

# The asymptotic behavior of fraudulent algorithms

Michel Benaïm<sup>†</sup> and Laurent Miclo<sup>‡</sup>

<sup>†</sup> Institut de Mathématiques  
Neuchâtel University

<sup>‡</sup> Toulouse School of Economics  
Institut de Mathématiques de Toulouse  
CNRS and University of Toulouse

## Abstract

Let  $U$  be a Morse function on a compact connected  $m$ -dimensional Riemannian manifold,  $m \geq 2$ , satisfying  $\min U = 0$  and let  $\mathcal{U} = \{x \in M : U(x) = 0\}$  be the set of global minimizers. Consider the stochastic algorithm  $X^{(\beta)} := (X^{(\beta)}(t))_{t \geq 0}$  defined on  $N = M \setminus \mathcal{U}$ , whose generator is  $U \Delta \cdot - \beta \langle \nabla U, \nabla \cdot \rangle$ , where  $\beta \in \mathbb{R}$  is a real parameter. We show that for  $\beta > \frac{m}{2} - 1$ ,  $X^{(\beta)}(t)$  converges a.s. as  $t \rightarrow \infty$ , toward a point  $p \in \mathcal{U}$  and that each  $p \in \mathcal{U}$  has a positive probability to be selected. On the other hand, for  $\beta < \frac{m}{2} - 1$ , the law of  $(X^{(\beta)}(t))$  converges in total variation (at an exponential rate) toward the probability measure  $\pi_\beta$  having density proportional to  $U(x)^{-1-\beta}$  with respect to the Riemannian measure.

**Keywords:** Global optimization, fraudulent stochastic algorithms, Morse functions, attractive and repulsive minimizers.

**MSC2020:** primary: 60J60, secondary: 58J65, 90C26, 65C05, 60F15, 60J35, 35K10, 37A50.

**Fundings:** the grants SNF 200020-219913, ANR-17-EURE-0010 and AFOSR-22IOE016 are acknowledged.

# 1 Introduction

Stochastic global minimization algorithms taking into account the a priori knowledge of the minimal value of the objective function  $U$  are called **fraudulent**, since this minimal value is often not available in practice. Nevertheless some of their interests are presented in [13], where such a procedure was introduced when  $U$  is a Morse function defined on a compact manifold  $M$  of dimension  $m \geq 2$ . The underlying stochastic process  $X^{(\beta)} := (X^{(\beta)}(t))_{t \geq 0}$ , taking values in  $M$ , comes with a real parameter  $\beta$  which can be tuned to increase the relative importance of  $U$  with respect to the injected randomness. Two quantities  $\beta_\vee \geq \beta_\wedge \in \mathbb{R}$  (depending explicitly on the eigenvalues of the Hessians of  $U$  at its global minima, see Remark 4 below) were introduced so that  $\beta > \beta_\vee$  implies the a.s. convergence of  $X^{(\beta)}(t)$  as  $t \rightarrow \infty$ , toward the global minima of  $U$  (and each of them attracts the algorithm with positive probability, when  $X^{(\beta)}(0)$  is not a global minima), while for  $\beta < \beta_\wedge$  the probability that  $X^{(\beta)}(t)$  a.s. converges toward a global minimum of  $U$  is zero.

Our goal here is to sharpen these result and describe completely the long term behavior of  $X^{(\beta)}$  for all  $\beta \neq \beta_0$ , where  $\beta_0 := \frac{m}{2} - 1$  is a universal (i.e independent of  $U$ ) critical value. We will show that for  $\beta > \beta_0$ ,  $X^{(\beta)}(t)$  a.s. converges toward a global minimizer of  $U$  and that each global minimizer has a positive probability to be selected. On the other hand, for  $\beta < \beta_0$ , the process converges in distribution toward a (unique) invariant distribution whose density (with respect to the Riemannian measure) is explicit. This result will be a consequence of the persistence/non-persistence approach presented in [4] and [2]. The paper is organized as follows. Section 2 sets the notation and presents the main results. Section 3 considers the situation where  $M$  is no longer a compact manifold but the Euclidean space  $\mathbb{R}^m$ . It allows to introduce the main ingredients of the proof in a simple setting. Section 4 is devoted to the proof of the main results. Certain additional points are discussed in appendix.

## 2 Notation and main result

We assume throughout that  $M$  is a compact connected Riemannian manifold having dimension  $m \geq 2$  and  $U : M \rightarrow \mathbb{R}$  is a smooth function such that (this is the fraudulent assumption):

$$\min_M U = 0.$$

The zero set of  $U$ ,

$$\mathcal{U} := \{p \in M : U(p) = 0\},$$

is then the set of global minimizers. We furthermore assume that every  $p \in \mathcal{U}$  is *non-degenerate*, meaning that the Hessian of  $U$  at  $p$  is non-degenerate. This assumption implies that  $\mathcal{U}$  is finite. In particular,  $N := M \setminus \mathcal{U}$  is a noncompact connected manifold.

Let  $L_\beta$  be the operator on  $\mathcal{C}^2(M)$  defined as

$$L_\beta := U \Delta \cdot - \beta \langle \nabla U, \nabla \cdot \rangle \quad (1)$$

where  $\Delta$ ,  $\langle \cdot, \cdot \rangle$  and  $\nabla$  stand for the Laplacian, scalar product and gradient associated to the Riemannian structure of  $M$ , and  $\beta \in \mathbb{R}$ .

A *diffusion process generated by (1)*, is a continuous-time Feller Markov process on  $M$ ,  $X^{(\beta)} = (X^{(\beta)}(t))_{t \geq 0}$ , with infinitesimal generator  $\mathcal{L}_\beta$  and domain  $D(\mathcal{L}_\beta) \subset \mathcal{C}^0(M)$  (see e.g. Le Gall [12], Section 6.2, for the definitions of Feller processes, domains and generators) such that for all  $f \in \mathcal{C}^2(M)$  :

$$f \in D(\mathcal{L}_\beta) \text{ and } \mathcal{L}_\beta f = L_\beta f.$$

Since the mapping  $\nabla U$  and  $\sqrt{U}$  are Lipschitzian, due to the non-degeneracy assumption of the zeroes of  $U$  for the latter, such a diffusion process exists. More details are given in the appendix. In addition,

given the initial distribution of  $X^{(\beta)}(0)$ , say  $\mu$ , the law of  $X^{(\beta)}$ ,  $\mathbb{P}_\mu^{(\beta)}$ , is uniquely determined by  $\mu$  and  $L_\beta$ . As usual, we write  $\mathbb{P}_x^{(\beta)}$  for  $\mathbb{P}_{\delta_x}^{(\beta)}$ . By a mild (but convenient) abuse of notation we may write  $\mathbb{P}_x(X^{(\beta)} \in \cdot)$  for  $\mathbb{P}_x^{(\beta)}(\cdot)$ . We also let  $P^{(\beta)} = (P_t^{(\beta)})_{t \geq 0}$  denote the semi-group induced by  $X^{(\beta)}$ . It is defined, as usual, by

$$\forall t \geq 0, \forall x \in M, \quad P_t^{(\beta)} f(x) := \mathbb{E}_x f(X^{(\beta)}(t))$$

for every measurable, bounded or nonnegative, map  $f : M \rightarrow \mathbb{R}$ .

The next proposition summarizes some basic properties of  $P^{(\beta)}$ . Its proof, which relies on classical results, is given in the appendix.

**Proposition 1** (i)  $P^{(\beta)}$  leaves  $N$  and  $\mathcal{U}$  invariant:

- For all  $t \geq 0$ ,  $P_t^{(\beta)} \mathbf{1}_N = \mathbf{1}_N$ .

(ii)  $P^{(\beta)}$  is Feller on  $M$  and strong-Feller on  $N$  :

- For all  $t \geq 0$  and  $f \in \mathcal{C}^0(M)$ ,  $P_t^{(\beta)}(f) \in \mathcal{C}^0(M)$ ;

- For all  $t > 0$  and  $f : N \rightarrow \mathbb{R}$  bounded measurable,  $P_t^{(\beta)}(f)$  is continuous on  $N$ .

Note that  $P^{(\beta)}$  is not strong Feller on  $M$ , as it can be seen by considering the indicator function of  $N$ . In order to state our main result we first associate, to each  $p \in \mathcal{U}$ , a certain Lyapunov exponent. Given a symmetric positive definite  $m \times m$  real matrix  $A$ , and  $\beta \in \mathbb{R}$ , define the probability measure  $\mu_{A,\beta}$  on  $\mathbb{S}^{m-1}$ , the unit sphere in  $\mathbb{R}^m$ , via

$$\forall \theta \in \mathbb{S}^{m-1}, \quad \mu_{A,\beta}(d\theta) = \frac{1}{Z(A, 1 + \beta)} \langle \theta, A\theta \rangle^{-1-\beta} \sigma(d\theta) \quad (2)$$

where,  $\sigma$  is the uniform probability measure on  $\mathbb{S}^{m-1}$ ,  $\langle \cdot, \cdot \rangle$  the Euclidean dot product (not to be confused with the Riemannian metric on  $M$ ) and  $Z(A, 1 + \beta)$  is the normalization constant.

Define the  $\beta$ -average eigenvalue of  $A$  as

$$\Lambda(A, \beta) = \int_{\mathbb{S}^{m-1}} \langle \theta, A\theta \rangle \mu_{A,\beta}(d\theta) = \frac{Z(A, \beta)}{Z(A, 1 + \beta)}. \quad (3)$$

Let  $\lambda_1(A) \leq \dots \leq \lambda_m(A)$  be the eigenvalues of  $A$ . Observe that  $\Lambda(A, \beta)$  only depends on these eigenvalues, because  $\sigma$  is invariant by orthogonal transformations and  $A$  is orthogonally conjugate to a diagonal matrix. Observe also that

$$\lambda_1(A) \leq \Lambda(A, \beta) \leq \lambda_m(A). \quad (4)$$

**Remark 2** Inequalities (4) are strict, except when  $\lambda_1(A) = \lambda_m(A)$ . Furthermore it can be shown (see the appendix Section 5.2) that for all numbers  $\lambda_- < \lambda < \lambda_+$ , there exists, for  $m$  sufficiently large, a  $m \times m$  definite positive matrix  $A$  such that  $\lambda_1(A) = \lambda_-$ ,  $\Lambda(A, \beta) = \lambda$  and  $\lambda_m(A) = \lambda_+$ .  $\square$

Given  $p \in \mathcal{U}$ , we let  $A_p$  denote the diagonal matrix whose entries  $0 < \lambda_1(p) \leq \dots \leq \lambda_m(p)$  are the eigenvalues of the Hessian of  $U$  at  $p$ . Set

$$\beta_0 := \frac{m}{2} - 1.$$

Our main result is the following.

**Theorem 3** Let  $x \in N$  and  $\beta \in \mathbb{R}$ .

(i) If  $\beta > \beta_0$ , then

$$\sum_{p \in \mathcal{U}} \mathbb{P}_x \left[ \limsup_{t \rightarrow +\infty} \frac{\ln(d(X^{(\beta)}(t), p))}{t} \leq -\Lambda(A_p, \beta)(\beta - \beta_0) \right] = 1,$$

where each term in the above sum is positive.

(ii) If  $\beta < \beta_0$ , then  $X^{(\beta)}$  has, on  $N$ , a unique invariant probability distribution given by

$$\pi_\beta(dx) := \frac{1}{C_\beta} U(x)^{-1-\beta} \ell(dx),$$

where  $C_\beta$  is a normalization constant and  $\ell(dx)$  stands for the Riemannian measure. Furthermore:

(a)  $X^{(\beta)}$  is positive recurrent on  $N$ , meaning that for all  $f \in L^1(\pi_\beta)$ ,  $\mathbb{P}_x$  a.s.,

$$\lim_{t \rightarrow +\infty} \frac{1}{t} \int_0^t f(X^{(\beta)}(s)) ds = \pi_\beta(f)$$

(b) There exist positive constants  $a, b, \chi$  (depending on  $\beta$ ) with  $\chi < \beta_0 - \beta$ , such that for all  $f : N \rightarrow \mathbb{R}$ , measurable,

$$|\mathbb{E}_x[f(X^{(\beta)}(t))] - \pi_\beta(f)| \leq \frac{ae^{-bt}}{d(x, \mathcal{U})^\chi} \|f\|_\chi,$$

where

$$\|f\|_\chi := \sup_{x \in N} |f(x)| d(x, \mathcal{U})^\chi.$$

(iii) If  $\beta = \beta_0$ , then, for every neighborhood  $O$  of  $\mathcal{U}$ ,  $\mathbb{P}_x$  a.s.,

$$\lim_{t \rightarrow +\infty} \frac{1}{t} \int_0^t \mathbf{1}_{\{X^{(\beta)}(s) \in O\}} ds = 1$$

**Remark 4** Theorem 3 is an improvement over the results of [13], which showed the a.s. convergence of  $X^{(\beta)}$  toward elements of  $\mathcal{U}$  (each being approached with a positive probability) only for  $\beta > \beta_\vee \geq \beta_0$  with

$$\beta_\vee := \max_{p \in \mathcal{U}} \frac{\sum_{l \in \llbracket m \rrbracket} \lambda_l(p)}{2\lambda_1(p)} - 1, \quad (5)$$

and the a.s. non-convergence of  $X^{(\beta)}$  toward elements of  $\mathcal{U}$  for  $\beta < \beta_\wedge \leq \beta_0$ , with

$$\beta_\wedge := \min_{p \in \mathcal{U}} \frac{\sum_{l \in \llbracket m \rrbracket} \lambda_l(p)}{2\lambda_m(p)} - 1. \quad (6)$$

□

**Remark 5** Here we restrict our attention to dimensions  $m \geq 2$ , so that  $N$  is connected. The case  $m = 1$  which corresponds to the circle is already treated in [13]. □

**Remark 6** By Theorem 3, the diffusion  $X^{(\beta)}$  on  $N$  is transient for  $\beta > \beta_0$  and positive recurrent if and only if  $\beta < \beta_0$ , due to the fact that  $\int_N U^{-1-\beta} d\ell = +\infty$  for  $\beta \geq \beta_0$ . By standard results (see e.g. Kliemann [10], Theorem 3.2 applied with  $C = N$ ), it is then either null recurrent or transient for  $\beta = \beta_0$ . It would be interesting to investigate this situation. □

### 3 Euclidean computations

This section considers a situation where the state space  $M$  is no longer a compact manifold but the Euclidean space  $\mathbb{R}^m$ , with  $m \geq 2$ . We state a theorem (Theorem 7 below) analogous to Theorem 3 (i). This result is interesting in itself, and its proof allows us to explain, in a simple framework, how to characterize the attractiveness/repulsivity of a global minimum. The main idea is to expand a critical point to a sphere, using polar decompositions, following [4].

Let  $U : \mathbb{R}^m \rightarrow \mathbb{R}_+$  be a smooth function with  $\min U = 0$ . We assume that for each  $p \in \mathcal{U} := U^{-1}(0)$ ,  $\text{Hess } U(p)$  is positive definite. In particular, points in  $\mathcal{U}$  are isolated and  $\mathcal{U}$  is therefore countable.

For any fixed  $\beta \in \mathbb{R}$ , as in (1), we are interested in the operator  $L_\beta$  defined on  $\mathcal{C}^2(\mathbb{R}^m)$ , via

$$\forall x \in \mathbb{R}^m, \quad L_\beta[f](x) := U(x)\Delta f(x) - \beta \langle \nabla U, \nabla f \rangle(x) \quad (7)$$

where  $\Delta$ ,  $\langle \cdot, \cdot \rangle$  and  $\nabla$ , respectively denote, the Euclidean Laplacian, scalar product and gradient. Throughout all this section  $\|x\| = \sqrt{\langle x, x \rangle}$  denotes the Euclidean norm of  $x$ .

Associated to (7) is the stochastic differential equation

$$dX^{(\beta)}(t) = -\beta \nabla U(X^{(\beta)}(t))dt + \sqrt{2U(X^{(\beta)}(t))}dB_t \quad (8)$$

where  $B = (B_t)_{t \geq 0}$  is a standard Brownian motion on  $\mathbb{R}^m$ .

By local Lipschitz continuity of  $\nabla U$  and  $\sqrt{U}$ , there exists, for each  $x \in \mathbb{R}^m$ , a unique solution  $X^{(\beta)} : [0, \tau^\infty) \rightarrow \mathbb{R}^m$  starting from  $x$ , (i.e.  $X^{(\beta)}(0) = x$ ). Here,  $0 < \tau^\infty \leq \infty$ , denotes the explosion time of  $X^{(\beta)}$  and is characterized by

$$\tau^\infty > t \Leftrightarrow \|X^{(\beta)}(t)\| < \infty.$$

The set  $\mathbb{R}^m \setminus \mathcal{U}$  is invariant, in the sense that for all  $t \geq 0, x \in \mathbb{R}^m \setminus \mathcal{U}$ ,

$$\mathbb{P}_x(X^{(\beta)}(t) \in \mathbb{R}^m \setminus \mathcal{U} | \tau^\infty > t) = 1.$$

The proof of this last point is the same as the proof of Proposition 1 (i) given in the appendix.

**Theorem 7** (i) Suppose  $\beta > \beta_0$ . Then, for all  $x \in \mathbb{R}^m \setminus \mathcal{U}$  and  $p \in \mathcal{U}$ ,

$$\mathbb{P}_x \left[ \limsup_{t \rightarrow +\infty} \frac{\ln(\|X^{(\beta)}(t) - p\|)}{t} \leq -\Lambda(A_p, \beta)(\beta - \beta_0) \right] > 0 \quad (9)$$

where  $A_p, \Lambda(A_p, \beta)$  are defined as in Section 2.

(ii) Suppose  $\beta > \beta_0$ , and in addition, that there exist positive constants  $\alpha, r$  (possibly depending on  $\beta$ ) such that

$$2\beta_0 U(x) - \beta \langle \nabla U(x), x \rangle \leq -\alpha \|x\|^2 \quad (10)$$

whenever  $\|x\| \geq r$ . Then,  $\mathcal{U}$  is finite and for all  $x \in \mathbb{R}^m \setminus \mathcal{U}$ ,

$$\sum_{p \in \mathcal{U}} \mathbb{P}_x \left[ \limsup_{t \rightarrow +\infty} \frac{\ln(\|X^{(\beta)}(t) - p\|)}{t} \leq -\Lambda(A_p, \beta)(\beta - \beta_0); \tau^\infty = \infty \right] = 1. \quad (11)$$

(iii) Suppose  $\beta < \beta_0$ . Then, for all  $p \in \mathcal{U}$  and  $x \in \mathbb{R}^m \setminus \{p\}$

$$\mathbb{P}_x \left[ \lim_{t \rightarrow \infty} X^{(\beta)}(t) = p \right] = 0.$$

**Remark 8** The condition (10) is given for its simplicity. However, the conclusion (11) holds true under the weaker assumption, implied by (10) (see Lemma 9 below), that  $X^{(\beta)}$  almost surely never explodes (i.e  $\tau^\infty = \infty$ ) and eventually enters a ball  $B(0, r)$  containing  $\mathcal{U}$  for some  $r > 0$ .  $\square$

The remainder of this section is devoted to the proof of Theorem 7. We first recall some classical facts about diffusion operators, see e.g. Bakry, Gentil and Ledoux [1]. The **carré du champ**  $\Gamma_L$  associated to a Markov generator  $L$  defined on an algebra  $\mathcal{A}(L)$  is the bilinear functional defined on  $\mathcal{A}(L) \times \mathcal{A}(L)$  via

$$\forall f, g \in \mathcal{A}(L), \quad \Gamma_L[f, g] := L[fg] - fL[g] - gL[f]$$

(we will denote  $\Gamma_L[f] := \Gamma_L[f, f]$ ).

The generator  $L$  is said to be of **diffusion**, if  $\mathcal{A}(L)$  is stable by composition with smooth functions and if we have

$$L[\varphi(f)] = \varphi'(f)L[f] + \frac{\varphi''(f)}{2}\Gamma_L[f] \quad (12)$$

for any  $f \in \mathcal{A}(L)$  and any function  $\varphi$  smooth on the image of  $f$ .

In this situation we also have, with the same notations,

$$\Gamma_L[\varphi(f)] = (\varphi'(f))^2\Gamma_L[f] \quad (13)$$

The Markov generator given in (7) is of diffusion with  $\mathcal{A}(L_\beta) = \mathcal{C}^2(\mathbb{R}^m)$ . The corresponding carré du champ is given by

$$\forall f \in \mathcal{C}^2(\mathbb{R}^m), \quad \Gamma_{L_\beta}[f] = 2U\|\nabla f\|^2. \quad (14)$$

Our first goal is to show that, under condition (10),  $X^{(\beta)}$  never explodes and always enter the ball  $B(0, r)$ . For all  $s \geq 0$ , we let

$$\tau_s = \inf\{t \geq 0 : \|X^{(\beta)}(t)\| \leq s\} \text{ and } \tau^s = \inf\{t \geq 0 : \|X^{(\beta)}(t)\| \geq s\}.$$

Note that these stopping times depend on  $\beta$ , but to shorten notation we omit this dependance in their definition.

**Lemma 9** *Under the condition (10),*

$$\mathbb{P}_x(\tau^\infty = \infty; \tau_r < \infty) = 1$$

for all  $x \in \mathbb{R}^m$  and  $r$  is as in (10).

**Proof**

Let  $V : \mathbb{R}^m \rightarrow \mathbb{R}$  be a smooth function coinciding with  $\ln(\|x\|^2)$  for  $\|x\| \geq r$ . Using the formulae (12) and (14) it comes that, for all  $\|x\| \geq r$ ,

$$L_\beta(V)(x) = \frac{2}{\|x\|^2} (2\beta_0 U(x) - \beta \langle \nabla U(x), x \rangle) \leq -2\alpha.$$

In particular, for all  $x \in \mathbb{R}^m$ ,  $L_\beta(V)(x) \leq C$  where  $C = \sup_{\{x \in \mathbb{R}^m : \|x\| \leq r\}} |L_\beta(V)(x)|$ . Thus, by Ito's formulae, for all  $k \geq 1$ ,

$$\ln(k^2)\mathbb{P}_x(\tau^k \leq t) \leq \mathbb{E}_x(V(X^{(\beta)}(t \wedge \tau^k)) = V(x) + \mathbb{E}_x \left[ \int_0^{t \wedge \tau^k} L_\beta[V](X^{(\beta)}(s)) ds \right] \leq V(x) + tC.$$

This shows that  $\mathbb{P}_x(\tau^k \leq t) \rightarrow 0$ , as  $k \rightarrow \infty$ . Hence  $\mathbb{P}_x(\tau^\infty < \infty) = 0$ .

Now, by Ito formulae again, the process  $(M_t)_{t \geq 0}$  defined as

$$M_t := V(X^{(\beta)}(t \wedge \tau_r)) - \ln(r^2) - \int_0^{t \wedge \tau_r} L_\beta V(X^{(\beta)}(s)) ds \geq 2\alpha(t \wedge \tau_r)$$

is a nonnegative  $\mathbb{P}_x$  local martingale. A nonnegative local martingale may not be a martingale but is always a supermartingale (Le Gall [12], Proposition 4.7). Thus  $2\alpha\mathbb{E}_x(t \wedge \tau_r) \leq \mathbb{E}_x(M_t) \leq V(x) - \ln(r^2)$ . Hence  $\mathbb{E}_x(\tau_r) < \infty$ .  $\blacksquare$

Our next goal is to investigate the behavior of  $X^{(\beta)}$  around a critical point  $p \in \mathcal{U}$ . Without loss of generality, we assume that  $p = \{0\}$ . We let  $A = \text{Hess } U(0)$ . Fix  $\epsilon \in (0, 1)$  small enough so that  $\mathcal{U} \cap B(0, \epsilon) = \{0\}$ . Write any  $x \in B(0, \epsilon) \setminus \{0\}$  under its polar decomposition  $x = \rho\theta$  with  $\rho \in (0, \epsilon)$  and  $\theta \in \mathbb{S}^{m-1}$ . This decomposition induces the mapping

$$P : \mathcal{C}^2(B(0, \epsilon)) \ni f \mapsto P[f] \in \mathcal{C}^2((0, \epsilon) \times \mathbb{S}^{m-1})$$

with

$$\forall (\rho, \theta) \in (0, \epsilon) \times \mathbb{S}^{m-1}, \quad P[f](\rho, \theta) := f(\rho\theta) \quad (15)$$

Endow  $\mathbb{S}^{m-1}$  with its usual Riemannian structure, inherited from  $\mathbb{R}^m$ , and denote  $\langle \cdot, \cdot \rangle_\theta$ ,  $\nabla_\theta$ ,  $\text{div}_\theta$  and  $\Delta_\theta$  the corresponding scalar product, gradient, divergence, and Laplace-Beltrami operator. Note that  $\langle \cdot, \cdot \rangle_\theta$  is just the restriction of  $\langle \cdot, \cdot \rangle$  to the tangent space of  $\mathbb{S}^{m-1}$  at  $\theta$ .

Classical computations in polar coordinates show that for any  $f, g \in \mathcal{C}^2(B(0, \epsilon))$ , we have on  $(0, \epsilon) \times \mathbb{S}^{m-1}$ ,

$$\begin{aligned} P[\langle \nabla f, \nabla g \rangle] &= \partial_\rho P[f] \partial_\rho P[g] + \frac{1}{\rho^2} \langle \nabla_\theta P[f], \nabla_\theta P[g] \rangle_\theta, \\ P[\Delta f] &= \partial_\rho^2 P[f] + \frac{m-1}{\rho} \partial_\rho P[f] + \frac{1}{\rho^2} \Delta_\theta P[f]. \end{aligned}$$

It leads us to introduce the operator  $L_\beta$  on  $\mathcal{C}^2((0, \epsilon) \times \mathbb{S}^{m-1})$  defined by

$$L_\beta \cdot := U \left( \partial_\rho^2 \cdot + \frac{m-1}{\rho} \partial_\rho \cdot + \frac{1}{\rho^2} \Delta_\theta \cdot \right) - \beta \left( (\partial_\rho U) \partial_\rho \cdot + \frac{1}{\rho^2} \langle \nabla_\theta U, \nabla_\theta \cdot \rangle_\theta \right) \quad (16)$$

where  $U := P[U]$ . Indeed, on  $\mathcal{C}^2(B(0, \epsilon))$ , we have the intertwining relation

$$L_\beta \circ P = P \circ L_\beta.$$

**Lemma 10** *The operator  $L_\beta$  extends to a diffusion operator, still denoted  $L_\beta$ , on  $\mathcal{C}^2([0, \epsilon) \times \mathbb{S}^{m-1})$ , whose associated diffusion process  $X^{(\beta)}$  leave  $\{0\} \times \mathbb{S}^{m-1}$  invariant. On  $\{0\} \times \mathbb{S}^{m-1}$ , identified with  $\mathbb{S}^{m-1}$ ,  $X^{(\beta)}$  admits for generator the operator  $G_\beta$  acting on  $\mathcal{C}^2(\mathbb{S}^{m-1})$  via*

$$\forall f \in \mathcal{C}^2(\mathbb{S}^{m-1}), \quad G_\beta[f] := \frac{1}{2} \Psi_A^{1+\beta} \text{div}_\theta (\Psi_A^{-\beta} \nabla_\theta f) \quad (17)$$

where  $\forall \theta \in \mathbb{S}^{m-1}$ ,  $\Psi_A(\theta) = \langle \theta, A\theta \rangle$ . Furthermore,  $G_\beta$  has a unique invariant probability measure on  $\mathbb{S}^{m-1}$ , given by  $\mu_{A, \beta}$  (see Equation (2)).

### Proof

Our assumptions on  $U$  imply that, uniformly over  $\theta \in \mathbb{S}^{m-1}$ ,

$$\lim_{\rho \rightarrow 0^+} \frac{U(\rho, \theta)}{\rho^2} = \frac{1}{2} \langle \theta, A\theta \rangle, \quad (18)$$

$$\lim_{\rho \rightarrow 0_+} \frac{\partial_\rho U(\rho, \theta)}{\rho} = \langle \theta, A\theta \rangle, \quad (19)$$

$$\lim_{\rho \rightarrow 0_+} \frac{\nabla_\theta U(\rho, \theta)}{\rho^2} = A\theta - \langle \theta, A\theta \rangle \theta. \quad (20)$$

Indeed, by the usual expansion of  $U$  around 0, we have

$$\begin{aligned} U(x) &= U(0) + \langle \nabla U(0), x \rangle + \frac{1}{2} \langle x, \text{Hess } U(0)x \rangle + o(\langle x, x \rangle) \\ &= \frac{1}{2} \langle x, Ax \rangle + o(\langle x, x \rangle) \end{aligned}$$

which translates into

$$U(\rho, \theta) = \frac{\rho^2}{2} \langle \theta, A\theta \rangle + o(\rho^2)$$

leading to the first announced limit (18). Similarly,

$$\begin{aligned} \nabla U(x) &= \nabla U(0) + \text{Hess } U(0)x + o(\sqrt{\langle x, x \rangle}) \\ &= Ax + o(\sqrt{\langle x, x \rangle}). \end{aligned}$$

At  $x = \rho\theta$  with  $\rho > 0$ ,  $\partial_\rho U(\rho, \theta)\theta$  is the radial part of  $\nabla U(x)$  and  $\nabla_\theta U(\rho, \theta)/\rho$  is the tangential part. It follows that

$$\begin{aligned} \partial_\rho U(\rho, \theta) &= \langle \nabla U(x), \theta \rangle, \\ \frac{\nabla_\theta U(\rho, \theta)}{\rho} &= \nabla U(x) - \partial_\rho U(\rho, \theta)\theta, \end{aligned}$$

and we get

$$\begin{aligned} \frac{\partial_\rho U(\rho, \theta)}{\rho} &= \langle \theta, A\theta \rangle + o(1), \\ \frac{\nabla_\theta U(\rho, \theta)}{\rho^2} &= A\theta - \langle \theta, A\theta \rangle \theta + o(1), \end{aligned}$$

leading to the wanted second and third results (19) and (20).

It follows that for any  $F \in \mathcal{C}^2([0, \epsilon] \times \mathbb{S}^{m-1})$ , we have, uniformly over  $\theta \in \mathbb{S}^{m-1}$ ,

$$\lim_{\rho \rightarrow 0_+} \mathbb{L}_\beta[F](\rho, \theta) = \frac{1}{2} \langle \theta, A\theta \rangle \triangle_\theta F(0, \theta) - \beta \langle A\theta - \langle \theta, A\theta \rangle \theta, \nabla_\theta F(0, \theta) \rangle_\theta \quad (21)$$

Denoting  $\mathbb{L}_\beta[F](0, \theta)$  the r.h.s. enables us to see  $\mathbb{L}_\beta$  as a diffusion operator on  $[0, \epsilon] \times \mathbb{S}^{m-1}$ , whose associated diffusion process  $X^{(\beta)}$  leaves  $\{0\} \times \mathbb{S}^{m-1}$  invariant, and such that on  $\{0\} \times \mathbb{S}^{m-1}$ , identified with  $\mathbb{S}^{m-1}$ , its generator coincides with the operator defined by

$$G_\beta(f) := \langle \theta, A\theta \rangle \left( \frac{1}{2} \triangle_\theta f(\theta) - \beta \langle b(\theta), \nabla_\theta f \rangle_\theta \right)$$

where

$$\forall \theta \in \mathbb{S}^{m-1}, \quad b(\theta) := \frac{A\theta - \langle \theta, A\theta \rangle \theta}{\langle \theta, A\theta \rangle} = \frac{1}{2} \nabla_\theta \ln(\langle \theta, A\theta \rangle).$$

It is easily checked that  $G_\beta$  can be rewritten under the divergence form given by (17). This divergence form implies that the probability measure  $\mu_{A, \beta}$  defined in (2) is invariant. By ellipticity of  $G_\beta$  there is no other invariant probability measure. ■



**Lemma 11** Suppose  $\beta > \beta_0$  and  $0 < \lambda < \Lambda(A, \beta)$ . There exists  $0 < \epsilon_0 \leq \epsilon$  with the property that for all  $0 < \eta \leq 1$ , there exists  $0 < \epsilon_1 < \epsilon_0$  such that for all  $\|x\| \leq \epsilon_1$ ,

$$\mathbb{P}_x \left[ \limsup_{t \rightarrow +\infty} \frac{\ln(\|X^{(\beta)}(t)\|)}{t} \leq -\lambda(\beta - \beta_0); \tau^{\epsilon_0} = \infty \right] = \mathbb{P}_x [\tau^{\epsilon_0} = \infty] \geq 1 - \eta. \quad (22)$$

If now,  $\beta < \beta_0$ , then for all  $x \in \mathbb{R}^m \setminus \{0\}$ ,

$$\mathbb{P}_x \left[ \lim_{t \rightarrow +\infty} \|X^{(\beta)}(t)\| = 0 \right] = 0.$$

### Proof

The proof follows from the stochastic persistence approach used in [4], [2]. Let  $V$  be the function defined on  $(0, \epsilon) \times \mathbb{S}^{m-1}$  via

$$V(\rho, \theta) := -\ln(\rho). \quad (23)$$

We claim that:

- [a]  $L_\beta[V]$  can be extended into a continuous function  $H_\beta$  on  $[0, \epsilon) \times \mathbb{S}^{m-1}$ ;
- [b]  $\Gamma_{L_\beta}[V]$  is bounded on  $(0, \epsilon) \times \mathbb{S}^{m-1}$ ; and
- [c]  $\mu_{A, \beta}[H_\beta(0, \cdot)] = \Lambda(A, \beta)(\beta - \beta_0)$ .

Using the form of  $L_\beta$  (equation (16)) and the equalities (18), (19), (a) holds true with

$$H_\beta(0, \theta) = (\beta - \beta_0) \langle \theta, A\theta \rangle \quad (24)$$

and (c) directly follows from the definition of  $\Lambda(A, \beta)$ . For (b), the definition of  $L_\beta$  and  $\Gamma_{L_\beta}$ , lead to

$$\forall f \in \mathcal{C}^2((0, \epsilon) \times \mathbb{S}^{m-1}), \quad \Gamma_{L_\beta}[f] = 2U \left( (\partial_\rho f)^2 + \frac{1}{\rho^2} |\nabla_\theta f|^2 \right).$$

Thus,

$$\Gamma_{L_\beta}[V] = 2 \frac{U(\rho, \theta)}{\rho^2}$$

which is bounded in view of (18). This concludes the proof of the claim.

If  $\beta > \beta_0$ ,  $\mu_{A, \beta}[H_\beta(0, \cdot)] > 0$  and the first assertion of the lemma follows from Theorem 5.4 in [4] (to be more precise, this follows from the proof of Theorem 5.4 in [4], because the formulation of Theorem 5.4 in [4] doesn't specify that  $\eta$  can be chosen arbitrary close to one). If  $\beta < \beta_0$ ,  $\mu_{A, \beta}[H_\beta(0, \cdot)] < 0$  and (see e.g. [2], Proposition 8.1 or the proof of Theorem 3.2 (iii) in [4]) there exist positive constants  $\epsilon_1 \leq \epsilon, C$  such that  $\mathbb{E}_x(\tau^{\epsilon_1}) \leq C |\ln(\|x\|)| < \infty$  for  $x \in B(0, \epsilon_1) \setminus \{0\}$ . ■

We can now conclude the proof of Theorem 7. We start with assertion (ii). Fix  $\beta > \beta_0$ . For  $n \in \mathbb{N}$  sufficiently large (so that  $\Lambda(A_p, \beta) > \frac{1}{n}$ ) and  $p \in \mathcal{U}$ , let  $\mathcal{E}_n(p)$  be the event defined as

$$\mathcal{E}_n(p) = \left\{ \limsup_{t \rightarrow +\infty} \frac{\ln(\|X^{(\beta)}(t) - p\|)}{t} \leq -(\Lambda(A_p, \beta) - \frac{1}{n})(\beta - \beta_0) \right\},$$

and let

$$\mathcal{E}_n = \bigcup_{p \in \mathcal{U}} \mathcal{E}_n(p).$$

The set  $\mathcal{U}$  is finite, since (10) cannot be satisfied by a point  $x \in \mathcal{U}$  and by consequence  $\mathcal{U}$  is included into the compact ball centered at 0 and of radius  $r$ . Thus there exists, by Lemma 11,  $\epsilon_1 > 0$  such that

$$\mathbb{P}_x(\mathcal{E}_n(p)) \geq \frac{1}{2}$$

for all  $x \in B(p, \epsilon_1)$  and all  $p \in \mathcal{U}$ . Let

$$\mathcal{U}_{\epsilon_1} := \bigcup_{p \in \mathcal{U}} B(p, \epsilon_1)$$

and  $\tau_{\mathcal{U}_{\epsilon_1}} := \inf\{t \geq 0 : X^{(\beta)}(t) \in \mathcal{U}_{\epsilon_1}\}$ . By ellipticity of  $L_\beta$  on  $\mathbb{R}^m \setminus \mathcal{U}$ ,  $\mathcal{U}_{\epsilon_1}$  is open and *accessible* from all  $x \in \mathbb{R}^m$ , in the sense that  $\mathbb{P}_x(X^{(\beta)}(t_x) \in \mathcal{U}_{\epsilon_1}) > 0$  for some  $t_x \geq 0$ . Thus, by Feller continuity and compactness of  $\overline{B(0, r)}$ , there exists  $\delta > 0$  such that

$$\mathbb{P}_x(\tau_{\mathcal{U}_{\epsilon_1}} < \infty) \geq \delta$$

for all  $x \in \overline{B(0, r)}$ . Combined with Lemma 9, this proves that  $\mathbb{P}_x(\tau_{\mathcal{U}_{\epsilon_1}} < \infty) \geq \delta$  for all  $x \in \mathbb{R}^m$ . Thus,

$$\mathbb{P}_x(\mathcal{E}_n) \geq \delta/2$$

for all  $x \in \mathbb{R}^m$ . The strong Markov property, implies that  $\mathbb{P}_x(\mathcal{E}_n) = 1$ . Hence

$$\mathbb{P}_x\left(\bigcap_n \mathcal{E}_n\right) = 1.$$

This concludes the proof of (ii).

We now pass to the proof of (i). Fix  $\beta > \beta_0$  and assume without loss of generality that  $p = \{0\}$ . Let  $\tilde{U} : \mathbb{R}^m \rightarrow \mathbb{R}^+$  be a smooth function which coincides with  $x \mapsto U(x)$  on a neighborhood of 0, with  $x \mapsto \|x\|^2$  for  $\|x\| \geq 1$ , and such that  $U^{-1}(0) = \{0\}$ . Let  $\tilde{X}^{(\beta)}$  be solution to the stochastic differential equation given by (8) with  $\tilde{U}$  instead of  $U$  and  $(B_t)_{t \geq 0}$  the Brownian motion governing  $X^{(\beta)}$ . The process  $\tilde{X}^{(\beta)}$  satisfies (11) because  $\tilde{U}$  satisfies (10). Thus, by Theorem 7 (ii) (applied to  $\tilde{X}^{(\beta)}$ ) and Lemma 11 (applied to  $X^{(\beta)}$ ), there exist  $0 < \epsilon_1 < \epsilon_0 \leq \epsilon$  such that whenever  $\|x\| \leq \epsilon_1$ ,

$$\begin{aligned} \frac{1}{2} &\leq \mathbb{P}_x[\tau^{\epsilon_0} = \infty] = \mathbb{P}_{x,x} \left[ \limsup_{t \rightarrow +\infty} \frac{\ln(\|\tilde{X}^{(\beta)}(t)\|)}{t} \leq -\Lambda(A, \beta)(\beta - \beta_0); \tau^{\epsilon_0} = \infty \right] \\ &= \mathbb{P}_x \left[ \limsup_{t \rightarrow +\infty} \frac{\ln(\|X^{(\beta)}(t)\|)}{t} \leq -\Lambda(A, \beta)(\beta - \beta_0); \tau^{\epsilon_0} = \infty \right]. \end{aligned}$$

Here  $\mathbb{P}_{x,x}$  stands for the law of  $(X^{(\beta)}, \tilde{X}^{(\beta)})$  starting from  $(x, x)$ . Since  $\mathbb{P}_x(\tau_{\epsilon_1} < \infty) > 0$  for all  $x \in \mathbb{R}^m \setminus \{0\}$ , the proof of (9) follows.

Finally (iii) is an immediate consequence of the second part of Lemma 11 (recall that 0 was an arbitrary point of  $\mathcal{U}$ , up to a translation).

## 4 Proof of Theorem 3

### 4.1 Proof of Theorem 3 (i)

The proof is similar to that of Theorem 7. We begin by proving a Riemannian version of Lemma 11. The proof of Theorem 3 (i) will then follow by an argument similar to that given at the end of Section 3.

Let  $y \in \mathcal{U}$  and let  $B_M(y, \epsilon)$  be the Riemannian ball with center  $y$  and radius  $\epsilon$ , where  $\epsilon > 0$  is sufficiently small so that

- the only critical point for  $U$  in  $B_M(y, \epsilon)$  is  $y$ ,
- the exponential mapping  $\exp_y : T_y M \rightarrow M$  is a diffeomorphism between the tangent ball  $B(0, \epsilon)$  of  $T_y M$  and  $B_M(y, \epsilon)$ .

Recall that the exponential mapping  $\exp_y : T_y M \rightarrow M$  associates to any tangent vector  $v \in T_y M$  the point  $x \in M$  which is the position at time 1 of the (constant speed) geodesic starting at time 0 from  $y$  with speed  $v$ .

Consider  $(e_1, e_2, \dots, e_m)$  an orthonormal basis of  $T_y M$  consisting of eigenvectors associated to the eigenvalues  $(\lambda_1, \lambda_2, \dots, \lambda_m)$  of the Hessian of  $U$  at the critical point  $y$ . A priori this Hessian is a bilinear form on  $T_y M$ , but the Euclidean structure of  $T_y M$  enables us to see it as a symmetric endomorphism on  $T_y M$ , and  $(\lambda_1, \lambda_2, \dots, \lambda_m)$  and  $(e_1, e_2, \dots, e_m)$  correspond to its spectral decomposition.

Let  $(v_1, v_2, \dots, v_m)$  be the coordinate system associated to  $(e_1, e_2, \dots, e_m)$  on  $B(0, \epsilon)$ . Such a coordinate system based on the exponential mapping is said to be a normal. From now on and until the end of this section, we identify a map  $f : B_M(y, \epsilon) \rightarrow \mathbb{R}$  with  $f \circ \exp_y : B(0, \epsilon) \rightarrow \mathbb{R}$ , and write  $f(v)$  for  $f \circ \exp_y(v)$ . Under this identification, the matrix corresponding to the Hessian at  $y$  admits the classical form

$$(\partial_{k,l} U(0))_{k,l \in \llbracket m \rrbracket}$$

where  $\partial_k$  is a shorthand for  $\frac{\partial}{\partial v_k}$ . The introduction of the lecture notes of Pennec [14] is a convenient reference for these assertions (a more thorough exposition can be found in the book of Gallot, Hulin and Lafontaine [7]).

A first interest of the normal coordinate system  $(v_1, v_2, \dots, v_m)$  on  $B(0, \epsilon)$  is that we can consider the corresponding polar decomposition as in the previous section: each  $v = (v_1, v_2, \dots, v_m) \in B(0, \epsilon) \setminus \{0\}$  can be uniquely written under the form  $\rho\theta$  with  $\rho \in (0, \epsilon)$  and  $\theta \in \mathbb{S}^{m-1}$ , where the basis  $(e_1, e_2, \dots, e_m)$  enables us to identify  $T_y M$  with  $\mathbb{R}^m$ .

Before going further, let us recall some other traditional notations and facts from Riemannian geometry. For any  $v \in B(0, \epsilon)$ , denote  $g(v) := (g_{k,l}(v))_{k,l \in \llbracket m \rrbracket}$  the matrix of the pull-back of the Riemannian metric: for any vectors  $b$  and  $\tilde{b}$  from  $T_{\exp_y(v)} M$ , identified with their coordinates  $(b_k)_{k \in \llbracket m \rrbracket}$  and  $(\tilde{b}_k)_{k \in \llbracket m \rrbracket}$  in the basis  $(\partial_k)_{k \in \llbracket m \rrbracket}$ , we have

$$\langle b, \tilde{b} \rangle_v = \sum_{k,l \in \llbracket m \rrbracket} g_{k,l}(v) b_k \tilde{b}_l$$

The determinant of  $g(v)$  and the inverse matrix  $g^{-1}(v)$  are respectively denoted  $|g|(v)$  and  $(g^{k,l}(v))_{k,l \in \llbracket m \rrbracket}$ . For any smooth function  $f$ , the expressions of its gradient and Laplacian are given by

$$\begin{aligned} \nabla f(v) &= \left( \sum_{l \in \llbracket m \rrbracket} g^{k,l}(v) \partial_l f(v) \right)_{k \in \llbracket m \rrbracket} \\ \Delta f(v) &= \frac{1}{\sqrt{|g|(v)}} \sum_{k,l \in \llbracket m \rrbracket} \partial_k \left( \sqrt{|g|} g^{k,l} \partial_l f \right) (v) \\ &= \sum_{k,l \in \llbracket m \rrbracket} g^{k,l}(v) \left( \partial_{k,l} f(v) - \sum_{j \in \llbracket m \rrbracket} \Gamma_{k,l}^j(v) \partial_j f(v) \right) \end{aligned}$$

where  $\Gamma_{k,l}^j(v)$  are the Christoffel symbols at  $v$ , see for instance the listing [16] (again we abuse notation in the r.h.s by identifying  $f$  with its formulation in the coordinate system  $v = (v_1, v_2, \dots, v_m)$ ).

A second interest of the normal coordinate system is that at 0, we recover the usual notions:  $g(0)$  is the identity matrix and the Christoffel symbols all vanishes at 0.

The above expressions lead to the following formulation of the generator  $L_\beta$  defined in (1):

$$L_\beta \cdot = U \sum_{k,l \in \llbracket m \rrbracket} g^{k,l} \left( \partial_{k,l} \cdot - \sum_{j \in \llbracket m \rrbracket} \Gamma_{k,l}^j \partial_j \cdot \right) - \beta \sum_{k,l \in \llbracket m \rrbracket} g^{k,l} \partial_k U \partial_l \cdot \quad (25)$$

Again we are slightly abusing notations by calling it  $L_\beta$  too, especially as we see it as only defined on  $\mathcal{C}^2(B(0, \epsilon))$ .

The associate carré du champ is given as

$$\Gamma_{L_\beta} \cdot = 2U \sum_{k,l \in \llbracket m \rrbracket} g^{k,l} \partial_k \cdot \partial_l \cdot \quad (26)$$

(this is a consequence of the algebraic relation  $\Gamma_{\partial_k \partial_l} \cdot = 2\partial_k \cdot \partial_l \cdot$ , even if  $\partial_k \partial_l$  is not a Markov generator, i.e. when  $k \neq l$ ).

Consider the mapping  $P$  associated in (15) to the polar decomposition. Since  $P$  is invertible from  $\mathcal{C}^2(B(0, \epsilon))$  to  $\mathcal{C}^2((0, \epsilon) \times \mathbb{S}^{m-1})$ , there is a unique diffusion generator  $\mathsf{L}_\beta$  acting on  $\mathcal{C}^2((0, \epsilon) \times \mathbb{S}^{m-1})$  such that

$$\mathsf{L}_\beta \circ P = P \circ L_\beta$$

To compute  $\mathsf{L}_\beta$ , let us write that for any  $v \in B(0, \epsilon) \setminus \{0\}$ ,

$$\begin{aligned} \rho &= \sqrt{\sum_{k \in \llbracket m \rrbracket} v_k^2} \\ \forall l \in \llbracket m \rrbracket, \quad \theta_l &= \frac{v_l}{\rho} \end{aligned}$$

It follows that for any  $k \in \llbracket m \rrbracket$ ,

$$\begin{aligned} \partial_k \rho &= \frac{v_k}{\rho} = \theta_k \\ \forall l \in \llbracket m \rrbracket, \quad \partial_k \theta_l &= \frac{\delta_{k,l}}{\rho} - \frac{v_l}{\rho^2} \partial_k \rho = \frac{1}{\rho} (\delta_{k,l} - \theta_k \theta_l) \end{aligned}$$

where  $\delta_{k,l}$  is the Kronecker symbol.

It follows that

$$\partial_k = \theta_k \partial_\rho + \frac{1}{\rho} \sum_{l \in \llbracket m \rrbracket} (\delta_{k,l} - \theta_k \theta_l) \partial_{\theta_l} \quad (27)$$

and by composition, for any  $k, l \in \llbracket m \rrbracket$ , we can also write  $\partial_{k,l}$  in terms of  $\partial_\rho$ ,  $\partial_\rho^2$ ,  $\partial_{\theta_i}$  and  $\partial_{\theta_i \theta_j}$ , for  $i, j \in \llbracket m \rrbracket$ . Replacing these expressions in (25), we get the formula for  $\mathsf{L}_\beta$  in terms of differentiations of order 1 and 2, with respect to  $\rho$  and the  $\theta_l$ ,  $l \in \llbracket m \rrbracket$ .

In order to apply the general method of [4] as in Section 3, we need to check the three facts respectively listed in the following lemmas.

**Lemma 12** *For any  $F \in \mathcal{C}^2([0, \epsilon) \times \mathbb{S}^{m-1})$ , we have, uniformly over  $\theta \in \mathbb{S}^{m-1}$ ,*

$$\lim_{\rho \rightarrow 0_+} L_\beta[F](\rho, \theta) = G_\beta[F(0, \cdot)](\theta)$$

where  $G_\beta$  is given in (17).

## Proof

For any  $v \in B(0, \epsilon)$ , define

$$\begin{aligned} \forall k, l \in \llbracket m \rrbracket, \quad \tilde{g}^{k,l}(v) &:= \delta_{k,l} \\ \forall j, k, l \in \llbracket m \rrbracket, \quad \tilde{\Gamma}_{k,l}^j(v) &:= 0 \end{aligned}$$

and in analogy with (25),

$$\tilde{L}_\beta \cdot = U \sum_{k,l \in \llbracket m \rrbracket} \tilde{g}^{k,l} \left( \partial_{k,l} \cdot - \sum_{j \in \llbracket m \rrbracket} \tilde{\Gamma}_{k,l}^j \partial_j \cdot \right) - \beta \sum_{k,l \in \llbracket m \rrbracket} \tilde{g}^{k,l} \partial_k U \partial_l \cdot$$

This operator coincides with the restriction of (7) to  $B(0, \epsilon)$ . It follows from (21) that uniformly over  $\theta \in \mathbb{S}^{m-1}$ ,

$$\lim_{\rho \rightarrow 0_+} \tilde{L}_\beta[F](\rho, \theta) = G_\beta[F(0, \cdot)](\theta)$$

where the operator  $\tilde{L}_\beta$  is such that  $\tilde{L}_\beta \circ P = P \circ \tilde{L}_\beta$ .

Thus to get the wanted result, it is sufficient to show that

$$\lim_{\rho \rightarrow 0_+} (L_\beta - \tilde{L}_\beta)[F](\rho, \theta) = 0 \quad (28)$$

This convergence is a consequence of the writing

$$\begin{aligned} (L_\beta - \tilde{L}_\beta)[F] &= U \sum_{k,l \in \llbracket m \rrbracket} (g^{k,l} - \tilde{g}^{k,l}) \left( \partial_{k,l} F - \sum_{j \in \llbracket m \rrbracket} \Gamma_{k,l}^j \partial_j F \right) - U \sum_{k,l,j \in \llbracket m \rrbracket} \tilde{g}^{k,l} (\Gamma_{k,l}^j - \tilde{\Gamma}_{k,l}^j) \partial_j F \\ &\quad - \beta \sum_{k,l \in \llbracket m \rrbracket} (g^{k,l} - \tilde{g}^{k,l}) \partial_k U \partial_l F \end{aligned}$$

and of the following facts, valid uniformly in  $\theta \in \mathbb{S}^{m-1}$  as  $\rho$  goes to  $0_+$ :

- According to (27), for any  $k, l \in \llbracket m \rrbracket$ ,  $\partial_k F$  is of order  $1/\rho$  and  $\partial_{k,l} F$  is of order  $1/\rho^2$ .
- Due to the regularity of  $g$  and of the Christoffel symbols, for any  $k, l \in \llbracket m \rrbracket$ ,  $g^{k,l} - \tilde{g}^{k,l}$  and  $\Gamma_{k,l}^j - \tilde{\Gamma}_{k,l}^j$  are of order  $\rho$ .
- By the assumption that  $y$  is a global minimal,  $U$  is of order  $\rho^2$  and  $\partial_k U$  is of order  $\rho$ , for any  $k \in \llbracket m \rrbracket$ .

■

We have seen in the previous section that  $G_\beta$  is reversible with respect to the probability measure  $\mu_{A,\beta}$  defined in (2), where here  $A := A_y$  is the diagonal matrix whose entries are the eigenvalues of the Hessian of  $U$  at  $y \in \mathcal{U}$ . To continue the method of [4], we also need the two following ingredients.

**Lemma 13** *Consider the function  $V$  defined on  $(0, \epsilon) \times \mathbb{S}^{m-1}$  via*

$$V(\rho, \theta) := -\ln(\rho).$$

*The function  $\Gamma_{L_\beta}[V]$  is bounded on  $(0, \epsilon) \times \mathbb{S}^{m-1}$  and the function  $L_\beta[V]$  can be extended into a continuous function  $H_\beta$  on  $[0, \epsilon) \times \mathbb{S}^{m-1}$  satisfying (24) and thus  $\mu_{A,\beta}[H_\beta(0, \cdot)] = \Lambda(A, \beta)(\beta - \beta_0)$ .*

## Proof

We have  $\Gamma_{L_\beta}[V] = P[\Gamma_{L_\beta}[V]]$  with  $V(v) = -\frac{1}{2} \ln(\sum_{k \in \llbracket m \rrbracket} v_k^2)$ , so it is sufficient to see that  $\Gamma_{L_\beta}[V]$  is bounded on  $B_M(y, \epsilon) \setminus \{y\}$ . Expanding  $U(v)$  near 0 in the normal coordinate system  $v = (v_1, v_2, \dots, v_d)$ , we get for  $v$  small

$$U(v) \sim \frac{1}{2} \sum_{l \in \llbracket m \rrbracket} \lambda_l v_l^2$$

Hence, using (26),

$$\Gamma_{L_\beta}[V](v) \sim \frac{\sum_{l \in \llbracket m \rrbracket} \lambda_l v_l^2}{(\sum_{l \in \llbracket m \rrbracket} v_l^2)^2} \sum_{k, l \in \llbracket m \rrbracket} g^{k, l}(0) v_k v_l = \frac{\sum_{l \in \llbracket m \rrbracket} \lambda_l v_l^2}{\sum_{l \in \llbracket m \rrbracket} v_l^2}.$$

(see also [13]).

This proves the wanted boundedness.

For the wanted convergence, in view of the computations of the previous section, it is sufficient to see that (28) holds with  $F$  replaced by  $V$ . Note that when applied to a function only depending on  $\rho$ , as  $V$ , (27) reduce to  $\partial_k = \theta_k \partial_\rho$ . It follows that  $\partial_k V$  is of order  $1/\rho$  and  $\partial_{k, l} V$  is of order  $1/\rho^2$ . This observation enables us to use the same arguments as in the end of the proof of Lemma 12 to conclude that (28) holds with  $F$  replaced by  $V$ .  $\blacksquare$

A Riemannian version of Lemma 11 follows directly from the preceding lemma, the proof being exactly the same as the proof of Lemma 11. The proof of Theorem 3 (i) then follows (almost) verbatim along the lines of the arguments given in the preceding section just after the proof of Lemma 11.

## 4.2 Proof of Theorem 3 (ii)

Let  $V : N \rightarrow \mathbb{R}, x \mapsto \ln(U(x)^{-\beta})$ . Observe that for all  $f \in C^2(N)$ ,

$$\operatorname{div}(e^V \nabla f) = e^V (\langle \nabla V, \nabla f \rangle + \Delta f) = U^{-\beta-1} L_\beta f.$$

Let  $C_c^2(N)$  be the set of  $f \in C^2(N)$  having compact support. Then, for all  $f \in C_c^2(N)$ ,

$$\int_N L_\beta f d\ell_\beta = 0,$$

where  $\ell_\beta$  is the measure on  $N$  defined as

$$\ell_\beta(dx) := U(x)^{-(1+\beta)} \ell(dx).$$

Let  $p \in \mathcal{U}$ . By Morse's lemma, there is a smooth chart at  $p$  such that, in this chart system,  $U$  writes  $x \mapsto \|x\|^2 = \sum_{i=1}^m x_i^2$ . Since the map  $x \mapsto \|x\|^{-2(\beta+1)}$  is locally integrable (i.e. in a neighborhood of  $0_{\mathbb{R}^m}$ ) if and only if  $2(\beta+1) < m$ , it comes that  $\int_N U(x)^{-(1+\beta)} \ell(dx) < \infty$  if and only if  $2(\beta+1) < m$ , that is  $\beta < \beta_0$ .

Assuming  $\beta < \beta_0$ , the probability measure

$$\pi_\beta(dx) := \frac{1}{C_\beta} \ell_\beta(dx)$$

(where  $C_\beta$  is a normalization constant) satisfies

$$\int_N L_\beta f d\pi_\beta = 0 \tag{29}$$

for all  $f \in C_c^2(N)$ . Observe that there is no evidence that the set  $C_c^2(N)$  is a core for  $\mathcal{L}_\beta$ , so that we cannot immediately deduce from (29) that  $\pi_\beta$  is an invariant probability measure of  $X^{(\beta)}$ . However, by Theorem 9.17 page 248 in Ethier and Kurtz [6] (originally due to Echeverria [5]) the following properties (a) – (d) below ensure that  $\pi_\beta$  is invariant:

- (a) The space  $N$  is a separable locally compact metric space (for which the space  $\hat{C}(N)$  of continuous function "vanishing at infinity" coincide with  $\{f \in C^0(M) : f|_{\mathcal{U}} = 0\}$ );
- (b) The set  $C_c^2(N)$  is an algebra dense in  $\hat{C}(N)$ ;
- (c) The operator  $L_\beta : C_c^2(N) \rightarrow \hat{C}(N)$ , satisfies the positive maximum principle;
- (d) The martingale problem for  $(L_\beta, C_c^2(N))$  is well-posed: for all  $x \in N$ ,  $\mathbb{P}_x^\beta$  (the law of  $X^{(\beta)}$  starting from  $X^{(\beta)}(0) = x$ ) is the unique probability on  $D([0, \infty), N)$  such that  $f(X(t)) - \int_0^t L_\beta f(X(s)) ds$  is a  $\mathbb{P}_x^\beta$ -martingale and  $\mathbb{P}_x^\beta[X(0) = x] = 1$ , where  $(X(t))_{t \geq 0}$  is the canonical process on  $D([0, \infty), N)$ .

Properties (a) – (c) are easy to verify. Property (d) follows from, on one hand, that for any  $\epsilon > 0$  sufficiently small, the stopped martingale problem on  $N_\epsilon := \{x \in M : U(x) \geq \epsilon\}$  is well-posed by uniform ellipticity of  $L^{(\beta)}$  on  $N_\epsilon$ , and on the other hand, that these localized martingale problems can next be extended to the whole state space  $N$ . For instance, corresponding precise statements are found in Ethier and Kurtz [6], see Theorem 5.4 page 199, providing the existence of a solution of the stopped martingale problem on the  $N_\epsilon$ , but also of the martingale problem on  $N$ , Theorem 4.1 page 182 for the uniqueness of stopped martingale problems on the  $N_\epsilon$ , and Theorem 6.2 page 217, for the deduction of the uniqueness of the solution of the martingale problem on  $N$  by localization.

•(ii)(a) : follows from the fact that a strong Feller process on a connected space having an invariant probability measure with full support, is positive recurrent (see e.g. [3], Corollary 7.10 for a statement on discrete time Markov chains and Proposition 4.58 (ii) for the application in continuous time). In particular, it is uniquely ergodic (i.e. its invariant probability measure is unique). Here the strong Feller property of  $X^{(\beta)}$  on  $N$  follows from Proposition 1.

•(ii)(b) : The following lemma is a consequence of Lemma 13 and the stochastic persistence approach exposed in [2], [4].

**Lemma 14** *Assume  $\beta < \beta_0$ . Then, there exist a continuous map  $W : N \rightarrow \mathbb{R}^+$ ,  $0 \leq \rho < 1$ ,  $\chi > 0$ ,  $\kappa \geq 0$  and  $T > 0$  such that*

(i)  $W(x) = d(x, \mathcal{U})^{-\chi}$  on a neighborhood of  $\mathcal{U}$ ,

(ii)  $P_T^{(\beta)} W \leq \rho W + \kappa$ .

### Proof

For  $y \in \mathcal{U}$ , and  $\epsilon > 0$  sufficiently small, let  $V_y : M \setminus \{y\} \rightarrow \mathbb{R}^+$  be a smooth map such that

$$P[V_y \circ \exp_y](\rho, \theta) = V(\rho, \theta) := -\ln(\rho)$$

whenever  $\rho < \epsilon$ , where, using the notation of Section 4.1,  $V : (0, \epsilon) \times \mathbb{S}^{m-1} \rightarrow \mathbb{R}$  is as in Lemma 13 and  $P$  is the mapping induced by the polar decomposition as in (15). Because  $\Gamma_{L_\beta}[V]$  is bounded on  $(0, \epsilon) \times \mathbb{S}^{m-1}$  and  $\mu_{A, \beta}[H_\beta(0, \cdot)] = \Lambda(A, \beta)(\beta - \beta_0) < 0$ , it is possible, for  $\epsilon$  sufficiently small, to find numbers  $\chi, T > 0, \kappa$  and  $0 \leq \rho < 1$  such that

$$P_T^{(\beta)}(e^{\chi V_y}) \leq \rho e^{\chi V_y} + \kappa$$

on  $M \setminus \{y\}$  (see [2], Proposition 8.2). The mapping  $W : N \rightarrow \mathbb{R}^+$ , defined as  $W(x) = \sum_{y \in \mathcal{U}} e^{\chi V_y}$  satisfies the conditions of the lemma. ■

By ellipticity of  $L^{(\beta)}$  on  $N$ , every point  $p \in N$  is an *accessible Doebelin* point for  $P_T^{(\beta)}$ . Combined with the preceding lemma this proves assertion (ii)(b) of Theorem 3 (see e.g. Theorem 8.15 in [3]).

### 4.3 Proof of Theorem 3 (iii)

It follows from compactness of  $M$  and Feller continuity of  $X^{(\beta)}$  that, with  $\mathbb{P}_x$  probability one, every limit point (for the weak\* topology) of the family

$$\left\{ \frac{1}{t} \int_0^t \delta_{X_s^{(\beta)}} ds \right\}_{t \geq 0}$$

is an invariant probability of  $X^{(\beta)}$  (see e.g. [3], Theorem 4.20 combined with Propositions 4.57 and 4.58). It then suffices to show that for  $\beta = \beta_0$ , every invariant probability of  $X^{(\beta_0)}$  is supported by  $\mathcal{U}$ , or equivalently, that every ergodic probability measure of  $X^{(\beta_0)}$  is a Dirac measure  $\delta_p$  for some  $p \in \mathcal{U}$ . We proceed by contradiction. Suppose that there exists an ergodic probability measure of  $X^{(\beta_0)}$ ,  $\mu$  with  $\mu(N) > 0$ . Then  $\mu(N) = 1$  (by invariance of  $N$ ) and, by ellipticity of  $X^{(\beta_0)}$  on  $N$ ,  $\mu$  is absolutely continuous with respect to  $\ell(dx)$ , hence also with respect to  $\ell_{\beta_0}(dx)$ . That is  $\mu(dx) = f(x)\ell_{\beta_0}(dx)$  with  $f \geq 0$  measurable and  $\ell_{\beta_0}[f] = 1$ . We claim that  $f$  is almost surely constant. This is in contradiction with the fact that  $\ell_{\beta_0}(N) = \infty$ . It remains to prove the claim. First assume that  $\|f\|_\infty = \sup_{x \in N} |f(x)| < \infty$ . Then,  $f \in L^2(\ell_{\beta_0})$  because  $\ell_{\beta_0}[f^2] = \mu[f] \leq \|f\|_\infty$ . Thus,

$$\ell_{\beta_0}[(P_t^{\beta_0} f - f)^2] = \ell_{\beta_0}[(P_t^{\beta_0} f)^2 + g]$$

where  $g := f^2 - 2fP_t^{\beta_0} f \in L^1(\ell_{\beta_0})$  and  $\ell_{\beta_0}[g] = -\mu[f]$ . Thus,

$$\ell_{\beta_0}[(P_t^{\beta_0} f - f)^2] = \ell_{\beta_0}[(P_t^{\beta_0} f)^2] - \mu(f) = \ell_{\beta_0}[(P_t^{\beta_0} f)^2 - f^2] \leq 0$$

where the last inequality follows from Jensen's inequality. This shows that  $\ell_{\beta_0}$ -almost surely,  $P_t^{\beta_0} f = f$ , and also  $\mu$ -almost surely. By ergodicity  $f$  is  $\mu$ -almost surely constant. Suppose now that  $\|f\|_\infty = \infty$ . Set  $f_n = \min\{f, n\}$  and  $\mu_n(dx) = f_n(x)\ell_{\beta_0}(dx)$ . For every Borel set  $A \subset N$ ,

$$\begin{aligned} (\mu_n P_t^{\beta_0})(A) &= \mu_n P_t^{\beta_0}(A \cap \{f \leq n\}) + (\mu_n P_t^{\beta_0})(A \cap \{f > n\}) \leq (\mu P_t^{\beta_0})(A \cap \{f \leq n\}) + n(\ell_{\beta_0} P_t^{\beta_0})(A \cap \{f > n\}) \\ &= \mu(A \cap \{f \leq n\}) + n\ell_{\beta_0}(A \cap \{f > n\}) = \mu_n(A). \end{aligned}$$

This shows that  $\mu_n$  is excessive, hence invariant because every finite excessive measure is invariant (see e.g. [3], Lemma 4.25). By what precedes,  $f_n$  is  $\mu$ -almost surely constant. Thus  $f$  is  $\mu$ -almost surely constant. This concludes the proof of the claim.

## 5 Appendix

### 5.1 The diffusion process generated by $L_\beta$ and Proposition 1

Here we briefly explain how the diffusion  $X^{(\beta)}$  can be constructed and give a proof of Proposition 1.

By Nash's embedding theorem, we can assume without loss of generality that  $M$  is a Riemannian submanifold of  $\mathbb{R}^n$  (equipped with its Euclidean scalar product  $\langle \cdot, \cdot \rangle$ ) for some  $n$  sufficiently large. For reasons that will become clear shortly, we write  $\nabla_M, \Delta_M, \text{div}_M$  the gradient, Laplacian, and divergence on  $M$ , and  $\nabla, \text{div}$ , the gradient and divergence on  $\mathbb{R}^n$ . If  $F$  is a smooth vector field on  $M$  and  $\tilde{F}$  a smooth globally integrable vector field on  $\mathbb{R}^n$  such that  $\tilde{F}|_M = F$ , then  $\tilde{F}$  and  $F$ , induce operators on  $C^1(\mathbb{R}^n)$  and  $C^1(M)$  respectively defined by:

$$\tilde{F}(\tilde{f})(x) = \langle \nabla \tilde{f}(x), \tilde{F}(x) \rangle = \frac{d(\tilde{f} \circ \Psi_t(x))}{dt} \Big|_{t=0}$$

for all  $\tilde{f} \in C^1(\mathbb{R}^n)$ , and  $x \in \mathbb{R}^n$ ;

$$F(f)(x) = \langle \nabla_M f(x), F(x) \rangle = \frac{d(f \circ \Psi_t(x))}{dt} \Big|_{t=0}$$



for all  $f \in C^1(M)$ , and  $x \in M$ . In both formulae,  $(\Psi_t^i)_{t \in \mathbb{R}}$  denotes the flow on  $\mathbb{R}^n$  induced by  $\tilde{F}$ .

A direct consequence of the right hand side equalities is that

$$\tilde{F}(\tilde{f})|_M = F(f) \quad (30)$$

for every  $f \in C^1(M)$  and  $\tilde{f} \in C^1(\mathbb{R}^n)$  such that  $f = \tilde{f}|_M$ .

Let  $(e_1, \dots, e_n)$  be the canonical basis of  $\mathbb{R}^n$ . For  $i = 1, \dots, n$  and  $x \in M$ , let  $E_i(x) \in T_x M$  be the orthogonal projection of  $e_i$  onto  $T_x M$ . Let  $\tilde{E}_i$  be a smooth vector field on  $\mathbb{R}^n$ , having compact support, such that  $\tilde{E}_i|_M = E_i$ . It is not hard to show that such a vector field exists. One can, for example, proceed as follows. Let  $\mathcal{M} \subset \mathbb{R}^n$  be a normal tubular neighborhood of  $M$ . Every point  $y \in \mathcal{M}$  writes uniquely  $y = x + v$  with  $x \in M$  and  $v \in T_x M^\perp$ . The map  $r : \mathcal{M} \ni x + v \mapsto x \in M$ , is a smooth retraction. It suffices to set  $\tilde{E}_i(x) = \eta(x)E_i(r(x))$  if  $x \in \mathcal{M}$  and  $\tilde{E}_i(x) = 0$  otherwise, where  $0 \leq \eta \leq 1$  is a smooth function with compact support in  $\mathcal{M}$  such that  $\eta|_M = 1$ .

The following, key property, is proved in Stroock [15], Section 4.2.1. For the reader's convenience we provide an alternative short proof.

**Lemma 15** *For every  $f \in C^2(M)$  and  $\tilde{f} \in C^2(\mathbb{R}^n)$ , such that  $f = \tilde{f}|_M$ , one has*

$$\sum_{i=1}^N \tilde{E}_i^2(\tilde{f})|_M = \Delta_M(f)$$

### Proof

Let  $F$  be a  $C^1$  vector field on  $M$ , and  $\tilde{F}$  a  $C^1$  vector field on  $\mathbb{R}^n$  such that  $\tilde{F}|_M = F$ . For all  $x \in \mathbb{R}^n$   $\text{div} \tilde{F}(x)$  equals the trace of the Jacobian matrix  $D\tilde{F}(x)$ , while for all  $x \in M$ ,  $\text{div}_M F(x)$  equals the trace of the  $d \times d$  matrix  $(\langle D\tilde{F}(x)u_i, u_j \rangle)_{i,j}$  where  $u_1, \dots, u_d$  is an (arbitrary) orthonormal basis of  $T_x M$ . This has the interesting consequence that

$$\text{div}_M(F) = \text{div}(F \circ r)|_M$$

where  $r : \mathcal{M} \rightarrow M$  is the retraction defined above. Let  $f \in C^2(M)$ . Then,

$$\nabla_M f = \sum_{i=1}^n \langle \nabla_M f, e_i \rangle e_i = \sum_{i=1}^n \langle \nabla_M f, E_i \rangle e_i = \sum_{i=1}^n E_i(f) e_i.$$

Thus,

$$\begin{aligned} \Delta_M f &:= \text{div}_M(\nabla_M f) = \text{div}(\nabla_M(f) \circ r)|_M = \sum_{i=1}^n \text{div}[(E_i(f) \circ r)e_i]|_M \\ &= \sum_{i=1}^n \langle \nabla(E_i(f) \circ r)|_M, e_i \rangle = \sum_{i=1}^n \langle \nabla_M E_i(f), e_i \rangle = \sum_{i=1}^n E_i^2(f). \end{aligned}$$

Here we have used the fact that  $\nabla(f \circ r)|_M = \nabla_M f$  for all  $f \in C^1(M)$ . ■

Now, let  $\tilde{U} : \mathbb{R}^n \rightarrow \mathbb{R}_+$  be a smooth function such that  $\tilde{U}|_M = U$ ,  $\sqrt{\tilde{U}}$  is Lipschitz and  $\nabla \tilde{U}$  has compact support. For instance  $\tilde{U}(x) = \eta(x)U(r(x)) + 1 - \eta(x)$  for  $x \in \mathcal{M}$  and  $\tilde{U}(x) = 1$  otherwise, where  $\eta, r$  are as above. Here, the Lipschitz continuity of  $\sqrt{\tilde{U}}$  follows from the fact that  $r$  is smooth and that, by assumption, the zeroes of  $U$  are non-degenerate.

Consider the stochastic differential equation on  $\mathbb{R}^n$  defined by

$$\begin{aligned} dX(t) &= \left(-\beta - \frac{1}{2}\right) \nabla \tilde{U}(X(t)) dt \\ &\quad + \sum_{i=1}^n \left( \frac{1}{2} \langle \nabla \tilde{U}(X(t)), \tilde{E}_i(X(t)) \rangle \tilde{E}_i(X(t)) + \tilde{U}(X(t)) D\tilde{E}_i(X(t)) \cdot \tilde{E}_i(X(t)) \right) dt \\ &\quad + \sqrt{2\tilde{U}(X(t))} \sum_{i=1}^n \tilde{E}_i(X(t)) dB^i(t) \end{aligned} \quad (31)$$

where  $B = (B^1(t), \dots, B^n(t))_{t \geq 0}$  is a  $n$ -dimensional Brownian motion with  $B(0) = 0$ .

Since the coefficients of (31) are globally Lipschitz and bounded, the following properties (a), (b), (c) are classical (see e.g. Le Gall [12], Theorems 8.3 and 8.7 for (a) and (b), and Kunita [11], Theorem 4.5.1 for (c)) :

- (a) For all  $x \in \mathbb{R}^n$ , there is a unique strong solution  $\mathbb{R}_+ \ni t \mapsto X^{(\beta, x)}(t)$  to (31) such that  $X^{(\beta, x)}(0) = x$ ,
- (b) The process  $\tilde{X}^{(\beta)} := (X^{(\beta, x)})_{x \in \mathbb{R}^n}$  is a Feller Markov process on  $\mathbb{R}^n$  whose generator  $\tilde{\mathcal{L}}_\beta$  contains  $\mathcal{C}_c^2(\mathbb{R}^n)$ , the set of compactly supported  $\mathcal{C}^2$  functions, in its domain and such that for all  $\tilde{f} \in \mathcal{C}_c^2(\mathbb{R}^n)$ ,

$$\begin{aligned} \tilde{\mathcal{L}}_\beta(\tilde{f}) &= -\beta \langle \nabla \tilde{U}, \nabla \tilde{f} \rangle - \frac{1}{2} \langle \nabla \tilde{U}, \nabla \tilde{f} \rangle + \frac{1}{2} \sum_{i=1}^n \tilde{E}_i[\tilde{U}] \tilde{E}_i[\tilde{f}] + \tilde{U} \sum_{i=1}^n \tilde{E}_i^2(\tilde{f}) \\ &= -\beta \nabla \tilde{U}(\tilde{f}) - \frac{1}{2} \nabla \tilde{U}(\tilde{f}) + \frac{1}{2} \sum_{i=1}^n \tilde{E}_i[\tilde{U}] \tilde{E}_i[\tilde{f}] + \tilde{U} \sum_{i=1}^n \tilde{E}_i^2(\tilde{f}) \end{aligned} \quad (32)$$

- (c) The map  $x \mapsto X^{(\beta, x)}(t)$  is an homeomorphism. In particular,

$$\forall t \geq 0, X^{(\beta, x)}(t) \in \mathbb{R}^n \setminus \mathcal{U} \Leftrightarrow \exists t \geq 0, X^{(\beta, x)}(t) \in \mathbb{R}^n \setminus \mathcal{U}.$$

Set  $S_i(x) = \sqrt{2\tilde{U}(x)} \tilde{E}_i(x)$ . On  $\mathbb{R}^n \setminus \mathcal{U}$ , (31) can be rewritten, using Stratonovich formalism, as

$$\begin{aligned} dX(t) &= \left( \left( -\beta - \frac{1}{2} \right) \nabla \tilde{U}(X(t)) + \frac{1}{2} \sum_{i=1}^n D S_i(X(t)) S_i(X(t)) \right) dt + \sum_{i=1}^n S_i(X(t)) dB^i(t) \\ &= \left( -\beta - \frac{1}{2} \right) \nabla \tilde{U}(X(t)) + \sum_{i=1}^n S_i(X(t)) \circ dB^i(t). \end{aligned} \quad (33)$$

The vector fields  $\nabla \tilde{U}$  and  $S_i$ 's being tangent to  $N$ , this latter expression shows that  $N$  (hence  $M$ ) is invariant for  $X^{(\beta)}$ . That is:

$$\forall t \geq 0, X^{(\beta, x)}(t) \in N \text{ (resp. } M) \Leftrightarrow \exists t \geq 0, X^{(\beta, x)}(t) \in N \text{ (resp. } M).$$

It then follows that  $X^{(\beta)} := (X^{(\beta, x)})_{x \in M}$  is a Feller Markov process on  $M$ , leaving  $N$  invariant, whose generator  $\mathcal{L}_\beta$  contains  $\mathcal{C}^2(M)$  in its domain and such that  $\mathcal{L}_\beta f = \tilde{\mathcal{L}}_\beta \tilde{f}|_M = L_\beta f$  for all  $f \in \mathcal{C}^2(M)$  and  $\tilde{f} \in \mathcal{C}^2(\mathbb{R}^n)$  such that  $\tilde{f}|_M = f$ . The last equalities follows from Lemma 15 and (32), since on  $M$  we have

$$\langle \nabla \tilde{U}, \nabla \tilde{f} \rangle = \sum_{i=1}^n \tilde{E}_i[\tilde{U}] \tilde{E}_i[\tilde{f}]$$

The strong Feller property on  $N$  follows from the ellipticity of  $L_\beta$  on  $N$  (see e.g. Ichihara and Kunita [8, 9], Lemma 5.1).

## 5.2 On Remark 2

Given  $0 < \lambda_- < \lambda_+$ , and  $m \geq 2$ , let  $D(\lambda_-, \lambda_+, m)$  be the set of diagonal matrices with entries  $\lambda_- = \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_{m-1} \leq \lambda_m = \lambda_+$ . The set  $\{\Lambda(A, \beta) : A \in D(\lambda_-, \lambda_+, m)\}$  is a compact interval  $[\lambda_-(m, \beta), \lambda_+(m, \beta)]$  (as the image by a continuous map of the compact connected set  $D(\lambda_-, \lambda_+, m)$ ) contained in  $[\lambda_-, \lambda_+]$ .

Let  $A \in D(\lambda_-, \lambda_+, m)$  be the matrix with entries  $\lambda_1 = \dots \lambda_{m-1} = \lambda_-$  and  $\lambda_m = \lambda_+$ . Then

$$Z(\beta, A) = \int [\lambda_+ \theta_m^2 + \lambda_- (1 - \theta_m^2)]^{-\beta} \sigma(d\theta) = \mathbb{E} \left[ \left( \frac{\lambda_+ X_m^2 + \lambda_- (\sum_{i=1}^{m-1} X_i^2)}{\sum_{i=1}^m X_i^2} \right)^{-\beta} \right],$$

where  $X_1, \dots, X_m$  are i.i.d.  $\mathcal{N}(0,1)$  random variables. By the strong law of large numbers and dominated convergence, this quantity converges, as  $m \rightarrow \infty$ , toward  $\lambda_-^{-\beta}$ . Thus  $\lim_{m \rightarrow \infty} \lambda_-(m, \beta) = \lambda_-$ . Similarly,  $\lim_{m \rightarrow \infty} \lambda_+(m, \beta) = \lambda_+$ .

### 5.3 On spherical integrals

In (23) we could have considered another function  $\mathbf{V}$ . Indeed, our first choice was

$$\tilde{\mathbf{V}} := -\ln(U)$$

since it seemed somewhat more “intrinsic” with respect to  $U$ . It can be shown similarly that the points [a] and [b] following (23) equally hold, with  $\mathbf{V}$  replaced by  $\tilde{\mathbf{V}}$  and  $\mathbf{H}_\beta$  by  $\tilde{\mathbf{H}}_\beta$  given on  $\{0\} \times \mathbb{S}^{m-1}$  by

$$\forall \theta \in \mathbb{S}^{m-1}, \quad \tilde{\mathbf{H}}_\beta(0, \theta) := -\text{tr}(A) + 2(1 + \beta) \frac{\langle \theta, A^2 \theta \rangle}{\langle \theta, A \theta \rangle}$$

where we recall that  $A := \text{Hess } U(0)$ .

the sign of the quantity  $\mu_{A,\beta}[\tilde{\mathbf{H}}_\beta(0, \cdot)]$  can then be used to discriminate between the attractiveness and repulsivity of 0. In particular  $\mu_{A,\beta}[\mathbf{H}_\beta(0, \cdot)]$  and  $\mu_{A,\beta}[\tilde{\mathbf{H}}_\beta(0, \cdot)]$  must have the same sign. We tried to prove directly (without success!) that

$$2(1 + \beta)\mu_{A,\beta}[\phi_A] > \text{tr}(A) \Leftrightarrow \beta > \frac{m}{2} - 1 \quad (34)$$

$$2(1 + \beta)\mu_{A,\beta}[\phi_A] < \text{tr}(A) \Leftrightarrow \beta < \frac{m}{2} - 1 \quad (35)$$

with

$$\forall \theta \in \mathbb{S}^{m-1}, \quad \phi_A(\theta) := \frac{\langle \theta, A^2 \theta \rangle}{\langle \theta, A \theta \rangle}$$

A by-product of our computations is thus to show the validity of (34) and (35), which look as natural bounds on the corresponding spherical integrals for any given definite positive matrix  $A$ .

## References

- [1] Dominique Bakry, Ivan Gentil, and Michel Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Cham, 2014.
- [2] Michel Benaïm. Stochastic Persistence. *arXiv*, 1806.08450, 2018.
- [3] Michel Benaïm and Tobias Hurth. *Markov Chains on Metric Spaces A Short Course*, Universitext. Springer, 2022.
- [4] Michel Benaïm and Edouard Strickler. Random switching between vector fields having a common zero. *Ann. Appl. Probab.*, 29(1):326–375, 2019.
- [5] Pedro Echeverria E. A criterion for invariant measures of Markov processes. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete.*, 61:1–16, 1982.
- [6] Stewart N. Ethier and Thomas G. Kurtz. *Markov Processes, Characterization and Convergence* John Wiley and Sons, 1986.
- [7] Sylvestre Gallot, Dominique Hulin, and Jacques Lafontaine. *Riemannian geometry*. Universitext. Springer-Verlag, Berlin, third edition, 2004.

- [8] Kanji Ichihara, and Hiroshi Kunita. *A classification of the second order degenerate elliptic operators and its probabilistic characterization*, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete.*,30:235–254, 1974.
- [9] Kanji Ichihara and Hiroshi Kunita. Supplements and corrections to the paper: A classification of the second order degenerate elliptic operators and its probabilistic characterization. *Z. Wahrscheinlichkeitstheor. Verw. Geb.*, 39:81–84, 1977.
- [10] Wolfgang Kliemann. *Recurrence and invariant measures for degenerate diffusions*, *Ann. Probab.*, 15(2):690–707, 1987.
- [11] Hiroshi Kunita. *Stochastic flows and stochastic differential equations*, volume 24 of Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge, 1990
- [12] Jean-François Le Gall. *Brownian motion, martingales, and stochastic calculus*, Graduate Texts in Mathematics, Springer, Cham, 2016.
- [13] Laurent Miclo. On the convergence of global-optimization fraudulent stochastic algorithms. Preprint available at <https://hal.science/hal-04094950>, April 2023.
- [14] Xavier Pennec. Probabilities and statistics on Riemannian manifolds : A geometric approach. Technical Report RR- 5093, inria-00071490, INRIA, 2004.
- [15] Daniel W. Stroock *An Introduction to the Analysis of Paths on a Riemmanian Manifold*, Mathematical Surveys and Monographs, Vol 74 American Mathematical Society, 2000.
- [16] Wikipedia contributors. List of formulas in Riemannian geometry — Wikipedia, the free encyclopedia. [https://en.wikipedia.org/w/index.php?title=List\\_of\\_formulas\\_in\\_Riemannian\\_geometry&oldid=1157164633](https://en.wikipedia.org/w/index.php?title=List_of_formulas_in_Riemannian_geometry&oldid=1157164633), 2023. [Online; accessed 15-October-2023].

michel.benaim@unine.ch

Institut de Mathématiques  
Faculté des sciences, Université de Neuchâtel  
11 rue E. Argand, 2000 Neuchâtel, Suisse

miclo@math.cnrs.fr

Toulouse School of Economics  
1, Esplanade de l’université  
31080 Toulouse cedex 6, France  
Institut de Mathématiques de Toulouse  
Université Paul Sabatier, 118, route de Narbonne  
31062 Toulouse cedex 9, France.