \mathscr{A}_\infty^D$

$$W(f)(x^*) = W\left(\sum_{i=1}^n t_i e_i\right)(x^*) = \langle W(e)(x^*), t\rangle = 0.$$

In other words, f vanishes to infinite order at x^*. But by a result of Aronzajn (see [2]), every nonzero eigenfunction of the Laplacian on a C^∞ manifold with C^∞ metric, never vanishes to infinite order. This proves the claim.

It remains to show that $\mathscr{A}_\infty(x)$ spans $\mathbb{R}^n$. Since $F_1(x), \ldots, F_N(x)$ span $T_x M$ for all x, there exist smooth real valued maps α_{ij}, $1 \leq i \leq m$, $1 \leq j \leq N$, defined on U such that for all $x \in U$ and $1 \leq j \leq N$

$$D_j(x) = \sum_{k=1}^{N} \alpha_{j,k}(x) F_k(x).$$

Thus

$$D_j(e)(x) = \sum_{i=1}^{N} \alpha_{j,i}(x) F_i(e)(x) \in span(\mathscr{A}_\infty(x)).$$

Now, for all $\psi, \xi \in \mathscr{C}^\infty(M)$ and all $H \in \mathscr{A}_\infty$, we have

$$H(\psi\xi)(x) = \psi(x)H(\xi)(x) + \xi(x)H(\psi)(x).$$

Thus,

$$\begin{aligned} D_i D_j(e)(x) &= \sum_{k=1}^{N} D_i(\alpha_{j,k})(x) F_k(e)(x) + \sum_{k=1}^{N} \alpha_{j,k}(x) D_i F_k(e)(x) \\ &= \sum_{k=1}^{N} D_i(\alpha_{j,k})(x) F_k(e)(x) + \sum_{k,l=1}^{N} \alpha_{j,k}(x)\alpha_{i,l}(x) F_l F_k(e)(x) \in span(\mathscr{A}_\infty(x)) \end{aligned}$$

By recursion, it comes that $\mathscr{A}_\infty^D(x) \subset span(\mathscr{A}_\infty(x))$ and since $\mathscr{A}_\infty^D(x)$ spans $\mathbb{R}^n$, so does $\mathscr{A}_\infty(x)$. □

Lemma 6 *Condition* (E') *holds.*

Proof Let Λ be the set of distinct eigenvalues of $\{e_1, \ldots, e_n\}$. For $\lambda \in \Lambda$ let $\{e_1^\lambda, \ldots, e_{n(\lambda)}^\lambda\} \subset \{e_1, \ldots, e_n\}$ be the set of eigenfunctions having eigenvalue λ and let $e^\lambda = (e_1^\lambda, \ldots, e_{n(\lambda)}^\lambda)$.

Let $x \in M$. By Lemma 5 there exist $W_1^\lambda, \ldots, W_{n(\lambda)}^\lambda \in \mathscr{A}_\infty$ such that the matrix

$$R_\lambda = (W_i^\lambda(e_j^\lambda)(x))_{1 \leqslant i,j \leqslant n(\lambda)} \tag{26}$$

has rank $n(\lambda)$.

Given a polynomial $P(x) = \sum_{j=0}^{k} \alpha_j x^j$, we let

$$P(\Delta_M) = \sum_{j=0}^{k} \alpha_j \Delta_M^j, \tag{27}$$

where Δ_M^j is the operator defined recursively by $\Delta_M^0 f = f$ and $\Delta_M^{j+1} f = \Delta_M^j(\Delta_M f)$ with $f \in \mathscr{C}^2(M)$. Note that for all $1 \leq i \leq n(\lambda)$

$$P(\Delta_M)(e_i^\lambda) = P(\lambda)e_i^\lambda. \tag{28}$$

Now let $P^\lambda(x) = \prod_{\alpha\in\Lambda;\alpha\neq\lambda}(x-\alpha)$. For $\lambda \in \Lambda$ and $i = 1, \ldots, n(\lambda)$, set

$$H_i^\lambda = W_i^\lambda P^\lambda(\Delta_M). \tag{29}$$

Then one has that $H_i^\lambda(e_j^\alpha)(x) = 0$ for $\alpha \neq \lambda$ and $H_i^\lambda(e_j^\lambda)(x) = P^\lambda(\lambda)W_i^\lambda(e_j^\lambda)(x)$. Thus, the matrix

$$H = (H_i^\lambda(e_j^\alpha)(x))_{\lambda\in\Lambda,\quad i=1,\ldots,n(\lambda)}$$

can, after a reordering if necessary, be written as a diagonal block matrix $(P^\lambda(\lambda)R_\lambda(x))_{\lambda\in\Lambda}$.

It is then easy to see that H has rank n. □

This later lemma combined with Lemmas 2 and 3 proves assertion (1).

Proof of assertions (2) and (3). Invariant probability measure and Harris Recurrence

Recall that a probability measure μ is invariant for the semi-group $(P_t)_{t\geqslant 0}$ if

$$\int_{\mathbb{M}} P_t f(y)\mu(dy) = \int_{\mathbb{M}} f(y)\mu(dy)$$

for all $f \in \mathscr{C}_0(\mathbb{M})$.

Existence of an invariant probability measure We will switch between the two notations $y \in \mathbb{M}$ and $(x, u) \in M \times \mathbb{R}^n$ which represent the same point. Setting

$$L^* = \frac{\sigma^2}{2}\Delta_M + \sum_{k=1}^n a_k u_k div_x(\nabla e_k(x).) - \sum_{k=1}^n e_k(x)\partial_{u_k}. \tag{30}$$

we then observe that

$$\begin{aligned}
L^*\varphi(y) &= \sum_{k=1}^n a_k u_k div_x(\nabla e_k(x)\varphi(y)) - \sum_{k=1}^n e_k(x)\partial_{u_k}\varphi(y) \\
&= \sum_{k=1}^n a_k u_k \lambda_k e_k(x)\varphi(y) + \sum_{k=1}^n e_k(x) a_k|\lambda_k|u_k\varphi(y) \\
&= 0.
\end{aligned}$$

By Propositions 2 and 3 together with Theorem 4, we get for $f \in \mathscr{C}_c^2(\mathbb{M})$

$$\int_{\mathbb{M}} (P_t f(y) - f(y))\mu(dy) = \int_0^t \int_{\mathbb{M}} \mathscr{L} P_s f(y)\mu(dy)ds$$
$$= \int_0^t \int_{\mathbb{M}} L P_s f(y)\varphi(dy)ds$$

Noting that for all $g, h \in \mathscr{C}_c^2(\mathbb{M})$

$$\int_{\mathbb{M}} Lg(y)h(y)dy = \int_{\mathbb{M}} g(y)L^* h(y)dy,$$

we obtain

$$\int_{\mathbb{M}} (P_t f(y) - f(y))\mu(dy) = \int_0^t \int_{\mathbb{M}} P_s f(y)L^*\varphi(dy)ds = 0$$

Since $\mathscr{C}_c^2(\mathbb{M})$ is dense in $\mathscr{C}_0(\mathbb{M})$ for $\|.\|_\infty$, it follows that $\mu(dy) = \varphi(y)dy$ is an invariant probability as stated.

Uniqueness of the invariant probability In order to do this, we begin by showing that μ is an ergodic probability; that is, if a subset $A \subset \mathbb{M}$ satisfies $P_t \mathbf{1}_A = \mathbf{1}_A\ \mu - a.s$ for all $t \geqslant 0$, then $\mu(A)$ is either 0 or 1.

Let us denote by f the function $P_t \mathbf{1}_A$. Then $f(y) \in \{0, 1\}$ for μ-almost $y \in \mathbb{M}$ and f is continuous by point 1 of Theorem 5. Since $\mathbb{M}$ is a connected space and μ has full support, it follows that f is either equal to 0 or 1; and therefore μ is ergodic.

Since two distinct ergodic probabilities are mutual singular, the strong Feller property imply that they must have disjoint support. Since μ has the whole space, which is connected, as support, the uniqueness of μ follows. The second part of the statement is Theorem 4.(i) in [21].

The proof that the process is Harris recurrent follows from the proof's lines of Proposition 5.1 in [21]; which also proves the third point.

5 Exponential decay in $L^2(\mu)$

The goal of this section is to prove the exponential decay in the $L^2(\mu)$ norm. The proof heavily relies on the hypocoercitivity method analyzed by M.Grothaus and P.Stilgenbauer in [16] whose roots lie in the series of paper [11], [12] and [15] initiated by Dolbeault, Mouhot and Schmeiser.

We emphasize that in the particular case where $M = \mathbb{S}^d, n = d + 1$ and $(e_j)_{j=1,\ldots,d+1}$ are the eigenfunctions associated to the first non-zero eigenvalue, our model coincides with the one studied in section 3 in [16].

For an operator T on some Hilbert space H, we denote by $D(T)$ its domain and T^* its adjoint.

We begin to recall the **Data (D)** and **Hypotheses (H1)–(H4)** introduced in [16]. For convenience we have chosen to replace certain hypotheses from [16] by slightly stronger ones (see the Remark 8 below) which are sufficient for our purpose.

Definition 5 (*The Data (D)*) Let H be a real Hilbert space and let (P_t) be a strongly continuous semigroup on H with generator $(\mathscr{L}, D(\mathscr{L}))$ and core $D \subset D(\mathscr{L})$. We suppose that

(i) There exist a closed symmetric operator $(S, D(S))$ and a closed antisymmetric operator $(A, D(A))$ such that $D \subset D(S) \cap D(A)$, $A(D) \subset D$ and $\mathscr{L}_{|D} = S_{|D} - A_{|D}$.
(iii) There exists a closed subspace $F \subset D(S)$ such that $S_{|F} = 0$ and $P(D) \subset D$ where P is the orthogonal projection $P : F \oplus F^{\perp} \to F : f + g \mapsto f$ for all $(f, g) \in F \times F^{\perp}$.

By density of $D \subset D(A)$, closedness of A and the fact that $P(D) \subset D \subset D(A)$, AP is closed and densely defined. Hence, by Von Neumann's Theorem, $(AP)^*AP$ is self-adjoint, closed and densely defined. Thus $(I + (AP)^*AP) : D((AP)^*AP) \to H$ is invertible with bounded inverse. Set

$$B_0 = (I + (AP)^*AP)^{-1}(AP)^* \quad \text{on} \quad D((AP)^*AP). \tag{31}$$

In the following we let $(,)_H$ denote the inner product on H and $\|\cdot\|_H$ the associated norm.

Definition 6 (*Hypotheses (H1)–(H4)*)

(H1) $PAP_{|D} = 0$
(H2) (Microscopic coercivity). There exists $\Lambda_1 > 0$ such that for all $f \in D \cap F^{\perp}$,

$$(-Sf, f)_H \geqslant \Lambda_1 \|f\|_H^2.$$

(H3) (Macroscopic coercivity). There exists $\Lambda_2 > 0$ such that for all $f \in D((AP)^*(AP)) \cap F$,

$$\|Af\|_H^2 \geqslant \Lambda_2 \|f\|_H^2. \tag{32}$$

(H4) (Boundedness of auxiliary operators). The operators $(B_0 S, D)$ and $(B_0 A(I - P), D)$ are bounded and there exists constants N_1 and N_2 such that for all $f \in F^{\perp} \cap D$
(H4, a)

$$\|B_0 Sf\|_H \leqslant N_1 \|f\|_H \tag{33}$$

and
(H4, b)

$$\|B_0 Af\|_H \leqslant N_2 \|f\|_H \tag{34}$$

.

If furthermore $(I - PA^2P)(D)$ is dense in H, then conditions $(H3)$ and $(H4, b)$ are implied by the following conditions, as shown by Corollary 2.13 and Proposition 2.15 in [16].

(H3') Equation (32) holds for all $f \in D \cap F$.
(H4') (b) For all $f \in D \cap F$

$$\|A^2 f\|_H \leqslant N_2 \|g\|_H \tag{35}$$

where $g = (I - PA^2P)f$.

Theorem 8 (Theorem 2 in [12], Theorem 1 in [11], Theorem 2.18 in [16]) *Assume that the assumptions of Definitions 5 and 6 hold. Then there exist constants $\kappa_1, \kappa_2 \in (0, \infty)$ explicitly computable such that for all $g \in H$ and $t \geqslant 0$,*

$$\|P_t g\|_H \leqslant \kappa_1 e^{-\kappa_2 t} \|g\|_H \tag{36}$$

Remark 7 Following the proof's line of section 3.4 in [11] and the beginning of the proof of Theorem 2.18 in [16], one obtains

$$\kappa_1 = \sqrt{\frac{1+\varepsilon_\eta}{1-\varepsilon_\eta}} \leqslant \sqrt{1+2\eta} \quad \text{and} \quad \kappa_2 = \varepsilon_\eta \frac{\Lambda_2}{4(1+\Lambda_2)}, \tag{37}$$

with

$$\varepsilon_\eta = \frac{\eta}{1+\eta} \frac{\varepsilon_0}{\max(1, \varepsilon_0)}, \quad \eta > 0 \tag{38}$$

and

$$\varepsilon_0 = \frac{2\Lambda_2\Lambda_1}{(1+\Lambda_2)(2+(1+N_1+N_2)^2)}. \tag{39}$$

Remark 8 In case (P_t) is a Markov semigroup with invariant probability μ, inducing a strongly continuous semigroup on $L^2(\mu)$, a natural choice for H is

$$L_0^2(\mu) = \left\{ f \in L^2(\mu) : \int f d\mu = 0 \right\}.$$

This choice will be adopted later. In this case, conditions $(D6)$ and $(D7)$ from [16] are automatically satisfied and Theorem 8 implies that for all $f \in L^2(\mu)$

$$\left\| P_t f - \int f d\mu \right\|_{L^2(\mu)} \leqslant \kappa_1 e^{-\kappa_2 t} \|f - \int f d\mu\|_{L^2(\mu)}.$$

5.1 Application to the proof of Theorem 6

Throughout we let

$$H = L_0^2(\mu) := \left\{ f \in L^2(\mathbb{M}, \mu) \ : \int_{\mathbb{M}} f(y)\mu(dy) = 0 \right\}$$

and

$$L_0^2(e^{-\Phi}) = \left\{ f \in L^2(\mathbb{R}^n, e^{-\Phi}) \ : \int_{\mathbb{R}^n} f(u) e^{-\Phi(u)} du = 0 \right\}$$

where μ and Φ are like in Definition 2. Both H and $L^2_0(e^{-\Phi})$ are equipped with the associated L^2 inner product and norm.

The map $\imath : L^2_0(e^{-\Phi}) \hookrightarrow H$ defined by $\imath(g)(x,u) = g(u)$ injects isometrically $L^2_0(e^{-\Phi})$ into H. We let

$$F = \imath(L^2_0(e^{-\Phi}))$$

and $P : F \oplus F^{\perp} \to F$ denote the orthogonal projection onto F. Alternatively P can be defined as

$$(Pf)(x,u) = \int_M f(x,u)\nu(dx). \tag{40}$$

Using the notation introduced in Sect. 3 we let (P_t) denote the semigroup defined by

$$P_t f(y) = \mathbb{E}(f(Y^y_t))$$

for every bounded Borel map $f : \mathbb{M} \to \mathbb{R}$; where (Y^y_t) stands for the solution to (11) with initial condition $Y^y_0 = y$.

Lemma 7 *(P_t) induces a strongly continuous contraction semigroup on H.*

Proof By invariance of μ and Jensen inequality P_t defines a bounded operator on H with norm less than 1 (as already proved in Remark 5).

Let $\varepsilon > 0$ and $f \in L^2(\mu)$. By density of $\mathscr{C}_0(\mathbb{M})$ in $L^2(\mu)$, there exists $g \in \mathscr{C}_0(\mathbb{M})$ such that $\|f - g\|_{L^2(\mu)} < \varepsilon$. Thus, by the contraction property

$$\begin{aligned}\|P_t f - f\|_{L^2(\mu)} &\leqslant \|P_t f - P_t g\|_{L^2(\mu)} + \|P_t g - g\|_{L^2(\mu)} + \|g - f\|_{L^2(\mu)}\\ &\leqslant 2\varepsilon + \|P_t g - g\|_{\infty}.\end{aligned}$$

Hence, by Feller continuity of (P_t) (see Lemma 1)

$$\limsup_{t\to 0} \|P_t f - f\|_{L^2(\mu)} \leqslant 2\varepsilon.$$

□

Remark 9 Note that the conclusion of Lemma 7 hold true for any Feller Markov semigroup having μ as invariant measure. This will be used later.

Let $(\mathscr{L}, D(\mathscr{L}))$ denote the infinitesimal generator of (P_t) (now seen as a strongly continuous semigroup on H) and let

$$D = \mathscr{C}^{\infty}_c(\mathbb{M}) \cap H.$$

Proposition 4 *There exist a closed symmetric operator $(S, D(S))$ and a closed anti-symmetric operator $(A, D(A))$ such that*

(i) *D is a core for S, A and $\mathscr{L}$ invariant under S, A, $\mathscr{L}$ and P.*
(ii) *$F \subset D(S)$ and $S|_F = 0$.*

(iii) *For all* $f \in D$

$$S(f) = \frac{\sigma^2}{2}\Delta_M f, \tag{41}$$

$$A(f) = -G_0(f) = \sum_{i=1}^{n} a_j u_j (\nabla e_j(x), \nabla_x f)_{TM} - e_j(x)\partial_{u_j} f \tag{42}$$

and

$$\mathscr{L} f = Lf = Sf - Af. \tag{43}$$

This later proposition shows that conditions of Definition 5 are fulfilled.

Let $\eta_1(M) = \eta_1$ denote the *spectral gap* of M. That is

$$\eta_1(M) := \inf\left\{\int_M |\nabla h|^2 \nu(dx)\ : h \in H^1(M),\ \int_M h^2\nu(dx) = 1,\ \int_M h\nu(dx) = 0\right\} \tag{44}$$

where $\|h\|^2 = (h, h)_{TM}$ and $(.,.)_{TM}$ is the scalar product on the tangent bundle. By a classical result in spectral geometry, compactness of M ensures that $\eta_1 > 0$ and equals the smallest non zero eigenvalue of $-\Delta_M$.

Proposition 5 *Hypotheses* (H1)–(H4) *in Definition* 6 *hold with*

$$\Lambda_1 = \frac{\eta_1\sigma^2}{2}, \quad \Lambda_2 = \min_{i=1,\dots,n} |\lambda_i| a_i,$$

$$N_1 = \frac{\sigma^2}{2}\sum_{j=1}^{n} |\lambda_j|,$$

and

$$N_2 = 2\frac{n}{\min\{|\lambda_j|, j = 1, \dots, n\}} \sup_{i=1,\dots,n} \|\nabla e_i\|_\infty^2 \sqrt{4 + \sum_{i=1}^{n} |\lambda_i| a_i + 4\|\sum_i e_i^2\|_\infty}$$

Remark 10 Since $N_1 \geqslant \frac{n\sigma^2}{2}\eta_1$, then $2\Lambda_1 < 2 + (1 + N_1 + N_2)^2$. Hence $\varepsilon_0 < 1$, where ε_0 is defined by (39).

5.2 Proof of Propositions 4 and 5

Proof of Proposition 4 We first recall some classical results that will be used throughout.

Proposition 6 (See e.g Corollary 1.6, Proposition 2.1, Proposition 3.1, Proposition 3.3 in [14]) *Let K be the generator of a strongly continuous contracting semi-group $(T_t)_t$ on some Banach space $\mathscr{H}$. Then*

1. *K is closed and densely defined.*

2. *The resolvent set of K contains $(0,\infty)$ and $(\lambda I-K)^{-1}g=\int_0^\infty e^{-\lambda t}T_t g\,dt$, for all $g\in\mathscr{H}$ and $\lambda>0$.*
3. *A subspace D of $D(K)$ is a core for K if and only if it is dense in $\mathscr{H}$ and $(\lambda I-K)(D)$ is dense in $\mathscr{H}$ for some $\lambda>0$.*
4. *Let D be a dense subset of $\mathscr{H}$ such that $D\subset D(K)$. If $T_t(D)\subset D$ for all $t\geqslant 0$, then D is a core for K.*

Similarly to (P_t), let (P_t^S) and (P_t^A) be the semigroups respectively induced by the following stochastic and ordinary differential equation on $\mathbb{M}$:

$$dY_t^S=\sum_{j=1}^N G_j(Y_t^S)\circ dB_t^j,$$

and

$$\frac{dY_t^A}{dt}=-G_0(Y_t^A). \tag{45}$$

Note that (P_t^A) is not merely a semigroup but a group of transformation defined as

$$P_t^A f(y)=(f\circ\psi_t)(y) \tag{46}$$

where $\{\psi_t\}$ is the flow induced by (45). The proofs given in Lemma 1, Proposition 2 and Remark 9 show that, not only (P_t) but also (P_t^S) and (P_t^A) are Feller, leave $\mathscr{C}_c^2(\mathbb{M})$ invariant and admit μ as invariant probability. Thus, by Remark 9 and Proposition 6 they induce strongly continuous semigroups on H whose generators, denoted S and A are closed, densely defined and admit $\mathscr{C}_c^2(\mathbb{M})\cap H$ as a core.

Since for all $f\in F$, $P_t^S f=f$, assertion (ii) of Proposition 4 is satisfied. Furthermore, the definition of $\mathscr{L}$, A and S easily imply assertion (iii) as well as invariance of D under the generators and under P. The end of the proof is given by the two following lemmas.

Lemma 8 *D is a core for $\mathscr{L}$, S and A.*

Proof Let G be one of the operators $\mathscr{L}$, S or A. It is easily checked that for all $f\in C_c^2(\mathbb{M})$

$$\|Af\|_{L^2(\mu)}\leq C\|\nabla f\|_\infty$$

and

$$\|Sf\|_{L^2(\mu)}\leq\frac{\sigma^2}{2}\|\Delta_M f\|_\infty$$

for some $C>0$ independent of f. Thus G maps continuously the space $\mathscr{C}_c^2(\mathbb{M})\cap H$ equipped with the $\mathscr{C}^2$ strong topology, into H. By standards approximation results $\mathscr{C}_c^\infty(\mathbb{M})$ is dense into $\mathscr{C}_c^2(\mathbb{M})$ for the $\mathscr{C}^2$ strong topology (see e.g. [19, Chapter 2]). Since $\mathscr{C}_c^2(\mathbb{M})\cap H$ is a core for G, $(I-G)(\mathscr{C}_c^2(\mathbb{M})\cap H)$ is dense in H (see Proposition 6). Thus $(I-G)(D)$ is dense in H and D is a core. □

Lemma 9 *S is symmetric and $A^* = -A$.*

Proof Let $f, g \in D$. Then

$$\begin{aligned}(Sf, g)_H &= \frac{\sigma^2}{2}\int\int_M (\Delta_M f) g \nu(dx) e^{-\Phi} d\xi = -\frac{\sigma^2}{2}\int\int_M (\nabla f, \nabla g)_{TM} \nu(dx) e^{-\Phi} d\xi \\ &= \frac{\sigma^2}{2}\int\int_M (\Delta_M g) f \nu(dx) e^{-\Phi} d\xi = (f, Sg)_H\end{aligned}$$

Since D is a core for S, this proves the symmetry of S.

For $f, g \in H$, we obtain from invariance of μ,

$$(P_t^A f, g)_H = \int_{\mathbb{M}} (f \circ \psi_t)(y) g(y) \mu(dy) = \int_{\mathbb{M}} f(\psi_t(y)) g(\psi_{-t} \circ \psi_t(y)) \mu(dy) \tag{47}$$

$$= \int_{\mathbb{M}} f(y)(g \circ \psi_{-t})(y) \mu(dy). \tag{48}$$

Hence $(P_t^A)^* = P_{-t}^A$. In particular, $((P_t^A)^*)$ is strongly continuous and admits $-A$ as infinitesimal generator. Now, when a semigroup and its adjoint are both strongly continuous, the generator of the adjoint equals the adjoint of the generator. This follows for instance from Theorem 1.5 in [27] combined with Proposition 6 2. Thus $A^* = -A$. □

Proof of Proposition 5 For all $f \in D$ let

$$A_j(f)(x, u) = a_j u_j (\nabla e_j(x), \nabla_x f)_{TM} - e_j(x) \partial_{u_j} f. \tag{49}$$

so that $Af = \sum_{j=1}^n A_j f$. Similarly to A, A_j enjoys the same properties as A. In particular, it leaves D invariant and is antisymmetric:

$$(A_j f, g)_{L^2(\mu)} = -(f, A_j g)_{L^2(\mu)}$$

for all $f, g \in D$.

Finally, we introduce the following operators

$$T = (I + (AP)^*(AP))^{-1} \quad \text{on } H \tag{50}$$

$$B_j = -T(PA_j) \quad \text{on } D \tag{51}$$

where I denotes the identity operator. Recall that B_0 was introduced to be the operator

$$B_0 = T(AP)^* \quad \text{on } D((AP)^* AP).$$

Hypothesis (H1) is immediate because for all $f \in D$, $A_j P f = -e_j(x) \partial_{u_j}(Pf)$ and $\int_M e_j(x) \nu(dx) = 0$, thus $PA_j Pf = 0$.

Hypothesis (H2) follows directly from the variational definition of the spectral gap (44). Indeed for all $f \in D \cap F^{\perp}$

$$-(Sf, f)_{L^2(\mu)} = -\frac{1}{2}\sigma^2 \int_{\mathbb{R}^n} \int_M (\Delta_M f) f \nu(dx) e^{-\Phi(u)} du$$
$$= \frac{1}{2}\sigma^2 \int_{\mathbb{R}^n} \int_M |\nabla_x f|^2 \nu(dx) e^{-\Phi(u)} du \geq \frac{\eta_1}{2}\sigma^2 \|f\|^2_{L^2(\mu)}.$$

For $k = 1, \ldots, n$ let

$$\alpha_k = |\lambda_k| a_k,$$

so that

$$\Phi(u) = \frac{1}{2}\sum_{k=1}^{n} \alpha_k u_k^2 + \ln(C(\Phi)).$$

Let (P_t^{OU}) denote the *Ornstein–Uhlenbeck* semi-group on $L_0^2(e^{-\Phi})$ defined as

$$P_t^{OU} f(u) = \int f\left(e^{-\mathsf{diag}(\alpha_i)t} u + \mathsf{diag}(\sqrt{1 - e^{-2\alpha_i t}})\xi\right) e^{-\Phi(\xi)} d\xi \tag{52}$$

or, equivalently, $P_t^{OU} f(u) = \mathbb{E}(f(U_t^u))$ where U_t^u is the solution to the linear equation on $\mathbb{R}^n$

$$dU_t^i = -\alpha_i U_t^i dt + \sqrt{2} dB_t^i, \quad i = 1 \ldots n,$$

with initial condition $U_0^u = u$ and independent Brownian motions $B^1, \ldots, B^n$.

Let L_{OU} denote the generator of (P_t^{OU}) on $L_0^2(e^{-\Phi})$. The set

$$\tilde{D} = \mathscr{C}_c^{\infty}(\mathbb{R}^n) \cap L_0^2(e^{-\Phi})$$

is a core[1] L_{OU} and for all $f \in \tilde{D}$

$$L_{OU} f = -\langle \nabla \Phi, \nabla f \rangle + \Delta f.$$

The next Lemma is similar to Corollary 2.13 and Proposition 3.13 in [16],

Lemma 10 (i) *For all* $f \in F$

$$PA^2 f = \iota \circ L_{OU} \circ \iota^{-1}(f)$$

[1] This is a classical result and can easily be verified as follows. Formula (52) shows that the set $C_b^{\infty}(\mathbb{R}^n)$ of bounded C^{∞} functions with bounded derivatives is stable under (P_t^{OU}); hence a Core by Proposition 6. Furthermore for each $f \in C_b^{\infty}(\mathbb{R}^n)$ it is easy to construct a sequence $f_n \in C_c^{\infty}(\mathbb{R}^n)$ such that $f_n \to f$ and $L_{OU} f_n \to L_{OU} f$ in $L^2(e^{-\Phi})$.

(ii) $(I - PA^2P)(D)$ *is dense in* H.
(iii) *(H3) holds with* $\Lambda_2 = \min\{\alpha_k : k = 1 \dots n\}$

Proof (i) Let $f \in F \cap D$. Then

$$\begin{aligned} A^2 f &= \sum_{k=1}^n A_k(Af) = \sum_{k=1}^n A_k \left(\sum_{j=1}^n A_j f \right) = \sum_{k=1}^n A_k (\sum_{j=1}^n e_j \partial_{u_j} f) \\ &= \sum_{k=1}^n \left[\left(\nabla e_k, \sum_{j=1}^n \nabla_x (e_j \partial_{u_j} f) \right)_{TM} a_k u_k - \partial_{u_k} \left(\sum_{j=1}^n e_j \partial_{u_j} f \right) e_k \right] \\ &= \sum_{k,j=1}^n \partial_{u_j} f a_k u_k (\nabla e_k, \nabla e_j)_{TM} - \sum_{k,j=1}^n (\partial^2_{u_j u_k} f) e_j e_k \end{aligned} \tag{53}$$

Therefore

$$\begin{aligned} PA^2 f &= \sum_{k,j=1}^n \partial_{u_j} f a_k u_k \int_M (\nabla e_k, \nabla e_j)_{TM} d\nu - \sum_{k,j=1}^n (\partial^2_{u_j u_k} f) \int_M e_j e_k d\nu \\ &= \sum_{j=1}^n \partial_{u_j} f a_j u_j |\lambda_j| - \sum_{j=1}^n (\partial_{u_j u_j} f) = \sum_{j=1}^n \partial_{u_j} f \alpha_j u_j - \sum_{j=1}^n (\partial^2_{u_j u_j} f). \end{aligned} \tag{54}$$

This proves the first assertion.

(ii) $(I - PA^2P)(D \cap F^\perp) = D \cap F^\perp$ is dense in $F^\perp$ because $F^\perp = (I-P)(H)$, $(I-P)(D) \subset D \cap F^\perp$ and D is dense. Also, $(I - PA^2P)(D \cap F) = \imath(I - L_{OU})(\tilde{D})$ is dense in F because, $\tilde{D}$ being a core for L_{OU}, $(I - L_{OU})(\tilde{D})$ is dense in $L_0^2(e^{-\Phi})$. This proves (ii).

(iii) Using antisymmetry of A, assertion (i) and the Poincaré inequality for the Gaussian measure $e^{-\Phi(u)}du$ (see e.g [1, chapter 1]) we get that for all $f \in F \cap D$,

$$\begin{aligned} \|Af\|_H^2 = \|APf\|_H^2 &= (-PA^2Pf, f)_H = (\imath(f), L_{OU}\imath(f))_{L_0^2(e^{-\Phi})} \\ &\geq \min(\alpha_i)\|\imath(f)\|_{L_0^2(e^{-\Phi})} = \min(\alpha_i)\|f\|_H^2. \end{aligned}$$

This proves (H3′), hence (H3). □

Lemma 11 *For* $f \in D \cap F$, *we have* $\|Af\|^2_{L^2(\mu)} = \sum_{k=1}^n \|A_k f\|^2_{L^2(\mu)} = \|\nabla f\|^2_{L^2(\mu)}$.

Proof Let $f \in D \cap F$. Since f does not depend on the x-variable, $A_j f = -e_j \partial_{u_j} f$. The result follows from the fact that the eigenfunctions $(e_j)_{j=1,\dots,n}$ are orthonormal in $L^2(M, dx)$.

The next Lemma is inspired from Lemma 2.4 in [16]

Lemma 12 *For $j = 1, \dots, n$ and $f \in D$,*

$$\|B_j f\|_H \leqslant \frac{1}{2}\|(I-P)f\|_H.$$

Proof The proof is quite similar to the proof of Lemma 2.4 in [16]. Let $f \in D$ and define $g = B_j f$. Thus $g \in D((AP)^*AP)$ and

$$-PA_j f = g + ((AP)^*AP)g. \tag{55}$$

Because $(I - PA^2P)(D)$ is dense in H (see Lemma 10(ii)), there exists a sequence $(g_n) \subset D$ such that

$$\lim_{n\to\infty} g_n - PA^2Pg_n = g + (AP)^*(AP)g. \tag{56}$$

Since $P(D), A(D) \subset D$, it follows from Lemma 2.2 in [16] that

$$-PA^2Pg_n = ((AP)^*(AP))g_n. \tag{57}$$

Thus, by continuity of T,

$$\lim_{n\to\infty} g_n = g \tag{58}$$

and from (57)

$$\lim_{n\to\infty} (AP)^*(AP)g_n = (AP)^*(AP)g. \tag{59}$$

Thus, taking the scalar product of (55) with respect to g_n on both side provides

$$\lim_{n\to\infty} -(PA_j f, g_n)_H - \|g_n\|_H^2 - \|APg_n\|_H^2 = 0.$$

Now, using successively antisymmetry of A_j, Cauchy Schwarz (and Young) inequalities and Lemma 12,

$$-(PA_j f, g_n)_H = ((I-P)f, A_jPg_n)_H \le \|(I-P)f\|_H\|A_jPg_n\|_H \tag{60}$$

$$\le \frac{1}{4}\|(I-P)f\|_H^2 + \|A_jPg_n\|_H^2 \le \frac{1}{4}\|(I-P)f\|_H^2 + \|APg_n\|_H^2 \tag{61}$$

Thus, letting n tends to ∞, leads to

$$\|g\|_H^2 \leqslant \frac{1}{4}\|(I-P)f\|_H^2. \tag{62}$$

□

Lemma 13 (H4 a) *holds with* $N_1 = \frac{\sigma^2}{2}\sum_{j=1}^n |\lambda_j|$.

Proof Let $f \in D \cap F^{\perp}$. Since $\int_{\mathbb{M}} A_j f(y)\mu(dy) = 0$, one has

$$-PAf = \sum_{j=1}^{n} -PA_j f = \sum_{j=1}^{n} P(a_j u_j (\nabla e_j, \nabla_x f)_{TM} - e_j \partial_{u_j} f)$$
$$= \sum_{j=1}^{n} \left[\int_M (\nabla e_j, \nabla_x f)_{TM} a_j u_j d\nu - \int_M e_j \partial_{u_j} f d\nu \right].$$

Since $S(D) \subset D$), then

$$-PASf = -\frac{\sigma^2}{2} \sum_{j=1}^{n} \left[\int_M (\nabla e_j, \nabla_x \Delta_M f)_{TM} a_j u_j d\nu - \int_M e_j \partial_{u_j} \Delta_M f d\nu \right].$$

Because

$$\int_M (\nabla e_j, \nabla_x \Delta_M f)_{TM} d\nu = -\int_M \Delta_M e_j \Delta_M f d\nu = -\lambda_j \int_M e_j \Delta_M f d\nu$$
$$= \lambda_j \int_M (\nabla e_j, \nabla_x f)_{TM} d\nu$$

and

$$\int_M e_j \partial_{u_j} \Delta_M f d\nu = \int_M e_j \Delta_M \partial_{u_j} f d\nu = \int_M \Delta_M e_j \partial_{u_j} f d\nu = \lambda_j \int_M e_j \partial_{u_j} f d\nu$$

for all $j = 1, \ldots, n$, it follows that

$$PASf = \frac{\sigma^2}{2} \sum_{j=1}^{n} \lambda_j (PA_j) f.$$

By antisymmetry of A (resp. A_j) and Lemma 2.2 in [16], for all g in D, $(AP)^* g = -PAg$ (resp. $(A_j P)^* f = -PA_j f$). Hence

$$B_0 Sf = T(AP)^* Sf = -TPASf = \frac{\sigma^2}{2} \sum_{j=1}^{n} \lambda_j B_j f.$$

Applying the triangle inequality, one has

$$\|B_0 Sf\|_{L^2(\mu)} \leqslant \frac{\sigma^2}{2} \sum_{j=1}^{n} |\lambda_j| \|B_j f\|_{L^2(\mu)}$$

and the result follows from Lemma 12. □

The following estimate can be compared with the a priori estimates obtained in [12] and discussed in Appendix $A1$ of [16] (lemmas A3, A4, A5, A7 and Proposition A6) for a more general elliptic equation. Note, however, that here we provide an elementary proof allowing precise estimates by making use of the Γ and Γ_2 operators combined with the specific form of L_{OU}.

Lemma 14 *Let $f \in \tilde{D}$ and*

$$g = (I - L_{OU})f. \tag{63}$$

Then

1. $\| |Hess(f)|_2 \|_{L^2(e^{-\Phi})} \leqslant 4 \|g\|_{L^2(e^{-\Phi})}$
2. $\| |\nabla \Phi|_2 . |\nabla f|_2 \|_{L^2(e^{-\Phi})} \leqslant 2\sqrt{4 + \sum_{i=1}^n \alpha_i} \|g\|_{L^2(e^{-\Phi})}$,

where $|.|_2$ stands for the usual Euclidean norm and $|Hess(f)|_2^2 = \sum_{ij} |\partial_{u_i u_j} f|^2$.

Proof From (63), we have $f = R_1 g$, where R_1 is the resolvent operator of L_{OU}. Thus

$$\|f\|_{L^2(e^{-\Phi})} \leqslant \|g\|_{L^2(e^{-\Phi})}$$

and

$$\|L_{OU} f\|_{L^2(e^{-\Phi})} \leqslant 2\|g\|_{L^2(e^{-\Phi})}.$$

Let Γ be the *"carré du champs"* operator defined by

$$\Gamma(\psi_1, \psi_2) = \frac{1}{2}[L_{OU}(\psi_1 \psi_2) - \psi_2 L_{OU} \psi_1 - \psi_1 L_{OU} \psi_2] \tag{64}$$

and

$$\Gamma_2(\psi) = \frac{1}{2}\Gamma(\psi, \psi) - \psi L_{OU} \psi). \tag{65}$$

It is known (see for instance Subsection 5.3.1 in [1]) that

(i) $\Gamma(f, f) = |\nabla f|_2^2$ and
(ii) $\Gamma_2(f) = |Hess(f)|_2^2 + \langle \nabla f, Hess(\Phi) \nabla f \rangle \geqslant |Hess(f)|_2^2$

by positive definiteness of $Hess(\Phi)$. Therefore, by invariance and reversibility of $e^{-\Phi(u)} du$,

$$\begin{aligned} \| |\nabla f|_2 \|_{L^2(e^{-\Phi})}^2 &= \int \Gamma(f, f) e^{-\Phi(u)} du \\ &= \int -f L_{OU} f e^{-\Phi(u)} du \\ &\leqslant \|f\|_{L^2(e^{-\Phi})} \|L_{OU} f\|_{L^2(e^{-\Phi})} \quad \text{(by the Cauchy-Schwarz inequality)} \\ &\leqslant 2\|g\|_{L^2(e^{-\Phi})}^2 \end{aligned} \tag{66}$$

and

$$\int \Gamma_2(f)e^{-\Phi(u)}du = \|L_{OU}f\|^2_{L^2(e^{-\Phi})} \leqslant 4\|g\|^2_{L^2(e^{-\Phi})}. \tag{67}$$

This last inequality implies (i). Set $h = |\nabla f|_2$ so that $\partial_{u_j}h = \frac{\partial_{u_j}f\partial^2_{u_j}f}{|\nabla f|_2}$. Following the line of the proof of Lemma A.18 in [33] and noting that $\Delta\Phi = \sum_{i=1}^n \alpha_i$, one obtains

$$\int |\nabla\Phi|_2^2 h^2 e^{-\Phi}du \leqslant \sum_{i=1}^n \alpha_i \int h^2 e^{-\Phi}du + 2\sqrt{(\int |\nabla\Phi|_2^2 h^2 e^{-\Phi}du)(\int |\nabla h|_2^2 e^{-\Phi}du)}. \tag{68}$$

Using the Young's inequality $2ab \leqslant \delta^2 a^2 + \frac{b^2}{\delta^2}$ with $\delta^2 = 1/2$, one has

$$\int |\nabla\Phi|_2^2 h^2 e^{-\Phi}du \leqslant 2\sum_{i=1}^n \alpha_i \int h^2 e^{-\Phi}du + 4\int |\nabla h|_2^2 e^{-\Phi}du. \tag{69}$$

Since

$$\begin{aligned} |\nabla h|_2^2 &= \sum_{j=1}^n \left(\frac{\partial_{u_j}f}{|\nabla f|_2}\right)^2 (\partial^2_{u_j}f)^2 \\ &\leqslant \sum_{j=1}^n (\partial^2_{u_j}f)^2 \\ &\leqslant |Hess(f)|_2^2, \end{aligned} \tag{70}$$

we obtain

$$\begin{aligned} \||\nabla\Phi|_2.|\nabla f|_2\|^2_{L^2(e^{-\Phi})} &\leqslant 2\left(\sum_{i=1}^n \alpha_i\right)\int |\nabla f|_2^2 e^{-\Phi}du + 4\int |Hess(f)|_2^2 e^{-\Phi}du \\ &\leqslant 4\left(\sum_{i=1}^n \alpha_i + 4\right)\|g\|^2_{L^2(e^{-\Phi})}. \end{aligned} \tag{71}$$

□

Corollary 2 *Hypothesis* (H4′) (b) *holds with*

$$N_2 = 2\frac{n}{\min\{|\lambda_j|,\ j = 1,\dots,n\}} \sup_{i=1,\dots,n} \|\nabla e_i\|^2_\infty \sqrt{4 + \sum_{i=1}^n \alpha_i + 4\|\sum_i e_i^2\|_\infty}$$

Proof Let $f \in F \cap D$. To shorten notation we identify f and $\iota^{-1}(f) \in \tilde{D}$. Then equation (53) and Cauchy–Schwarz inequality implies

$$
\begin{aligned}
|A^2 f| &\leqslant \sum_{j,k=1}^{n} |\partial_{u_j} f||\nabla e_j|_M \frac{\alpha_k}{|\lambda_k|} u_k |\nabla e_k|_M + |\sum_{k,j=1}^{n} (\partial_{u_j u_k} f) e_j e_k| \\
&\leqslant \left(\sum_{j=1}^{n} \partial_{u_j} f||\nabla e_j|_M\right)\left(\sum_{k=1}^{n} (\alpha_k u_k)|\nabla e_k|_M\right)\lambda_* + |\sum_{k,j=1}^{n} (\partial_{u_j u_k} f) e_j e_k| \\
&\leqslant n\lambda_* \sqrt{\sum_{i=1}^{n} (\partial_{u_i} f|\nabla e_i|_M)^2} \sqrt{\sum_{i=1}^{n} (\alpha_i u_i)^2 |\nabla e_i|_M^2} + |Hess(f)|_2 \left(\sum_i e_i^2\right) \\
&\leqslant n\lambda_* \sup_i \|\nabla e_i\|_\infty^2 |\nabla f|_2 |\nabla \Phi|_2 + |Hess(f)|_2 \|\sum_i e_i^2\|_\infty,
\end{aligned}
$$

where $\lambda_* = \frac{1}{\min\{|\lambda_j|, j=1,\ldots,n\}}$. The result then follows from the preceding lemma. □

6 Exponential decay in the total variation norm

The idea for proving the exponential decay in total variation consists on translating our problem to a setting for which the arguments used for the exponential decay in $L^2(\mu)$ remain valid.

Let $z_0 \in \mathbb{M}$. Since for all $t > 0$, $P_t(z_0, dz) = p_t(z_0, z)dz$ where $p_t(z_0, .)$ is a smooth function and that the invariant probability μ has a smooth density φ, one has

$$
\|P_t(z_0, dz) - \mu(dz)\|_{TV} = \int_{\mathbb{M}} |p_t(z_0, z) - \varphi(z)| dz.
$$

Because $\varphi > 0$, we can define a function $h(t, z_0, .)$ by

$$
h(t, z_0, z) = \frac{p_t(z_0, z)}{\varphi(z)}
$$

By Proposition 1, $P_t(z_0, dz)$ has a compact support, ie $p_t(z_0, .)$ has a compact support. Hence so does $h(t, z_0, .)$. Moreover the smoothness of φ and $p_t(z_0, .)$ implies the smoothness of $h(t, z_0, .)$. Consequently, $h(t, z_0, .) \in L^2(\mathbb{M}, \mu)$ and

$$
\begin{aligned}
\int |p_t(z_0, z) - \varphi(z)| dz = \int |h(t, z_0, z) - 1| \mu(dz) \\
\leqslant \left(\int (h(t, z_0, z) - 1)^2 \mu(dz)\right)^{\frac{1}{2}} \\
= \|h(t, z_0, .) - 1\|_{L^2(\mu)}. \qquad (72)
\end{aligned}
$$

Since $\int_{\mathbb{M}} h(t, z_0, y)\mu(dy) = 1$ for all t and z_0, we have a similar formulation to the one of Theorem 5.

So, in order to give the exponential rate of convergence, we will show that $h(t, z_0, .)$ is solution to the abstract Cauchy problem $\partial_t u(t) = \mathscr{L}_2 u(t)$ in $L^2(\mu)$ where $\mathscr{L}_2$ is an operator for which the arguments used for $\mathscr{L}$ remain valid.

In the following, we denote by h_t (resp. p_t) the function $h_t(z_0, .)$ (resp. $p_t(z_0, .)$)

Since $\partial_t p_t(z_0, .) = L^*(p_t(z_0, .))$ by Theorem 3.(iii) in [21] (recall that L^* is defined by (30)), then

$$\begin{aligned}
\partial_t h_t &= \frac{\partial_t p_t}{\varphi} \\
&= \frac{L^*(p_t)}{\varphi} \\
&= \frac{\sigma^2}{2}\Delta_M h_t + \sum_{k=1}^n a_k u_k \frac{div_x(\nabla e_k(x) p_t)}{\varphi} - \sum_{k=1}^n \frac{\partial_{u_k} p_t}{\varphi} e_k(x) \qquad (73)
\end{aligned}$$

Because $\partial_{u_k}\varphi = -a_k u_k |\lambda_k| \varphi$,

$$-\frac{\partial_{u_k} p_t}{\varphi} = -\partial_{u_k} h_t + a_k u_k |\lambda_k| h_t.$$

Moreover,

$$\frac{div_x(\nabla e_k(x) p_t)}{\varphi} = \Delta_M(e_k) h_t + (\nabla e_k(x), \nabla_x h_t)_{TM}.$$

Hence,

$$\begin{aligned}
\partial_t h_t &= \frac{\sigma^2}{2}\Delta_M h_t + \sum_{k=1}^n a_k u_k (\nabla e_k(x), \nabla_x h_t)_{TM} - \sum_{k=1}^n \partial_{u_k} h_t e_k(x) \\
&=: L_2 h_t.
\end{aligned}$$

Thus, $h_t = T(t-1)h_1$, where $T(t)$ is the semi-group whose infinitesimal generator restricted to $C_c^\infty(\mathbb{M})$ is L_2. Because

$$L_2 = S + \sum_{k=1}^n A_k,$$

whereas

$$L = S - \sum_{k=1}^n A_k,$$

L_2 is the adjoint operator of L in $L^2(\mu)$. So all the arguments used for proving Theorem 6 for L work for L_2. Applying Theorem 6 to L_2 with $g_t = h_{t+1}$ gives the result.

Appendix: A deterministic study

In this Appendix, we study on $\mathbb{S}^1 \times \mathbb{R}^2$ the ODE

$$\begin{cases} \dot{X}_t = (\sin(X_t)U_t - \cos(X_t)V_t) \\ \dot{U}_t = \cos(X_t) \\ \dot{V}_t = \sin(X_t) \end{cases} \tag{74}$$

in order to prove Theorem 3. Since the vectorial field F defined by

$$F(X, U, V) = \begin{pmatrix} (\sin(X)U - \cos(X)V) \\ \cos(X) \\ \sin(X) \end{pmatrix} \tag{75}$$

is smooth and sub-linear,it induces a smooth flow $\psi : \mathbb{R} \times (\mathbb{S}^1 \times \mathbb{R}^2) \to \mathbb{S}^1 \times \mathbb{R}^2$. A first and important observation is

Proposition 7 *If the initial condition for the ODE* (74) *is*

$$(X_0, U_0, V_0) = (X_0, \cos(X_0), \sin(X_0)),$$

then

$$\psi_t(X_0, U_0, V_0) = (X_0, \cos(X_0)(t+1), \sin(X_0)(t+1))\ \forall t \in \mathbb{R}.$$

In particular, the line

$$\{(X, Y, Z) \in \mathbb{S}^1 \times \mathbb{R}^2 \ :\ X = X_0,\ \exists t \in \mathbb{R}\ \ \textit{such that}\ (Y, Z) = (\cos(X_0)t, \sin(X_0)t)\}$$

is invariant under ψ.

Proof By the hypothesis, we have $\dot{X}(0) = 0$. Hence $X(t) = X_0$ for all $t \in \mathbb{R}$. Therefore, $U(t) = \cos(X_0)(t+1)$ and $V(t) = \sin(X_0)(t+1)$

An immediate consequence is

Corollary 3 *If* $\dot{X}(0) > 0$ (*respectively* $\dot{X}(0) < 0$)*, then* $\dot{X}(t) > 0$ (*respectively* $\dot{X}(t) < 0$) *for all* t.

Proof We proceed by contradiction. Hence, by continuity of $\dot{X}$, there exists t_0 such that $\dot{X}(t_0) = 0$. Then the two last Propositions imply that $\dot{X}(t) = 0$ for all t. In particular $\dot{X}(0) = 0$, which is a contradiction.

Let

$$\begin{pmatrix} x \\ u \\ v \end{pmatrix} = \Xi\left(\begin{pmatrix} X \\ U \\ V \end{pmatrix}\right) = \begin{pmatrix} X \\ \cos(X)U + \sin(X)V \\ -\sin(X)U + \cos(X)V \end{pmatrix}. \tag{76}$$

Note that (u, v) is obtained from (U, V) by a rotation of angle $-X$. Then, in the new variable, the ODE (74) becomes the ODE

$$\dot{x}(t) = -v(t) \tag{77}$$

$$\begin{cases} \dot{u}(t) = 1 - v(t)^2 \\ \dot{v}(t) = u(t)v(t) \end{cases} \tag{78}$$

Let

$$H(u, v) = \begin{cases} \frac{1}{2}(u^2 + v^2 - \log(v^2)), & \text{if } v \neq 0, \\ \infty, & \text{if } v = 0. \end{cases} \tag{79}$$

Proposition 8 *The function H is a first integral for the ODE (78).*

Proof Let $v_0 \neq 0$. Deriving H with respect to t and applying the chain rule, we obtain

$$\begin{aligned} \frac{d}{dt} H(u, v) &= (u\dot{u} + v\dot{v}) - \frac{\dot{v}}{v} \\ &= (u - uv^2 - vuv) - u \\ &= 0 \end{aligned}$$

□

Note that H is convex, reaches its global minimum in $(0, \pm 1)$ and takes the value $1/2$ at these points.

For $c \in [1/2, \infty[$, let

$$H_c^+ = H^{-1}(c) \cap \{v > 0\}, \quad H_c^- = H^{-1}(c) \cap \{v < 0\}$$

and set $H_\infty = \{v = 0\}$. Then, we define $\mathbb{T}_c^\alpha = \mathbb{S}^1 \times H_c^\alpha$ for $\alpha \in \{+, -\}$ and $T_\infty = \mathbb{S}^1 \times H_\infty$.

Since the function H is strictly convex on $\{v > 0\}$ and $\{v < 0\}$, we observe that $T_{1/2}^\alpha$ is a closed curve, T_c^α a torus and T_∞ a cylinder.

A first result is

Proposition 9 *Let $(x(t), u(t), v(t))$ be a solution of the ODE defined by (77) and (78).*

(i) *$\mathbb{T}_{1/2}^\alpha$ is a periodic orbit with period 2π, $\alpha \in \{+, -\}$*
(ii) *On T_∞, the dynamic takes the form $(x(t), u(t), v(t)) = (x(0), u(0) + t, 0)$.*

For $c > 1/2$, let T_c be the period of (78) on H_c^α

(iii) *If $\frac{x(T_c)}{2\pi} \in \mathbb{Q}$, then every trajectory on T_c^α is periodic with period qT_c if the irreducible fraction of $\frac{x(T_c)}{2\pi}$ writes $\frac{p}{q}$.*
(iv) *If $\frac{x(T_c)}{2\pi} \notin \mathbb{Q}$, then every trajectory on $\mathbb{S}^1 \times H^{-1}(c)$ is dense either on T_c^+ or T_c^-.*

Proof Points (i) and (ii) follow immediately from (77), (78) and the function H.

Without loss of generality, we assume that $x(0) = 0$. Let $c > 1/2$. Because for $m \in \mathbb{N}^*$, we have

$$\begin{aligned} x(mT_c) &= \int_0^{mT_c} \dot{x}(t)dt = -\int_0^{mT_c} v(t)dt \\ &= -m\int_0^{T_c} v(t)dt \\ &= m\int_0^{T_c} \dot{x}(t)dt, \\ &= mx(T_c) \end{aligned} \tag{80}$$

we obtain that when $(u(t), v(t))$ is back to its initial condition, then $x(t)$ does a rotation of angle $x(T_c)$. Hence if $\frac{x(T_c)}{2\pi} = \frac{p}{q}$, with $q \in \mathbb{N}^*$, $p \in \mathbb{Z}$ and such that the fraction is irreducible, then

$$\begin{aligned} 2p\pi &= qx(T_c) \\ &= x(qT_c). \end{aligned}$$

This proves (iii).

If $\frac{x(T_c)}{2\pi} \notin \mathbb{Q}$, then $(x(qT_c))_{q\in\mathbb{N}}$ is dense on $\mathbb{S}^1$. Now, assume without lost of generality that $v(0) < 0$ and let T be the first time such that $x(T) = 2\pi$. We claim that $(u(nT), v(nT))_{n\in\mathbb{N}}$ is dense on H_c^-. Indeed, if it is not the case, then it is periodic since H_c^- is a closed simple curve. This implies that $(x(t), u(t), v(t))$ is periodic with period n_0T. Thus, there exists $q \in \mathbb{N}$ such that $n_0T = qT_c$. Therefore, by (80), we have $2n_0\pi = x(qT_c) = qx(T_c)$; so that $\frac{x(T_c)}{2\pi} = \frac{n_0}{q}$. This is a contradiction.

The density of $(x(qT_c))_{q\in\mathbb{N}}$ on $\mathbb{S}^1$ and the one of $(u(nT), v(nT))_{n\in\mathbb{N}}$ on H_c^- implies the density of $((x(t), u(t), v(t)))_{t\geqslant 0}$ on T_c^-. This proves (iv). □

From now, we assume without lost of generality that $v(0) < 0$ (the case $v(0) > 0$ being symmetric). In order to derive properties of $c \mapsto T_c$ (see Proposition (9)), we change the time scale by use of $t \mapsto x(t)$. This is possible because it is strictly increasing. We denote by y the inverse function of x. Since we have assumed that $x(0) = 0$, it follows that $y(0) = 0$.

Set $u_2(t) = u(y(t))$ and $v_2(t) = v(y(t))$. Therefore (u_2, v_2) is solution to the ODE

$$\begin{cases} \dot{u}_2(t) = \left(v_2(t) - \frac{1}{v_2(t)}\right) \\ \dot{v}_2(t) = -u_2(t) \end{cases} \tag{81}$$

with initial condition $(u(0), v(0))$. Observe that H is still a first integral for this system.

Proposition 10 *Let $(x(t), u(t), v(t))$ be a solution to the ODE defined by equation (77) with initial condition $(0, u_0, v_0)$ and let $(t, u_2(t), v_2(t))$ where $(u_2(t), v_2(t))$ is the solution to the ODE defined by Eq. (81) with initial condition (u_0, v_0).*

Then $(x(t), u(t), v(t))$ *is periodic in* $\mathbb{S}^1 \times \mathbb{R}^2$ *iff* $(t, u_2(t), v_2(t))$ *is periodic in* $\mathbb{S}^1 \times \mathbb{R}^2$.

Further, if T *is the period of* $(x(t), u(t), v(t))$*, then* $x(T)$ *is the period of* $(t, u_2(t), v_2(t))$.

Proof Straightforward. □

Denote by $T_{c,2}$ the period of $(u_2(t), v_2(t))$, where $c = H(u_2(0), v_2(0)) > 1/2$. Then

$$T_{c,2} = x(T_c). \tag{82}$$

An immediate consequence of Propositions 9 and 10 is that $(t, u_2(t), v_2(t))$ is periodic if and only if

$$\frac{T_{c,2}}{2\pi} \in \mathbb{Q}. \tag{83}$$

In the rest of this Appendix, we study the *"period-function"*

$$f : (1/2, +\infty) \to \mathbb{R}_+ : c \mapsto T_{c,2}. \tag{84}$$

First notice that (0, 1) and (0, −1) are stationary points for the ODE (81).

Let $(u_0, v_0) \in \mathbb{R} \times (0, \infty)$. By symmetry of H along the line $v_2 = 0$, what follow remains true for $v_0 < 0$.

Set $c = H(u_0, v_0)$. Since H is a first integral, then $H(u_2(t), v_2(t)) = c$ for all t.

Using the fact that $\dot{v}_2 = -u_2$, we have that

$$\frac{1}{2}\dot{v}_2^2 + \left(\frac{v_2^2}{2} - \log(v_2)\right) = c. \tag{85}$$

Set $\phi(v) = (\frac{v^2}{2} - \log(v))$. Below (Fig. 5) is given its graph.

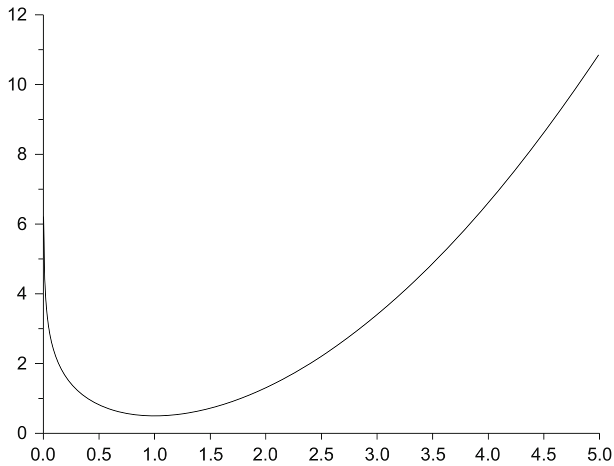

Fig. 5 Graph of the function ϕ

Since the curve $H^{-1}(c)$ is symmetric along the line $u_2 = 0$, we have that

$$\frac{T_{c,2}}{2} = \int_{c_1}^{c_2} \frac{dv}{\sqrt{2(c-\phi(v))}}, \tag{86}$$

i.e.

$$T_{c,2} = \sqrt{2}\int_{c_1}^{c_2} \frac{dv}{\sqrt{(c-\phi(v))}}, \tag{87}$$

where $0 < c_1 < 1 < c_2 < \infty$ are the roots of the function $v \mapsto \phi(v) - c$.

Denote by h the inverse function of ϕ restricted to $[1, \infty)$ and by g the inverse function of ϕ restricted to $(0, 1)$. By a change of variable, we then obtain

$$\int_{1}^{c_2} \frac{dv}{\sqrt{(c-\phi(v))}} = \int_{\frac{1}{2}}^{c} \frac{h'(v)dv}{\sqrt{(c-v)}} \tag{88}$$

and

$$\int_{c_1}^{1} \frac{dv}{\sqrt{(c-\phi(v))}} = -\int_{\frac{1}{2}}^{c} \frac{g'(v)dv}{\sqrt{(c-v)}}. \tag{89}$$

Therefore

$$f(c) = T_{c,2} = \sqrt{2}\int_{\frac{1}{2}}^{c} \frac{(h'-g')(v)}{\sqrt{(c-v)}}dv = \int_{\mathbb{R}} \Lambda(v)A(c-v)dv = (\Lambda * A)(c), \tag{90}$$

where $*$ stands for the convolution product, $\Lambda(v) = \sqrt{2}(h' - g')(v)\mathbf{1}_{v>1/2}$ and $A(v) = \frac{1}{\sqrt{v}}\mathbf{1}_{v>0}$.

Hence

$$f'(c) = (\Lambda * A')(c). \tag{91}$$

Since $g(v) \in (0, 1)$ and $h(v) > 1$ for $v \in (1/2, c)$, then $g'(v) = \frac{1}{\phi'(g(v))} < 0$ and $h'(v) = \frac{1}{\phi'(h(v))} > 0$. Using the fact that $A'(v) = -\frac{1}{2}\mathbf{1}_{v>0}\frac{1}{\sqrt{v^3}}$, we have

$$f'(c) < 0 \quad \text{for all} \quad 1/2 < c < \infty. \tag{92}$$

Our next goal is now to study the limiting behaviour $c \to 1/2$ and $c \to \infty$

Lemma 15 *Let $c > 1/2$ and let c_1 and c_2 the two roots of the function $v \mapsto \phi(v) - c$. Then*

$$T_{c,2} \geqslant 2\sqrt{2}[\sqrt{\frac{c_1}{1+c_1}} + \sqrt{\frac{c_2}{1+c_2}}].$$

Proof By convexity of ϕ, we have $\frac{\phi(v)-\phi(c_1)}{v-c_1} \geqslant \phi'(c_1)$. Hence

$$\sqrt{c-\phi(v)} \leqslant \sqrt{-\phi'(c_1)}\sqrt{v-c_1}.$$

Therefore

$$\int_{c_1}^{1} \frac{dv}{\sqrt{c-\phi(v)}} \geqslant \frac{1}{\sqrt{-\phi'(c_1)}} \int_{c_1}^{1} \frac{dv}{\sqrt{v-c_1}} = 2\frac{\sqrt{1-c_1}}{\sqrt{-\phi'(c_1)}}.$$

Since $-\phi'(v) = \frac{1}{v} - v$, $-\phi'(c_1) = (1-c_1^2)/c_1$ and thus

$$\int_{c_1}^{1} \frac{dv}{\sqrt{c-\phi(v)}} \geqslant 2\sqrt{\frac{c_1}{(1+c_1)}}.$$

Once again convexity of ϕ implies $\frac{\phi(c_2)-\phi(v)}{c_2-v} \leqslant \phi'(c_2)$, so that $c-\phi(v) \leqslant \phi'(c_2)(c_2-v)$. By proceeding as above, we obtain

$$\int_{1}^{c_2} \frac{dv}{\sqrt{c-\phi(v)}} \geqslant 2\sqrt{\frac{c_2}{(1+c_2)}}.$$

Hence

$$f(c) = T_{c,2} = \sqrt{2}[\int_{c_1}^{1} \frac{dv}{\sqrt{c-\phi(v)}} + \int_{1}^{c_2} \frac{dv}{\sqrt{c-\phi(v)}}] \geqslant 2\sqrt{2}[\sqrt{\frac{c_1}{1+c_1}} + \sqrt{\frac{c_2}{1+c_2}}].$$

□

Lemma 16 $\lim_{c\to 1/2} f(c) = \sqrt{2}\pi$.

Proof We have $c_1, c_2 \to 1$ as $c \to 1/2$. Thus, it implies that $\log(v) \approx (v-1) - \frac{1}{2}(v-1)^2$ for $v \in (c_1, c_2)$ and therefore

$$\phi(v) = \frac{1}{2}(v-1+1)^2 - \log(v) \approx \frac{1}{2} + (v-1)^2.$$

But

$$\begin{aligned}\int_{c_1}^{c_2} \frac{dv}{\sqrt{c-\frac{1}{2}-(v-1)^2}} &= \frac{1}{\sqrt{c-\frac{1}{2}}} \int_{c_1-1}^{c_2-1} \frac{dv}{\sqrt{1-(v/\sqrt{c-\frac{1}{2}})^2}} \\ &= \int_{\frac{c_1-1}{\sqrt{c-1/2}}}^{\frac{c_2-1}{\sqrt{c-1/2}}} \frac{du}{\sqrt{1-u^2}} \\ &= \arcsin\left(\frac{c_2-1}{\sqrt{c-1/2}}\right) + \arcsin\left(\frac{1-c_2}{\sqrt{c-1/2}}\right)\end{aligned}$$

Since for c sufficiently close to $1/2$, $c = \phi(1+c_j-1) \approx \frac{1}{2} + (c_j-1)^2$, then $\lim_{c\to 1/2} \frac{|c_j-1|}{\sqrt{c-\frac{1}{2}}} = 1$, $j = 1, 2$.

Thus, $\lim_{c\to 1/2}\int_{c_1}^{c_2}\frac{dv}{\sqrt{c-(v-1)^2}}=\pi$ and therefore

$$\lim_{c\to 1/2} f(c) = \lim_{c\to 1/2} T_{c,2} = \lim_{c\to 1/2} \sqrt{2}\int_{c_1}^{c_2} \frac{dv}{\sqrt{c-\frac{1}{2}-(v-1)^2}} = \sqrt{2}\pi. \tag{93}$$

□

Remark 11 One can prove that $\sqrt{2}\pi$ is the period of the orbits from the linear ODE

$$\begin{cases} \dot{u}(t) = 2v(t) \\ \dot{v}(t) = -u(t). \end{cases} \tag{94}$$

But this is nothing else than the linearized system at (0, 1) from the ODE (81).

Summarizing all these information concerning $T_{c,2}$, we obtain

Proposition 11 *The "period-function" $f : (1/2, \infty) \to \mathbb{R}_+ : c \mapsto T_{c,2}$ is continuous, decreasing, bounded from below by $2\sqrt{2}$ and converge to $\sqrt{2}\pi$ when c tends to 1/2.*

Proof The decreasing property comes from (92) whereas the continuity follows from (90). While c_1 converges to 0 and $\frac{c_2}{1+c_2}$ converges to 1 when c tends to ∞, then Lemma 15 combined with the decreasing property implies that $f(c) \geqslant 2\sqrt{2}$ for all $c > 1/2$.

Since f is decreasing, then $\sup_{c>1/2} f(c) = \lim_{c\to 1/2} f(c) = \sqrt{2}\pi$. Below (Fig. 6) is the graph of the period-function. □

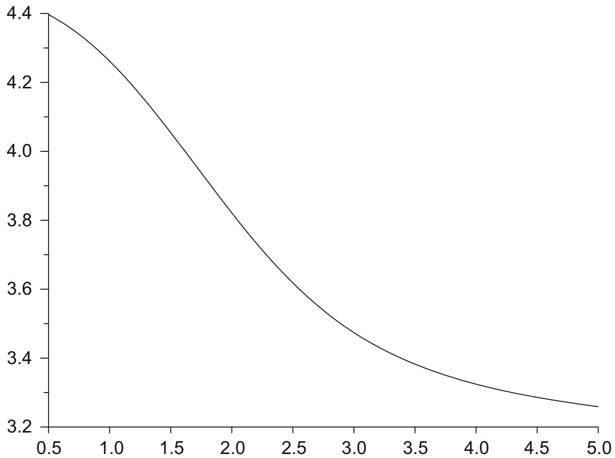

Fig. 6 Graph of the function $c \mapsto T_{c,2}$

References

1. Ané, C., Blachère, S., Chafaï, D., Fougères, P., Gentil, F., Malrieu, F., Roberto, C., Scheffer, G.: Sur les inégalités de Sobolev logaritheoremique, Panorama et synthèse, no. 10. SMF (2000)
2. Aronszajn, N.: A unique continuation theorem for solutions of elliptic partial differential equations or inequalities of second order. J. Math. Pures Appl. **9**(36), 235–249 (1957)
3. Azéma, J., Duflo, M., Revuz, D.: Mesure Invariante des processus de markov récurrent, Convergence of Probability Measures. Séminaire de probabilités (Strasbourg) **3**, 24–33 (1969)
4. Benaïm, M., Ledoux, M., Raimond, O.: Self-interacting diffusions. Probab. Theory Related Fields **122**, 1–41 (2002)
5. Benaïm, M., Raimond, O.: Self-interacting diffusions II: convergence in law. Ann. Inst. H. Poincaré **6**, 1043–1055 (2003)
6. Benaïm, M., Raimond, O.: Self-interacting diffusions III: symmetric interactions. Ann. Probab. **33**(5), 1716–1759 (2005)
7. Cranston, M., Le Jan, Y.: Self-attracting diffusions: two case studies. Math. Ann. **303**, 87–93 (1995)
8. Cranston, M., Mountford, T.: The strong law of large number for a Brownian polymer. Ann. Probab. **24**(3), 1300–1323 (1996)
9. Da Prato, G.: Kolmogorov Equations for Stochastic PDEs. Springer, New York (2004)
10. Da Prato, G., Röckner, M.: Cores for generators of some Markov semigroups. In: Proceedings of the Centennial Conference "Alexandra Muller" Mathematical Seminar, vol. 87, Iasi, p. 97 (2011)
11. Dolbeault, J., Klar, A., Mouhot, C., Schmeiser, C.: Exponential rate of convergence to equilibrium for a model describing fiber lay-down processes. Appl. Math. Res. Express **2**, 165–175 (2013)
12. Dolbeault, J., Mouhot, C., Schmeiser, C.: Hypocoercitivity for linear kinetic equations conserving mass. Trans. Am. Math. Soc. (2015). ISSN 0002-9947
13. Durrett, R.T., Rogers, L.C.G.: Asymptotic behavior of Brownian polymers. Probab. Theory Related Fields **92**(3), 337–349 (1992)
14. Ethier, S.N., Kurtz, T.G.: Markov Processes: Characterization and Convergence. Wiley, New yORK (1986)
15. Grothaus, M., Klar, A., Maringer, J., Stilgenbauer, P.: Geometry, mixing properties and hypocoercitivity of a degenerate diffusion arising in technical textile industry (2012, arXiv preprint)
16. Grothaus, M., Stilgenbauer, P.: Hypocoercitivity for Kolmogorov backward evolution equations and applications. J. Funct. Anal. **267**, 3515–3556 (2014)
17. Herrmann, S., Roynette, B.: Boundedness and convergence of some self-attracting diffusions. Math. Ann. **325**(1), 81–96 (2003)
18. Herrmann, S., Scheutzow, M.: Rate of convergence of some self-attracting diffusions. Stoch. Process. Appl. **111**(1), 41–55 (2004)
19. Hirsch, M.W.: Differential Topology, Graduate Texts in Mathematics, vol. 33. Springer, Berlin (1976)
20. Hsu, Elton P.: Stochastic Analysis on Manifold. American Mathematical Society, Providence (2002)
21. Ichihara, K., Kunita, H.: A classification of the second order degenerate elliptic operators and its probabilistic characterization. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **30**(3), 235–254 (1974)
22. Kallenberg, O.: Foundations of Modern Probability. Springer, Berlin (1997)
23. Kleptsyn, V., Kurtzmann, A.: Ergodicity of self-attracting motions. Electron. J. Probab. **17**(50), 1–37 (2012)
24. Kurtzmann, A.: The ODE method for some self-interacting diffusions on $\mathbb{R}^d$. Ann. Inst. Henri Poincaré Probab. Stat. **46**(3), 618–643 (2010)
25. Mountford, T., Tarrès, P.: An asymptotic result for Brownian polymer. Ann. Inst. Henri Poincaré Probab. Stat. **44**(3), 29–46 (2008)
26. Pemantle, R.: A survey of random processes with reinforcement. Probab. Survey **4**, 1–76 (2007)
27. Phillips, R.S.: The adjoint semi-group. Pac. J. Math. **5**(2), 269–283 (1955)
28. Portegies, J.W.: Embeddings of riemannian manifolds with heat kernels and eigenfunctions. Commun. Pure Appl. Math. **69**(3), 478–518 (2016)
29. Raimond, O.: Self attracting diffusions: case of the constant interaction. Probab. Theory Related Fields **107**(2), 177–196 (1996)
30. Raimond, O.: Self-interacting diffusions: a simulated annealing version. Probab. Theory Related Fields **144**, 247–279 (2009)
31. Revuz, D., Yor, M.: Continuous Martingales and Brownian Motion. Springer, North-Holland Mathematical Library, Amsterdam (1999)

32. Tarrès, P., Tóth, B., Valkó, B.: Diffusivity bounds for 1D Brownian polymers. Ann. Probab. **40**(2), 695–713 (2012)
33. Villani, C.: Hypocoercivity. Mem. Am. Math. Soc. **202**(950) (2009)