Large Deviations for Diffusion Models of Adaptive Dynamics

Nicolas Champagnat

October 23, 2003

Abstract

In order to study punctualism in the Darwinian evolution of a monomorphic population, we prove a large deviation principle for a diffusion model of adaptive dynamics constructed in Champagnat [2]. Because this diffusion process has degenerate and non-Lipschitz diffusion part, and non-continuous drift part, classical methods do not apply. We have to extend methods of Doss and Priouret [7] to achieve this. The biological motivation of punctualism leads us to examine whether this diffusion process a.s. never reaches particular points of the trait space — called evolutionary singularities — or not. We then apply these results to the problem of exit from a domain, which is the key question for punctualism: what are the exit point and the exit time from an attracting domain of the trait space containing a steady state of the unperturbed dynamics.

1 Introduction

The Darwinian evolution of a population, in which individuals are characterized by quantitative traits (such as height or time to maturity), results from birth and death processes involving mutation, and selection via the ecology of the system. The theoretical approach to this phenomenon has been initiated in the early 90s by Hofbauer and Sigmund [9], Marrow *et al.* [12] and Metz *et al.* [13]. This new approach to the so-called adaptive dynamics of an ecological system has thrown new light on fundamental issues of evolutionary biology: the origin and maintenance of genetic polymorphism within a population (Metz *et al.* [14]); the process of adaptive radiation whereby new species evolve (Dieckmann and Doebeli [4]).

Among the evolutionary phenomena that have not yet been interpreted in this framework, is the phenomenon of punctualism. Punctualism, as observed in the fossil record or experimental evolution, is a pattern of population states that alternates periods of evolutionary equilibrium with periods of rapid change (Rand and Wilson [16]). Typically, adaptive dynamics models describe the dynamics of an evolving population as a stochastic process in the trait space. We will be concerned here with a diffusion model of adaptive dynamics that has been obtained in Champagnat [2]. The description of the phenomenon of punctualism will appear as a consequence of a large deviation principle for this process.

Let us define this diffusion process. We will assume that the quantitative trait characterizing individuals belongs to a convex open subset $\mathcal{X}$ of $\mathbb{R}^d$. A population is called monomorphic when all its individuals hold the same trait value.

Adaptive dynamics models are based on two biological hypotheses: mutations are very rare (evolutionary and ecological time scales are separated, see [14]), and the population is large, so that changes in the population size are nearly deterministic. The hypothesis of very large population is the most unrealistic: it leads to models in which evolution is possible only in particular directions in the trait space, determined by the gradient of the fitness function (the function measuring the selective advantage of a mutant trait in a given population, see [13] and [14]). On the contrary, in finite populations, any mutant, even deleterious ones, could settle by chance in the resident population. In [2], we propose a model compatible with this requirement: this is a diffusion process allowing evolution in any direction of the trait space.

In the monomorphic case, in which we will be concerned in this paper, this process is weak solution on $\overline{\mathcal{X}}$ to the stochastic differential equation

$$dX_t^\varepsilon = (b(X_t^\varepsilon) + \varepsilon\tilde{b}(X_t^\varepsilon))dt + \sqrt{\varepsilon}\sigma(X_t^\varepsilon)dW_t, \tag{1}$$

where $b(x)$ and $\tilde{b}(x)$ are in $\mathbb{R}^d$, $\sigma(x)$ is a $d \times d$ symmetric real matrix, and $\varepsilon > 0$ is a small parameter scaling the size of the mutation jumps. The parameters b, $\tilde{b}$ and $\sigma\sigma^* = a$ are expressed in terms of individual biological parameters, listed below.

Let us denote the state of a monomorphic population by $x \in \mathcal{X}$.

$p(x, h)dh$ is the law of $h = x' - x$, where x' is a mutant trait born from an individual with trait x. $p(x, h)$ is defined on $\mathcal{X} \times \mathbb{R}^d$ and is an even function of h for any $x \in \mathcal{X}$ (*i.e.* $p(x, h)dh$ is a measure on $\mathbb{R}^d$ symmetric with respect to 0). Since x' must be in the trait space $\mathcal{X}$, the support of $h \mapsto p(x, h)$ is a subset of

$$\mathcal{X} - x = \{y - x; y \in \mathcal{X}\}.$$

The function $p : x \mapsto p(x, h)dh$ from $\mathcal{X}$ to the set of probability measures on $\mathbb{R}^d$ is extended to $\overline{\mathcal{X}}$ by setting $p(x) = \delta_0$ when x is in the boundary $\partial\mathcal{X}$ of $\mathcal{X}$, where δ_0 is the Dirac measure at 0.

$g(x', x)$ is a function from $\overline{\mathcal{X}} \times \overline{\mathcal{X}}$ to $\mathbb{R}$ that can be expressed in terms of biological parameters (see [2]), and that measures the *fitness*, *i.e.* the selective advantage (or disadvantage), of a single individual with trait x' in a monomorphic population made of individuals holding trait x. For biological reasons, we assume that $g = 0$ on the boundary of $\overline{\mathcal{X}} \times \overline{\mathcal{X}}$. This function satisfies the fundamental property of fitness functions:

$$\forall x \in \mathcal{X}, \quad g(x, x) = 0. \tag{2}$$

Note that the assumption that $h \mapsto p(x, h)$ is even is required for the construction of solutions to (1) in [2]. This assumption is almost always made in adaptive dynamics models (see *e.g.* Dieckmann and Law [5] or Kisdi [11]). See [2] for a discussion of the difficulties arising when $p(x, \cdot)$ is asymmetric.

We will assume that g is $\mathcal{C}^2$ on $\overline{\mathcal{X}} \times \overline{\mathcal{X}}$, and we will denote by $\nabla_1 g$ the gradient vector of $g(x', x)$ with respect to the first variable x', and by $H_{i,j}g$ the

Hessian matrix of $g(x', x)$ with respect to the i^{th} and j^{th} variables ($1 \leq i, j \leq 2$). Then, we can define the parameters b, $\tilde{b}$ and a of the SDE (1) as follows: let $b(x) = (b_1(x), \dots, b_d(x))$, $\tilde{b}(x) = (\tilde{b}_1(x), \dots, \tilde{b}_d(x))$, and $a(x) = (a_{kl}(x))_{1 \leq k,l \leq d}$. Then, as obtained in [2], for $x \in \overline{\mathcal{X}}$

$$b_k(x) = \int_{\mathbb{R}^d} h_k [\nabla_1 g(x,x) \cdot h]_+ p(x,h) dh,$$

$$\tilde{b}_k(x) = \begin{cases} \dfrac{1}{2} \displaystyle\int_{\{h \cdot \nabla_1 g(x,x) > 0\}} h_k (h^* H_{1,1} g(x,x) h) p(x,h) dh & \\ & \text{if} \quad \nabla_1 g(x,x) \neq 0, \\ 0 & \text{otherwise,} \end{cases} \tag{3}$$

$$\text{and} \quad a_{kl}(x) = \int_{\mathbb{R}^d} h_k h_l [h \cdot \nabla_1 g(x,x)]_+ p(x,h) dh.$$

Since $x \mapsto p(x,h)dh$ has been extended by δ_x on $\partial\mathcal{X}$, $b(x) = \tilde{b}(x) = 0$, and $a(x) = 0$ for $x \in \partial\mathcal{X}$. Let us also define

$$b^\varepsilon = b + \varepsilon \tilde{b}.$$

Weak existence for solutions to (1) has been shown in [2] under weaker assumptions than the one we will use here.

Let us also define three sets used throughout the paper:

$$\Gamma = \partial\mathcal{X} \cup \{x \in \mathcal{X} : \nabla_1 g(x,x) = 0\}, \tag{4}$$

$$\forall \alpha > 0, \ \Gamma_\alpha = \{x \in \mathcal{X} : d(x, \Gamma) \geq \alpha\}, \tag{5}$$

$$\text{and} \quad \forall c > 0, \ \tilde{\Gamma}_c = \{x \in \mathcal{X} : \forall s \in \mathbb{R}^d, \ s^* a(x) s \geq c\|s\|^2\} \tag{6}$$

The set Γ is called the set of *evolutionary singularities*, and, since $b = \tilde{b} = a = 0$ on Γ, points of Γ are possible rest points of solutions to (1).

This paper begins with some regularity results for the parameters a, b and $\tilde{b}$ (section 2). In section 3 we construct precisely the solution X^ε to (1) used in the following, we make observations about uniqueness and the strong Markov property of X^ε, and we give conditions under which the process X^ε a.s. never reaches Γ. In section 4, we adapt methods of Doss and Priouret [7], inspired from Azencott [1] to prove the main result of this paper: a Wentzell-Freidlin-like large deviation principle [8] for the paths of X^ε as $\varepsilon \to 0$. Finally (section 5), we apply this result to the problem of exit of an attracting domain (see [8]), which is the key question for interpreting the biological phenomenon of punctualism.

Notations

- $\mathcal{C}(I, E)$ (resp. $\mathcal{C}_x(I, E)$, $\mathcal{C}^{ac}(I, E)$, $\mathcal{C}^{ac}_x(I, E)$) is the set of continuous functions from an interval I of $\mathbb{R}_+$ containing 0 to a subset E of $\mathbb{R}^d$ (resp. with value $x \in E$ at 0, resp. absolutely continuous, resp. absolutely continuous with value x at 0).

- $\|\cdot\|$ denotes, for vectors of $\mathbb{R}^d$, the Euclidean norm, for real $d \times d$ matrices, the subordinate Euclidean norm $\|A\| = \sup_{\|h\|=1} \|Ah\|$ (remind that for symmetrical positive matrices, this norm equals the spectral norm — the greatest eigenvalue), and for functions, the $\mathbb{L}^\infty$ norm.

- $d(x,\Gamma)$ is the Euclidean distance from a point $x \in \mathbb{R}^d$ to a set $\Gamma \subset \mathbb{R}^d$.
- For $\varphi \in \mathcal{C}([0,T],\mathbb{R}^d)$ and $0 \le a < b \le T$, define

$$\|\varphi\|_{a,b} = \sup_{a \le t \le b} \|\varphi(t)\|, \tag{7}$$

and

$$B_b(\varphi,\delta) = \{\tilde{\varphi} \in \mathcal{C}([0,T],\mathbb{R}^d) : \|\tilde{\varphi} - \varphi\|_{0,b} \le \delta\}. \tag{8}$$

When $a = 0$ and $b = T$, $\|\cdot\|_{0,T}$ is the usual norm of uniform convergence in $\mathcal{C}([0,T],\mathbb{R}^d)$, and when $b = T$, $B_T(\varphi,\delta)$ is the usual closed ball centered at φ with radius δ in $\mathcal{C}([0,T],\mathbb{R}^d)$ for the norm of uniform convergence.

2 Study of a, b and $\tilde{b}$

Let us first define properly the function $\sigma(x)$ appearing in (1) from $a(x)$ defined in (3). We will use the notation $\mathcal{S}_+$ for the set of symmetrical non-negative $d \times d$ real matrices, and for any $c > 0$,

$$\mathcal{S}_c = \{a \in \mathcal{S}_+ : \forall s \in \mathbb{R}^d,\ s^* a s \ge c\|s\|^2\}. \tag{9}$$

Then

Proposition 2.1 *For any $a \in \mathcal{S}_+$, there exists a unique $\sigma \in \mathcal{S}_+$ such that $\sigma^2 = a$ (since σ is symmetrical, this can rewrite as $\sigma\sigma^* = a$). Let us call ζ the function from $\mathcal{S}_+$ to $\mathcal{S}_+$ that maps a on σ. Then ζ is Hölder with exponent $1/2$ on $\mathcal{S}_+$, and Lipschitz on $\mathcal{S}_c$ for any $c > 0$.*

Proof The construction of σ from a is standard: find an orthonormal basis of $\mathbb{R}^d$ where a is diagonal, put the square root of its elements in a new diagonal matrix, and express it back in the first basis. It follows from this change of basis that the uniqueness of σ is equivalent to the uniqueness of a solution $S \in \mathcal{S}_+$ to $S^2 = D$, where D is a diagonal matrix with non-negative diagonal elements. Two symmetrical commuting matrices diagonalize in the same orthonormal basis. Since $SD = S^3 = DS$, it is now easy to establish the uniqueness of σ.

Let $\tilde{\zeta} : \mathcal{S}_+ \to \mathcal{S}_+$ be defined by $\tilde{\zeta}(a) = a^2$. Then $\zeta \circ \tilde{\zeta} = \tilde{\zeta} \circ \zeta = \mathrm{Id}_{\mathcal{S}_+}$. The differential of $\tilde{\zeta}$ at a writes $d_a\tilde{\zeta}(h) = ah + ha$ and, when it is invertible, its inverse is the differential $d_{a^2}\zeta$ of ζ at a^2. Let $(ah+ha)_{ij}$ denote the i,j coefficient of the matrix $ah + ha$. Using the symmetry of a and h, an easy calculation gives that

$$\sum_{i,j} h_{ij}(ah+ha)_{ij} = \sum_{i=1}^{d} h_i^* a h_i + \sum_{j=1}^{d} h_j^* a h_j \tag{10}$$

where h_i is the i^{th} column (and, by symmetry, row) of the symmetrical matrix h. Let $\|\cdot\|$ be the spectral norm on $\mathcal{S}_+$. If $a \in \mathcal{S}_c$, the quantity (10) is greater than $2c\sum_i \|h_i\|^2 \ge Kc\|h\|^2$ for some constant K, and it is obviously smaller than $K'\|h\|\|ah+ha\|$ for another constant K'.

Hence, $\|ah + ha\| \ge \frac{Kc}{K'}\|h\|$ for any $h \in \mathcal{S}_+$ and $a \in \mathcal{S}_c$. In particular, $\|h\| = \|d_a\tilde{\zeta} \circ d_{a^2}\zeta(h)\| \ge \frac{Kc}{K'}\|d_{a^2}\zeta(h)\|$, which shows that ζ has a bounded differential on $\mathcal{S}_{c^2}$ for all $c > 0$. Therefore, ζ is Lipschitz on $\mathcal{S}_c$ for any $c > 0$.

The proof that ζ is Hölder on $\mathcal{S}_+$ is taken from Serre [17]. Let us consider two matrices a and b in $\mathcal{S}_+$. Then, there exists an orthonormal basis in which a is diagonal, and another one where b is diagonal. So, there are two orthogonal matrices U and V, and two diagonal matrices D_1 and D_2 with respective non-negative diagonal elements $\lambda_1, \ldots, \lambda_d$ and $\mu_1, \ldots, \mu_d$, such that $a = UD_1U^*$ and $b = UVD_2V^*U^*$.

Define $A = D_1$ and $B = VD_2V^*$, and denote by $\|\cdot\|_F$ the Schur-Frobenius norm on $d\times d$ matrices, given by $\|M\|_F = \sqrt{\sum_{i,j} m_{ij}^2}$ where $M = (m_{ij})_{1\leq i,j\leq d}$. Then,

$$\begin{aligned}
\|B-A\|_F^2 &= \sum_{i,j}\left(\sum_k v_{ik}\mu_k v_{jk} - \lambda_i\delta_{ij}\right)^2 \\
&= \sum_{i,j}\sum_{k,l} v_{ik}v_{il}v_{jk}v_{jl}\mu_k\mu_l - 2\sum_i\sum_k v_{ik}^2\lambda_i\mu_k + \sum_i \lambda_i^2 \\
&= \sum_k \mu_k^2 - 2\sum_i\sum_k v_{ik}^2\lambda_i\mu_k + \sum_i\lambda_i^2,
\end{aligned}$$

where δ_{ij} is the Kronecker delta symbol, and where we used the fact that $V = (v_{ij})_{1\leq i,j\leq d}$ is an orthogonal matrix to obtain the last line. This can be rewritten as

$$\|B-A\|_F^2 = \sum_{i,j} v_{ij}^2(\lambda_i - \mu_j)^2.$$

Now, observe that $\zeta(a) = U\zeta(A)U^* = U\zeta(D_1)U^*$, that $\zeta(b) = U\zeta(B)U^* = UV\zeta(D_2)V^*U^*$, and that $\zeta(D_1)$ and $\zeta(D_2)$ are diagonal matrices with respective elements $\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_d}$ and $\sqrt{\mu_1}, \ldots, \sqrt{\mu_d}$. Then, it follows from the fact that $(\sqrt{\lambda} - \sqrt{\mu})^2 \leq |\lambda - \mu|$ for $\lambda \geq 0$ and $\mu \geq 0$, and from the Cauchy-Schwartz inequality, that

$$\begin{aligned}
\|\zeta(A) - \zeta(B)\|_F^4 = \left(\sum_{i,j} v_{ij}^2(\sqrt{\lambda_i} - \sqrt{\mu_j})^2\right)^2 &\leq \left(\sum_{i,j} v_{ij}^2|\lambda_i - \mu_j|\right)^2 \\
&\leq \sum_{i,j} v_{ij}^2 \sum_{i,j} v_{ij}^2(\lambda_i - \mu_j)^2 = d\|A-B\|_F^2.
\end{aligned}$$

It remains to observe that, for any orthogonal matrix U and for any matrix M, $\|UMU^*\|_F = \|M\|_F$ to finally obtain that $\|\zeta(a) - \zeta(b)\|_F \leq d^{1/4}\sqrt{\|a-b\|_F}$ for any a and b in $\mathcal{S}_+$. □

Let us list here all the hypotheses needed in the following:

Hypotheses 2.1

(H1) *$g(x',x)$ is $\mathcal{C}^2$ with respect to the first variable x', $\nabla_1 g$ and $H_{1,1}g$ are bounded and Lipschitz on $\mathcal{X}^2$. For biological reasons, we will also assume that $\nabla_1 g(x,x) = 0$ when $x \in \partial\mathcal{X}$ (this fact will only be used in Theorem 3.2 of section 3.3).*

(H2) *$p(x,h)$ is Lipschitz with respect to the first variable x in the following* *sense: there exists a positive function m defined on $\mathbb{R}^d$ such that*

$$\forall (x,x')\in\mathcal{X}^2,\ \forall h\in\mathbb{R}^d,\ |p(x,h)-p(x',h)|\leq \|x-x'\|m(h), \tag{11}$$

$$\int_{\mathbb{R}^d}(1\vee\|h\|^3)m(h)dh<+\infty \tag{12}$$

$$\text{and}\quad \forall x\in\mathcal{X},\ \int_{\|h\|\geq d(x,\partial\mathcal{X})}(1\vee\|h\|^3)p(x,h)dh\leq Cd(x,\partial\mathcal{X}) \tag{13}$$

for some constant C. Note that (11) and (12) imply that $p(x,h)dh$ has finite third-order moment for all $x\in\mathcal{X}$. We will moreover assume that this third-order moment is bounded by a constant M_3 on $\mathcal{X}$.

(H2') *Assume (H2) with the additional assumption that for all $\alpha>0$, there exists a function $\bar{m}_\alpha:\mathbb{R}_+\to\mathbb{R}_+$ such that for all $x\in\Gamma_\alpha$,*

$$p(x,h)\leq\bar{m}_\alpha(\|h\|)$$
$$\text{and}\quad \int_{\mathbb{R}^d}\|h\|^3\bar{m}_\alpha(\|h\|)dh<+\infty. \tag{14}$$

(H3) *For all $\alpha>0$,* $\inf_{d(x,\partial\mathcal{X})\geq\alpha,\ u,v\in\mathbb{R}^d:\|u\|=\|v\|=1}\int|h\cdot u|^2|h\cdot v|p(x,h)dh>0.$

(H4) *For all $\alpha>0$, $\inf_{x\in\Gamma_\alpha}\|\nabla_1 g(x,x)\|>0$, where Γ_α has been defined in (5).*

When $\mathcal{X}=\mathbb{R}^d$, (H2) is true for example if $p(x,h)dh$ is Gaussian for all $x\in\mathcal{X}$ with covariance matrix $K(x)$ uniformly non-degenerate (*i.e.* in $\mathcal{S}_c$ for a given $c>0$), bounded by some constant $C>0$ and Lipschitz on $\mathbb{R}^d$. This fact can easily be proved by bounding the differential of the function $K\mapsto(\det K)^{-1/2}\exp(-x^*K^{-1}x/2)$ on $\{K\in\mathcal{S}_c:\|K\|\leq C\}$.

Note that condition (13) is a technical condition compatible with the fact that $p(x,\cdot)$ has been extended by δ_0 on $\partial\mathcal{X}$. The following lemma, ensured by assumption (H2), and the fact that $p(x,h)dh$ has finite and uniformly bounded third-order moment are the only conditions that will be necessary in the following. One could replace (H2) with any condition ensuring these facts.

Lemma 2.1 *For any continuous function $f:\mathbb{R}^d\to\mathbb{R}$ Lipschitz in a neighborhood of 0 and such that for any $h\in\mathbb{R}^d$, $|f(h)|\leq K(\|h\|^3\vee 1)$ for some constant K, the function*

$$x\mapsto\begin{cases}\int_S f(h)p(x,h)dh & \text{when } x\in\mathcal{X}\\ \int_S f(h)\delta_0(dh)=f(0) & \text{when } x\in\partial\mathcal{X}\end{cases} \tag{15}$$

is globally Lipschitz on $\overline{\mathcal{X}}$.

Proof of Lemma 2.1 Let us call ϕ the function (15). The fact that ϕ is globally Lipschitz on $\mathcal{X}$ follows easily from assumptions (11) and (12).

Fix $x \in \mathcal{X}$ and $x' \in \partial\mathcal{X}$. Decomposing $\mathbb{R}^d$ as $B(d(x,\partial\mathcal{X})) \cup B(d(x,\partial\mathcal{X}))^c$, where $B(\alpha) = \{h \in \mathbb{R}^d : \|h\| \leq \alpha\}$ yields, for x sufficiently close to x',

$$\begin{aligned}|\phi(x) - \phi(x')| &\leq \int_{\mathbb{R}^d} |f(0) - f(h)| p(x,h)dh \\ &\leq \int_{B(d(x,\partial\mathcal{X}))^c} (K + K(1 \vee \|h\|^3))p(x,h)dh \\ &\quad + \sup_{\|h\| \leq d(x,\partial\mathcal{X})} |f(0) - f(h)| \int_{B(d(x,\partial\mathcal{X}))} p(x,h)dh \\ &\leq 2KCd(x,\partial\mathcal{X}) + K'd(x,\partial\mathcal{X}),\end{aligned}$$

where the last inequality follows from (13) and from the fact that f is K'-Lipschitz in a neighborhood of 0. This completes the proof of Lemma 2.1 □

Remark 2.1 *This lemma is actually true under the following more general assumption: let ρ be the Kantorovich metric (see Rachev [15]) on the set of probability measures on $\mathbb{R}^d$ with finite third-order moments, defined by*

$$\rho(P_1, P_2) = \inf \int_{S^2} c(x,y) R(dx,dy),$$

where the infimum is taken over the set of measures $R(dx,dy)$ with marginals $P_1(dx)$ and $P_2(dy)$, and where $c(x,y) = \|x-y\| \max\{\|x\|^3, \|y\|^3, 1\}$. Then we could have assumed that the application $p : x \mapsto p(x,h)dh$ is Lipschitz for this metric ρ. We have chosen a simpler presentation to avoid a formalism which is not necessary in usual applications, and which does not give much benefit in this case of probability measures absolutely continuous with respect to the Lebesgue measure.

Assumption (H2') is necessary to control the probability measures $p(x,h)dh$ uniformly in any direction of $\mathbb{R}^d$ and uniformly in x. It will only be used to prove that $\tilde{b}$ is locally Lipschitz on $\mathcal{X} \setminus \Gamma$, and we will make use of this fact only in sections 3 and 5.

(H3) is only a technical condition needed to control the non-degeneracy of the matrix a. It means in fact that $p(x,h)dh$ gives a sufficient mass uniformly in any directions around x. Since the support of $p(x,h)dh$ is a subset of $\mathcal{X} - x$, this is possible only for x not too close to $\partial\mathcal{X}$. That is why the condition $d(x,\partial\mathcal{X}) \geq \alpha$ is necessary. For example, this assumption is true if there exists $\alpha > 0$ and $\beta > 0$ such that for all h and for x such that $d(x,\partial\mathcal{X}) < \alpha$, $p(x,h) \geq \beta \mathbf{1}_{\{\|x-h\| \leq d(x,\partial\mathcal{X}/2\}}$.

(H4) is obviously true if (H1) is true and $\overline{\mathcal{X}}$ is compact (then Γ_α is compact and the infimum of $\|\nabla_1 g(x,x)\|$ on this set is attained, and hence is positive). So, in the case where $\mathcal{X}$ is not bounded, (H4) only states that $\nabla_1 g(x,x)$ does not converge too fast to 0 when $\|x\| \to +\infty$.

Let us prove

Proposition 2.2

(i) *Assume (H1) and (H2). Then a and b are Lipschitz and bounded on $\mathcal{X}$, and $\tilde{b}$ is bounded on $\mathcal{X}$ and continuous on $\mathcal{X} \setminus \Gamma$. Under the additional assumptions (H2') and (H4), for all $\alpha > 0$, $\tilde{b}$ is Lipschitz on Γ_α, where Γ_α has been defined in (5).*

(ii) *The matrix a is symmetrical and non-negative on $\mathcal{X}$, $a(x) = 0$ if $x \in \Gamma$, and $a(x)$ is definite positive if $x \in \mathcal{X} \setminus \Gamma$. Proposition 2.1, allows us to define*

$$\forall x \in \overline{\mathcal{X}}, \quad \sigma(x) = \zeta(a(x)),$$

so that $\sigma\sigma^* = a$. Under assumptions (H1) and (H2), σ is Hölder with exponent $1/2$ on $\overline{\mathcal{X}}$.

(iii) *Assume (H3) and (H4). Then, $\forall \alpha > 0$, $\exists c > 0$ such that $\Gamma_\alpha \subset \tilde{\Gamma}_c$, where $\tilde{\Gamma}_c$ has been defined in (6). Under the additional assumptions (H1) and (H2), σ is Lipschitz on Γ_α for all $\alpha > 0$.*

Proof of (i) a, b and $\tilde{b}$ are trivially bounded.

By assumption (H1), $\nabla_1 g$ is K-Lipschitz on $\overline{\mathcal{X}}^2$ for some constant K. For any x and x' in $\mathcal{X}$, and for $1 \leq k \leq d$,

$$\begin{aligned}|b_k(x) - b_k(x')| \leq &\left| \int_{\mathbb{R}^d} h_k([\nabla_1 g(x,x) \cdot h]_+ - [\nabla_1 g(x',x') \cdot h]_+) p(x,h) dh \right| \\ &+ \left| \int_{\mathbb{R}^d} h_k [\nabla_1 g(x,x) \cdot h]_+ (p(x,h) - p(x',h)) dh \right|.\end{aligned}$$

Using the fact that $|[a]_+ - [b]_+| \leq |a - b|$, the first term of the right-hand side is less than $K(2\|x - x'\|)M_2$, where M_2 is a bound, given by (H2), for the second-order moments of $p(x,h)dh$ for $x \in \mathcal{X}$. The second term can be bounded by $K\|\nabla_1 g(x,x)\|\|x - x'\|$ because of Lemma 2.1. Since, by (H1), $\nabla_1 g$ is bounded on $\mathcal{X}^2$, it follows that b is Lipschitz on $\mathcal{X}$. A similar computation, using (13), extends this result to $\overline{\mathcal{X}}$. Similarly, a is Lipschitz on $\overline{\mathcal{X}}$.

Fix $\alpha > 0$ and x and x' in Γ_α. Define $S = \{h \in \mathbb{R}^d : h \cdot \nabla_1 g(x,x) > 0\}$ and $S' = \{h : h \cdot \nabla_1 g_i(x',x') > 0\}$. Then, it follows from (11) that

$$\begin{aligned}&|\tilde{b}_k(x) - \tilde{b}_k(x')| \\ &\quad \leq \frac{1}{2} \left| \int_{S \cap S'} h_k [h^*(H_{1,1} g(x,x) - H_{1,1} g(x',x'))h] p(x',h) dh \right. \\ &\qquad + \int_S h_k (h^* H_{1,1} g(x,x) h)(p(x,h) - p(x',h)) dh \\ &\qquad - \int_{S \cap S'^c} h_k (h^* H_{1,1} g(x,x) h) p(x',h) dh \\ &\qquad \left. - \int_{S^c \cap S'} h_k (h^* H_{1,1} g(x',x') h) p(x',h) dh \right| \\ &\quad \leq \frac{1}{2} \left(K\|x - x'\| \int_{\mathbb{R}^d} \|h\|^3 p(x',dh) \right. \\ &\qquad + \|H_{1,1} g(x,x)\| \|x - x'\| \int_{\mathbb{R}^d} \|h\|^3 m(h) dh \\ &\qquad \left. + \int_{(S \cap S'^c) \cup (S^c \cap S')} K\|h\|^3 p(x',h) dh \right),\end{aligned} \tag{16}$$

where $K > 0$ is such that $H_{1,1} g$ is K-Lipschitz and bounded by K. It follows from (H1) and (H2) that the first two terms of the right-hand side are bounded by a constant times $\|x-x'\|$. Note that the set $S\cap S'^c=\{h: h\cdot\nabla_1 g(x,x)>0,\ h\cdot\nabla_1 g(x',x')\leq 0\}$ converges to $\emptyset$ as $x\to x'$, and that $S^c\cap S'$ converges to $\{h: h\cdot\nabla_1 g(x,x)=0\}$, which has Lebesgue measure 0 since $\nabla_1 g(x,x)\neq 0$ ($x\notin\Gamma$), as $x\to x'$. So, by the dominated convergence Theorem, the last term of the right-hand side of (16) converges to 0 as $x\to x'$, and $\tilde{b}$ is continuous on $\mathcal{X}\setminus\Gamma$.

Under assumptions (H2') and (H4), in order to prove that $\tilde{b}$ is Lipschitz on Γ_α, it suffices to find $C>0$ such that

$$\int_{(S\cap S'^c)\cup(S^c\cap S')}\|h\|^3 p(x',h)dh=\int_{(S\cap S'^c)\cup(S^c\cap S')}\|h\|^3\bar{m}_\alpha(\|h\|)dh\leq C\|x-x'\|.$$

Since $\nabla_1 g$ is K-Lipschitz, $\|\nabla_1 g(x,x)-\nabla_1 g(x',x')\|\leq 2K\|x-x'\|$, so

$$h\cdot\nabla_1 g(x',x')>0\Rightarrow h\cdot\nabla_1 g(x,x)>-2K\|h\|\|x-x'\|.$$

This means that

$$S^c\cap S'\subset\{h: h\cdot\nabla_1 g(x,x)\in[-2K\|h\|\|x-x'\|,0]\}.$$

Now, by (H4), $\|\nabla_1 g(x,x)\|\geq c$ for some $c>0$ depending only on α, therefore

$$S^c\cap S'\subset\{h:\cos(h,\nabla_1 g(x,x))\in[-2K\|x-x'\|/c,0]\}.$$

Then, using the spherical coordinates change of variable,

$$\begin{aligned}&\int_{S^c\cap S'}\|h\|^3\bar{m}_\alpha(\|h\|)dh\\&\quad\leq\int_T\int_{-\pi/2}^{\pi/2}\cdots\int_{-\pi/2}^{\pi/2}\int_0^{+\infty}r^3\bar{m}_\alpha(r)|J(r,\theta_1,\dots,\theta_{d-2},\varphi)|drd\theta_1\dots d\theta_{d-2}d\varphi,\end{aligned}$$

where $T=\{\theta\in[-\pi,\pi]:\cos\theta\in[-2K\|x-x'\|/c,0]\}$, and J is the Jacobian of the change of variable of spherical coordinates. It trivially follows from the explicit development of the determinant that $|J(r,\theta_1,\dots,\theta_{d-2},\varphi)|\leq d!r^{d-1}$, so (14) implies that for $\|x-x'\|$ sufficiently small,

$$\begin{aligned}\int_{S^c\cap S'}\|h\|^3\bar{m}_\alpha(\|h\|)dh&\leq C\int_T\int_0^{+\infty}r^{d+2}\bar{m}_\alpha(r)drd\varphi\\&\leq C\left(\int_{-\arccos\left(-\frac{2K}{c}\|x-x'\|\right)}^{-\frac{\pi}{2}}d\varphi+\int_{\frac{\pi}{2}}^{\arccos\left(-\frac{2K}{c}\|x-x'\|\right)}d\varphi\right)\\&\leq C\frac{K}{c}\|x-x'\|,\end{aligned}$$

where the constant C may change from line to line, and where we used the fact that arccos is Lipschitz in the neighborhood of 0 for the last inequality. The same estimate for the set $S\cap S'^c$ completes the proof of (i).

Proof of (ii) a is obviously symmetrical, and $\forall s=(s_1,\dots,s_d)\in\mathbb{R}^d$, using the symmetry of $p(x,h)dh$, an easy calculation (the second line is obtained by the change of variable $h'=-h$) shows that

$$\begin{aligned}s^*a(x)s&=\int_{\mathbb{R}^d}(h\cdot s)^2[h\cdot\nabla_1 g(x,x)]_+p(x,h)dh\\&=\frac{1}{2}\int_{\mathbb{R}^d}(h\cdot s)^2|h\cdot\nabla_1 g(x,x)|p(x,h)dh.\end{aligned}$$

This is non-negative for all $s \in \mathbb{R}^d$, and is non zero if $s \neq 0$ and $x \notin \Gamma$. The fact that σ is Hölder is a trivial consequence of Proposition 2.1 and of the fact that a is Lipschitz, proved in (i).

Proof of (iii) Fix $\alpha > 0$, $x \in \Gamma_\alpha$, and $s = (s_1, \ldots, s_d) \in \mathbb{R}^d$. Denote by u and v the unit vectors of $\mathbb{R}^d$ such that $s = \|s\|u$ and $\nabla_1 g(x,x) = \|\nabla_1 g(x,x)\|v$. Then

$$\begin{aligned} s^* a(x) s &= \frac{1}{2}\|s\|^2 \|\nabla_1 g(x,x)\| \int_{\mathbb{R}^d} |h \cdot u|^2 |h \cdot v| p(x, dh) \\ &\geq C_\alpha \|s\|^2 \|\nabla_1 g(x,x)\| \end{aligned} \tag{17}$$

where C_α is given by (H3), and it follows from (H4) that there exists some $c > 0$ such that $a(x) \in \mathcal{S}_c$ for all $x \in \Gamma_\alpha$ (where $\mathcal{S}_c$ has been defined in (9)), *i.e.* $\Gamma_\alpha \subset \tilde{\Gamma}_c$. Because of Proposition 2.1, it is now simple to prove that σ is Lipschitz on Γ_α. □

3 Construction, uniqueness and strong Markov property of X^ε

Since no standard method to prove uniqueness of solutions to SDEs applies directly to our particular case, we will in this section construct a particular solution, identify the difficulty for uniqueness and for the strong Markov property, and give some conditions solving this difficulty, both in dimension $d = 1$, and for greater dimensions.

3.1 Construction of a particular solution to (1)

Under assumptions (H1) and (H2), and assuming that points $x \in \mathcal{X}$ such that $\nabla_1 g(x,x) = 0$ are isolated in $\mathbb{R}^d$, the weak existence of solutions to the SDE (1) in $\overline{\mathcal{X}}$ has been shown in [2] using an approximation technique. Since we cannot show in general any uniqueness result for this SDE, let us give a precise construction of the process that we will study in the rest of this paper.

Weak existence ensures that for all $\varepsilon > 0$, there is a Brownian motion $(\Omega^\varepsilon, \mathcal{F}^\varepsilon_t, W^\varepsilon, \mathbf{P}^\varepsilon)$, with respect to which can be constructed a solution X^ε to (1) with given (random) initial state $X_0 \in \overline{\mathcal{X}}$. Define the stopping time

$$\tau = \inf\{t \geq 0, X^\varepsilon_t \in \Gamma\}.$$

Then $\tilde{X}^\varepsilon_t = X^\varepsilon_{t\wedge\tau}$ is also a solution to (1): for any $t \geq 0$, we can write

$$\begin{aligned} \tilde{X}^\varepsilon_t = X^\varepsilon_{t\wedge\tau} &= X_0 + \int_0^{t\wedge\tau} b^\varepsilon(X^\varepsilon_s) ds + \sqrt{\varepsilon} \int_0^{t\wedge\tau} \sigma(X^\varepsilon_s) dW_s \\ &= X_0 + \int_0^t b^\varepsilon(\tilde{X}^\varepsilon_s) ds + \sqrt{\varepsilon} \int_0^t \sigma(\tilde{X}^\varepsilon_s) dW_s, \end{aligned}$$

since, by (3), $b(x) = \tilde{b}(x) = 0$ and $a(x) = 0$ as soon as $x \in \Gamma$.

Under additional assumptions (H2'), (H3) and (H4), by Proposition 2.2 (i) and (iii), the coefficients σ, b and $\tilde{b}$ are Lipschitz on Γ_α for any $\alpha > 0$, so strong existence and uniqueness of solutions to (1) hold for $t \leq \tau_\alpha = \inf\{t \geq 0, X^\varepsilon \notin$

$\Gamma_\alpha\}$ for any $\alpha > 0$, *i.e.* for $t \leq \tau$. This proves that $\tilde{X}^\varepsilon$ is a strong solution to (1) on $\mathbb{R}_+$.

In the remaining of this paper, we will study this particular solution $\tilde{X}^\varepsilon$. Let us for convenience denote by X^ε this process.

3.2 Uniqueness and strong Markov property

No standard technique applies directly to prove the uniqueness in law of solutions to (1). This comes from the fact that σ degenerates at points of Γ, *and* that $\tilde{b}$ is not continuous at these points. Uniqueness is known to hold when only one of these difficulties arises, but the combination of both of them leads to great difficulties. Moreover, the strong Markov property for solutions to SDEs is known to be linked to the uniqueness of solutions to the corresponding martingale problem. Here, we are only able to solve these questions under particular assumptions.

Proposition 3.1 *Assume (H1), (H2'), (H3) and (H4), and consider a Brownian motion* $(\Omega, \mathcal{F}, W, \mathbf{P})$. *We have seen in the previous paragraph that there is strong existence of a solution* X^ε *to (1) on* $(\Omega, \mathcal{F}, W, \mathbf{P})$ *with initial state* x. *Let us denote by* $\mathbf{P}_x$ *its law.*

(a) *Suppose that*

$$\mathbf{P}_x(\tau = \infty) = 1. \tag{18}$$

Then, there is strong uniqueness of solutions to (1) with initial state x.

(b) *For any* $\mathcal{F}_t$*-stopping time* $S < \tau$, *for any* $t > 0$, *for any Borel set* $B \subset \mathcal{B}(\mathbb{R}^d)$ *and for any* $x \in \mathcal{X} \setminus \Gamma$,

$$\mathbf{P}_x(X^\varepsilon_{S+t} \in B|\mathcal{F}_S) = \mathbf{P}_x(X^\varepsilon_{S+t} \in B|X^\varepsilon_S). \tag{19}$$

(c) *Assume (18) for all* $x \in \mathcal{X} \setminus \Gamma$. *Then* X^ε *satisfies the strong Markov property with respect to the canonical filtration* $(\mathcal{F}_t, t \geq 0)$ *associated to the Brownian motion* W.

Proof Let us first prove (a). Under assumptions (H1), (H2'), (H3) and (H4), we have seen in the last paragraph that there is strong existence of X^ε, and that there is strong uniqueness of solutions to (1) with initial state x for $t \leq \tau$. Since $\mathbf{P}_x(\tau = \infty) = 1$, this is actually true for $t \in \mathbb{R}$.

Let us come to the proof of (b). Fix $S < \tau$, $t > 0$, B and x as in the statement of (b). It follows from the fact that X^ε is constant after time τ that (19) is equivalent to

$$\mathbf{P}(X^\varepsilon_{(S+t)\wedge\tau} \in B|\mathcal{F}_S) = f(X^\varepsilon_S) \tag{20}$$

for some Lebesgue-measurable function f from $\mathbb{R}$ to $\mathbb{R}$.

For any $\alpha > 0$, let us define $X^{\varepsilon,\alpha}$ as the strong solution to the SDE

$$dX^{\varepsilon,\alpha}_t = (b(X^{\varepsilon,\alpha}_t) + \varepsilon\tilde{b}_\alpha(X^{\varepsilon,\alpha}_t))dt + \sqrt{\varepsilon}\sigma_\alpha(X^{\varepsilon,\alpha}_t)dW_t,$$

on $(\Omega, \mathcal{F}, W, \mathbf{P})$ with initial state x, where $\tilde{b}_\alpha$ and σ_α are bounded and (globally) Lipschitz functions on $\mathbb{R}^d$ such that $\forall x \in \Gamma_\alpha$, $\tilde{b}_\alpha(x) = \tilde{b}(x)$ and $\sigma_\alpha(x) = \sigma(x)$

(such functions exist since σ and $\tilde{b}$ are both Lipschitz on Γ_α). Note that, if $t \leq \tau_\alpha = \inf\{t \geq 0 : X^\varepsilon_t \notin \Gamma_\alpha\}$, $X^{\varepsilon,\alpha}_t = X^\varepsilon_t$ $\mathbf{P}$-a.s.

Since b, $\tilde{b}_\alpha$ and σ_α are bounded and Lipschitz, $X^{\varepsilon,\alpha}$ is a strong Markov process, so, for any $\alpha > 0$, there is a Lebesgue-measurable function f_α from $\mathbb{R}^d$ to $[0, 1]$ such that

$$\mathbf{P}(X^{\varepsilon,\alpha}_{(S+t)\wedge\tau} \in B|\mathcal{F}_S) = f_\alpha(X^{\varepsilon,\alpha}_S).$$

Therefore,

$$\begin{aligned}\mathbf{P}(\tau_\alpha > S,\ X^{\varepsilon,\alpha}_{(S+t)\wedge\tau} \in B|\mathcal{F}_S) &= \mathbf{1}_{\{\tau_\alpha > S\}}\mathbf{P}(X^{\varepsilon,\alpha}_{(S+t)\wedge\tau} \in B|\mathcal{F}_S) \\ &= \mathbf{1}_{\{\tau_\alpha > S\}} f_\alpha(X^\varepsilon_S).\end{aligned} \tag{21}$$

Now, since, as $\alpha \to 0$, $\tau_\alpha \to \tau > S$ a.s. and $X^{\varepsilon,\alpha}_{(S+t)\wedge\tau} \to X^\varepsilon_{(S+t)\wedge\tau}$ a.s., it follows from the dominated convergence Theorem that

$$\mathbf{P}(\tau_\alpha > S,\ X^{\varepsilon,\alpha}_{(S+t)\wedge\tau} \in B|\mathcal{F}_S) \longrightarrow \mathbf{P}(X^\varepsilon_{(S+t)\wedge\tau} \in B|\mathcal{F}_S) \quad \text{a.s.}$$

when $\alpha \to 0$. For the same reason, $\mathbf{1}_{\{\tau_\alpha \leq S\}} f_\alpha(X^\varepsilon_S) \to 0$ a.s. when $\alpha \to 0$. Combining these two facts in equation (21), it follows that $f_\alpha(X_S)$ converges almost surely to a function which is $\sigma(X_S)$-measurable (as a limit of $\sigma(X_S)$-measurable function), and which is a.s. equal to $\mathbf{P}(X^\varepsilon_{(S+t)\wedge\tau} \in B|\mathcal{F}_S)$. This completes the proof of (20).

Finally, under the assumption (18), (c) is a trivial consequence of (b). □

3.3 The dimension 1 case

As we seen above, the uniqueness and the strong Markov property of X^ε rely on the fact that $\mathbf{P}_x(\tau = \infty) = 1$. It is possible, in dimension $d = 1$, to give conditions under which this is true. In this case, an elementary calculation gives the following formulas for a, b and $\tilde{b}$:

$$\begin{aligned} b(x) &= \frac{M_2(x)}{2}\partial_1 g(x,x) \\ \tilde{b}(x) &= \frac{M_3(x)}{4}\mathrm{sign}[\partial_1 g(x,x)]\partial^2_{1,1} g(x,x) \\ a(x) &= \frac{M_3(x)}{2}|\partial_1 g(x,x)|, \\ \text{where} \quad M_k(x) &= \int_{\mathbb{R}} |h|^k p(x,h)dh \\ \text{and} \quad \mathrm{sign}(x) &= \begin{cases} -1 & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ 1 & \text{if } x > 0, \end{cases} \end{aligned}$$

where $\partial_i g(x,x)$ denotes the partial derivative of $g(x,y)$ with respect to the i^{th} variable, and $\partial^2_{i,j} g(x,x)$ the second partial derivative of $g(x,y)$ with respect to the i^{th} and j^{th} variables ($i, j = 1$ or 2).

Theorem 3.1 *Assume (H1), (H2), that $d = 1$, and that g is $\mathcal{C}^3$ with bounded third-order derivatives. Let X^ε be a solution to (1) starting at $x \notin \Gamma$. Define*

$c = \sup\{y \in \Gamma, y < x\}$, $c' = \inf\{y \in \Gamma, y > x\}$, and let τ be the stopping time $\inf\{t \geq 0, X^\varepsilon \in \Gamma\} = \inf\{t \geq 0, X^\varepsilon \in \{c, c'\}\}$. Assume that c and c' do not belong to $\partial\mathcal{X}$, that $-\infty < c < c' < \infty$, and that $\partial^2_{1,1}g(c,c) + \partial^2_{1,2}g(c,c) \neq 0$ and $\partial^2_{1,1}g(c',c') + \partial^2_{1,2}g(c',c') \neq 0$. Then we can define

$$\begin{aligned} \alpha &:= \frac{\partial^2_{1,1}g(c,c)}{\partial^2_{1,1}g(c,c) + \partial^2_{1,2}g(c,c)} = \frac{2\partial^2_{1,1}g(c,c)}{\partial^2_{1,1}g(c,c) - \partial^2_{2,2}g(c,c)} \\ \beta &:= \frac{\partial^2_{1,1}g(c',c')}{\partial^2_{1,1}g(c',c') + \partial^2_{1,2}g(c',c')} = \frac{2\partial^2_{1,1}g(c',c')}{\partial^2_{1,1}g(c',c') - \partial^2_{2,2}g(c',c')}. \end{aligned} \tag{22}$$

These equalities follow from the fact that $\partial^2_{1,1}g + 2\partial^2_{1,2}g + \partial^2_{2,2}g = 0$, obtained by differentiating equation (2). Then, we distinguish four cases:

(a) *If $\alpha \geq 1$ and $\beta \leq -1$, then $\mathbf{P}(\tau = \infty) = 1$ and the process X^ε is recurrent in (c, c').*

(b) *If $\alpha \geq 1$ and $\beta > -1$, then $\mathbf{P}(\tau < \infty) = 1$ and $\mathbf{P}(\lim_{t\to\tau} X^\varepsilon_t = c') = 1$.*

(c) *If $\alpha < 1$ and $\beta \leq -1$, then $\mathbf{P}(\tau < \infty) = 1$ and $\mathbf{P}(\lim_{t\to\tau} X^\varepsilon_t = c) = 1$.*

(d) *If $\alpha < 1$ and $\beta > -1$, then $\mathbf{P}(\tau < \infty) = 1$, $\mathbf{P}(\lim_{t\to\tau} X^\varepsilon_t = c) > 0$ and $\mathbf{P}(\lim_{t\to\tau} X^\varepsilon_t = c') = 1 - \mathbf{P}(\lim_{t\to\tau} X^\varepsilon_t = c) > 0$.*

Remark 3.1 *The conditions $\partial^2_{1,1}g(c,c) + \partial^2_{1,2}g(c,c) \neq 0$ and $\partial^2_{1,1}g(c',c') + \partial^2_{1,2}g(c',c') \neq 0$ are only technical. A higher order calculation is possible if one of them does not hold.*

Remark 3.2 *When $c = -\infty$ or $c' = \infty$, the calculation below depends on technical properties of g and M_k, and no simple general result can be stated.*

Remark 3.3 *The biological theory of adaptive dynamics gives a classification of evolutionary singularities, depending on the values of $\partial^2_{1,1}g$ and $\partial^2_{2,2}g$ at these points. Here, the condition $\alpha \geq 1$ corresponds, when $\partial^2_{1,1}g(c,c) - \partial^2_{2,2}g(c,c) > 0$, to the case $\partial^2_{1,1}g(c,c) + \partial^2_{2,2}g(c,c) \geq 0$, which corresponds in the biological terminology (see* e.g. *Diekmann [6]) to a converging stable strategy with mutual invasibility, which includes the evolutionary* branching *condition; and when $\partial^2_{1,1}g(c,c) - \partial^2_{2,2}g(c,c) < 0$, to the case $\partial^2_{1,1}g(c,c) + \partial^2_{2,2}g(c,c) \leq 0$, which corresponds biologically to a repelling strategy without mutual invasibility.*

Proof of Theorem 3.1 We will here use the classical methods of removal of drift of Engelbert and Schmidt and the explosion criterion of Feller (see Karatzas and Shreve [10]). They can be applied to X^ε, considered as a process with value in (c, c') killed when it reaches c or c', under the following assumptions, obviously fulfilled by our process:

$$\forall x \in (c,c'),\ \sigma(x) > 0,$$
$$\text{and } \forall x \in (c,c'),\ \exists \delta > 0 \text{ such that } \int_{x-\delta}^{x+\delta} \frac{1 + |b^\varepsilon(y)|}{\varepsilon\sigma^2(y)} dy < \infty.$$

These methods involve the two following functions, defined for a fixed $\gamma \in (c, c')$:

$$
\begin{aligned}
p(x) &= \int_\gamma^x \exp\left[-2\int_\gamma^y \frac{b^\varepsilon(z)dz}{\varepsilon\sigma^2(z)}\right] dy,\ \forall x \in (c, c'),\\
\text{and}\quad v(x) &= \int_\gamma^x p'(y)\int_\gamma^y \frac{2dz}{\varepsilon p'(z)\sigma^2(z)}dy,\ \forall x \in (c, c').
\end{aligned}
\tag{23}
$$

Then, as can be seen in [10] pp. 345 (Proposition 5.5.22), 348 (Theorem 5.5.29) and 351 (Proposition 5.5.32), the statements about the limit of the process X^ε_t when $t \to \tau$ and about the recurrence of X^ε depend on whether $p(x)$ is finite or not when $x \to c$ and c' (point (a) corresponds to $p(c+) = -\infty$ and $p(c'-) = +\infty$, point (b) to $p(c+) = -\infty$ and $p(c'-) < +\infty$, etc...), and the statements about τ depends on whether $v(x) < \infty$ or not when $x \to c$ and c' ($\mathbf{P}(\tau = \infty) = 1$ if and only if $v(c+) = v(c'-) = \infty$, and $\mathbf{P}(\tau < \infty) = 1$ if and only if $v(c+) < \infty$ and $v(c'-) < \infty$, or $v(c+) < \infty$ and $p(c'-) = +\infty$, or $v(c'-) < \infty$ and $p(c+) = -\infty$).

So let us compute these limits.

$$
\frac{b^\varepsilon(x)}{\varepsilon\sigma^2(x)} = \frac{b^\varepsilon(x)}{\varepsilon a(x)} = \frac{M_2(x)}{\varepsilon M_3(x)}\mathrm{sign}[\partial_1 g(x,x)] + \frac{1}{2}\frac{\partial^2_{1,1}g(x,x)}{\partial_1 g(x,x)}. \tag{24}
$$

So, for $x < y < \gamma$, the quantity inside the exponential appearing in the definition of p writes

$$
\int_y^\gamma \frac{2M_2(z)}{\varepsilon M_3(z)}\mathrm{sign}[\partial_1 g(z,z)]dz + \int_y^\gamma \frac{\partial^2_{1,1}g(z,z)}{\partial_1 g(z,z)}dz.
$$

Since $c \notin \partial\mathcal{X}$, the first term is bounded for $c < y < \gamma$ (by assumption (H2), M_3 is positive and continuous on $[c, \gamma]$, so it is bounded away of 0 on this interval), so we only have to study the second term. When $y \to c$, $\frac{\partial^2_{1,1}g(z,z)}{\partial_1 g(z,z)} \sim \frac{\partial^2_{1,1}g(c,c)}{(z-c)(\partial^2_{1,1}g(c,c)+\partial^2_{1,2}g(c,c))} = \frac{\alpha}{z-c}$. So, if $\alpha \neq 0$, when $y \to c$,

$$
\int_y^\gamma \frac{\partial^2_{1,1}g(z,z)}{\partial_1 g(z,z)}dz \sim \int_y^\gamma \frac{\alpha}{z-c}dz = -\alpha(\log(y-c) - \log(\gamma - c)).
$$

Consequently,

$$
\exp\left[-2\int_\gamma^y \frac{b^\varepsilon(z)dz}{\varepsilon\sigma^2(z)}\right] = \exp[-\alpha\log(y-c)(1+o(1))] = (y-c)^{-\alpha+o(1)}.
$$

$p(x)$ has a finite limit when $x \to c$ if and only if the integral of the quantity above on (c, γ) is convergent. So, if $\alpha < 1$, $p(c+) > -\infty$ (the case $\alpha = 0$ leads obviously to a finite integral), and if $\alpha > 1$, $p(c+) = -\infty$. The case $\alpha = 1$ depends on a development of $\frac{\partial^2_{1,1}g(z,z)}{\partial_1 g(z,z)}$ to a higher order. An easy computation, shows that when $\alpha = 1$,

$$
\frac{\partial^2_{1,1}g(z,z)}{\partial_1 g(z,z)} = \frac{1}{x} - \frac{\partial^3_{1,1,2}g(c,c) + \partial^3_{1,2,2}g(c,c)}{\partial^2_{1,1}g(c,c)} + o(1),
$$

so that

$$
p'(y) = \exp\left[-2\int_\gamma^y \frac{b^\varepsilon(z)dz}{\varepsilon\sigma^2(z)}\right] = \exp[-\log(y-c) + C + o(1)] = \frac{e^{C+o(1)}}{y-c}, \tag{25}
$$

which gives an infinite value to $p(c+)$.

The same computation gives the required result when $x \to c'$.

Now let us compute the limit of v at c and c'. Since $p(c'-) = \infty \Rightarrow v(c'-) = \infty$ and $p(c+) = -\infty \Rightarrow v(c+) = \infty$ (see [10] page 348), we only have to deal with the cases $\alpha < 1$ and $\beta > -1$.

The higher-order estimation (25) above can be obtained for any value of α, and gives

$$p'(y) = \exp\left[-2\int_{\gamma}^{y} \frac{b^{\varepsilon}(z)dz}{\varepsilon\sigma^2(z)}\right] = \frac{e^{C+o(1)}}{(y-c)^{\alpha}}, \tag{26}$$

so, for some constant C

$$\frac{2}{\varepsilon p'(z)a(z)} \sim C(z-c)^{\alpha-1},$$

since

$$a(z) = \frac{M_3(z)}{2}|\partial_1 g(z,z)| \sim \frac{M_3(c)}{2}|\partial^2_{1,1}g(c,c) + \partial^2_{1,2}g(c,c)|(z-c). \tag{27}$$

If $\alpha < 0$, when $y \to c$, $p'(y)\int_y^{\gamma} \frac{2dz}{\varepsilon p'(z)a(z)} \sim -Cp'(y)(y-c)^{\alpha}$ is bounded on (c,γ) by (26), and so $v(c+) < \infty$. If $\alpha = 0$, $p'(y)\int_y^{\gamma} \frac{2dz}{\varepsilon p'(z)a(z)} \sim C\log(y-c)$, which has a finite integral on (c,γ), so $v(c+) < \infty$. Finally, if $0 < \alpha < 1$, $\int_y^{\gamma} \frac{2dz}{\varepsilon p'(z)a(z)}$ is bounded, so $v(c+) < \infty$ is equivalent to the convergence of the integral $\int_c^{\gamma} p'(y)dy$, which holds since $p'(y) \sim \frac{C}{(y-a)^{\alpha}}$ and $\alpha < 1$. □

In the case where c or c' belong to $\partial\mathcal{X}$, what changes in the calculation above? Assume for example that $c \in \partial\mathcal{X}$. The problem is that $M_2(x)$ and $M_3(x)$ are not bounded away from zero in the neighborhood of c. Indeed, the support of $p(x,h)dh$ is a subset of $\mathcal{X} - x$ which is symmetrical with respect to 0, so it is a subset of $(-(x-c), x-c)$, which converges to $\{0\}$ when $x \to c$. So the quantity $2\frac{b(x)}{a(x)} = \frac{M_2(x)}{\varepsilon M_3(x)}\text{sign}[\partial_1 g(x,x)]$ appearing in the equation (24) may not be bounded in the neighborhood of c.

If $x \in \mathcal{X}$, the support of $p(x,h)dh$ is a subset of $(-(x-c), x-c)$, so $\int |h|^3 p(x,h)dh \le (x-c)\int |h|^2 p(x,h)dh$, *i.e.*

$$\frac{M_2(x)}{M_3(x)} \ge \frac{1}{x-c}.$$

Since $\Gamma \cap (c,c') = \emptyset$, the sign of $\partial_1 g(x,x)$ is constant on (c,c'), equal to the sign of $\partial^2_{1,1}g(c,c) + \partial^2_{1,2}g(c,c)$ (by expanding $\partial_1 g(x,x)$ in the neighborhood of c), that is not null by assumption. Let us call s this sign.

If $s = +1$, then, for $c < y < \gamma$, there is a constant $C > 0$ such that

$$\exp\left[\frac{2}{\varepsilon}\int_y^{\gamma} \frac{M_2(z)}{M_3(z)}\text{sign}[\partial_1 g(z,z)]dz\right] \ge \frac{C}{(y-c)^{2/\varepsilon}}.$$

Combining this fact with the estimations obtained in the proof of Theorem 3.1 for the term $\frac{\partial^2_{1,1}g(z,z)}{\partial_1 g(z,z)}$, we see that

$$p'(y) = \exp\left[-2\int_{\gamma}^{y} \frac{b^{\varepsilon}(z)dz}{\varepsilon\sigma^2(z)}\right] \ge \frac{C}{(y-c)^{\alpha+2/\varepsilon}}, \tag{28}$$

and we finally obtain that $p(c+) > -\infty$ if $\alpha + \frac{2}{\varepsilon} < 1$.

In the case where $s = -1$, we obtain

$$p'(y) \exp\left[-2\int_\gamma^y \frac{b^\varepsilon(z)dz}{\varepsilon\sigma^2(z)}\right] \leq \frac{C}{(y-c)^{\alpha-2/\varepsilon}}, \tag{29}$$

and so $p(c+) = -\infty$ if $\alpha - \frac{2}{\varepsilon} \geq 1$.

Observe that if, for example,

$$\exists \rho > 0,\ \exists 0 < \theta < 1,\ \forall x \in \mathcal{X},\ p(x,h) \geq \rho \mathbf{1}_{B(0,\theta(x-c))}, \tag{30}$$

we easily obtain that $\frac{M_2(x)}{M_3(x)} \geq \frac{C}{x-c}$, and, consequently, the converse inequalities in (28) and (29) hold. So, under this assumption, if $s = +1$, $p(c+) > -\infty \Leftrightarrow \alpha + \frac{2}{\varepsilon} < 1$, and if $s = -1$, $p(c+) > -\infty \Leftrightarrow \alpha - \frac{2}{\varepsilon} < 1$.

Concerning the limit of the function v at c, observe that assumption (H2) implies that M_3 is a Lipschitz function, so there is some constant K such that $M_3(x) \leq K(x-c)$. In equation (27), this gives that $a(z) \leq C(z-c)^2$ for some constant C.

Then, in the case $s = +1$, if we assume (30), we obtain

$$\frac{2}{\varepsilon p'(z)a(z)} \geq C(z-c)^{\alpha-2+2/\varepsilon}.$$

Thus, if $\alpha + 2/\varepsilon < 1$,

$$p'(y)\int_y^\gamma \frac{2dz}{\varepsilon p'(z)a(z)} \geq \frac{C}{(y-c)^{\alpha+2/\varepsilon}} - \frac{C'}{y-c},$$

so the first term of the right-hand side has a finite integral on (c,γ), and the second term has a divergent integral, so $v(c+) = \infty$.

In the case where $s = -1$, the same calculation can be made replacing $\alpha + 2/\varepsilon$ by $\alpha - 2/\varepsilon$, and gives that $v(c+) = \infty$ when $\alpha - 2/\varepsilon < 1$.

Let us collect all these results in the following theorem:

Theorem 3.2 *With the same notations and assumptions as in Theorem 3.1, except that $c \in \partial\mathcal{X}$ and $c' \notin \partial\mathcal{X}$, and with the notation $s = sign[\partial^2_{1,1}g(c,c) + \partial^2_{1,2}g(c,c)]$,*

(a) *If $s = -1$, $\alpha - 2/\varepsilon \geq 1$ and $\beta \leq -1$, then the process X^ε is recurrent in (c,c'). The same holds if $s = 1$, $\alpha + 2/\varepsilon \geq 1$ and $\beta \leq -1$, under the additional assumption (30). In both cases, $\mathbf{P}(\tau = \infty) = 1$.*

(b) *If $s = -1$, $\alpha - 2/\varepsilon \geq 1$ and $\beta > -1$, then $\mathbf{P}(\lim_{t\to\tau} X^\varepsilon_t = c') = 1$. The same holds if $s = 1$, $\alpha + 2/\varepsilon \geq 1$ and $\beta > -1$, under the additional assumption (30). In both cases, $\mathbf{P}(\tau < \infty) = 1$.*

(c) *If $s = 1$, $\alpha + 2/\varepsilon < 1$ and $\beta \leq -1$, then $\mathbf{P}(\lim_{t\to\tau} X^\varepsilon_t = c) = 1$. The same holds if $s = -1$, $\alpha - 2/\varepsilon \geq 1$ and $\beta \leq -1$, under the additional assumption (30). In both cases, under assumption (30), $\mathbf{P}(\tau = \infty) = 1$.*

(d) *If $s = 1$, $\alpha + 2/\varepsilon < 1$ and $\beta > -1$, then $\mathbf{P}(\lim_{t\to\tau} X^\varepsilon_t = c) > 0$ and $\mathbf{P}(\lim_{t\to\tau} X^\varepsilon_t = c') = 1 - \mathbf{P}(\lim_{t\to\tau} X^\varepsilon_t = c) > 0$. The same holds if $s = -1$, $\alpha - 2/\varepsilon \geq 1$ and $\beta > -1$, under the additional assumption (30). In both cases, $0 < \mathbf{P}(\tau < \infty) < 1$.*

Remark 3.4 *A similar result can be stated in the case where $c \notin \partial\mathcal{X}$ and $c' \in \partial\mathcal{X}$, and in the case where c and c' are both in $\partial\mathcal{X}$.*

3.4 The dimension $d \geq 2$ case

Let us turn now to the case $d \geq 2$. If we restrict ourselves to the case $\mathcal{X} = \mathbb{R}^d$, it is possible to give conditions under which $\mathbf{P}_x(\tau = \infty) = 1$, based on a comparison of $d(X^\varepsilon, \Gamma)$ with Bessel processes:

Theorem 3.3 *Assume (H1), (H2'), (H3), (H4), that $\mathcal{X} = \mathbb{R}^d$, that g is $\mathcal{C}^2$ on $\mathbb{R}^d \times \mathbb{R}^d$ and that the points of Γ are isolated. Let τ denote the stopping time $\inf\{t \geq 0, X_t^\varepsilon \in \Gamma\}$. For any $y \in \Gamma$, let $\mathcal{U}_y$ be a neighborhood y, and take two constants $a^y > 0$ and $a_y > 0$ such that a is a^y-Lipschitz on $\mathcal{U}_y$ with respect to the spectral norm $\|\cdot\|$, and such that $\forall x \in \mathcal{U}_y,\ \forall v \in \mathbb{R}^d,\ s^* a(x) s \geq a_y \|s\|^2 \|x - y\|$. Define also*

$$\tilde{b}_y = \inf_{x \in \mathcal{U}_y \setminus \{y\}} \frac{x-y}{\|x-y\|} \cdot \tilde{b}(x)$$
$$\text{and} \quad \tilde{b}^y = \sup_{x \in \mathcal{U}_y \setminus \{y\}} \frac{x-y}{\|x-y\|} \cdot \tilde{b}(x).$$

Then

(a) *If for any $y \in \Gamma$, $\frac{\tilde{b}_y + d a_y/2}{a^y} \geq 1$, then, for any $x \notin \Gamma$, $\mathbf{P}_x(\tau = \infty) = 1$ and $\mathbf{P}_x(\lim_{t \to +\infty} X_t^\varepsilon \in \Gamma) = 0$.*

(b) *If there exists $y \in \Gamma$ such that $\frac{\tilde{b}^y + d a^y/2}{a_y} < 1$, then, $\mathbf{P}_x(\lim_{t \to \tau} X_t^\varepsilon = y) > 0$ for any $x \notin \Gamma$.*

Remark 3.5 *In the case where $\mathcal{X} \neq \mathbb{R}^d$, the method of this proof applies only in the case where $x \mapsto d(x, \partial\mathcal{X})$ is $\mathcal{C}^2$ on $\{x \in \mathcal{X} : d(x, \partial\mathcal{X}) < \alpha\}$ for some $\alpha > 0$ (this holds in particular when $\partial\mathcal{X}$ is compact and $\mathcal{C}^2$). This gives conditions for $\mathbf{P}(\tau = \infty) = 1$ involving sharp constants governing the behaviour of a, b and $\tilde{b}$ near $\partial\mathcal{X}$.*

Before proving Theorem 3.3, let us give some bounds for the constants involved in this Theorem:

Proposition 3.2 *Assume (H1), (H2), (H3), that $\mathcal{X} = \mathbb{R}^d$, that g is $\mathcal{C}^2$ on $\mathbb{R}^d \times \mathbb{R}^d$ and that the points of Γ are isolated. Fix $y \in \Gamma$, and fix $\alpha > 0$ such that $B(y, \alpha) \cap \Gamma = \{y\}$. Define*

$$C = \inf_{u, v \in \mathbb{R}^d : \|u\| = \|v\| = 1} \int |h \cdot u|^2 |h \cdot v| p(x, h) dh.$$

$C > 0$ by (H3). Let M_3 be a bound for the third-order moment of $p(x, h)dh$ on $\mathcal{X}$, given by (H2). Denote by D the differential $H_{1,1}g(y, y) + H_{1,2}g(y, y)$ of $x \mapsto \nabla_1 g(x, x)$ at y and denote by λ^y (resp. λ_y) the greatest (resp. the smallest) eigenvalue of D^*D. Observe that $\lambda^y \geq \lambda_y \geq 0$ (D^*D is a positive symmetrical matrix), and that $\lambda_y > 0$ if and only if the kernel of D is $\{0\}$. Suppose that this

is true. Then, for any $\delta > 0$ there exists a neighborhood $\mathcal{U}_y$ of y such that, in the statement of Theorem 3.3, we can take

$$a^y = M_3\sqrt{\lambda^y} + \delta, \quad a_y = C\sqrt{\lambda_y} - \delta,$$
$$\tilde{b}^y < \frac{M_3}{2}\|H_{1,1}g(y,y)\| + \delta \quad \text{and} \quad \tilde{b}_y > -\frac{M_3}{2}\|H_{1,1}g(y,y)\| - \delta.$$

Remark 3.6 *There are cases where $\tilde{b}_y \geq 0$. This holds for example when $D = cId$ for some constant $c > 0$ and $H_{1,1}g(y,y)$ is a positive symmetrical matrix, or $D = -cId$ and $H_{1,1}g(y,y)$ is a negative matrix (and also for cases sufficiently close to these two ones). These facts will appear clearly in the proof of Proposition 3.2.*

Proof of Proposition 3.2 Let us begin with $\tilde{b}^y$ and $\tilde{b}_y$: it follows from the definition (3) of $\tilde{b}$ that for $x \neq y$,

$$\frac{x-y}{\|x-y\|} \cdot \tilde{b}(x) = \int_{\{\nabla_1 g(x,x)\cdot h>0\}} \left(\frac{x-y}{\|x-y\|} \cdot h\right)(h^* H_{1,1}g(x,x)h)p(x,h)dh, \quad (31)$$

and, because of (H1), the quantity inside the integral can be bounded by $\|h\|^3[\|H_{1,1}g(y,y)\| + O(\|x-y\|)]p(x,h)$. So

$$\begin{aligned}\frac{x-y}{\|x-y\|} \cdot \tilde{b}(x) &\leq [\|H_{1,1}g(y,y)\| + O(\|x-y\|)] \int_{\{\nabla_1 g(x,x)\cdot h>0\}} \|h\|^3 p(x,h)dh \\ &= \frac{M_3}{2}[\|H_{1,1}g(y,y)\| + O(\|x-y\|)].\end{aligned}$$

Hence, for any $\delta > 0$, $\tilde{b}^y$ can be made smaller than $\frac{M_3}{2}\|H_{1,1}g(y,y)\| + \delta$ if we choose $\mathcal{U}_y$ sufficiently small. Similarly, $\tilde{b}_y > -\frac{M_3}{2}\|H_{1,1}g(y,y)\| - \delta$ if $\mathcal{U}_y$ is sufficiently small.

To prove Remark 3.6, is suffices to notice that, if $H_{1,1}g(y,y)$ is symmetrical positive, the quantity inside the integral (31) is positive for all h such that $(x-y)\cdot h > 0$, and that $D = c\mathrm{Id}$ implies that $\nabla_1 g(x,x) \sim c(x-y)$ when $x \to y$. So the set $\{\nabla_1 g(x,x) \cdot h > 0\} \setminus \{(x-y)\cdot h > 0\}$ converges to $\emptyset$, and we can conclude thanks to the dominated convergence Theorem.

It follows from equation (17) in the proof of Proposition 2.2, that $\forall s \in \mathbb{R}^d$ and $\forall x \in \mathbb{R}^d$

$$C\|s\|^2\|\nabla_1 g(x,x)\| \leq s^* a(x) s \leq M_3\|s\|^2\|\nabla_1 g(x,x)\|.$$

Considering an orthonormal basis of $\mathbb{R}^d$ in which D^*D is diagonal, one can easily see that $\lambda_y\|v\|^2 \leq \|Dv\|^2 = v \cdot D^*Dv \leq \lambda^y\|v\|^2$ for any $v \in \mathbb{R}^d$. It remains to observe that $\nabla_1 g(x,x) \sim D(x-y)$ when $x \to y$ to obtain the required bounds for a^y and a_y. □

Proof of Theorem 3.3 Fix $y \in \Gamma$. It will be more convenient in this proof to reduce, by translation, to the case $y = 0$. By assumption, to this point of Γ is associated a neighborhood $\mathcal{U}_0$ of 0 and four constants $a_0 > 0$, $a^0 > 0$, $\tilde{b}_0$ and

$\tilde{b}^0$. A standard computation using the Itô formula gives that $\forall t < \tau$,

$$\|X_t^\varepsilon\| = \|x\| + \int_0^t \frac{1}{\|X_s^\varepsilon\|}\Bigg[X_s^\varepsilon \cdot (b(X_s^\varepsilon) + \varepsilon\tilde{b}(X_s^\varepsilon)) \\ + \frac{\varepsilon}{2}\mathrm{Tr}(a(X_s^\varepsilon)) - \frac{\varepsilon}{2}\frac{(X_s^\varepsilon)^*}{\|X_s^\varepsilon\|}a(X_s^\varepsilon)\frac{X_s^\varepsilon}{\|X_s^\varepsilon\|}\Bigg]ds + M_t,$$

where

$$M_t := \sqrt{\varepsilon}\int_0^t \frac{(X_s^\varepsilon)^*}{\|X_s^\varepsilon\|}\sigma(X_s^\varepsilon)dW_s,$$

and where Tr is the trace operator on $d \times d$ matrices. Since $b(0) = \tilde{b}(0) = a(0) = 0$ $(0 \in \Gamma)$, this relation is in fact true for all $t \geq 0$, if we intend that the products $\infty \times 0$ appearing inside the integrals are 0.

Since σ is bounded, M_t is a $\mathbb{L}^2$-martingale in $\mathbb{R}$ with quadratic variation

$$\langle M\rangle_t = \varepsilon\int_0^t \frac{(X_s^\varepsilon)^*}{\|X_s^\varepsilon\|}a(X_s^\varepsilon)\frac{X_s^\varepsilon}{\|X_s^\varepsilon\|}ds. \tag{32}$$

It follows from the Dubins-Schwartz Theorem that for any $t \geq 0$, $M_t = B_{\langle M\rangle_t}$, where B is a one-dimensional Brownian motion. Because of Proposition 2.2 (ii), $\langle M\rangle$ is strictly increasing for $t < \tau$, constant after τ, and it is $\mathcal{C}^1$ on $[0,\tau)$ with bounded derivative $\varepsilon\frac{(X_s^\varepsilon)^*}{\|X_s^\varepsilon\|}a(X_s^\varepsilon)\frac{X_s^\varepsilon}{\|X_s^\varepsilon\|}$.

Define the time change $T_t = \inf\{s \geq 0 : \langle M\rangle_s > t\}$ for all $t \geq 0$. If $t < \langle M\rangle_\infty = \lim_{t\to\infty}\langle M\rangle_t$, then $T_t < \infty$ and $\langle M\rangle_{T_t} = t$. For $t < \langle M\rangle_\infty$, define $Y_t = X_{T_t}^\varepsilon$. An easy change of variable shows that for $t < \langle M\rangle_\infty$,

$$\|Y_t\| = \|x\| + \int_0^t c(Y_s)ds + B_t,$$

where

$$c(z) = \|z\|\frac{z\cdot(b(z) + \varepsilon\tilde{b}(z)) + \mathrm{Tr}(a(z))/2}{\varepsilon z^* a(z) z} - \frac{1}{2\|z\|}.$$

Using the constants defined in the statement of Theorem 3.3, the fact that b is K-Lipschitz on $\mathbb{R}^d$, and the fact that $\mathrm{Tr}(a) = \sum_{i=1}^d e_i^* a e_i$, where e_i is the i^{th} vector of the canonical basis of $\mathbb{R}^d$, one easily obtains that, for $z \in \mathcal{U}_0$,

$$\forall z \in \mathbb{R}^d,\ c_1(\|z\|) < c(z) < c_2(\|z\|),$$

where, for $u > 0$,

$$c_1(u) = \left(\frac{da_0 + \tilde{b}_0}{a^0} - \frac{1}{2}\right)\frac{1}{u} - \frac{2K}{\varepsilon a_0}$$
$$\text{and}\quad c_2(u) = \left(\frac{da^0 + \tilde{b}^0}{a_0} - \frac{1}{2}\right)\frac{1}{u} + \frac{2K}{\varepsilon a_0}.$$

Let ρ be small enough for $B(\rho) := \{x \in \mathbb{R}^d : \|x\| \leq \rho\} \subset \mathcal{U}_0$ and $\Gamma \cap B(2\rho) = \{0\}$ to hold, and define $\tau_\rho := \inf\{t \geq 0 : \|X_t^\varepsilon\| = \rho\}$ and $\tau_0 = \inf\{t \geq 0 : X_t^\varepsilon = 0\}$. We intend to prove the following lemma.

Lemma 3.1
(a) *If* $\frac{\tilde{b}_0+da_0/2}{a^0} \geq 1$*, then, for all* $x \in B(\rho) \setminus \{0\}$, $\mathbf{P}_x(\tau_\rho < \tau_0) = 1$.

(b) *If* $\frac{\tilde{b}^0+da^0/2}{a_0} < 1$*, then, there exists a constant* $c > 0$ *such that, for all* $x \in B(\rho/2) \setminus \{0\}$, $\mathbf{P}_x(\{\tau_0 < \tau_\rho\} \cup \{\tau_0 = \tau_\rho = \infty$ *and* $\lim_{t\to+\infty} X^\varepsilon_t = 0\}) \geq c$.

Together with the incomplete strong Markov property of Proposition 3.1 (b), part (a) of this lemma easily implies Theorem 3.3 (a), and part (b) implies Theorem 3.3 (b) if we can prove that for any $x \in \mathcal{X} \setminus \Gamma$, $\mathbf{P}_x(\tau_{\rho/2} < \infty) > 0$. This can be proved as follows.

Fix $x \in \mathcal{X} \setminus \Gamma$. If $x \in B(\rho/2)$, there is nothing to prove, so let us assume that $x \notin B(\rho/2)$, and let $\alpha < d(x, \Gamma) \wedge (\rho/4)$. Let ϕ be a $\mathcal{C}^1$ function from an interval $[0, T]$ to Γ_α such that $\phi(0) = x$ and $\|\phi(T)\| = \alpha$. Remember the definition of the process $X^{\varepsilon,\alpha/2}$ in the proof of Proposition 3.1. This process has uniformly Lipschitz and bounded drift and diffusion parts, and its diffusion part is uniformly non-degenerate. It is well-known that, for such a process, $\mathbf{P}_x(\|X^{\varepsilon,\alpha/2} - \phi\|_{0,T} \leq \alpha/2) > 0$ (this can be seen as a consequence of the Girsanov's formula). Since $X^{\varepsilon,\alpha/2}_t = X^\varepsilon_t$ for any $t \leq \inf\{t \geq 0 : d(X^\varepsilon_t) \leq \alpha/2\}$, this implies that $\mathbf{P}_x(\exists t \in [0, T] : X^\varepsilon_t \in B(3\alpha/2)) > 0$, which yields $\mathbf{P}_x(\tau_{\rho/2} < \infty) > 0$, as required. □

Proof of Lemma 3.1 Define the processes Z^1 and Z^2 strong solutions in $(0, \infty)$ to the SDEs

$$Z^i_t = \|x\| + \int_0^t c_i(Z^i_s)ds + B_t$$

for $i = 1, 2$, and stopped when they reach 0. As strong solutions, these processes can be constructed on the same probability space than X^ε (and Y). Define also the stopping times (remind that Y_t is defined only for $t < \langle M\rangle_\infty$)

$$\begin{aligned} \theta_\rho &= \inf\{t \geq 0 : \|Y_t\| = \rho\} \wedge \langle M\rangle_\infty, \\ \theta_0 &= \inf\{t \geq 0 : Y_t = 0\} \wedge \langle M\rangle_\infty, \\ \text{for } i = 1, 2,\ \theta^i &= \inf\{t \geq 0 : Z^i = 0\} \\ \text{and} \quad \theta^{i,\rho} &= \inf\{t \geq 0 : Z^i = \rho\}. \end{aligned}$$

It follows from the definition of Y that $T_{\theta_\rho} = \tau_\rho$ and $T_{\theta_0} = \tau_0$. Then

Lemma 3.2 *Almost surely,* $\forall t < \theta_\rho$, $Z^1 \leq \|Y_t\| \leq Z^2_t$.

Proof of Lemma 3.2 Observe that for $t < \theta_\rho \wedge \theta^1$,

$$\|Y_t\| - Z^1_t = \int_0^t (c(Y_s) - c_1(Z^1_s))ds.$$

If there exists $t_0 \in (0, \theta_\rho \wedge \theta^1)$ such that $\|Y_{t_0}\| = Z^1_{t_0}$, then $(\|Y\| - Z^1)'(t_0) = c(Y_{t_0}) - c_1(Z^1_{t_0}) = c(Y_{t_0}) - c_1(\|Y_{t_0}\|) > 0$, and therefore, $\|Y_t\| > Z^1_t$ for $t > t_0$ in a neighborhood of t_0. Consequently, $Z^1_t \leq \|Y_t\|$ for any $t < \theta_\rho \wedge \theta^1$. Since $Z^1_t = 0$ for $t \geq \tau_1$, this inequality holds for $t < \theta_\rho$. The proof of the other inequality is similar. □

The processes Z^1 and Z^2 are Bessel processes with an additional drift, and we can actually prove that

Lemma 3.3
(a) Z^1 *is recurrent in* $(0, +\infty)$ *if and only if* $\frac{\tilde{b}_0 + da_0/2}{a^0} \geq 1$.

(b) $\mathbf{P}(\theta^2 < \infty) > 0$ *if and only if* $\frac{\tilde{b}^0 + da^0/2}{a_0} < 1$.

Proof of Lemma 3.3 The proof relies on the same functions p and v than in the proof of Theorem 3.1. They are defined by equation (23), where b^ε has to be replaced in our case by c_i, and εa by 1. For the process Z^1, if we fix $\gamma > 0$, then, for any $x > 0$,

$$p(y) = \int_\gamma^y \exp\left[-2\int_\gamma^u c_1(z)dz\right] du$$
$$= -\int_y^\gamma \exp\left[2k\int_u^\gamma \frac{dz}{z} - k'(\gamma - u)\right] du = -C\int_y^\gamma u^{-2k} e^{k'u} du,$$

where we have used the constants $k = \frac{\tilde{b}_0 + da_0/2}{a^0} - \frac{1}{2}$ and $k' = \frac{4K}{\varepsilon a_0}$. Consequently, $p(0+) = -\infty$ if and only if $2k \geq 1$, and $p(+\infty) = +\infty$, which yields (a). A similar computation for Z^2 gives that $p(0+) > -\infty$ if and only if $\frac{\tilde{b}^0 + da^0/2}{a_0} < 1$, which completes the proof of Lemma 3.3. □

We are now able to prove Lemma 3.1. Assume first that $\frac{\tilde{b}_0 + da_0/2}{a^0} \geq 1$, and fix $x \in B(\rho) \setminus \{0\}$. Then, by Lemma 3.3 (a), $\theta^1 = \infty$ a.s. and there exists a.s. $t < \infty$ such that $Z^1_t = \rho$. Therefore, Lemma 3.1 (a) follows from the fact that, by Lemma 3.2, $\forall t \geq 0$, $\|X^\varepsilon_{T_{\langle M \rangle_t}}\| = \|X^\varepsilon_t\| > \|Z^1_{\langle M \rangle_t}\|$.

The proof of Lemma 3.1 (b) is more delicate. Fix $x \in B(\rho/2)$ and assume that $\frac{\tilde{b}^0 + da^0/2}{a_0} < 1$. Define $A = \{\tau_0 < \tau_\rho\} \cup \{\tau_0 = \tau_\rho = \infty$ and $\lim_{t\to+\infty} X^\varepsilon_t = 0\}$. First, it is sufficient to prove that $\mathbf{P}_x(A) \geq \mathbf{P}_x(\theta^2 < \theta^{2,\rho})$. Indeed, if, for $u > 0$, $\bar{\mathbf{P}}_u$ is the law of Z^2 with initial state u, then, since Z^2 is strong Markov, $\mathbf{P}_x(\theta^2 < \theta^{2,\rho}) = \bar{\mathbf{P}}_{\|x\|}(\theta^2 < \theta^{2,\rho})$ is greater than $\bar{\mathbf{P}}_{\rho/2}(\theta^2 < \theta^{2,\rho})$. This is, by Lemma 3.3 (b), a positive constant which can be taken as the positive constant c involved in Lemma 3.1 (b).

So, let us prove that $\mathbf{P}_x(A) \geq \mathbf{P}_x(\theta^2 < \theta^{2,\rho})$. The set $\{\theta^2 < \theta^{2,\rho}\}$ can be decomposed as $B \cup C \cup D$, where

$$B = \{\theta^2 < \theta^{2,\rho} \text{ and } \langle M \rangle_\infty = \infty\},$$
$$C = \{\theta^2 < \theta^{2,\rho},\ \langle M \rangle_\infty < \infty \text{ and } \lim_{t\to\infty} X^\varepsilon_t = 0\}$$
$$\text{and} \quad D = \{\theta^2 < \theta^{2,\rho},\ \langle M \rangle_\infty < \infty \text{ and } \exists \alpha > 0,\ \limsup_{t\to\infty} \|X^\varepsilon_t\| \geq \alpha\}.$$

We will prove that $B \cup C \subset A$, and that $\mathbf{P}(D) = 0$, which will complete the proof of Theorem 3.3.

Observe first that, by Lemma 3.2, $\theta^2 < \theta^{2,\rho}$ implies that, for all $t \geq 0$, $\|X^\varepsilon_t\| = \|Y_{\langle M \rangle_t}\| \leq \|Z^2_{\langle M \rangle_t}\| < \rho$ (the problem is actually to prove $\theta^2 \leq \langle M \rangle_\infty$ a.s.). This implies easily that $B \subset \{\tau_0 < \tau_\rho\} \subset A$ and that $C \subset A$.

Now, let us assume that $\mathbf{P}(D) > 0$. Then, there exists $\alpha > 0$ such that

$$\delta := \mathbf{P}(\theta^2 < \theta^{2,\rho},\ \langle M \rangle_\infty < \infty,\ \limsup_{t\to\infty} \|X^\varepsilon_t\| \geq \alpha) > 0.$$

Define for any $t > 0$ the stopping time $\tau_{\alpha,t} = \inf\{s \geq t : \|X_s^\varepsilon\| \geq \alpha\}$. For any $t > 0$, $\mathbf{P}(\theta^2 < \theta^{2,\rho},\ \langle M\rangle_\infty < \infty,\ \tau_{\alpha,t} < \infty) \geq \delta$. Then

$$\forall t > 0,\ \exists T < \infty,\ \mathbf{P}(\theta^2 < \theta^{2,\rho},\ \langle M\rangle_\infty < \infty,\ \tau_{\alpha,t} < T) \geq \delta/2.$$

We will obtain a contradiction from this statement thanks to the following lemma:

Lemma 3.4 *Given an a.s. finite stopping time S, and $\varepsilon < 1$, for any $h \in (0,1)$*

$$\mathbf{E}\left[\sup_{0<u<h} \|X_{S+u}^\varepsilon - X_S^\varepsilon\|^2\right] \leq 10C^2 h,$$

where C is a bound for b, $\tilde{b}$ and σ on $\mathcal{X}$.

Proof of Lemma 3.4 This is a straightforward consequence of the inequality

$$\|X_{S+u}^\varepsilon - X_S^\varepsilon\|^2 \leq 2\left(\int_S^{S+u} \|b(X_s^\varepsilon) + \varepsilon\tilde{b}(X_s^\varepsilon)\| ds\right)^2 + 2\sqrt{\varepsilon}\left\|\int_S^{S+u} \sigma(X_s^\varepsilon) dW_s\right\|^2$$

and of Doob's inequality. □

Set $h = \delta\alpha^2/160C^2$. Then, it follows from Lemma 3.4 and Tchebichev's inequality that for any a.s. finite stopping time S

$$\mathbf{P}\left(\sup_{0<u<h} \|X_{S+u}^\varepsilon - X_S^\varepsilon\| > \frac{\alpha}{2}\right) \leq \frac{\delta}{4}.$$

This inequality applied to the finite stopping time $\tau_{\alpha,t} \wedge T$ yields the first line of the following inequality, and its forth line makes use of a constant $C > 0$ such that $s^* a(x) s \geq C\|s\|^2$ for any $s \in \mathbb{R}^d$ and $x \in \Gamma_{\alpha/2}$ (given by Proposition 2.2 (iii)) and of the formula (32) for $\langle M\rangle$.

$$\begin{aligned}
\frac{\delta}{4} &\leq \mathbf{P}\left(\theta^2 < \theta^{2,\rho},\ \langle M\rangle_\infty < \infty,\ \tau_{\alpha,t} < T,\ \sup_{0<u<h} \|X_{\tau_{\alpha,t}\wedge T+u}^\varepsilon - X_{\tau_{\alpha,t}\wedge T}^\varepsilon\| \leq \frac{\alpha}{2}\right) \\
&\leq \mathbf{P}\left(\theta^2 < \theta^{2,\rho},\ \langle M\rangle_\infty < \infty,\ \sup_{0<u<h} \|X_{\tau_{\alpha,t}+u}^\varepsilon - X_{\tau_{\alpha,t}}^\varepsilon\| \leq \frac{\alpha}{2}\right) \\
&\leq \mathbf{P}\left(\theta^2 < \theta^{2,\rho},\ \langle M\rangle_\infty < \infty,\ \inf_{0<u<h} \|X_{\tau_{\alpha,t}+u}^\varepsilon\| \geq \frac{\alpha}{2}\right) \\
&\leq \mathbf{P}\left(\langle M\rangle_\infty < \infty,\ \langle M\rangle_{\tau_{\alpha,t}+h} - \langle M\rangle_{\tau_{\alpha,t}} \geq \varepsilon C h\right).
\end{aligned}$$

So

$$\mathbf{P}\left(\langle M\rangle_\infty < \infty,\ \langle M\rangle_\infty - \langle M\rangle_t \geq \varepsilon C h\right) \geq \frac{\delta}{4}$$

holds for any $t > 0$, which is impossible. □

4 Large deviations for X^ε as $\varepsilon \to 0$

This result will be obtained using a transfer technique to carry the LDP from the family $\{\sqrt{\varepsilon}W\}_{\varepsilon>0}$, where W is a standard d-dimensional Brownian motion

(Schilder's Theorem, see for example Dembo and Zeitouni [3]) to the family $\{X^\varepsilon\}_{\varepsilon>0}$, where X^ε is the solution to the SDE (1) defined in section 3.1. The method of the proof, adapted from Doss and Priouret [7], consists in obtaining a function S that maps in some sense the paths of $\sqrt{\varepsilon}W$ on the corresponding paths of X^ε when ε is small.

4.1 Statement of the result

Let us first recall Shilder's Theorem:

Theorem 4.1 *Let W be a d-dimensional standard Brownian motion. Then, given $T>0$, the family of processes $\{\sqrt{\varepsilon}W\}_{\varepsilon>0}$ satisfies a large deviation principle on $\mathcal{C}([0,T],\mathbb{R}^d)$ equipped with the uniform norm $\|\cdot\|_{0,T}$ defined in (7), with good rate function (*i.e. *lower semicontinuous with compact level sets)*

$$J_T(\varphi)=\begin{cases}\frac{1}{2}\int_0^T\|\dot\varphi(t)\|^2dt & \text{if}\quad \varphi\in\mathcal{C}_0^{ac}([0,T],\mathbb{R}^d)\\ +\infty & \text{otherwise,}\end{cases}$$

where $\dot\varphi$ denotes the derivative of φ. Namely, for any open subset O and closed subset C of $\mathcal{C}([0,T],\mathbb{R}^d)$,

$$\begin{aligned}\liminf_{\varepsilon\to 0}\varepsilon\ln\mathbf{P}(\sqrt{\varepsilon}W\in O)&\geq-\inf_{\varphi\in O}J_T(\varphi)\\ \text{and}\quad\limsup_{\varepsilon\to 0}\varepsilon\ln\mathbf{P}(\sqrt{\varepsilon}W\in C)&\leq-\inf_{\varphi\in C}J_T(\varphi).\end{aligned}$$

Let us now give the expression of the rate function involved in our result. Fix $T>0$ and $x\in\overline{\mathcal{X}}$, and define

$$\begin{aligned}\forall\psi\in\mathcal{C}\left([0,T],\overline{\mathcal{X}}\right),\quad & t_\psi=\inf\{t\in[0,T]:\psi(t)\in\Gamma\}\wedge T\\ \text{and}\quad \tilde{\mathcal{C}}_x^{ac}([0,T],\overline{\mathcal{X}})&=\{\psi\in\mathcal{C}_x^{ac}([0,T],\overline{\mathcal{X}})\text{ constant on }[t_\psi,T]\}.\end{aligned}$$

Then, we can define for $\psi\in\mathcal{C}([0,T],\overline{\mathcal{X}})$

$$I_{T,x}(\psi)=\begin{cases}\dfrac{1}{2}\displaystyle\int_0^{t_\psi}[\dot\psi(t)-b(\psi(t))]^*a^{-1}(\psi(t))[\dot\psi(t)-b(\psi(t))]dt\\ \qquad\qquad\qquad\qquad\text{if}\quad\psi\in\tilde{\mathcal{C}}_x^{ac}([0,T],\overline{\mathcal{X}})\\ +\infty\qquad\qquad\qquad\text{otherwise.}\end{cases}\tag{33}$$

The inverse matrix $a^{-1}(x)$ of $a(x)$ is, by Proposition 2.2 (ii), defined for $x\notin\Gamma$, so the quantity inside the integral is well defined. Moreover, since $a(x)$ is a non-negative symmetrical matrix, this quantity is positive, so the integral is well-defined and $I_{T,x}(\psi)$ belongs to $\mathbb{R}_+\cup\{+\infty\}$. When $t_\psi=T$, $I_{T,x}(\psi)$ takes the classical form of rate functions for diffusion processes. Note also that $\tilde b$ does not appear in these expressions. This comes from the fact that b^ε uniformly converges to b when $\varepsilon\to 0$.

Now, we can state

Theorem 4.2 *Assume (H1), (H2), (H3) and (H4). Suppose also that the points $x\in\mathcal{X}$ such that $\nabla_1 g(x,x)=0$ are isolated points of $\mathbb{R}^d$. For any $y\in\overline{\mathcal{X}}$, let $\{X^{\varepsilon,y}\}_{\varepsilon>0}$ be the solution to (1) with initial state y constructed in section 3.1, and let $\mathbf{P}_y^\varepsilon$ denote its law. Fix $T>0$ and $x\in\overline{\mathcal{X}}$. Then, for any open subset O of*

$\mathcal{C}\left([0,T],\overline{\mathcal{X}}\right)$, and for any closed subset C of $\mathcal{C}\left([0,T],\overline{\mathcal{X}}\right)$ such that $\mathcal{C}^1_x([0,T],\mathcal{X}\setminus\Gamma)$ is dense in $C\cap\mathcal{C}_x([0,T],\overline{\mathcal{X}})$,

$$\liminf_{\varepsilon\rightarrow 0,y\rightarrow x}\varepsilon\ln\mathbf{P}^\varepsilon_y(O)=\liminf_{\varepsilon\rightarrow 0,y\rightarrow x}\varepsilon\ln\mathbf{P}(X^{\varepsilon,y}\in O)\geq-\inf_{\psi\in O}I_{T,x}(\psi) \tag{34}$$

$$\limsup_{\varepsilon\rightarrow 0,y\rightarrow x}\varepsilon\ln\mathbf{P}^\varepsilon_y(C)=\limsup_{\varepsilon\rightarrow 0,y\rightarrow x}\varepsilon\ln\mathbf{P}(X^{\varepsilon,y}\in C)\leq-\inf_{\psi\in C}I_{T,x}(\psi). \tag{35}$$

Remark 4.1 *If we take $y=x$ in (34) and (35), we recover the classical form of large deviations bounds. The more general form of Theorem 4.2 is necessary to handle the problem of exit from a domain of section 5.*

Remark 4.2 *Since we cannot prove that the upper bound holds for any closed set C, Theorem 4.2 is actually an incomplete large deviation principle. This comes from the degeneracy of a at points of Γ and the fact that σ is not Lipschitz near Γ. Note also that the fact that the process X^ε stays constant when it reaches Γ plays an important role in the proof. Yet, let us emphasize that the condition on C that we obtain covers* all *the closed sets of interest in the usual applications of large deviations (in particular the closed sets involved in the problem of exit from a domain, see section 5).*

Before proving this theorem, let us prove two corollaries. The first one will be useful in section 5:

Corollary 4.1 *Assume the conditions of Theorem 4.2. Then, for any compact set $K\subset\overline{\mathcal{X}}$, for any open $O\subset\mathcal{C}([0,T],\overline{\mathcal{X}})$, and for any closed $C\subset\mathcal{C}([0,T],\overline{\mathcal{X}})$ such that for any $y\in K$, $\mathcal{C}^1_y([0,T],\mathcal{X}\setminus\Gamma)$ is dense in $C\cap\mathcal{C}_y([0,T],\overline{\mathcal{X}})$,*

$$\liminf_{\varepsilon\rightarrow 0}\varepsilon\ln\inf_{y\in K}\mathbf{P}(X^{\varepsilon,y}\in O)\geq-\sup_{y\in K}\inf_{\psi\in O}I_{T,y}(\psi),$$

$$\text{and}\quad\limsup_{\varepsilon\rightarrow 0}\varepsilon\ln\sup_{y\in K}\mathbf{P}(X^{\varepsilon,y}\in C)\leq-\inf_{y\in K,\ \psi\in C}I_{T,y}(\psi).$$

This result can be easily deduced from Theorem 4.2 (for details, see the proof of Corollary 5.6.15 in Dembo and Zeitouni [3]).

The second corollary of Theorem 4.2 states that X^ε converges as $\varepsilon\rightarrow 0$ to a deterministic function, solution to a differential equation called in the biological litterature *canonical equation of adaptive dynamics* (see [5] and [14]).

Corollary 4.2 *Assume the conditions of Theorem 4.2. Then, for any $T>0$ and $x\in\overline{\mathcal{X}}$, $X^{\varepsilon,x}$ converges in probability as $\varepsilon\rightarrow 0$ to the solution ϕ of $\dot{\phi}=b(\phi)$ with initial state x on $[0,T]$.*

Proof of Corollary 4.2 This follows immediately from (35) applied to the set $C_\eta=\{\psi:\|\psi-\phi\|_{0,T}\geq\eta\}$ for $\eta>0$ (where $\|\cdot\|_{a,b}$ has been defined in (7)), if we can prove that $\inf_{\psi\in C_\eta}I_{T,x}(\psi)>0$ for each $\eta>0$. This can be proved as below:

For any $\psi\in C_\eta$ such that $I_{0,T}(\psi)<\infty$, there exists $t\in[0,T]$ such that $\|\psi_t-\phi_t\|\geq\eta$. Note that ϕ (resp. ψ) is constant after the time t_ϕ (resp. t_ψ) where it reaches Γ, and that $b(x)=0$ for $x\in\Gamma$. Note also that, by

Proposition 2.2 (i), there exists $K > 0$ such that b is K-Lipschitz, and a is bounded by K on $\overline{\mathcal{X}}$. The Cauchy-Schwartz inequality yields

$$
\begin{aligned}
\|\psi_t - \phi_t\|^2 &\leq 2\left\|\int_0^{t\wedge t_\psi}(\dot\psi_s - b(\psi_s))ds\right\|^2 + 2\left\|\int_0^t (b(\psi_s) - b(\phi_s))ds\right\|^2 \\
&\leq 2T\int_0^{t_\psi}\|\dot\psi_s - b(\psi_s)\|^2 ds + 2T\int_0^t \|b(\phi_s) - b(\psi_s)\|^2 ds \\
&\leq \frac{2T}{K}\int_0^{t_\psi}(\dot\psi_s - b(\psi_s))^* a^{-1}(\psi_s)(\dot\psi_s - b(\psi_s))ds \\
&\quad + 2TK^2\int_0^t \|\phi_s - \psi_s\|^2 ds \\
&\leq \frac{2T}{K} I_{T,x}(\psi) + 2TK^2\int_0^t \|\phi_s - \psi_s\|^2 ds,
\end{aligned}
$$

and, by Gronwall's Lemma, $I_{T,x}(\psi) \geq K\eta^2 e^{-2T^2K^2}/2T > 0$. □

Finally, let us make observations about the lower semicontinuity of $I_{T,x}$. Define, when $d \geq 2$,

$$
\tilde I_{T,x}(\psi) = \begin{cases} \frac{1}{2}\int_0^T \mathbf{1}_{\{\psi(t)\notin\Gamma\}}[\dot\psi(t) - b(\psi(t))]^* a^{-1}(\psi(t))[\dot\psi(t) - b(\psi(t))]dt \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{if} \quad \psi \in \mathcal{C}_x^{ac}([0,T],\overline{\mathcal{X}}) \\ +\infty \qquad\qquad\qquad\qquad\qquad\quad \text{otherwise,} \end{cases} \tag{36}
$$

and when $d = 1$, define $\tilde I_{T,x}$ by the same formula, except that the condition $\psi \in \mathcal{C}_x^{ac}([0,T],\overline{\mathcal{X}})$ is replaced by $\psi \in \mathcal{C}_x^{ac}([0,T],\overline{C_x})$, where C_x is the connected component of $\mathcal{X}\setminus\Gamma$ containing x.

Define also for $u \geq 0$

$$
\Phi(u) = \{\varphi \in \mathcal{C}_0([0,T],\mathbb{R}^d) : J_T(\varphi) \leq u\} \tag{37}
$$

$$
\Psi(u) = \{\psi \in \mathcal{C}_x([0,T],\overline{\mathcal{X}}) : I_{T,x}(\psi) \leq u\} \tag{38}
$$

$$
\tilde\Psi(u) = \{\psi \in \mathcal{C}_x([0,T],\overline{\mathcal{X}}) : \tilde I_{T,x}(\psi) \leq u\}. \tag{39}
$$

Note that Schilder's Theorem (Theorem 4.1) implies that $\Phi(u)$ is compact for any $u \geq 0$. Then

Proposition 4.1 *Assume the conditions of Theorem 4.2. Then, for any $u \geq 0$, $\tilde\Psi(u)$ is closed (*i.e.* $\tilde I_{T,x}$ is lower semicontinuous), and $\overline{\Psi(u)} \subset \tilde\Psi(u)$. Assume additionally that there exists an isolated point y of Γ such that g is $\mathcal{C}^2$ at y, and the differential $D = H_{1,1}g(y,y) + H_{1,2}g(y,y)$ of $x \mapsto \nabla_1 g(x,x)$ at y has a null kernel. Then, if $x \notin \Gamma$, $\Psi(u) \subsetneq \overline{\Psi(u)}$, and so $I_{T,x}$ is* not *lower semicontinuous.*

Remark 4.3 *It is always possible to obtain a large deviation principle with a* rate *function (lower semincontinuous) from a large deviation principle with a non-lower semincontinuous "rate" function: if we define the function $\bar I_{T,x}$ by $\bar I_{T,x}(\psi) = \liminf_{\hat\psi\to\psi} I_{T,x}(\hat\psi)$, it is easy to prove that $\bar I_{T,x}$ is lower semicontinuous, that $\bar I_{T,x} \leq I_{T,x}$, and that for any open set O in $\mathcal{C}_x([0,T],\overline{\mathcal{X}})$,*

$\inf_{\psi \in O} \bar{I}_{T,x}(\psi) = \inf_{\psi \in O} I_{T,x}(\psi)$. So (34) and (35) hold with $\bar{I}_{T,x}$ instead of $I_{T,x}$.

Proposition 4.1 shows that $\tilde{I}_{T,x}$ is a good candidate for $\bar{I}_{T,x}$, since $\tilde{I}_{T,x} \leq I_{T,x}$ and $\tilde{I}_{T,x}$ is lower semicontinuous. Unfortunately, we are not able to prove that, for any open set O in $\mathcal{C}_x([0,T],\overline{\mathcal{X}})$, $\inf_{\psi \in O} \tilde{I}_{T,x}(\psi) = \inf_{\psi \in O} I_{T,x}(\psi)$. This is actually equivalent to the equality $\overline{\Psi(u)} = \tilde{\Psi}(u)$. Whether this is true or not is not clear.

Proof of Proposition 4.1 Proposition 4.1 relies on the following lemma, which is Proposition 3.1 of Doss and Priouret [7]:

Lemma 4.1 *Let $\hat{\sigma}$ be a bounded and Lipschitz function from $\mathbb{R}^d$ to $\mathcal{S}_c$ for some $c > 0$, and let $\hat{b}$ be a bounded and Lipschitz function from $\mathbb{R}^d$ to $\mathbb{R}^d$. Define $\hat{I}_T$ on $\mathcal{C}^{ac}([0,T],\mathbb{R}^d)$ as follows:*

$$\hat{I}_T(\psi) = \frac{1}{2}\int_0^T [\dot{\psi}(t) - \hat{b}(\psi(t))]^* \hat{a}^{-1}(\psi(t))[\dot{\psi}(t) - \hat{b}(\psi(t))]dt,$$

where $\hat{a} = \hat{\sigma}\hat{\sigma}^*$. Then $\hat{I}_T$ is lower semicontinuous on $\mathcal{C}^{ac}([0,T],\mathbb{R}^d)$ for the norm of uniform convergence. Moreover, for any compact set K and positive u, $\{\psi \in \mathcal{C}([0,T],\mathbb{R}^d) : \hat{I}_T(\psi) \leq u, \psi(0) \in K\}$ is compact.

Let (ψ_n) be a sequence of functions of $\tilde{\Psi}(u)$ uniformly converging to a function $\psi \in \mathcal{C}_x([0,T],\overline{\mathcal{X}})$. Note that, in the case where $d = 1$, since $\tilde{I}_{T,x}(\psi_n) \leq u$ implies that $\psi_n \in \mathcal{C}_x^{ac}([0,T],\overline{C_x})$, we actually have $\psi \in \mathcal{C}_x^{ac}([0,T],\overline{C_x})$.

For any $\delta > 0$, define $K_\delta = \{t \in [0,T] : d(t, \{t \in [0,T] : \psi(t) \in \Gamma\}) \geq \delta\}$. K_δ is a compact set, made of a finite union of intervals (since between each interval, there is at least a distance of 2δ). By compactness, there exists $\alpha > 0$ such that, for all $t \in K_\delta$, $d(\psi(t),\Gamma) \geq \alpha$. Consequently, for n sufficiently large, $d(\psi_n(t),\Gamma) \geq \alpha/2$ for all $t \in K_\delta$. Define $\hat{a} = a + \chi\mathrm{Id}$ with χ Lipschitz, $\chi \equiv 0$ on $\Gamma_{\alpha/2}$, and $\chi \equiv 1$ on $\Gamma_{\alpha/4}$. Then, by Proposition 2.2 (i) and (iii), $\hat{a}$ is Lipschitz and uniformly non-degenerate and b is Lipschitz, and, therefore, $\hat{\sigma} = \zeta(\hat{a})$ and $\hat{b} = b$ satisfy the assumptions of Lemma 4.1.

Let $[s,t]$ be a maximal interval included in K_δ. Since $\hat{\sigma} = \sigma$ on $\Gamma_{\alpha/2}$, if we replace $[0,T]$ by $[s,t]$ in the statement of Lemma 4.1, we obtain

$$\begin{aligned}\frac{1}{2}\int_s^t &[\dot{\psi}(v) - b(\psi(v))]^* a^{-1}(\psi(v))[\dot{\psi}(v) - b(\psi(v))]dv \\ &\leq \liminf \frac{1}{2}\int_s^t [\dot{\psi}_n(v) - b(\psi_n(v))]^* a^{-1}(\psi_n(v))[\dot{\psi}_n(v) - b(\psi_n(v))]dv.\end{aligned}$$

Consequently,

$$\begin{aligned}\frac{1}{2}\int_{K_\delta} &[\dot{\psi}(v) - b(\psi(v))]^* a^{-1}(\psi(v))[\dot{\psi}(v) - b(\psi(v))]dv \\ &\leq \liminf \frac{1}{2}\int_{K_\delta} [\dot{\psi}_n(v) - b(\psi_n(v))]^* a^{-1}(\psi_n(v))[\dot{\psi}_n(v) - b(\psi_n(v))]dv \leq u.\end{aligned}$$

Finally, since K_δ converges to $\{t \in [0,T] : \psi(t) \notin \Gamma\}$ when $\delta \to 0$, it follows from the monotone convergence Theorem that

$$\tilde{I}_{T,x}(\psi) = \frac{1}{2}\int_{\psi(t)\notin\Gamma} [\dot{\psi}(v) - b(\psi(v))]^* a^{-1}(\psi(v))[\dot{\psi}(v) - b(\psi(v))]dv \leq u,$$

and, therefore, $\tilde{\Psi}(u)$ is closed.

Since $\tilde{I}_{T,x} \leq I_{T,x}$, this implies immediately that $\overline{\Psi(u)} \subset \tilde{\Psi}(u)$.

Now assume that $x \notin \Gamma$ and that there exists an isolated point y of Γ such that g is $\mathcal{C}^2$ at y, and the differential $D = H_{1,1}g(y,y) + H_{1,2}g(y,y)$ of $x \mapsto \nabla_1 g(x,x)$ at y has a null kernel. By translation, we can suppose that $y = 0$. Then, Proposition 3.2 implies that there exists a neighborhood $\mathcal{N}_0$ of 0 and a constant $a_0 > 0$ such that for all $s \in \mathbb{R}^d$ and $x \in \mathcal{N}_0$, $s^* a(x) s \geq a_0 \|x\| \|s\|^2$, *i.e.* each eigenvalue of $a(x)$ is greater than $a_0\|x\|$. Therefore, for all $s \in \mathbb{R}^d$ and $x \in \mathcal{N}_0$,

$$s^* a^{-1}(x) s \leq \frac{\|s\|^2}{a_0 \|x\|}. \tag{40}$$

Firstly, take $x_0 \in \mathcal{X} \setminus \Gamma$ such that the segment $(0, x_0]$ is included in $\mathcal{X} \setminus \Gamma$ and in $\mathcal{N}_0$, and define for $0 \leq t \leq T$

$$\psi(t) = \left(1 - \frac{2t}{T}\right)^2 x_0,$$

and for all $n \geq 1$

$$\psi_n(t) = \begin{cases} \psi(t) & \text{if} \quad t \in \left[0, \frac{T}{2} - \frac{1}{n}\right] \cup \left[\frac{T}{2} + \frac{1}{n}, T\right] \\ \psi\left(\frac{T}{2} - \frac{1}{n}\right) & \text{otherwise.} \end{cases}$$

Since $\psi(T/2 - 1/n) = \psi(T/2 + 1/n)$, ψ_n is continuous and piecewise differentiable. Note that the values of ψ and ψ_n belong to the segment $[0, x_0]$, that $\psi(t) \notin \Gamma$ except if $t = T/2$, and that $\psi_n(t) \notin \Gamma$ for any $t \in [0, T]$. Therefore, $I_{T,x_0}(\psi) = \infty$, and $I_{T,x_0}(\psi_n) < \infty$. Since ψ_n uniformly converges to ψ, it suffices to prove that $\liminf I_{T,x_0}(\psi_n) < \infty$ in order to prove that $\Psi(u) \subsetneq \overline{\Psi}(u)$.

It follows from (40) and from the fact that b is K-Lipschitz (Proposition 2.2), that

$$\begin{aligned} \tilde{I}_{T,x_0}(\psi) &\leq \frac{1}{2a_0} \int_0^T \frac{\|(1 - 2t/T) 2x_0/T + b(\psi(t))\|^2}{\|\psi(t)\|} dt \\ &\leq \frac{1}{2a_0} \int_0^T \frac{2(1 - 2t/T)^2 4\|x_0\|^2/T^2 + 2K^2\|\psi(t)\|^2}{\|\psi(t)\|} dt \\ &\leq \frac{1}{2a_0} \int_0^T \left(\frac{8}{T^2}\|x_0\| + 2K^2\|\psi(t)\|\right) dt < \infty. \end{aligned} \tag{41}$$

Therefore, for all $n \geq 1$,

$$\begin{aligned} I_{T,x_0}(\psi_n) &\leq \tilde{I}_{T,x_0}(\psi) + \frac{1}{2a_0} \int_{T/2-1/n}^{T/2+1/n} \frac{\|b(\psi_n(t))\|^2}{\|\psi_n(t)\|} dt \\ &\leq \tilde{I}_{T,x_0}(\psi) + \frac{1}{2a_0} \int_{T/2-1/n}^{T/2+1/n} K^2 \|\psi_n(t)\| dt, \end{aligned}$$

which is bounded.

It remains to observe that for an arbitrary $x \notin \Gamma$, since $\mathcal{X}$ is connected and the points of Γ are isolated in $\mathcal{X}$, there exists $\phi \in \mathcal{C}^1([0,T], \mathcal{X} \setminus \Gamma)$ such that $\phi(0) = x$ and $\phi(T) = x_0$, and there exists $\alpha > 0$ such that all the values of ϕ belong to Γ_α. Since, by Proposition 2.2 (iii), a is uniformly non-degenerate on

Γ_α, $I_{T,x}(\phi) < \infty$. One can concatenate ϕ and ψ to obtain a function $\tilde{\psi}$ defined on $[0, 2T]$ such that $\tilde{I}_{2T,x}(\tilde{\psi}) < \infty$ and $I_{2T,x}(\tilde{\psi}) = \infty$, and the same procedure as above ends easily the proof of Proposition 4.1 for general x. □

4.2 Proof of Theorem 4.2

Let us first give some notations. For any $\varepsilon > 0$ and $y \in \overline{\mathcal{X}}$, $X^{\varepsilon,y}$ is a weak solution to (1), defined on some filtered probability space $(\Omega^{\varepsilon,y}, \mathcal{F}_t^{\varepsilon,y}, W^{\varepsilon,y}, \mathbf{P}^{\varepsilon,y})$, where $W^{\varepsilon,y}$ is a standard d-dimensional $\mathbf{P}^{\varepsilon,y}$-Brownian motion.

Let us define the function S that transfers the LDP for the Brownian motion to the processes X^ε: for any $\varphi \in \mathcal{C}_0^{ac}([0,T], \mathbb{R}^d)$, let $S(\varphi)$ be the solution on $[0,T]$ to

$$S(\varphi)_t = x + \int_0^t b(S(\varphi)_s)ds + \int_0^t \sigma(S(\varphi)_s)\dot{\varphi}_s ds, \tag{42}$$

obtained in the following way: by Proposition 2.2 (i) and (iii), b and σ are bounded and locally Lipschitz on $\mathcal{X} \setminus \Gamma$, so, by the Cauchy-Lipschitz Theorem, there is local existence and uniqueness in $\mathcal{X} \setminus \Gamma$ of an absolutely continuous function solution to $\dot{y} = b(y) + \sigma(y)\dot{\varphi}$. This defines properly $S(\varphi)$ until the time $t_{S(\varphi)}$ where it reaches Γ. In the case where $t_{S(\varphi)} < T$, set $S(\varphi)_t = S(\varphi)_{t_{S(\varphi)}}$ for $t_{S(\varphi)} \leq t \leq T$. Since for any $x \in \Gamma$, $b(x) = 0$ and $\sigma(x) = 0$, this function $S(\varphi)$ is actually a solution to (42) on $[0,T]$. Hence, we have defined properly the function S from $\mathcal{C}_0^{ac}([0,T], \mathbb{R}^d)$ to $\mathcal{C}_x^{ac}\left([0,T], \overline{\mathcal{X}}\right)$.

The proof of Theorem 4.2 is based on the following three lemmas. Their proof is postponed after the proof of the theorem.

The first lemma precises in which sense the function S maps the paths of $\sqrt{\varepsilon}W$ to the paths of $X^{\varepsilon,y}$ when ε is small and y is close to x.

Lemma 4.2

(i) *Fix $\varphi \in \mathcal{C}_0^{ac}([0,T], \mathbb{R}^d)$ such that $\psi := S(\varphi)$ takes no value in Γ and such that $J_T(\varphi) < +\infty$. Then, $\forall \eta > 0$, $\forall R > 0$, $\exists \delta > 0$ such that*

$$\limsup_{\varepsilon \to 0, y \to x} \varepsilon \ln \mathbf{P}_y^\varepsilon \left(\|X^{\varepsilon,y} - S(\varphi)\|_{0,T} \geq \eta, \|\sqrt{\varepsilon}W^{\varepsilon,y} - \varphi\|_{0,T} \leq \delta \right) \leq -R. \tag{43}$$

(ii) *With the same φ as in (i), $\forall \delta > 0$, $\forall R > 0$, $\exists \eta > 0$ such that*

$$\limsup_{\varepsilon \to 0, y \to x} \varepsilon \ln \mathbf{P}_y^\varepsilon \left(\|X^{\varepsilon,y} - S(\varphi)\|_{0,T} \leq \eta, \|\sqrt{\varepsilon}W^{\varepsilon,y} - \varphi\|_{0,T} \geq \delta \right) \leq -R. \tag{44}$$

(iii) *Fix $\tilde{\varphi} \in \mathcal{C}_0^{ac}([0,T], \mathbb{R}^d)$ such that $\psi_t := S(\tilde{\varphi})_t \in \Gamma$ for some $t \in [0,T]$. Define $\varphi_t = \tilde{\varphi}_t$ for $t < t_\psi$ and $\varphi_t = \tilde{\varphi}_{t_\psi}$ for $t_\psi \leq t \leq T$. Then $S(\varphi) = S(\tilde{\varphi}) = \psi$. Suppose that $J_T(\varphi) < +\infty$. Then, $\forall \eta > 0$, $\forall R > 0$, $\exists \delta > 0$ such that (43) holds.*

Let us briefly comment this lemma. In Doss and Priouret [7], $b^\varepsilon = b + \varepsilon\tilde{b}$ and σ are both supposed Lipschitz on $\mathcal{X}^n$, which is, by Proposition 2.2, only true for b in our case. Moreover, the process X^ε stays constant after the time where it reaches Γ. Because of these difficulties, the method of [7] has to be adapted in order to obtain (i). A more careful study is necessary to obtain (iii), and this only gives the lower bound (34). In order to establish (35), we have firstly to adapt the method of the proof of (i) to prove (ii), and then to obtain the upper bound thanks to a different estimate.

The second lemma specifies the usual relation between S, $I_{T,x}$ and J_T for transfers of large deviation principles.

Lemma 4.3

(i) *For all* $\psi \in \mathcal{C}_x\left([0,T],\overline{\mathcal{X}}\right)$, $I_{T,x}(\psi) = \inf\{J_T(\varphi), S(\varphi) = \psi\}$, *and when* $I_{T,x}(\psi) < +\infty$, *there is unique* $\varphi \in \mathcal{C}_0^{ac}([0,T],\mathbb{R}^d)$ *that realizes this infimum, and this function is constant after* t_ψ.

(ii) $\mathcal{C}^1([0,T],\mathcal{X}\setminus\Gamma)$ *is dense in* $S(\{J_T < \infty\})$.

The last lemma gives a uniform exponential tightness estimate.

Lemma 4.4 *Define for any* $k > 0$ *and* $y \in \overline{\mathcal{X}}$ *the compact set*

$$K_k^y = \left\{\psi \in \mathcal{C}_y([0,T],\overline{\mathcal{X}}) : \forall l \geq k,\ \omega\left(\psi, \frac{1}{l^3}\right) \leq \frac{1}{l}\right\}, \tag{45}$$

where $\omega(\psi,\delta) = \sup_{|t-s|\leq\delta}\|\psi(t)-\psi(s)\|$. *Then, there exists* k_0 *and* ε_0, *such that for all* $y \in \overline{\mathcal{X}}$, $k \geq k_0$ *and* $\varepsilon \leq \varepsilon_0$,

$$\varepsilon \ln \mathbf{P}(X^{\varepsilon,y} \notin K_k^y) \leq -\frac{k}{64d\Sigma^2}, \tag{46}$$

where Σ *is a bound for* σ *on* $\overline{\mathcal{X}}$.

All the preliminary steps required for the proof of Theorem 4.2 have now been completed.

Proof of Theorem 4.2 (34) The lower bound (34) for any open set O is classically equivalent to the fact that $\forall\psi \in \mathcal{C}_x([0,T],\overline{\mathcal{X}})$ and $\forall\eta > 0$,

$$\liminf_{\varepsilon\to 0, y\to x} \varepsilon \ln \mathbf{P}^{\varepsilon,y}(\|X^{\varepsilon,y} - \psi\|_{0,T} \leq \eta) \geq -I_{T,x}(\psi). \tag{47}$$

Fix ψ and η as above, and assume that $I_{T,x}(\psi) < +\infty$ (otherwise, there is nothing to prove). By Lemma 4.3 (i), there is a unique $\varphi \in \mathcal{C}_0^{ac}([0,T],\mathbb{R}^d)$ such that $S(\varphi) = \psi$ and $u := J_T(\varphi) = I_{T,x}(\psi)$. Choose $R > u$. If the image of ψ has empty intersection with Γ, apply Lemma 4.2 (i). Otherwise, apply Lemma 4.2 (iii). In both cases, there exists $\delta > 0$ such that

$$\limsup_{\varepsilon\to 0, y\to x} \varepsilon \ln \mathbf{P}^{\varepsilon,y}\left(\|X^{\varepsilon,y} - \psi\|_{0,T} \geq \eta, \|\sqrt{\varepsilon}W^{\varepsilon,y} - \varphi\|_{0,T} \leq \delta\right) \leq -R.$$

Write

$$\begin{aligned}\mathbf{P}^{\varepsilon,y}(\|\sqrt{\varepsilon}W^{\varepsilon,y} - \varphi\|_{0,T} \leq \delta) &\leq \mathbf{P}^{\varepsilon,y}(\|X^{\varepsilon,y} - \psi\|_{0,T} < \eta) \\ &+ \mathbf{P}^{\varepsilon,y}(\|X^{\varepsilon,y} - \psi\|_{0,T} \geq \eta, \|\sqrt{\varepsilon}W^{\varepsilon,y} - \varphi\|_{0,T} \leq \delta)\end{aligned}$$

and observe that $\mathbf{P}^{\varepsilon,y}(\|\sqrt{\varepsilon}W^{\varepsilon,y} - \varphi\|_{0,T} \leq \delta)$ is independent of ε and y. Take the lim inf of ε times the log of both sides of this inequality: using Schilder's

Theorem on the second line, we obtain

$$\begin{aligned}
-u = -J_T(\varphi) &\le -\inf\{J_T(\tilde\varphi), \tilde\varphi \in B_T(\varphi,\delta)\} \\
&\le \liminf_{\varepsilon\to 0, y\to x} \varepsilon \ln \mathbf{P}^{\varepsilon,y}(\|\sqrt{\varepsilon}W^{\varepsilon,y} - \varphi\|_{0,T} < \delta) \\
&\le \sup\Big\{ \liminf_{\varepsilon\to 0, y\to x} \varepsilon \ln \mathbf{P}^{\varepsilon,y}(\|X^{\varepsilon,y} - \psi\|_{0,T} < \eta), \\
&\qquad \liminf_{\varepsilon\to 0, y\to x} \varepsilon \ln \mathbf{P}^{\varepsilon,y}(\|X^{\varepsilon,y} - \psi\|_{0,T} \ge \eta, \|\sqrt{\varepsilon}W^{\varepsilon,y} - \varphi\|_{0,T} \le \delta)\Big\} \\
&\le \sup\left\{ \liminf_{\varepsilon\to 0, y\to x} \varepsilon \ln \mathbf{P}^{\varepsilon,y}(\|X^{\varepsilon,y} - \psi\|_{0,T} < \eta), -R\right\},
\end{aligned}$$

and since $R > u$, (47) is established.

Proof of Theorem 4.2 (35) Let us first study the case where $x \in \Gamma$. In this case, $\forall\varepsilon > 0$ and $\forall t \ge 0$, $X^\varepsilon_t = x$, and $I_{T,x}(\psi) = +\infty$ as soon as $\psi \not\equiv x$, so (35) is trivial when $x \in C$. Assume that $x \notin C$. Then, there exists $\eta > 0$ such that $B_T(x,\eta) \cap C = \emptyset$, where $B_T(\varphi,\eta)$ has been defined in (8). In particular, if $\|y - x\| < \eta$, there is no function in C with initial value y, so $\mathbf{P}(X^{\varepsilon,y} \in C) = 0$. This yields (35) when $x \in \Gamma$.

Now, fix $x \notin \Gamma$. Let us first establish (35) for particular compact sets.

Let K be a non-empty compact set of $\mathcal{C}\left([0,T], \overline{\mathcal{X}}\right)$ such that $S(\{J_T < +\infty\})$ is dense in K_x, where $K_x := K \cap \mathcal{C}_x([0,T], \overline{\mathcal{X}})$. By Lemma 4.3 (i), $S(\{J_T < +\infty\}) = \{I_{T,x} < +\infty\}$, and so $u := \inf\{I_{T,x}(\psi), \psi \in K\} < +\infty$.

Fix $\rho > 0$. For any $\psi \in K \cap S(\{J_T < +\infty\})$, by Lemma 4.3 (i), there exists a unique $\varphi \in \mathcal{C}^{ac}_0([0,T], \mathbb{R}^d)$ such that $S(\varphi) = \psi$ and $I_{T,x}(\psi) = J_T(\varphi) < \infty$. We intend to use Lemma 4.2 (ii), which holds only if ψ takes no value in Γ. So we have to introduce $\alpha_\psi > 0$ such that

$$\frac{1}{2}\int_{t_\psi - \alpha_\psi}^{t_\psi} \|\dot\varphi_s\|^2 ds < \frac{\rho}{2},$$

so that $J_T(\varphi) \le J_{t_\psi - \alpha_\psi}(\varphi) + \rho/2$ (by Lemma 4.3 (i), $\dot\varphi_t = 0$ for $t > t_\psi$), and, since $J_{t_\psi - \alpha_\psi}$ is lower semicontinuous, there exists $\delta_\psi > 0$ such that

$$\forall \tilde\varphi \in B_{t_\psi - \alpha_\psi}(\varphi, \delta_\psi), \quad J_{t_\psi - \alpha_\psi}(\tilde\varphi) \ge J_{t_\psi - \alpha_\psi}(\varphi) - \frac{\rho}{2} \ge J_T(\varphi) - \rho, \tag{48}$$

where $B_t(\varphi,\delta)$ has been defined in (8).

Since $\psi_t \notin \Gamma$ for any $t \in [0, t_\psi - \alpha_\psi]$, we can apply Lemma 4.2 (ii) with $T = t_\psi - \alpha_\psi$, $\delta = \delta_\psi$ and $R > u$: there exists $\eta_\psi > 0$ such that

$$\limsup_{\varepsilon\to 0, y\to x} \varepsilon \ln \mathbf{P}^{\varepsilon,y}\left(\|X^{\varepsilon,y} - \psi\|_{0,t_\psi-\alpha_\psi} \le \eta_\psi, \|\sqrt{\varepsilon}W^{\varepsilon,y} - \varphi\|_{0,t_\psi-\alpha_\psi} \ge \delta_\psi\right) \le -R.$$

Since we have assumed that $K_x = \overline{K_x \cap S(\{J_T < +\infty\})}$, we can write

$$K_x \subset \bigcup_{\psi \in K_x \cap S(\{J_T < +\infty\})} B_T(\psi, \eta_\psi),$$

so, from the compactness of K_x follows the existence of a finite number of functions $\psi_1, \dots, \psi_n$ in $K_x \cap S(\{J_T < +\infty\})$ such that

$$K_x \subset \bigcup_{i=1}^{n} B_T(\psi_i, \eta_i),$$

where we wrote η_i instead of η_{ψ_i}. It easily follows from the compactness of K that there exists a neighborhood $\mathcal{N}_x$ of x such that

$$K_{\mathcal{N}_x} \subset \bigcup_{i=1}^{n} B_T(\psi_i, \eta_i),$$

where $K_{\mathcal{N}_x} = \{\psi \in K : \psi(0) \in \mathcal{N}_x\}$. Define

$$U = \bigcup_{i=1}^{n} B_{t_i-\alpha_i}(\varphi_i, \delta_i),$$

where $t_i = t_{\psi_i}$, $\alpha_i = \alpha_{\psi_i}$ and $\delta_i = \delta_{\psi_i}$, and where φ_i is the function satisfying $S(\varphi_i) = \psi_i$ and $I_{Tx}(\psi_i) = J_T(\varphi_i)$.

Then, it remains to write for $y \in \mathcal{N}_x$

$$\begin{aligned}
\mathbf{P}^{\varepsilon,y}(X^{\varepsilon,y} \in K) &\le \mathbf{P}^{\varepsilon,y}(\sqrt{\varepsilon}W^{\varepsilon,y} \in U) + \mathbf{P}^{\varepsilon,y}(\sqrt{\varepsilon}W^{\varepsilon,y} \notin U,\ X^{\varepsilon,y} \in K_{\mathcal{N}_x}) \\
&\le \sum_{i=1}^{n} \mathbf{P}^{\varepsilon,y}(\sqrt{\varepsilon}W^{\varepsilon,y} \in B_{t_i-\alpha_i}(\varphi_i, \delta_i)) \\
&\quad + \sum_{i=1}^{n} \mathbf{P}^{\varepsilon,y}(\|X^{\varepsilon,y} - \psi_i\|_{0,T} \le \eta_i, \sqrt{\varepsilon}W^{\varepsilon,y} \notin U) \\
&\le \sum_{i=1}^{n} \mathbf{P}^{\varepsilon,y}(\|\sqrt{\varepsilon}W^{\varepsilon,y} - \varphi_i\|_{0,t_i-\alpha_i} < \delta_i) \\
&\quad + \sum_{i=1}^{n} \mathbf{P}^{\varepsilon,y}(\|X^{\varepsilon,y} - \psi_i\|_{0,t_i-\alpha_i} \le \eta_i, \|\sqrt{\varepsilon}W^{\varepsilon,y} - \varphi_i\|_{0,t_i-\alpha_i} \ge \delta_i),
\end{aligned}$$

and to observe that, by Schilder's Theorem and (48),

$$\begin{aligned}
&\limsup_{\varepsilon\to 0, y\to x} \varepsilon \ln \mathbf{P}^{\varepsilon,y}(\|\sqrt{\varepsilon}W^{\varepsilon,y} - \varphi_i\|_{0,t_i-\alpha_i} \le \delta_i) \\
&\qquad \le - \inf_{\varphi \in B_{t_i-\alpha_i}(\varphi_i,\delta_i)} J_{t_i-\alpha_i}(\varphi) \le -J_T(\varphi_i) + \rho,
\end{aligned}$$

to obtain

$$\begin{aligned}
&\limsup_{\varepsilon\to 0, y\to x} \varepsilon \ln \mathbf{P}^{\varepsilon,y}(X^{\varepsilon,y} \in K) \\
&\quad \le \sup\Big\{ \sup_{1\le i\le n} \limsup_{\varepsilon\to 0, y\to x} \varepsilon \ln \mathbf{P}^{\varepsilon,y}(\|\sqrt{\varepsilon}W^{\varepsilon,y} - \varphi_i\|_{0,t_i-\alpha_i} < \delta_i), \\
&\sup_{1\le i\le n} \limsup_{\varepsilon\to 0, y\to x} \varepsilon \ln \mathbf{P}^{\varepsilon,y}(\|X^{\varepsilon,y} - \psi_i\|_{0,t_i-\alpha_i} \le \eta_i, \|\sqrt{\varepsilon}W^{\varepsilon,y} - \varphi_i\|_{0,t_i-\alpha_i} \ge \delta_i) \Big\} \\
&\quad \le \sup\Big\{ - \inf_{1\le i\le n} (J_T(\varphi_i) - \rho), -R \Big\} \\
&\quad \le \sup\{ -\inf\{I_{T,x}(\psi), \psi \in K\} + \rho, -R\} \le -u + \rho.
\end{aligned}$$

Since this holds for any $\rho > 0$, the proof of (35) for the set K is complete.

Now, let C be a closed subset of $\mathcal{C}([0,T],\overline{\mathcal{X}})$ such that $\mathcal{C}^1_x([0,T],\mathcal{X}\setminus\Gamma)$ is dense in $C\cap\mathcal{C}_x([0,T],\overline{\mathcal{X}})$. Define the compact set

$$\begin{aligned}K_k &= \{\psi\in\mathcal{C}([0,T],\overline{\mathcal{X}}) : \|\psi(0)-x\|\leq 1, \forall l\geq k, \omega(\psi,1/k^3)\leq 1/k\}\\ &= \cup_{\|y-x\|\leq 1} K^y_k.\end{aligned}$$

The following exponential tightness estimate follows trivially from Lemma 4.4:

$$\limsup_{\varepsilon\to 0, y\to x}\varepsilon\ln\mathbf{P}(X^{\varepsilon,y}\notin K_k)\leq -k/64d\Sigma^2. \tag{49}$$

The exponential tightness method requires to decompose C as $(C\cap(K_k)^c)\cup(C\cap K_k)$, to use inequality (49) to bound the probability of the first set, and to use the large deviations upper bound for compact sets to bound the probability of the second set. Unfortunately, $S\{(J_T<\infty\})$ may not be dense in $C\cap K_k\cap\mathcal{C}_x([0,T],\overline{\mathcal{X}})$. However, (49) still holds if we replace K_k by any bigger compact set. Therefore, we will introduce a compact set $\tilde{K}_k\supset K_k$ so that $S\{(J_T<\infty\})$ should be dense in $C\cap\tilde{K}_k\cap\mathcal{C}_x([0,T],\overline{\mathcal{X}})$.

Let us construct $\tilde{K}_k$ as follows. $C\cap K_k\cap\mathcal{C}_x([0,T],\overline{\mathcal{X}})$ is compact, so it is separable. Let ψ_n be a sequence of functions dense in this set. For all $n\geq 0$, $\psi_n\in C$, so, by assumption, there exists a sequence $(\psi_{n,p})_{p\geq 0}$ in $C\cap\mathcal{C}^1_x([0,T],\mathcal{X}\setminus\Gamma)$ converging to ψ_n. We can moreover assume that $\|\psi_{n,p}-\psi_n\|_{0,T}\leq 2^{-p}$ for all $p\geq 0$. Let us define

$$\tilde{K}_k = K_k\cup\left(\bigcup_{n\geq 0}\{\psi_{n,p} : p\geq n\}\right),$$

and let us prove that $\tilde{K}_k$ is compact. Let (ϕ_m) be a sequence of $\tilde{K}_k$, and let us extract a converging subsequence. The only problem is when $\{m:\phi_m\in K_k\}$ is finite, and when for all $n\geq 0$, $\{m:\phi_m\in\{\psi_{n,p}:p\geq n\}\}$ is finite. In this case, there exists two (strictly) increasing sequences of integers (α_m) and (β_m) such that for all $m\geq 0$, $\phi_{\alpha_m}\in\{\psi_{\beta_m,p}:p\geq\beta_m\}$. For all $m\geq 0$, ψ_{β_m} belongs to the compact set $C\cap K_k\cap\mathcal{C}_x([0,T],\overline{\mathcal{X}})$, so, extracting a subsequence from (α_m) and (β_m), we can suppose that $\psi_{\beta_m}\to\psi\in C\cap K_k\cap\mathcal{C}_x([0,T],\overline{\mathcal{X}})$. Then

$$\|\phi_{\alpha_m}-\psi\|_{0,T}\leq\|\phi_{\alpha_m}-\psi_{\beta_m}\|_{0,T}+\|\psi_{\beta_m}-\psi\|_{0,T}\leq 2^{-\beta_m}+\|\psi_{\beta_m}-\psi\|,$$

which converge to 0 when $m\to\infty$. Hence $\phi_{\alpha_m}\to\psi$, and $\tilde{K}_k$ is compact.

Moreover, since $\tilde{K}_k\supset K_k$, it follows from (49) that

$$\limsup_{\varepsilon\to 0, y\to x}\varepsilon\ln\mathbf{P}(X^{\varepsilon,y}\notin\tilde{K}_k)\leq -k/64d\Sigma^2, \tag{50}$$

and $\tilde{K}_k$ has been constructed in such a way that $\mathcal{C}^1_x([0,T],\mathcal{X}\setminus\Gamma)$ is dense in $C\cap\tilde{K}_k\cap\mathcal{C}_x([0,T],\overline{\mathcal{X}})$. So, by Lemma 4.3 (ii), $S(\{J_T<\infty\})$ is dense in $C\cap\tilde{K}_k\cap\mathcal{C}_x([0,T],\overline{\mathcal{X}})$, and we can apply the first part of the proof to the compact set $C\cap\tilde{K}_k$:

$$\limsup_{\varepsilon\to 0, y\to x}\varepsilon\ln\mathbf{P}(X^{\varepsilon,y}\in C\cap\tilde{K}_k)\leq -\inf\{I_{T,x}(\psi),\psi\in C\cap\tilde{K}_k\}.$$

Together with (50), this yields for sufficiently large k

$$\limsup_{\varepsilon\to 0, y\to x} \varepsilon \ln \mathbf{P}(X^{\varepsilon,y} \in C) \leq \sup \left\{ \limsup_{\varepsilon\to 0, y\to x} \varepsilon \ln \mathbf{P}(X^{\varepsilon,y} \in C \cap \tilde{K}_k), \right.$$
$$\left. \limsup_{\varepsilon\to 0, y\to x} \varepsilon \ln \mathbf{P}(X^{\varepsilon,y} \notin \tilde{K}_k) \right\}$$
$$\leq - \inf_{\psi \in C} I_{T,x}(\psi),$$

and the proof of (35) is complete □

The proofs of Lemmas 4.2 and 4.3 are based on the following lemma, generalization of classical exponential inequalities for stochastic integrals.

Let $\mathcal{M}_{d,d}$ denote the set of real $d \times d$ matrices, and let $\|\cdot\|$ be the norm on this vector space defined by $\|M\| = \sup_{\|\zeta\|=1} \|M\zeta\|$. Then,

Lemma 4.5 *Let Y_t be a $\mathcal{F}_t$-martingale with values in $\mathbb{R}^d$ on a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$, and suppose that its quadratic covariation process $\langle Y \rangle_t$ satisfies $\sup_{t\leq T} \|\langle Y \rangle_t\| \leq A$. Let τ be a $\mathcal{F}_t$ stopping time, and let $Z : \mathbb{R}_+ \times \Omega \to \mathcal{M}_{d,d}$ be a progressively measurable process such that $\sup_{t\leq\tau} \|Z_t^*\| \leq B$ (where Z_t^* is the transpose matrix of Z_t). Then for any $R > 0$,*

$$\mathbf{P}\left(\sup_{t\leq T} \left\| \int_0^{t\wedge\tau} Z_s dY_s \right\| \geq R \right) \leq 2d \exp\left(-\frac{R^2}{2dTAB^2} \right).$$

Proof of Lemma 4.5 For any $v \in \mathbb{R}^d$, let $M(v)$ be the exponential local martingale defined for $t > 0$ by

$$M_t(v) = \exp\left[v^* \int_0^{t\wedge\tau} Z_s dY_s - \frac{1}{2} \int_0^{t\wedge\tau} (v^* Z_s) d\langle Y \rangle_s (v^* Z_s)^* \right].$$

Since $\int_0^{t\wedge\tau} (v^* Z_s) d\langle Y \rangle_s (v^* Z_s)^* \leq TAB^2 \|v\|^2$, by Novikov's criterion, $M_t(v)$ is actually a martingale.

Then Doob's inequality gives that if $\|v\| = 1$ and $\lambda > 0$,

$$\mathbf{P}\left(\sup_{0\leq t\leq T} v^* \int_0^{t\wedge\tau} Z_s dY_s \geq \frac{R}{\sqrt{d}} \right)$$
$$\leq \mathbf{P}\left(\sup_{0\leq t\leq T} M_t(\lambda v) \geq \exp\left(\frac{\lambda R}{\sqrt{d}} - \frac{\lambda^2 AB^2 T}{2} \right) \right)$$
$$\leq \exp\left(-\frac{\lambda R}{\sqrt{d}} + \frac{\lambda^2 AB^2 T}{2} \right).$$

The infimum of the right-hand side is obtained when $\lambda = \frac{R}{\sqrt{d}AB^2T}$, which gives

$$\mathbf{P}\left(\sup_{0\leq t\leq T} v^* \int_0^{t\wedge\tau} Z_s dY_s \geq \frac{R}{\sqrt{d}} \right) \leq \exp\left(-\frac{R^2}{2dAB^2T} \right).$$

Finally, let us consider an orthonormal basis $\{v_1, \dots, v_d\}$ of $\mathbb{R}^d$. Then

$$\begin{aligned}\mathbf{P}\left(\sup_{t\le T}\left\|\int_0^{t\wedge\tau} Z_s dY_s\right\| \ge R\right) &= \mathbf{P}\left(\sup_{t\le T}\sum_{i=1}^d\left(v_i\cdot\int_0^{t\wedge\tau} Z_s dY_s\right)^2 \ge R^2\right)\\ &\le \sum_{i=1}^d \mathbf{P}\left(\sup_{t\le T}\left|v_i^*\int_0^{t\wedge\tau} Z_s dY_s\right| \ge \frac{R}{\sqrt{d}}\right)\\ &\le 2d\exp\left(-\frac{R^2}{2dTAB^2}\right)\end{aligned}$$

which is exactly the required result. □

Proof of Lemma 4.2 Let φ be as in any point of Lemma 4.2. We will first use the Girsanov's Theorem to restrict ourselves to the case $\varphi = 0$.

Define on $(\Omega^{\varepsilon,y}, \mathcal{F}_T^{\varepsilon,y})$ the probability measure $\tilde{\mathbf{P}}^{\varepsilon,y}$ by

$$\frac{d\tilde{\mathbf{P}}^{\varepsilon,y}}{d\mathbf{P}^{\varepsilon,y}} = \exp\left(\frac{1}{\sqrt{\varepsilon}}\int_0^t \dot\varphi_s dW_s^{\varepsilon,y} - \frac{1}{2\varepsilon}\int_0^t \|\dot\varphi_s\|^2 ds\right) := Z_t.$$

Since in all cases $J_T(\varphi) = 1/2\int_0^T \|\dot\varphi_t\|^2 dt < +\infty$, by Novikov's criterion, $(Z_t)_{t\le T}$ is a $\mathbf{P}^{\varepsilon,y}$-martingale, and it follows from Girsanov's Theorem that $\tilde W_t^{\varepsilon,y} := W_t^{\varepsilon,y} - \varphi_t/\sqrt{\varepsilon}$ is a $\tilde{\mathbf{P}}^{\varepsilon,y}$-Brownian motion for $t \le T$. If we denote by $\tilde X^{\varepsilon,y}$ the process $X^{\varepsilon,y}$ on $(\Omega^{\varepsilon,y}, \mathcal{F}_t^{\varepsilon,y}, \tilde W^{\varepsilon,y}, \tilde{\mathbf{P}}^{\varepsilon,y})$, then, $\tilde{\mathbf{P}}^{\varepsilon,y}$-a.s., for $t \le T$

$$\tilde X_t^{\varepsilon,y} = y + \int_0^t (b^\varepsilon(\tilde X_s^{\varepsilon,y}) + \sigma(\tilde X_s^{\varepsilon,y})\dot\varphi_s)ds + \sqrt{\varepsilon}\int_0^t \sigma(\tilde X_s^{\varepsilon,y})d\tilde W_s^{\varepsilon,y}. \tag{51}$$

Let $F^{\varepsilon,y}$ denote the event $\{\|X^{\varepsilon,y} - S(\varphi)\|_{0,T} \ge \eta,\ \|\sqrt{\varepsilon}W^{\varepsilon,y} - \varphi\|_{0,T} \le \delta\} = \{\|\tilde X^{\varepsilon,y} - S(\varphi)\|_{0,T} \ge \eta,\ \|\sqrt{\varepsilon}\tilde W^{\varepsilon,y}\|_{0,T} \le \delta\}$. It follows from Cauchy-Schwartz's Theorem that

$$\mathbf{P}^{\varepsilon,y}(F^{\varepsilon,y}) = \int \mathbf{1}_{F^{\varepsilon,y}}\frac{d\mathbf{P}^{\varepsilon,y}}{d\tilde{\mathbf{P}}^{\varepsilon,y}}d\tilde{\mathbf{P}}^{\varepsilon,y} \le \left(\tilde{\mathbf{P}}^{\varepsilon,y}(F^{\varepsilon,y})\right)^{\frac12}\left(\int\left(\frac{d\mathbf{P}^{\varepsilon,y}}{d\tilde{\mathbf{P}}^{\varepsilon,y}}\right)^2 d\tilde{\mathbf{P}}^{\varepsilon,y}\right)^{\frac12}. \tag{52}$$

Since $W_t^{\varepsilon,y} = \tilde W_t^{\varepsilon,y} + \varphi_t/\sqrt{\varepsilon}$, we can write

$$\begin{aligned}\left(\frac{d\mathbf{P}^{\varepsilon,y}}{d\tilde{\mathbf{P}}^{\varepsilon,y}}\right)^2 &= \exp\left(-\frac{2}{\sqrt{\varepsilon}}\int_0^T \dot\varphi_s d\tilde W_s^{\varepsilon,y} - \frac{1}{\varepsilon}\int_0^T \|\dot\varphi_s\|^2 ds\right)\\ &= \exp\left(\int_0^T\left(-\frac{2\dot\varphi_s}{\sqrt{\varepsilon}}\right)d\tilde W_s^{\varepsilon,y} - \frac12\int_0^T\left\|\frac{2\dot\varphi_s}{\sqrt{\varepsilon}}\right\|^2 ds\right)\\ &\quad\times\exp\left(\frac{1}{\varepsilon}\int_0^T \|\dot\varphi_s\|^2 ds\right).\end{aligned}$$

The first term in the product of the right-hand side is a $\tilde{\mathbf{P}}^{\varepsilon,y}$-martingale (by Novikov's criterion), and the second term is equal to $\exp(2J_T(\varphi)/\varepsilon)$. Therefore, (52) yields

$$\varepsilon\ln\mathbf{P}^{\varepsilon,y}(F^{\varepsilon,y}) \le \frac{\varepsilon}{2}\ln\tilde{\mathbf{P}}^{\varepsilon,y}(F^{\varepsilon,y}) + J_T(\varphi).$$

So, in order to complete the proof of Lemma 4.2, it suffices to prove the following lemma:

Lemma 4.6 *The three points of Lemma 4.2 hold when (43) and (44) are replaced respectively by*

$$\limsup_{\varepsilon\to 0, y\to x} \varepsilon \ln \tilde{\mathbf{P}}^{\varepsilon,y}\left(\|\tilde{X}^{\varepsilon,y} - S(\varphi)\|_{0,T} \geq \eta, \|\sqrt{\varepsilon}\tilde{W}^{\varepsilon,y}\|_{0,T} \leq \delta\right) \leq -R \quad (53)$$

and

$$\limsup_{\varepsilon\to 0, y\to x} \varepsilon \ln \tilde{\mathbf{P}}^{\varepsilon,y}\left(\|\tilde{X}^{\varepsilon,y} - S(\varphi)\|_{0,T} \leq \eta, \|\sqrt{\varepsilon}\tilde{W}^{\varepsilon,y}\|_{0,T} \geq \delta\right) \leq -R. \quad (54)$$

In order to keep notations simple, we will write throughout the proof of this lemma W instead of $\tilde{W}^{\varepsilon,y}$ and $\mathbf{P}$ instead of $\tilde{\mathbf{P}}^{\varepsilon,y}$, for events involving the process $\tilde{X}^{\varepsilon,y}$ solution to (51).

Lemma 4.6 relies on the following lemma, deduced from Lemma 4.5 adapting the proof of Lemma 1.3 in Doss and Priouret [7].

Lemma 4.7 *Let $\tilde{X}^{\varepsilon,y}$ be defined by (51) with φ satisfying $J_T(\varphi) < \infty$. Let Y_t be a $\mathcal{F}_t^{\varepsilon,y}$-martingale such that $\sup_{t\leq T}\|\langle Y\rangle_t\| \leq A$, let τ be a $\mathcal{F}_t^{\varepsilon,y}$-stopping time, and let ξ be uniformly continuous bounded function defined on $\overline{\mathcal{X}}$. Then, $\forall\eta > 0$, $\forall R > 0$, $\exists\delta > 0$, $\exists\varepsilon_0 > 0$ both depending on Y only through A, and both independant of τ, such that $\forall y \in \overline{\mathcal{X}}$, $\forall\varepsilon < \varepsilon_0$,*

$$\varepsilon \ln \mathbf{P}\left(\left\|\sqrt{\varepsilon}\int_0^{\cdot\wedge\tau} \xi(\tilde{X}_s^{\varepsilon,y})dY_s\right\|_{0,T} \geq \eta,\ \|\sqrt{\varepsilon}Y\|_{0,T} \leq \delta\right) \leq -R. \quad (55)$$

Proof of Lemma 4.7 We use a discretization technique: for any $p \in \mathbb{N}$, let us define $\tilde{X}_t^{\varepsilon,y,p} = \tilde{X}_{k2^{-p}}^{\varepsilon,y}$, where $k \in \mathbb{N}$ is such that $k \leq t2^{-p} < k+1$. Given $\gamma > 0$, $p \geq 1$ and $\delta > 0$, to be determined next, we can write

$$\left\{\left\|\sqrt{\varepsilon}\int_0^{\cdot\wedge\tau} \xi(\tilde{X}_s^{\varepsilon,y})dY_s\right\|_{0,T} \geq \eta,\ \|\sqrt{\varepsilon}Y\|_{0,T} \leq \delta\right\} \subset A^\varepsilon \cup B^\varepsilon \cup C^\varepsilon,$$

where

$$A^\varepsilon = \{\|\tilde{X}^{\varepsilon,y} - \tilde{X}^{\varepsilon,y,p}\|_{0,\tau} \geq \gamma\},$$

$$B^\varepsilon = \left\{\|\tilde{X}^{\varepsilon,y} - \tilde{X}^{\varepsilon,y,p}\|_{0,\tau} \leq \gamma,\ \left\|\sqrt{\varepsilon}\int_0^{\cdot\wedge\tau}[\xi(\tilde{X}_s^{\varepsilon,y}) - \xi(\tilde{X}_s^{\varepsilon,y,p})]dY_s\right\|_{0,T} \geq \frac{\eta}{2}\right\}$$

$$\text{and}\quad C^\varepsilon = \left\{\left\|\sqrt{\varepsilon}\int_0^{\cdot\wedge\tau} \xi(\tilde{X}_s^{\varepsilon,y,p})dY_s\right\|_{0,T} \geq \frac{\eta}{2},\ \|\sqrt{\varepsilon}Y\|_{0,T} \leq \delta\right\}.$$

We will choose first γ such that $\mathbf{P}(B^\varepsilon)$ is sufficiently small, then choose $p \geq 1$ to control $\mathbf{P}(A^\varepsilon)$, and finally choose $\delta > 0$ such that $C^\varepsilon = \emptyset$.

Firstly, let us apply Lemma 4.5 with $Z_t = \sqrt{\varepsilon}[\xi(\tilde{X}_t^{\varepsilon,y}) - \xi(\tilde{X}_t^{\varepsilon,y,p})]$. If we define $B_\gamma := \sup_{\|x-y\|\leq\gamma}\|\xi(x) - \xi(y)\|$, then, on B^ε, for $t \leq \tau$, $\|Z_t^*\| \leq \sqrt{\varepsilon}B_\gamma$. Therefore, Lemma 4.5 gives

$$\mathbf{P}(B^\varepsilon) \leq 2d\exp\left(-\frac{\eta^2/4}{2dTA\varepsilon B_\gamma^2}\right).$$

Now, since ξ is uniformly continuous, $B_\gamma \to 0$ when $\gamma \to 0$. Therefore, for $\varepsilon < 1$, $\varepsilon \ln \mathbf{P}(B^\varepsilon)$ can be made smaller than $-2R$ choosing γ small enough.

Secondly, $\gamma > 0$ being fixed as above, equation (51) yields

$$\begin{aligned}
&\mathbf{P}(\|\tilde{X}^{\varepsilon,y} - \tilde{X}^{\varepsilon,y,p}\|_{0,\tau} \geq \gamma) \\
&\quad \leq \sum_{k=0}^{T2^p-1} \mathbf{P}\left(\sup_{k2^{-p} \leq t \leq (k+1)2^{-p}} \left\| \int_{k2^{-p}\wedge\tau}^{t\wedge\tau} \sqrt{\varepsilon}\sigma(\tilde{X}_s^{\varepsilon,y})dW_s \right\| \geq \frac{\gamma}{2} \right) \\
&\qquad + \sum_{k=0}^{T2^p-1} \mathbf{P}\left(\sup_{k2^{-p} \leq t \leq (k+1)2^{-p}} \left\| \int_{k2^{-p}\wedge\tau}^{t\wedge\tau} \left[b^\varepsilon(\tilde{X}_s^{\varepsilon,y}) + \sigma(\tilde{X}_s^{\varepsilon,y})\dot{\varphi}_s \right] ds \right\| \geq \frac{\gamma}{2} \right) \\
&\quad \leq \sum_{k=0}^{T2^p-1} \mathbf{P}\left(\sup_{k2^{-p} \leq t \leq (k+1)2^{-p}} \left\| \int_{k2^{-p}\wedge\tau}^{t\wedge\tau} \sqrt{\varepsilon}\sigma(\tilde{X}_s^{\varepsilon,y})dW_s \right\| \geq \frac{\gamma}{2} \right) \\
&\qquad + \sum_{k=0}^{T2^p-1} \mathbf{P}\left(C2^{-p} + C2^{-p/2}\sqrt{u} \geq \frac{\gamma}{2} \right),
\end{aligned}$$

where C is a bound for b^ε and σ, where $u = 2J_T(\varphi) = \int_0^T \|\dot{\varphi}_s\|^2 ds < +\infty$, and where we have used the Cauchy-Schwartz inequality to obtain $\sqrt{u}$ in the last line of this inequality. For p big enough, the second sum of the right-hand side equals 0. For the first sum, Lemma 4.5 with $\tau = T = 2^{-p}$, $Y = W$, $A = 1$, $R = \gamma/2$ and $B = \sqrt{\varepsilon}C$ gives that, for $0 \leq k < T2^p$,

$$\mathbf{P}\left(\sup_{k2^{-p} \leq t \leq (k+1)2^{-p}} \left\| \int_{k2^{-p}\wedge\tau}^{t\wedge\tau} \sqrt{\varepsilon}\sigma(\tilde{X}_s^{\varepsilon,y})dW_s \right\| \geq \frac{\gamma}{2} \right) \leq 2d \exp\left(-\frac{\gamma^2/4}{2d2^{-p}C^2\varepsilon} \right).$$

Therefore, taking p large enough, $\varepsilon \ln \mathbf{P}(A^\varepsilon) \leq -2R$ for all $\varepsilon < 1$.

Finally, with $p \geq 1$ and $\gamma > 0$ as above, for $t \leq T$,

$$\sqrt{\varepsilon} \int_0^{t\wedge\tau} \xi(\tilde{X}_s^{\varepsilon,y,p})dY_s = \sum_{i=0}^{T2^p-1} \sqrt{\varepsilon}\xi(\tilde{X}^{\varepsilon,y}_{i2^{-p}\wedge\tau})[Y_{(i+1)2^{-p}\wedge t\wedge\tau} - Y_{i2^{-p}\wedge t\wedge\tau}].$$

On C^ε, $\|\sqrt{\varepsilon}Y\|_{[0,T]} \leq \delta$, so, for $t \leq T$,

$$\left\| \sqrt{\varepsilon} \int_0^{t\wedge\tau} \xi(\tilde{X}_s^{\varepsilon,y,p})dY_s \right\| \leq \sum_{i=0}^{T2^p-1} 2\delta C,$$

where C is a bound for ξ. Hence $C^\varepsilon = \emptyset$ as soon as $\delta < \eta 2^{-(p+2)}/CT$.

We obtain that $\varepsilon \ln \mathbf{P}(A^\varepsilon \cup B^\varepsilon \cup C^\varepsilon) \leq \varepsilon \ln 2 - 2R$, which yields (55) as soon as $\varepsilon < R/\ln 2 \wedge 1$.

This argument is true for any $y \in \overline{\mathcal{X}}$ and for any stopping time τ. It remains to observe that A is the only information about Y that we used to estimate $\mathbf{P}(B^\varepsilon)$, that Y does not appear in A^ε, and that no assumption about Y is necessary to obtain $C^\varepsilon = \emptyset$. Hence, the constant A is the only information about Y required to obtain δ and ε_0. □

Now, let us prove Lemma 4.6.

Proof of Lemma 4.6 (i) The function $\psi = S(\varphi)$ does not take any value in Γ on $[0,T]$, so there exists $\alpha > 0$ such that $\forall t \in [0,T]$, $\psi_t \in \Gamma_\alpha$. Suppose without loss of generality that $\eta < \alpha/2$, and define for $y \in \overline{\mathcal{X}}$

$$\tau^{\varepsilon,y} = \inf\{t : d(\tilde{X}^{\varepsilon,y}_t, \Gamma) \leq \alpha/2\} \wedge T.$$

When $\tau^{\varepsilon,y} < T$, $\|\tilde{X}^{\varepsilon,y}_{\tau^{\varepsilon,y}} - S(\varphi)_{\tau^{\varepsilon,y}}\| \geq d(S(\varphi)_{\tau^{\varepsilon,y}}, \Gamma) - d(\tilde{X}^{\varepsilon,y}_{\tau^{\varepsilon,y}}, \Gamma) \geq \alpha/2 > \eta$, so, in any case,

$$\|\tilde{X}^{\varepsilon,y} - S(\varphi)\|_{[0,T]} \geq \eta \Rightarrow \|\tilde{X}^{\varepsilon,y} - S(\varphi)\|_{[0,\tau^{\varepsilon,y}]} \geq \eta.$$

Consequently, (53) will be proved if we find $\delta > 0$ such that

$$\limsup_{\varepsilon\to 0, y\to x} \varepsilon \ln \mathbf{P}(\|\tilde{X}^{\varepsilon,y} - S(\varphi)\|_{0,\tau^{\varepsilon,y}} \geq \eta,\ \|\sqrt{\varepsilon}W\|_{0,T} \leq \delta) \leq -R.$$

Now, take C such that σ is C-Lipschitz on $\Gamma_{\alpha/2}$, b is C-Lipschitz and $\tilde{b}$ is bounded by C on $\overline{\mathcal{X}}$ (see Proposition 2.2). It follows from (51) that, for $t \leq \tau^{\varepsilon,y}$,

$$\begin{aligned}
&\|\tilde{X}^{\varepsilon,y}_t - S(\varphi)_t\| \\
&\quad \leq \sqrt{\varepsilon}\left\|\int_0^t \sigma(\tilde{X}^{\varepsilon,y}_s)dW_s\right\| + \varepsilon\int_0^t \|\tilde{b}(\tilde{X}^{\varepsilon,y}_s)\|ds + \|x-y\| \\
&\qquad + \int_0^t \|b(\tilde{X}^{\varepsilon,y}_s) - b(S(\varphi)_s)\|ds + \int_0^t \|\sigma(\tilde{X}^{\varepsilon,y}_s) - \sigma(S(\varphi)_s)\| \times \|\dot{\varphi}_s\|ds \\
&\quad \leq \sqrt{\varepsilon}\left\|\int_0^t \sigma(\tilde{X}^{\varepsilon,y}_s)dW_s\right\| + \varepsilon CT + \|x-y\| \\
&\qquad + C\int_0^t (1 + \|\dot{\varphi}_s\|)\|\tilde{X}^{\varepsilon,y}_s - S(\varphi)_s\|ds.
\end{aligned}$$

Remind that $u := \int_0^T \|\dot{\varphi}_s\|^2 ds < +\infty$. Gronwall's Lemma and the Cauchy-Schwartz inequality yield for $t \leq \tau^{\varepsilon,y}$

$$\begin{aligned}
&\|\tilde{X}^{\varepsilon,y}_t - S(\varphi)_t\| \\
&\quad \leq \left[\sqrt{\varepsilon}\left\|\int_0^t \sigma(\tilde{X}^{\varepsilon,y}_s)dW_s\right\| + \varepsilon CT + \|x-y\|\right] \exp\left[C\left(T + \sqrt{uT}\right)\right].
\end{aligned}$$

Therefore, it suffices to find $\delta > 0$ such that

$$\limsup_{\varepsilon\to 0, y\to x} \varepsilon \ln \mathbf{P}\left(\sqrt{\varepsilon}\left\|\int_0^t \sigma(\tilde{X}^{\varepsilon,y}_s)dW_s\right\|_{0,\tau^{\varepsilon,y}} \geq \eta\beta,\ \sqrt{\varepsilon}\|W\|_{0,T} \leq \delta\right) \leq -R,$$

where $\beta = \exp[-C(T + \sqrt{uT})]/2$. This is an direct consequence of Lemma 4.7 with $Y = W$, $A = 1$, $\xi = \sigma$ and $\tau = \tau^{\varepsilon,y}$ (by Proposition 2.2 (ii), σ is uniformly Hölder, so ξ is uniformly continuous).

Proof of Lemma 4.6 (ii) As above, take $\alpha > 0$ such that $\forall t \in [0,T]$, $S(\varphi)_t \in \Gamma_\alpha$. Fix $\eta \leq \alpha/2$. Then, on the event $\{\|\tilde{X}^{\varepsilon,y} - S(\varphi)\|_{0,T} \leq \eta\}$, for any $t \in [0,T]$, $\tilde{X}^{\varepsilon,y}_t \in \Gamma_{\alpha/2}$. Take C such that b is C-Lipschitz, $\tilde{b}$ is bounded by C, and σ is C-Lipschitz on $\Gamma_{\alpha/2}$. It follows from equation (51) that, on the event $\{\|\tilde{X}^{\varepsilon,y} - S(\varphi)\|_{[0,T]} \leq \eta\}$, for any $t \in [0,T]$,

$$\begin{aligned}\sqrt{\varepsilon}\left\|\int_0^t \sigma(\tilde{X}_s^{\varepsilon,y})dW_s\right\| &= \left\|\tilde{X}_t^{\varepsilon,y} - S(\varphi)_t + y - x - \int_0^t [\sigma(\tilde{X}_s^{\varepsilon,y}) - \sigma(S(\varphi)_s)]\dot{\varphi}_s ds \right.\\ &\qquad \left. - \int_0^t [b(\tilde{X}_s^{\varepsilon,y}) - b(S(\varphi)_s)]ds - \varepsilon\int_0^t \tilde{b}(\tilde{X}_s^{\varepsilon,y})ds\right\| \\ &\leq \|\tilde{X}_t^{\varepsilon,y} - S(\varphi)_t\| + \|x-y\| + C\int_0^T (1+\|\dot{\varphi}_s\|)\|\tilde{X}_s^{\varepsilon} - S(\varphi)_s\|ds + \varepsilon CT \\ &\leq \eta + \|x-y\| + C\eta(T+\sqrt{uT}) + \varepsilon CT \leq \eta(2+2CT+C\sqrt{uT})\end{aligned}$$

for $\varepsilon < \eta$ and $\|x-y\| \leq \eta$. Therefore, using the notation $\beta = 2+2CT+C\sqrt{uT}$,

$$\{\|\tilde{X}^{\varepsilon,y} - S(\varphi)\|_{0,T} \leq \eta,\ \|\sqrt{\varepsilon}W\|_{0,T} \geq \delta\} \subset \tag{56}$$
$$\left\{\forall t \in [0,T],\ \tilde{X}_t^{\varepsilon,y} \in \Gamma_{\frac{\alpha}{2}},\ \sqrt{\varepsilon}\left\|\int_0^t \sigma(\tilde{X}_s^{\varepsilon,y})dW_s\right\|_{0,T} \leq \eta\beta,\ \sqrt{\varepsilon}\|W\|_{0,T} \leq \delta\right\}.$$

Define

$$\begin{aligned}\tau^{\varepsilon,y} &= \inf\{t : d(\tilde{X}_t^{\varepsilon,y},\Gamma) \leq \alpha/2\} \wedge T,\\ Y_t^{\varepsilon,y} &= \int_0^t \sigma(\tilde{X}_s^{\varepsilon,y})dW_s,\\ \xi &= \chi\sigma^{-1},\end{aligned}$$

where χ is a Lipschitz function from $\overline{\mathcal{X}}$ to $[0,1]$ such that $\chi(x)=0$ if $d(x,\Gamma) \leq \alpha/4$ and $\chi(x) = 1$ if $d(x,\Gamma) \geq \alpha/2$. So $\xi(x) = \sigma^{-1}(x)$ if $x \in \Gamma_{\alpha/2}$. With these notations, a small computation shows that (56) rewrites

$$\{\|\tilde{X}^{\varepsilon,y} - S(\varphi)\|_{0,T} \leq \eta,\ \|\sqrt{\varepsilon}W\|_{0,T} \geq \delta\} \subset$$
$$\left\{\sqrt{\varepsilon}\|Y^{\varepsilon,y}\|_{0,T} \leq \eta\beta,\ \sqrt{\varepsilon}\left\|\int_0^{t\wedge\tau^{\varepsilon,y}} \xi(\tilde{X}_s^{\varepsilon,y})dY_s^{\varepsilon,y}\right\|_{0,T} \geq \delta\right\}$$

and (54) is now a direct consequence of Lemma 4.7: ξ is Lipschitz and bounded on $\overline{\mathcal{X}}$ (by Proposition 2.2 (iii), $\Gamma_{\alpha/4} \subset \tilde{\Gamma}_c$ for some $c > 0$, so $\sigma(x) \in \mathcal{S}_{\sqrt{c}}$ for $x \in \Gamma_{\alpha/4}$, and it remains to observe that the inverse matrix application is Lipschitz and bounded on $\mathcal{S}_{\sqrt{c}}$), and for any $t \leq \tau$, $\langle Y^{\varepsilon,y}\rangle_t = \int_0^t a(\tilde{X}_s^{\varepsilon,y})ds$ which is bounded, by Proposition 2.2 (i), by a constant A independent of y and ε.

Proof of Lemma 4.6 (iii) In Lemma 4.6 (iii), φ is defined from $\tilde{\varphi}$ by $\varphi_t = \tilde{\varphi}_t$ for $t \leq t_\psi$, and $\varphi_t = \tilde{\varphi}_{t_\psi}$ otherwise (*i.e.* $\dot{\varphi}_t = 0$ for $t > t_\psi$), where $\psi = S(\tilde{\varphi})$. Then $\psi = S(\varphi) = S(\tilde{\varphi})$ since $S(\varphi)$ does not depend on φ_t for $t \geq t_\psi$. By the Cauchy-Schwartz inequality, $\int_0^{t_\psi} \|\dot{\varphi}_s\|ds \leq (2TJ_T(\varphi))^{1/2} < +\infty$, so there exists $\rho > 0$ small enough such that

$$\int_{t_\psi-\rho}^{t_\psi} \|\dot{\varphi}_s\|ds \leq \frac{\eta e^{-CT}}{8C}, \tag{57}$$

where C is a constant bounding b, $\tilde{b}$ and σ, and such that b is C-Lipschitz.

Distinguishing when $\|\tilde{X}^{\varepsilon,y} - \psi\|_{0,t_\psi-\rho} \geq \eta e^{-CT}/4$ or not, we can write

$$\{\|\tilde{X}^{\varepsilon,y} - \psi\|_{0,T} \geq \eta,\ \|\sqrt{\varepsilon}W\|_{0,T} \leq \delta\} \subset D^{\varepsilon,y} \cup E^{\varepsilon,y},$$

where

$$D^{\varepsilon,y} = \left\{\|\tilde{X}^{\varepsilon,y} - \psi\|_{0,t_\psi-\rho} \leq \frac{\eta e^{-CT}}{4},\ \|\tilde{X}^{\varepsilon,y} - \psi\|_{t_\psi-\rho,T} \geq \eta, \|\sqrt{\varepsilon}W\|_{0,T} \leq \delta\right\}$$

$$\text{and}\quad E^{\varepsilon,y} = \left\{\|\tilde{X}^{\varepsilon,y} - \psi\|_{0,t_\psi-\rho} \geq \frac{\eta e^{-CT}}{4},\ \|\sqrt{\varepsilon}W\|_{0,t_\psi-\rho} \leq \delta\right\}.$$

Part (i) of Lemma 4.6 shows that $\mathbf{P}(E^{\varepsilon,y})$ has the required exponential decay if δ is small enough. Let us estimate the probability of $D^{\varepsilon,y}$.

It follows from (51) and from the fact that $\dot{\varphi}_t = 0$ for $t > t_\psi$ that, for any $t \geq t_\psi - \rho$

$$\|\tilde{X}^{\varepsilon,y}_t - \psi_t\| \leq \|\tilde{X}^{\varepsilon,y}_{t_\psi-\rho} - \psi_{t_\psi-\rho}\| + \sqrt{\varepsilon}\left\|\int_{t_\psi-\rho}^t \sigma(\tilde{X}^{\varepsilon,y}_s)dW_s\right\|$$
$$+C\int_{t_\psi-\rho}^t \|\tilde{X}^{\varepsilon,y}_s - \psi_s\|ds + \varepsilon CT + \int_{t_\psi-\rho}^{t_\psi\wedge t} \|\sigma(\tilde{X}^{\varepsilon,y}_s) - \sigma(\psi_s)\| \times \|\dot{\varphi}_s\|ds.$$

On the event $D^{\varepsilon,y}$, the first term of the right-hand side is smaller than $\eta e^{-CT}/4$, and, since σ is bounded by C, the last term is smaller than $2C\int_{t_\psi-\rho}^{t_\psi}\|\dot{\varphi}\|ds$, which is smaller than $\eta e^{-CT}/4$ by (57). Moreover, we can suppose ε small enough to have $\varepsilon CT \leq \eta e^{-CT}/4$. So, on the event $D^{\varepsilon,y}$, by Gronwall's Lemma, for $t \geq t_\psi - \rho$,

$$\|\tilde{X}^{\varepsilon,y}_t - \psi_t\| \leq \left[\frac{3}{4}\eta e^{-CT} + \sqrt{\varepsilon}\left\|\int_{t_\psi-\rho}^t \sigma(\tilde{X}^{\varepsilon,y}_s)dW_s\right\|\right] e^{CT}.$$

Since $\|\tilde{X}^{\varepsilon,y} - \psi\|_{t_\psi-\rho,T} \geq \eta$ on $D^{\varepsilon,y}$, we finally can write

$$D^{\varepsilon,y} \subset$$
$$\left\{\left\|\sqrt{\varepsilon}\int_{t_\psi-\rho}^{\cdot} \sigma(\tilde{X}^{\varepsilon,y}_s)dW_s\right\|_{t_\psi-\rho,T} \geq \frac{\eta e^{-CT}}{4},\ \|\sqrt{\varepsilon}(W_\cdot - W_{t_\psi-\rho})\|_{t_\psi-\rho,T} \leq 2\delta\right\}$$

and (53) is now a consequence of Lemma 4.7. □

Proof of Lemma 4.3 Take $\psi \in \tilde{\mathcal{C}}^{ac}_x([0,T],\overline{\mathcal{X}})$, and assume that there exists $\varphi \in \mathcal{C}^{ac}_0([0,T],\mathbb{R}^d)$ such that $S(\varphi) = \psi$. Then for any $t \in [0,T]$

$$\dot{\psi}_t = b(\psi_t) + \sigma(\psi_t)\dot{\varphi}_t.$$

For $t < t_\psi$, $\psi_t \notin \Gamma$, so $\sigma(\psi_t)$ is invertible, and

$$\dot{\varphi}_t = \sigma^{-1}(\psi_t)[\dot{\psi}_t - b(\psi_t)]. \tag{58}$$

This defines uniquely the function φ on $[0,t_\psi)$, and φ_{t_ψ} exists if and only if $\sigma^{-1}(\psi_t)[\dot{\psi}_t - b(\psi_t)]$ is $\mathbb{L}^1$ on $[0,t_\psi]$. Assume this is true. Then, since for $t \geq t_\psi$,

$\psi_t \in \Gamma$, and so $b(\psi_t) = \sigma(\psi_t) = 0$, $S(\varphi)_t = \psi_t = \psi_{t_\psi}$ for $t > t_\psi$ whatever is the function φ_t on $(t_\psi, T]$. Consequently,

$$\begin{aligned} J_T(\varphi) &= \frac{1}{2}\int_0^T \|\dot{\varphi}_t\|^2 dt \\ &\geq \frac{1}{2}\int_0^{t_\psi} \|\dot{\varphi}_t\|^2 dt = \frac{1}{2}\int_0^{t_\psi} \|\sigma^{-1}(\psi_t)[\dot{\psi}_t - b(\psi_t)]\|^2 dt = I_{T,x}(\psi), \end{aligned}$$

and, when $I_{T,x}(\psi) < +\infty$, a solution φ to $S(\varphi) = \psi$ satisfies $J_T(\varphi) = I_{T,x}(\psi)$ if and only if φ is constant for $t \geq t_\psi$.

Conversely, when $\sigma^{-1}(\psi_t)[\dot{\psi}_t - b(\psi_t)]$ is not $\mathbb{L}^1$ on $[0, t_\psi]$ or ψ does not belong to $\tilde{\mathcal{C}}^{ac}_x([0,T], \overline{\mathcal{X}})$, there is no solution to $S(\varphi) = \psi$. Moreover, in this case $I_{T,X}(\psi) = +\infty$, so the proof of (i) is completed.

Since by Proposition 2.2 (iii) σ is uniformly non-degenerate on Γ_α for any $\alpha > 0$, it follows trivially from (58) that $\mathcal{C}^1([0,T], \mathcal{X} \setminus \Gamma) \subset S(\{J_T < \infty\})$. (ii) is now clear. □

Proof of Lemma 4.4 It follows from (1) that, for any $y \in \overline{\mathcal{X}}$, $s > 0$ and $t \in [0,T]$,

$$\|X^{\varepsilon,y}_{t+s} - X^{\varepsilon,y}_t\| \leq Cs + \sqrt{\varepsilon}\left\|\int_t^{t+s} \sigma(X^{\varepsilon,y}_u) dW_u\right\|$$

where C is a bound for b^ε (for $\varepsilon < 1$). So, for a given $h > 0$, we can apply Lemma 4.5 with $R \geq Ch$ to obtain

$$\mathbf{P}\left(\sup_{0\leq s\leq h} \|X^{\varepsilon,y}_{t+s} - X^{\varepsilon,y}_t\| \geq R\right) \leq 2d\exp\left(-\frac{(R-Ch)^2}{2dh\varepsilon\Sigma^2}\right),$$

where Σ is a bound for σ. Writing this for $t = ih$ for $0 \leq i < T/h$, we easily deduce that, for $R \geq Ch$,

$$\mathbf{P}\left(\omega(X^\varepsilon, h) \geq 2R\right) \leq 2d\left(\frac{T}{h} + 1\right)\exp\left(-\frac{(R-Ch)^2}{2d\varepsilon\Sigma^2 h}\right), \tag{59}$$

where $\omega(\psi, h)$ has been defined in the statement of Lemma 4.4.

For any $l \geq 1$, set $R_l = 1/2l$ and $h_l = 1/l^3$. Then, for sufficiently large l, $R_l \geq Ch_l$ and

$$\frac{(R_l - Ch_l)^2}{2d\varepsilon\Sigma^2 h_l} = \frac{(\sqrt{l} - 2C/l^{3/2})^2}{8d\varepsilon\Sigma^2} \geq \frac{(\sqrt{l}/2)^2}{8d\varepsilon\Sigma^2} = \frac{l}{32d\varepsilon\Sigma^2}. \tag{60}$$

Now, observe that the set K^y_k, defined in (45) in the statement of Theorem 4.2 satisfies

$$K^y_k = \{\psi \in \mathcal{C}_y([0,T], \mathcal{X}) : \forall l \geq k, \omega(\psi, h_l) \leq 2R_l\}.$$

This is a compact set by Ascoli's Theorem, and a simple computation, using (59) and (60) shows that $\varepsilon \ln \mathbf{P}(X^{\varepsilon,y} \notin K^y_k) \leq -k/64d\Sigma^2$ for sufficiently large k. □

5 Application to the problem of exit from a domain

When evolution in a population comes to an equilibrium, two interesting long-term biological phenomena may happen. The first one is the *evolutionary branching*, the appearance of two (or more) distinct subpopulations evolving in different directions of the trait space, which may eventually lead to speciation, and the second one is the *punctualism*, the quick evolution of the whole population from an evolutionary equilibrium to another, due to a large mutation or to successive invasions of deleterious mutants in the population. Many paleontological data, and experiments on artificial evolution, seem to reflect this kind of rapid evolution (see Rand and Wilson [16]). In this section, we propose to study the phenomenon of punctualism by answering the following questions: in which direction a rapid evolution is more likely to occur, and how long does it take to happen?

In adaptive dynamics models, evolutionary equilibria (also called *evolutionary singularities*) have to be understood as equilibria of the unperturbed dynamic (corresponding to the case $\varepsilon = 0$) $\dot{\phi} = b(\phi)$, called *canonical equation of adptive dynamics* (see [2] and [5]). These equilibria are the points of Γ. Because of Corollary 4.2, when ε is small, $X^{\varepsilon,x}$ is close to the solution of the canonical equation with initial state x with high probability. Yet, the diffusion part of $X^{\varepsilon,x}$ may imply that $X^{\varepsilon,x}$ almost surely leaves in finite time any bounded domain G containing an evolutionary singularity. The result of this section gives estimates for this time, and precises where the exit occurs in ∂G.

We will assume $d \geq 2$. Otherwise, the problem has little interest: the process $X^{\varepsilon,y}$ can exit from an open domain $G = (c, c')$ of $\mathcal{X}$ containing a unique point x of Γ, with $y > x$, only from the right side c' of G, and the probability of reaching x before c' can be computed explicitly using Proposition 5.5.22 of Karatzas and Shreve [10].

In this section, it will be sometimes more convenient to use Markov processes' notations: in some cases, we will write X^ε instead of $X^{\varepsilon,x}$ and the initial state of X^ε will be specified as indices in the probability and expectation symbols: $\mathbf{P}_x$ and $\mathbf{E}_x$.

Let us follow the treatment of section 5.7 of Dembo and Zeitouni [3]. Fix an open bounded domain $G \subset \mathcal{X}$, and suppose that the boundary of G is smooth enough for

$$\tau^\varepsilon = \inf\{t > 0 : X^\varepsilon_t \in \partial G\}$$

to be a well-defined stopping time. Assume also that

Hypotheses 5.1

(Ha) *The unique stable equilibrium point in G of the d-dimensional ordinary differential equation*

$$\dot{\phi} = b(\phi) \tag{61}$$

is at $0 \in G$, and

$$\phi(0) \in G \Rightarrow \forall t > 0, \phi(t) \in G \text{ and } \lim_{t\to\infty} \phi(t) = 0.$$

(Hb) *For any $\varepsilon > 0$ and $y \in G \setminus \{0\}$,* $\mathbf{P}_y \left(\lim_{t\to\infty} X^\varepsilon_t = 0 \right) = 0.$

(Ha) states that the domain G is an *attracting* domain, and we have given in Theorem 3.3 (section 3.4) conditions under which (Hb) holds. Remind that, by Proposition 3.1 (b), assumption (Hb) implies that X^ε is strong Markov as long as it stays inside G.

Define

$$V(y,z,t)=\inf_{\{\psi\in\mathcal{C}([0,t],\overline{\mathcal{X}}):\psi(0)=y,\psi(t)=z\}} I_{t,y}(\psi),$$

which is, heuristically, the cost of forcing $X^{\varepsilon,y}$ to be at z at time t. Define also

$$V(y,z)=\inf_{t>0} V(y,z,t).$$

Observe that the *quasi-potential* of Freidlin and Wentzell [8] $V(0,z)$ has no interest in our setting, since for all $t>0$, $V(0,z,t)=\infty$ if $z\neq 0$ ($X^{\varepsilon,0}$ is constant, equal to 0). Instead, let us define

$$\bar{V}(0,z):=\liminf_{\rho\to 0}\inf_{y\in S(\rho)} V(y,z),$$

where $S(\rho)=\{y\in\mathbb{R}^d:\|y\|=\rho\}$ is the sphere of $\mathbb{R}^d$ centered at 0 and with radius δ. Let also $B(\rho)$ be the closed ball in $\mathbb{R}^d$ centered at 0 and with radius ρ. Note that the lim inf above is in fact an increasing limit, since, if $\rho<\rho'$ and if $z\notin B(\rho')$, one can obtain from a path from $S(\rho)$ to z a new path from $S(\rho')$ to z by "cutting" the beginning of the former path.

The treatment to follow is guided by the heuristics that, as $\varepsilon\to 0$, X^ε wanders around 0 for an exponentially long time, during which its chance of hitting a closed set $N\subset\partial G$ is determined by $\inf_{z\in N}\bar{V}(0,z)$. Any excursion off the stable point 0 has an overwhelmingly high probability of being pulled back near 0, and it is not the time spent near any part of ∂G that matters but the *a priori* chance for a direct, fast exit due to a rare segment in the Brownian motion's path.

Three other assumptions are required:

Hypotheses 5.2

(Hc) *All the trajectories of the deterministic system (61) with initial value in* ∂G *converge to* 0 *as* $t\to\infty$.

(Hd) $\bar{V}:=\inf_{z\in\partial G}\bar{V}(0,z)<\infty$.

(He) g *is* $\mathcal{C}^2$ *at* 0 *and the differential* $D:=H_{1,1}g(0,0)+H_{1,2}g(0,0)$ *of the function* $x\mapsto\nabla_1 g(x,x)$ *at* 0 *has a null kernel.*

Assumption (Hc) prevents consideration of situations in which ∂G is the characteristic boundary of the domain of attraction of 0. Assumption (Hd) is natural, since otherwise all points on ∂G are equally unlikely on the large deviation scale. We have already encountered an assumption similar to (He) in Theorem 3.3 and Proposition 3.2 (section 3.4). It allows to bound below the eigenvalues of $a(x)$ for x near 0.

Now, let us state

Theorem 5.1

(a) *Assume (H1), (H2'), (H3), (H4), (Ha), (Hb), (Hd) and (He). Then, for all* $x\in G\setminus\{0\}$ *and* $\delta>0$,

$$\lim_{\varepsilon\to 0}\mathbf{P}_x(\tau^\varepsilon>e^{(\bar{V}-\delta)/\varepsilon})=1. \qquad (62)$$

(b) *Assume (H1), (H2'), (H3), (H4) and (Ha)–(He). If N is a closed subset of ∂G and if $\inf_{z\in N}\bar V(0,z)>\bar V$, then for any $x\in G\setminus\{0\}$,*

$$\lim_{\varepsilon\to 0}\mathbf{P}_x(X^\varepsilon_{\tau^\varepsilon}\in N)=0. \tag{63}$$

In particular, if there exists $z^*\in\partial G$ such that $\bar V(0,z^*)<\bar V(0,z)$ for all $z\in\partial G\setminus\{z^*\}$, then

$$\forall\delta>0,\forall x\in G\setminus\{0\},\quad \lim_{\varepsilon\to 0}\mathbf{P}_x(\|X^\varepsilon_{\tau^\varepsilon}-z^*\|<\delta)=1. \tag{64}$$

Note that assumption (Hc) is not necessary to prove (a).

Remark 5.1 *Part (a) of this kind of results usually includes an upper bound for τ^ε. We are not able to achieve this because of the singularity at 0 of the process X^ε: we are only able to obtain an uniform exponential lower bound on $\mathbf{P}_x(\tau^\varepsilon\leq T)$ for $x\in G\setminus B(\rho)$, and not for any $x\in G\setminus\{0\}$. This uniform lower estimate is crucial in the classical proofs of the fact that $\mathbf{P}_x(\tau^\varepsilon<e^{(\bar V+\delta)/\varepsilon})\to 1$.*

Remark 5.2 *Some additional work might allow to generalize part (b) of Theorem 5.1 to more general domains including several points of Γ, or including ω-limit sets (such as cycles) of the differential equation (61). We refer the interested reader to chapter 6 of the book of Freidlin and Wentzell [8].*

The proof of a similar result in Dembo and Zeitouni [3] (Theorem 5.7.11 and Corollary 5.7.16) is based on the strong Markov property for X^ε, which holds in our case for quite general stopping times thanks to assumption (Hb) and Proposition 3.1 (b), and on several lemmas, which have to be adapted to our degenerate case. Some of them will be very close to the lemmas of [3], and some of them will require a different treatment. In particular, part (a) of Theorem 5.1 will be obtained in a very similar way than in [3], whereas part (b) has to be obtained without using any upper bound on τ^ε.

The first lemma is a continuity result for V.

Lemma 5.1 *Assume (H1), (H2'), (H3), (H4) and (He). For any $\delta>0$, there exists $\rho>0$ small enough such that*

$$\sup_{x,y\in B(\rho)\setminus\{0\}}\ \inf_{t\in[0,1]}V(x,y,t)<\delta \tag{65}$$

and

$$\sup_{\{x,y:\inf_{z\in\partial G}(\|y-z\|+\|x-z\|)\leq\rho\}}\ \inf_{t\in[0,1]}V(x,y,t)<\delta. \tag{66}$$

The second lemma states that the diffusion wanders in G for an arbitrary long time without hitting a small neighborhood of 0 with an exponentially negligible probability. Assumption (Hc) is necessary to prove this lemma.

Lemma 5.2 *Assume (H1), (H2'), (H3), (H4), (Ha)–(He). Let*

$$\sigma_\rho:=\inf\{t\geq 0: X^\varepsilon\in B(\rho)\cup\partial G\},$$

for ρ small enough to have $B(\rho)\subset G$ (mind that σ_ρ depends on ε; we do not mention it to keep notations simple). Then

$$\lim_{t\to\infty}\limsup_{\varepsilon\to 0}\varepsilon\ln\sup_{x\in G}\mathbf{P}_x(\sigma_\rho>t)=-\infty.$$

The third lemma gives a uniform lower bound on the probability of an exit from G starting from a small sphere around 0 before hitting an ever smaller sphere.

Lemma 5.3 *Assume (H1), (H2'), (H3), (H4), (Ha), (Hb), (Hd) and (He). Then*

$$\lim_{\rho\to 0}\liminf_{\varepsilon\to 0}\varepsilon\ln\inf_{y\in S(2\rho)}\mathbf{P}_y(X^\varepsilon_{\sigma_\rho}\in\partial G)\geq -\bar{V}=-\inf_{z\in\partial G}\bar{V}(0,z).$$

The following upper bound relates our quasi-potential $\bar{V}(0,\cdot)$ with the probability that an excursion starting from a small sphere around 0 hits a given subset of ∂G before hitting an even smaller sphere.

Lemma 5.4 *Assume (H1), (H2'), (H3), (H4), (Ha)–(He). For any closed set $N\subset\partial G$,*

$$\lim_{\rho\to 0}\limsup_{\varepsilon\to 0}\varepsilon\ln\sup_{y\in S(2\rho)}\mathbf{P}_y(X^\varepsilon_{\sigma_\rho}\in N)\leq-\inf_{z\in N}\bar{V}(0,z)$$

The following lemma is used to extend the upper bound to hold for every initial point $x\in G$.

Lemma 5.5 *Assume (H1), (H2'), (H3), (H4) and (Ha). For every $\rho>0$ such that $B(\rho)\subset G$ and all $x\in G$,*

$$\lim_{\varepsilon\to 0}\mathbf{P}_x(X^\varepsilon_{\sigma_\rho}\in B(\rho))=1.$$

Finally, we need a uniform estimate stating that over short time intervals, the process X^ε with initial state x has an exponentially negligible probability of getting too far from x.

Lemma 5.6 *Assume (H1), (H2'), (H3), (H4), (Ha), (Hb), (Hd) and (He). For every $\rho>0$ and every $c>0$, there exists a constant $T(c,\rho)<\infty$ such that*

$$\limsup_{\varepsilon\to 0}\varepsilon\ln\sup_{x\in G}\mathbf{P}_x\left(\|X^\varepsilon-x\|_{0,T(c,\rho)}\geq\rho\right)<-c.$$

Let us briefly comment these lemmas. All of them except Lemmas 5.3 are similar to those of [3], but their proofs all require some modifications. Lemma 5.1 will be proved in a different way than in [3], and the new Lemma 5.3 will allow to prove Theorem 5.1 (b), but is not sufficient to prove the usual upper bound on τ^ε.

We will first prove Theorem 5.1, and postpone the proof of all the preceding lemmas to the end of the section.

Proof of Theorem 5.1 (a) We will first prove (a) under the additional assumption (Hc). Let $\rho>0$ be small enough such that $B(2\rho)\subset G$ (ρ will be specified later). Let $\theta_0=0$ and for $m=0,1,\dots$ define the stopping times

$$\begin{aligned}\tau_m&=\inf\{t\geq\theta_m: X^\varepsilon_t\in B(\rho)\cup\partial G\},\\ \theta_{m+1}&=\inf\{t>\tau_m: X^\varepsilon_t\in S(2\rho)\},\end{aligned}$$

with the convention that $\theta_{m+1} = \infty$ if $X^\varepsilon_{\tau_m} \in \partial G$. Each interval $[\tau_m, \tau_{m+1}]$ represents one significant excursion off $B(\rho)$. Note that, necessarily, $\tau^\varepsilon = \tau_m$ for some integer m.

Moreover, assumption (Hb) implies that $\theta_{m+1} < \infty$ as soon as $X^\varepsilon_{\tau_m} \in B(\rho)$. This can be proved as follows. Since for any $\alpha > 0$, X^ε is a diffusion with bounded drift part and uniformly non-degenerate diffusion part in $B(2\rho) \cap \Gamma_{\alpha/2}$, X^ε has an uniformly positive probability to reach $S(2\rho)$ before $S(\alpha/2)$ starting from any point of $S(\alpha)$. Hence, by the strong Markov property of Proposition 3.1 (b), for all $x \in S(\rho)$, $\mathbf{P}_x(\theta_1 < +\infty | \limsup_{t\to+\infty} \|X^\varepsilon_t\| \geq \alpha) = 1$. Since assumption (Hb) implies that for all $x \in S(\rho)$, $\mathbf{P}_x(\limsup_{t\to+\infty} \|X^\varepsilon_t\| \geq \alpha) \to 1$ when $\alpha \to 0$, we actually have, for all $x \in S(\rho)$, $\mathbf{P}_x(\theta_1 < +\infty) = 1$. By the strong Markov property, this provides almost surely the implication $X^\varepsilon_{\tau_m} \in B(\rho) \Rightarrow \theta_{m+1} < \infty$.

For $\bar{V} = 0$, the lower bound on τ^ε is an easy consequence of Lemmas 5.5 and 5.6. Hence, assume hereafter that $\bar{V} > 0$, and fix $\delta > 0$ arbitrarily small. Note that ∂G is a closed set and choose $\rho > 0$ small enough as needed by Lemma 5.4 for

$$\limsup_{\varepsilon\to 0} \varepsilon \ln \sup_{y \in S(2\rho)} \mathbf{P}_y(X^\varepsilon_{\sigma_\rho} \in \partial G) < -\bar{V} + \frac{\delta}{2}$$

to hold. Now, let $c = \bar{V}$ and let $T_0 = T(c, \rho)$ be as determined by Lemma 5.6. Then, there exists $\varepsilon_0 > 0$ such that for all $\varepsilon \leq \varepsilon_0$ and all $m \geq 1$,

$$\sup_{x \in G\setminus\{0\}} \mathbf{P}_x(\tau^\varepsilon = \tau_m) \leq \sup_{y \in S(2\rho)} \mathbf{P}_y(X^\varepsilon_{\sigma_\rho} \in \partial G) \leq e^{-(\bar{V}-\delta/2)/\varepsilon}$$

and

$$\sup_{x \in G\setminus\{0\}} \mathbf{P}_x(\theta_m - \tau_{m-1} \leq T_0) \leq \sup_{x \in G} \mathbf{P}_x(\|X^\varepsilon - x\|_{0,T_0} \geq \rho) \leq e^{-(\bar{V}-\delta/2)/\varepsilon}.$$

The event $\{\tau^\varepsilon \leq kT_0\}$ implies that either one of the mutually exclusive events $\{\tau^\varepsilon = \tau_m\}$ for $0 \leq m \leq k$ occurs, or else that at least one of the first k excursions $[\tau_m, \tau_{m+1}]$ off $B(\rho)$ is of length at most T_0. Thus, utilizing the preceding estimates, for all $x \in G \setminus \{0\}$ and any integer k,

$$\begin{aligned}\mathbf{P}_x(\tau^\varepsilon \leq kT_0) &\leq \sum_{m=0}^{k} \mathbf{P}_x(\tau^\varepsilon = \tau_m) + \sum_{m=1}^{k} \mathbf{P}_x(\theta_m - \tau_{m-1} \leq T_0) \\ &\leq \mathbf{P}_x(\tau^\varepsilon = \tau_0) + 2ke^{-(\bar{V}-\delta/2)/\varepsilon}.\end{aligned}$$

Recall the identity $\{\tau^\varepsilon = \tau_0\} \equiv \{X^\varepsilon_{\sigma_\rho} \notin B(\rho)\}$ and apply the preceding inequality with $k = [T_0^{-1} e^{(\bar{V}-\delta)/\varepsilon}] + 1$ to obtain, for small enough ε,

$$\mathbf{P}_x(\tau^\varepsilon \leq e^{(\bar{V}-\delta)/\varepsilon}) \leq \mathbf{P}_x(\tau^\varepsilon \leq kT_0) \leq \mathbf{P}_x(X^\varepsilon_{\sigma_\rho} \notin B(\rho)) + 4T_0^{-1} e^{-\delta/2\varepsilon}.$$

By Lemma 5.5, the left side of this inequality converges to 0 as $\varepsilon \to 0$; hence, the proof of (62) is complete.

It remains to study the case where assumption (Hc) is removed. In this case, let $G^{-\rho} := \{x \in G : d(x, \partial G) > \rho\}$. Observe that $G^{-\rho}$ are open sets for which assumption (Hc) holds. Therefore, (a) is true for these sets. Moreover, the stopping times τ^ε related to $G^{-\rho}$ are increasing when ρ decreases to 0. The announced lower bound on τ^ε results easily from this fact and from the continuity of the quasi-potential $\bar{V}$ at any point of ∂G, implied by (66). □

Proof of Theorem 5.1 (b) Fix a closed set $N \subset G$ such that $\bar{V}_N := \inf_{z\in N} \bar{V}(0,z) > \bar{V}$ (if $\bar{V}_N = \infty$, then simply use throughout the proof an arbitrary large constant as $\bar{V}_N$). Fix $\eta > 0$ such that $\eta < (\bar{V}_N - \bar{V})/3$, and set $\rho > 0$ and $\varepsilon_0 > 0$ as needed in Lemmas 5.3 and 5.4 for

$$\inf_{y\in S(2\rho)} \mathbf{P}_y(X^\varepsilon_{\sigma_\rho} \in \partial G) \geq e^{-(\bar{V}+\eta)/\varepsilon}, \quad \forall \varepsilon \leq \varepsilon_0 \tag{67}$$

and

$$\sup_{y\in S(2\rho)} \mathbf{P}_y(X^\varepsilon_{\sigma_\rho} \in N) \leq e^{-(\bar{V}_N-\eta)/\varepsilon}, \quad \forall \varepsilon \leq \varepsilon_0$$

to hold. Fix $y \in B(\rho)$ and remind the definition of the stopping times τ_m and θ_m in the proof of Theorem 5.1 (a). Observe that

$$\mathbf{P}_y(X^\varepsilon_{\tau^\varepsilon} \in N) \leq \mathbf{P}_y(\tau^\varepsilon > \tau_l) + \sum_{m=1}^{l} \mathbf{P}_y(\tau^\varepsilon = \tau_m \text{ and } X^\varepsilon_{\tau_m} \in N). \tag{68}$$

Firstly, let us bound the second term: for $m \geq 1$, $y \in B(\rho)$ and $\varepsilon \leq \varepsilon_0$, thanks to the strong Markov property of Proposition 3.1 (b),

$$\begin{aligned}\mathbf{P}_y(\tau^\varepsilon = \tau_m \text{ and } X^\varepsilon_{\tau_m} \in N) &= \mathbf{P}_y(\tau^\varepsilon > \tau_{m-1})\mathbf{P}_y(X^\varepsilon_{\tau_m} \in N | \tau^\varepsilon > \tau_{m-1}) \\ &= \mathbf{P}_y(\tau^\varepsilon > \tau_{m-1})\mathbf{E}_y[\mathbf{P}_{X^\varepsilon_{\theta_m}}(X^\varepsilon_{\sigma_\rho} \in N)|\tau^\varepsilon > \tau_{m-1}] \\ &\leq \sup_{x\in S(2\rho)} \mathbf{P}_x(X^\varepsilon_{\sigma_\rho} \in N) \leq e^{-(\bar{V}_N-\eta)/\varepsilon}.\end{aligned}$$

Secondly, let us bound the first term of the right-hand side of (68): for $l \geq 1$ and $y \in B(\rho)$,

$$\mathbf{P}_y(\tau^\varepsilon > \tau_l) = \mathbf{E}_y[\mathbf{P}_{X^\varepsilon_{\theta_1}}(\tau^\varepsilon > \tau_{l-1})] \leq \sup_{x\in S(2\rho)} \mathbf{P}_x(\tau^\varepsilon > \tau_{l-1}). \tag{69}$$

Now, for $x \in S(2\rho)$ and $k \geq 1$,

$$\begin{aligned}\mathbf{P}_x(\tau^\varepsilon > \tau_k) &= [1 - \mathbf{P}_x(\tau^\varepsilon = \tau_k | \tau^\varepsilon > \tau_{k-1})]\mathbf{P}_x(\tau^\varepsilon > \tau_{k-1}) \\ &= [1 - \mathbf{E}_x[\mathbf{P}_{X^\varepsilon_{\theta_k}}(X^\varepsilon_{\sigma_\rho} \in \partial G)|\tau^\varepsilon > \tau_{k-1}]]\mathbf{P}_x(\tau^\varepsilon > \tau_{k-1}) \\ &\leq (1-q)\mathbf{P}_x(\tau^\varepsilon > \tau_{k-1}),\end{aligned}$$

where $q := \inf_{y\in S(2\rho)} \mathbf{P}_y(X^\varepsilon_{\sigma_\rho} \in \partial G) \geq e^{-(\bar{V}+\eta)/\varepsilon}$ by (67). Iterating over $k = 1, 2, \ldots$ gives for $k \geq 0$

$$\sup_{y\in S(2\rho)} \mathbf{P}_x(\tau^\varepsilon > \tau_k) \leq (1-q)^k.$$

In (69), this yields for all $l \geq 1$ and $y \in B(\rho)$

$$\mathbf{P}_y(\tau^\varepsilon > \tau_l) \leq (1-q)^{l-1}.$$

Putting together these estimates in (68) gives finally for all $y \in B(\rho)$ and $\varepsilon \leq \varepsilon_0$

$$\mathbf{P}_y(X^\varepsilon_{\tau^\varepsilon} \in N) \leq \left(1 - e^{-\frac{\bar{V}+\eta}{\varepsilon}}\right)^{l-1} + l e^{-\frac{\bar{V}_N-\eta}{\varepsilon}}.$$

Choosing $l = [2e^{(\bar{V}+2\eta)/\varepsilon}]$, for ε small enough, $l - 1 > e^{(\bar{V}+2\eta)/\varepsilon}$, and hence

$$\mathbf{P}_y(X^\varepsilon_{\tau^\varepsilon} \in N) \leq \left[\left(1 - \frac{1}{u_\varepsilon}\right)^{u_\varepsilon}\right]^{e^{\eta/\varepsilon}} + 2e^{\frac{\bar{V}-\bar{V}_N+3\eta}{\varepsilon}},$$

where $u_\varepsilon := e^{(\bar{V}+\eta)/\varepsilon} \to +\infty$. So $(1 - 1/u_\varepsilon)^{u_\varepsilon} \to 1/e$, and, finally, $\mathbf{P}_y(X^\varepsilon_{\tau^\varepsilon} \in N) \to 0$ for $y \in B(\rho)$ (recall that $0 < \eta < (\bar{V}_N - \bar{V})/3$). The proof of (63) is now completed by combining Lemma 5.5 and the inequality

$$\mathbf{P}_x(X^\varepsilon_{\tau^\varepsilon} \in N) \leq \mathbf{P}_x(X^\varepsilon_{\sigma_\rho} \notin B(\rho)) + \sup_{y \in B(\rho)} \mathbf{P}_y(X^\varepsilon_{\tau^\varepsilon} \in N).$$

Applying (63) to the set $N = \{z \in \partial G : \|z - z^*\| \geq \delta\}$ and observing that Lemma 5.1 (66) implies the continuity of $z \mapsto \bar{V}(0, z)$ on ∂G, we easily obtain (64).
□

Proof of Lemma 5.1 (65) Fix $\delta > 0$, let ρ be small enough for $B(\rho) \subset G$ to hold, and fix x and y in $B(\rho) \setminus \{0\}$. In order to be clear, we will use the complex notation for the coordinates of points of the plane of $\mathbb{R}^d$ containing 0, x and y, and we will assume that $x = r \in \mathbb{R}$ and $y = r'e^{i\theta}$, with $0 < r, r' \leq \rho$. Define $\psi \in \mathcal{C}([0,1], B(\rho))$ by

$$\psi(t) = \begin{cases} (1-(3t)^2)r + (3t)^2\rho & \text{if } \ 0 \leq t \leq 1/3 \\ \rho e^{i\theta(3t-1)} & \text{if } \ 1/3 \leq t \leq 2/3 \\ (1-(3-3t)^2)r'e^{i\theta} + (3-3t)^2\rho e^{i\theta} & \text{if } \ 2/3 \leq t \leq 1. \end{cases}$$

Then $\psi(0) = x$ and $\psi(1) = y$, and $\psi(t) \in B(\rho) \setminus \{0\}$ for any $t \in [0,1]$. Moreover, for $0 \leq t \leq 1/3$, $\psi(t) = r + 9t^2(\rho - r)$, so that $\|\psi(t)\| \geq 9t^2(\rho - r)$, and, similarly, for $2/3 \leq t \leq 1$, $\|\psi(t)\| \geq 9(1-t)^2(\rho - r')$. A calculation similar to equation (41) in the proof of Proposition 4.1 gives, with the same K, $\mathcal{N}_0$ and a_0 as therein, if $B(\rho) \subset \mathcal{N}_0$,

$$\begin{aligned} I_{1,x}(\psi) &\leq \frac{1}{2a_0}\left(\int_0^{1/3} \frac{2(18t(\rho-r))^2 + 2K^2\|\psi(t)\|^2}{\|\psi(t)\|} dt \right. \\ &\quad + \int_{1/3}^{2/3} \frac{2(3\theta\rho)^2 + 2K^2\|\psi(t)\|^2}{\|\psi(t)\|} \\ &\quad \left. + \int_{2/3}^{1} \frac{2(18(1-t)(\rho-r'))^2 + 2K^2\|\psi(t)\|^2}{\|\psi(t)\|} dt \right) \\ &\leq \frac{1}{2a_0}\left(\int_0^{1/3} (648(\rho - r) + 2K^2\|\psi(t)\|)dt + \int_{1/3}^{2/3} (18\theta^2 + 2K^2)\rho dt \right. \\ &\quad \left. + \int_{2/3}^{1} (648(\rho - r') + 2K^2\|\psi(t)\|)dt \right) \\ &\leq \frac{2(216 + 2K^2/3)\rho + (6\theta^2 + 2K^2/3)\rho}{2a_0}. \end{aligned}$$

Consequently, for sufficiently small $\rho > 0$ not depending on x and y in $B(\rho) \setminus \{0\}$, $I_{1,x}(\psi) \leq \delta/2$, which yields (65). □

Proof of Lemma 5.1 (66) Fix $\delta > 0$, let ρ be small enough for $\{x : d(x, \partial G) \leq 3\rho\} \subset \mathcal{X} \setminus \Gamma$ to hold, and fix x and y such that there exists $z \in \partial G$ with $\|x - z\| + \|y - z\| \leq \rho$. Let us denote by Γ^r the r-enlargement of Γ: $\Gamma^r = \{x : d(x, \Gamma) \leq r\}$. Then x and y belong to $\mathcal{X} \setminus \Gamma^{2\rho}$, and $\|x - y\| \leq \rho$, so the segment $[x, y]$ is included in $\mathcal{X} \setminus \Gamma^\rho$.

For any $t_0 > 0$, define $\psi_{t_0} \in \mathcal{C}([0, t_0], G^{2\rho} \setminus \Gamma^\rho)$ by

$$\psi_{t_0}(t) = \left(1 - \frac{t}{t_0}\right) x + \frac{t}{t_0} y$$

for $0 \leq t \leq t_0$. Then $\psi_{t_0}(0) = x$ and $\psi_{t_0}(t_0) = y$.

Remind that assumption (H3) implies that there exists a constant $C > 0$ such that for all $s \in \mathbb{R}^d$ and $x \in \mathcal{X}$, $s^* a(x) s \geq C \|\nabla_1 g(x, x)\| \|s\|^2$, *i.e.* $s^* a^{-1}(x) s \leq \|s\|^2 / C \|\nabla_1 g(x, x)\|$. From the fact that, on the closure of $G^{2\rho} \setminus \Gamma^\rho$, which is a compact set, $\|\nabla_1 g(x, x)\|$ never vanishes, it follows the existence of a constant C' such that for all $s \in \mathbb{R}^d$ and $x \in G^{2\rho} \setminus \Gamma^\rho$, $s^* a^{-1}(x) s \leq C' \|s\|^2$.

It follows from this estimate that

$$\begin{aligned} I_{t_0,x}(\psi_{t_0}) &\leq \frac{C'}{2} \int_0^{t_0} (2\|\dot{\psi}_{t_0}(t)\|^2 + 2\|b(\psi(t)_{t_0})\|^2) dt \\ &\leq C' \int_0^{t_0} \left(\frac{\rho^2}{t_0^2} + B^2\right) dt \\ &\leq C' \left(\frac{\rho^2}{t_0} + B^2 t_0\right), \end{aligned}$$

where B is a bound for b on $\overline{\mathcal{X}}$. The infimum of the right-hand side is obtained for $t_0 = \rho / B$, and gives

$$I_{\rho/B,x}(\psi_{\rho/B}) \leq 2BC'\rho,$$

which converges to 0 when $\rho \to 0$. It remains to observe that $t_0 = \rho / B \leq 1$ for ρ small enough to complete the proof of (66). □

Proof of Lemma 5.2 If $x \in B(\rho)$, then $\sigma_\rho = 0$ and the lemma trivially holds. Otherwise, consider the closed sets

$$C_t = \mathcal{C}([0, t], \overline{G \setminus B(\rho)}),$$

and observe that, for $x \in G$, the event $\{\sigma_\rho > t\}$ is contained in $\{X^\varepsilon \in C_t\}$. Corollary 4.1 yields, for all $t > 0$,

$$\limsup_{\varepsilon \to 0} \varepsilon \ln \sup_{x \in \overline{G \setminus B(\rho)}} \mathbf{P}_x(X^\varepsilon \in C_t) \leq - \inf_{\psi \in C_t} I_t(\psi),$$

where throughout this proof $I_t(\psi)$ stands for $I_{t,\psi(0)}(\psi)$. Hence, in order to complete the proof of the lemma, it suffices to show that

$$\lim_{t \to \infty} \inf_{\psi \in C_t} I_t(\psi) = \infty. \tag{70}$$

Let ψ^x denote the trajectory of (61) starting at $x \in \overline{G \setminus B(\rho)}$. By assumptions (Ha) and (Hc), ψ^x hits $S(\rho/3)$ in a finite time, denoted T_x. Moreover, by the uniform Lipschitz continuity of b and Gronwall's lemma, there exists an open neighborhood $\mathcal{W}_x$ of x such that, for all $y \in \mathcal{W}_x$, the path ψ^y hits $S(2\rho/3)$ before T_x. Extracting a finite cover of the compact set $\overline{G \setminus B(\rho)}$ by such sets, it follows that there exists $T < \infty$ such that for all $y \in \overline{G \setminus B(\rho)}$, ψ^y hits $S(2\rho/3)$ before time T.

Assume now that (70) does not hold true. Then , for some $M < \infty$ and every integer n, there exists $\psi^n \in C_{nT}$ such that $I_{nT}(\psi^n) \le M$. Consequently, for some $\psi^{n,k} \in C_T$,

$$M \ge I_{nT}(\psi^n) = \sum_{k=1}^{n} I_T(\psi^{n,k}) \ge n \min_{1 \le k \le n} I_T(\psi^{n,k}).$$

Hence, there exists a sequence $\phi^n \in C_T$ with $\lim I_T(\phi^n) = 0$. It follows from the fact that a is uniformly non-degenerate and uniformly Lipschitz on $\overline{G \setminus B(\rho)}$ and from Lemma 4.1, that the set $\{\phi \in C_T : I_T(\phi) \le 1\}$ is compact, so the sequence ϕ^n has a limit point ϕ^* in C_T, and that I_T is lower semicontinuous on C_T, and therefore, $I_T(\phi^*) = 0$. Consequently, ϕ^* is a trajectory of (61) staying inside $\overline{G \setminus B(\rho)}$ on $[0, T]$, which yields a contradiction with the definition of T. □

Proof of Lemma 5.3 Fix $\eta > 0$ and let $\rho > 0$ be small enough for $B(2\rho) \subset G$ and for Lemma 5.1 to hold with $\delta = \eta/3$ and 2ρ instead of ρ. Note that the definition of $\bar{V}(0, z)$ yields the inequality $\inf_{y \in S(2\rho)} V(y, z) \le \bar{V}(0, z)$ as soon as $z \notin B(2\rho)$. Then, by (66) and assumption (Hd), there exists $x \in S(2\rho)$, $z \notin \overline{G}$, $T_1 < \infty$ and $\psi \in \mathcal{C}([0, T_1], \mathcal{X})$ such that $\psi(0) = x$, $\psi(T_1) = z$ and $I_{T_1,x}(\psi) \le \bar{V} + \eta/3$. Moreover, by properly "cutting" the beginning of the path ψ, we can suppose that for all $t > 0$, $\psi(t) \notin B(2\rho)$. Since $z \in \mathcal{X} \setminus \overline{G}$, the constant $\Delta := d(z, \partial G \cup \partial \mathcal{X})$ is (strictly) positive.

Thanks to (65), for any $y \in S(2\rho)$, there exists a continuous path ψ^y of length $t_y \le 1$ such that $\psi^y(0) = y$, $\psi^y(t_y) = x$, and $I_{t_y,y}(\psi^y) \le \eta/3$. Moreover, the construction of such a function in the proof of Lemma 5.1 allows us to assume that $\|\psi^y(t)\| = 2\rho$ for all $t \in [0, t_y]$. Let ϕ^y denote the path obtained by concatenating ψ^y and ψ (in that order) and extending the resulting function to be of length $T_0 = T_1 + 1$ by following the trajectory of (61) after reaching z. Since the latter path does not contribute to the rate function, it follows that $I_{T_0,y}(\phi^y) \le \bar{V} + 2\eta/3$.

Consider the set

$$O := \bigcup_{y \in S(2\rho)} \left\{ \psi \in \mathcal{C}([0, T_0], \overline{\mathcal{X}}) : \|\psi - \phi^y\|_{0,T_0} < \frac{\Delta \wedge \rho}{2} \right\}.$$

Observe that O is an open subset of $\mathcal{C}([0, T_0], \overline{\mathcal{X}})$ that contains the functions $\{\phi^y\}_{y \in S(2\rho)}$. Therefore, by Corollary 4.1,

$$\begin{aligned}\liminf_{\varepsilon \to 0} \varepsilon \ln \inf_{y \in S(2\rho)} \mathbf{P}_y(X^\varepsilon \in O) &\ge - \sup_{y \in S(2\rho)} \inf_{\psi \in O} I_{T_0,y}(\psi) \\ &\ge - \sup_{y \in S(2\rho)} I_{T_0,y}(\phi^y) > -(\bar{V} + \eta).\end{aligned}$$

If $\psi \in O$, then ψ reaches the open ball of radius $\Delta/2$ centered at z before hitting $B(\rho)$, so ψ hits ∂G before hitting $B(\rho)$. Hence, for $X_0^\varepsilon = y \in S(2\rho)$, the event $\{X^\varepsilon \in O\}$ is contained in $\{X^\varepsilon_{\sigma_\rho} \in \partial G\}$, and the proof is completed. □

Proof of Lemma 5.4 Let us first notice that the fact that $V(x,z) \leq V(x,y) + V(y,z)$ for all x, y and z in $\mathcal{X}$ implies that

$$\forall y, z \in \mathcal{X}, \quad \bar{V}(0,z) \leq \bar{V}(0,y) + V(y,z). \tag{71}$$

Fix a closed set $N \subset \partial G$, fix $\delta > 0$ and define $\bar{V}_N^\delta := (\inf_{z\in N} \bar{V}(0,z) - \delta) \wedge 1/\delta$. Then, it follows from (71) and Lemma 5.1 (65) that, for $\rho > 0$ small enough,

$$\inf_{y\in S(2\rho), z\in N} V(y,z) \geq \inf_{z\in N} \bar{V}(0,z) - \sup_{y\in S(2\rho)} \bar{V}(0,y) \geq \bar{V}_N^\delta.$$

Moreover, by Lemma 5.2, there exists $T < \infty$ large enough for getting

$$\limsup_{\varepsilon\to 0} \varepsilon \ln \sup_{y\in S(2\rho)} \mathbf{P}_y(\sigma_\rho > T) < -\bar{V}_N^\delta.$$

Consider the following closed subset of $\mathcal{C}([0,T], \overline{\mathcal{X}})$:

$$C := \{\psi \in \mathcal{C}([0,T], \overline{\mathcal{X}}) : \exists t \in [0,T] \text{ such that } \psi(t) \in N\}.$$

Note that C obviously satisfies the assumptions of Corollary 4.1, and that

$$\inf_{y\in S(2\rho), \psi\in C} I_{y,T}(\psi) \geq \inf_{y\in S(2\rho), z\in N} V(y,z) \geq \bar{V}_N^\delta.$$

Thus

$$\limsup_{\varepsilon\to 0} \varepsilon \ln \sup_{y\in S(2\rho)} \mathbf{P}_y(X^\varepsilon \in C) \leq -\inf_{y\in S(2\rho), \psi\in C} I_{T,y}(\psi) \leq -\bar{V}_N^\delta.$$

Since $\mathbf{P}_y(X^\varepsilon_{\sigma_\rho} \in N) \leq \mathbf{P}_y(\sigma_\rho > T) + \mathbf{P}_y(X^\varepsilon \in C)$, it follows that

$$\limsup_{\varepsilon\to 0} \varepsilon \ln \sup_{y\in S(2\rho)} \mathbf{P}_y(X^\varepsilon_{\sigma_\rho} \in N) \leq -\bar{V}_N^\delta.$$

Taking $\delta \to 0$ completes the proof of the lemma. □

Proof of Lemma 5.5 Let ρ be small enough for $B(\rho) \subset G$. For $x \in B(\rho)$, there is nothing to prove. Thus, fix $x \in G \setminus B(\rho)$, let ϕ denote the trajectory of (61) with initial state x, and let $T := \inf\{t : \phi(t) \in S(\rho/2)\}$. Because of assumption (Ha), $T < \infty$ and there exists a positive distance between $\{\phi(t)\}_{t\leq T}$ and ∂G. Let $\Delta := \rho \wedge d(\{\phi(t)\}_{t\leq T}, \partial G)$, then

$$\|X^{\varepsilon,x} - \phi\|_{0,T} \leq \Delta/2 \Rightarrow X^{\varepsilon,x}_{\sigma_\rho} \in B(\rho).$$

By the uniform Lipschitz continuity of b, for $t \leq T$,

$$\|X^{\varepsilon,x}_t - \phi(t)\| \leq K \int_0^t \|X^{\varepsilon,x}_s - \phi(s)\| ds + \varepsilon BT + \sqrt{\varepsilon} \left\| \int_0^t \sigma(X^{\varepsilon,x}_s) dW_s \right\|,$$

where B is a bound for $\tilde{b}$. Hence, by Gronwall's lemma,

$$\|X^{\varepsilon,x} - \phi\|_{0,T} \leq \sqrt{\varepsilon} e^{KT} \left(\sqrt{\varepsilon} BT + \left\| \int_0^\cdot \sigma(X^{\varepsilon,x}_s) dW_s \right\|_{0,T} \right),$$

and, by Lemma 4.5, for sufficiently small $\varepsilon > 0$,

$$\begin{aligned}\mathbf{P}_x(X^\varepsilon_{\sigma_\rho} \in \partial G) &\leq \mathbf{P}_x(\|X^\varepsilon - \phi\|_{0,T} > \Delta/2) \\ &\leq \mathbf{P}_x\left(\left\|\int_0^\cdot \sigma(X^\varepsilon_s)dW_s\right\|_{0,T} > \frac{\Delta}{4\sqrt{\varepsilon}}e^{-KT}\right) \\ &\leq 2d\exp\left(-\frac{\Delta^2 e^{-2KT}}{32d\varepsilon T\Sigma^2}\right) \to_{\varepsilon\to 0} 0,\end{aligned}$$

where Σ is a uniform bound for σ. □

Proof of Lemma 5.6 Let B be a uniform bound for b and $\tilde{b}$. For any $t \leq \rho/4B$, and for any $x \in G$, (1) yields

$$\|X^{\varepsilon,x}_t - x\| \leq \frac{\rho}{4} + \varepsilon\frac{\rho}{4} + \sqrt{\varepsilon}\left\|\int_0^t \sigma(X^{\varepsilon,x}_s)dW_s\right\|.$$

Therefore, for any $\varepsilon \leq 1$ and any $t \leq \rho/4B$, by Lemma 4.5,

$$\begin{aligned}\mathbf{P}_x(\|X^\varepsilon - x\|_{0,t} \geq \rho) &\leq \mathbf{P}_x\left(\sqrt{\varepsilon}\left\|\int_0^\cdot \sigma(X^\varepsilon_s)dW_s\right\|_{0,t} \geq \frac{\rho}{2}\right) \\ &\leq 2d\exp\left(-\frac{\rho^2}{8\varepsilon dt\Sigma^2}\right),\end{aligned}$$

where Σ is a uniform bound for σ. Therefore,

$$T(c,\rho) = \frac{\rho}{4B} \wedge \frac{\rho^2}{8dT\Sigma^2 c}$$

is an appropriate constant for Lemma 5.6. □

References

[1] R. Azencott, *Grandes déviations et applications*, Summer school of St Flour VII, 78, Lecture Notes in Math, Springer Verlag (1980).

[2] N. Champagnat, *Convergence and existence for polymorphic adaptive dynamics jump and degenerate diffusion models*, prépublication Université Paris X – Nanterre No. 03/7 (2003).

[3] A. Dembo and O. Zeitouni, *Large Deviation Techniques and Applications*, second edition, Jones & Bartlett Publishers (1993).

[4] U. Dieckmann and M. Doebeli, *On the origin of species by sympatric speciation*, Nature **400**, 354-357 (1999).

[5] U. Dieckmann and R. Law, *The dynamical theory of coevolution: a derivation from stochastic ecological processes*, Journal of Mathematical Biology **34**, 579-612 (1996).

[6] O. Diekmann, *A beginners guide to adaptive dynamics*, in A. Margheri, C. Rebelo and F. Zondin (ed.), Summer School on Mathematical Biology, CIM, Lisboa, Portugal (2002).

[7] H. Doss and P. Priouret, *Remarques sur les petites perturbations des systèmes dynamiques*, Séminaire de Probabilités XVII, 313-370, Lecture Notes in Mathematics, Springer (1983).

[8] M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems*. Springer-Verlag, Berlin (1984).

[9] J. Hofbauer and R. Sigmund, *Adaptive dynamics and evolutionary stability.* Applied Mathematics Letters **3**, 75-79 (1990).

[10] I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus.* Springer-Verlag, New York (1988).

[11] É. Kisdi, *Evolutionary branching under asymmetric competition*, Journal of Theoretical Biology **198**, 149-162 (1999).

[12] P. Marrow, R. Law and C. Cannings, *The coevolution of predator-prey interactions: ESSs and Red Queen dynamics* Proc. R. Soc. Lond. B **250**, 133-141 (1992).

[13] J.A.J Metz, R.M. Nisbet and S.A.H. Geritz, *How should we define 'fitness' for general ecological scenarios?*. Trends in Ecology and Evolution **7**, 198-202 (1992).

[14] J.A.J. Metz, S.A.H. Geritz, G. Meszéna, F.A.J. Jacobs and J.S. van Heerwaarden, *Adaptive Dynamics, a geometrical study of the consequences of nearly faithful reproduction*, in S.J. van Strien & S.M. Verduyn Lunel (ed.), *Stochastic and Spatial Structures of Dynamical Systems*, North Holland, Amsterdam, 183-231 (1996).

[15] S. T. Rachev, *Probability Metrics and the Stability of Stochastic Models*, John Wiley & Sons (1991).

[16] D.A. Rand and H.B. Wilson, *Evolutionary catastrophes, punctuated equilibria and gradualism in ecosystem evolution.* Proc. Roy. Soc. London B **253**, 137-141 (1993).

[17] D. Serre, *Matrices: Theory and Applications*, Springer-Verlag (2002).