HAL
open science

Exponential convergence to quasi-stationary distribution for multi-dimensional diffusion processes

Nicolas Champagnat, Abdoulaye Coulibaly-Pasquier, Denis Villemonais

To cite this version:

Nicolas Champagnat, Abdoulaye Coulibaly-Pasquier, Denis Villemonais. Exponential convergence to quasi-stationary distribution for multi-dimensional diffusion processes. 2016. hal-01293622v1

HAL Id: hal-01293622

https://hal.science/hal-01293622v1

Preprint submitted on 25 Mar 2016 (v1), last revised 16 Jan 2018 (v2)

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

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

Exponential convergence to quasi-stationary distribution for multi-dimensional diffusion processes

Nicolas Champagnat[1,2,3], Koléhè Abdoulaye Coulibaly-Pasquier[1,2], Denis Villemonais[1,2,3]

March 25, 2016

Abstract

We consider diffusion processes killed at the boundary of Riemannian manifolds. The aim of the paper if to provide two different sets of assumptions ensuring the exponential convergence in total variation norm of the distribution of the process conditioned not to be killed. Our first criterion makes use of two sided estimates and applies to general Markov processes. Our second criterion is based on gradient estimates for the semi-group of diffusion processes.

Keywords: Diffusions on Riemannian manifolds; diffusions in a bounded domain; absorption at the boundary; quasi-stationary distribution; Q-process; uniform exponential mixing; two-sided estimates; gradient estimates.

2010 Mathematics Subject Classification. Primary: 60J60; 37A25; 60B10; 60F99. Secondary: 60J75; 60J70

1 Introduction

We consider a Markov process X evolving in a Riemannian compact manifold (M, ρ) of dimension $d \geq 1$ with boundary ∂M, such that, when it hits ∂M, X is killed and immediately sent to a cemetery point $\partial \notin M$. We

[1]IECL, Université de Lorraine, Site de Nancy, B.P. 70239, F-54506 Vanduvre-ls-Nancy Cedex, France

[2]CNRS, IECL, UMR 7502, Vandœuvre-lès-Nancy, F-54506, France

[3]Inria, TOSCA team, Villers-lès-Nancy, F-54600, France.

E-mail: Nicolas.Champagnat@inria.fr, Kolehe.Coulibaly@univ-lorraine.fr, Denis.Villemonais@univ-lorraine.fr

are interested in providing sufficient criteria ensuring exponential mixing for the distribution of X conditioned not to be killed when it is observed. More precisely, our goal is to prove the existence of a probability measure α on M and of positive constants $C, \gamma > 0$ such that, for all initial distribution μ on M,

$$\|\mathbb{P}_\mu(X_t \in \cdot \mid t < \tau_\partial) - \alpha(\cdot)\|_{TV} \leq Ce^{-\gamma t}, \ \forall t \geq 0, \tag{1.1}$$

where $\tau_\partial = \inf\{t \geq 0, \ X_t = \partial\}$ is the killing time of X, $\mathbb{P}_\mu$ is the law of X with initial distribution μ and $\|\cdot\|_{TV}$ is the total variation norm on finite signed measures. In the whole paper we will consider diffusions such that $\mathbb{P}_x(\tau_\partial < \infty) = 1$ and $\mathbb{P}_x(t < \tau_\partial) > 0$, for all $x \in M$ and all $t \geq 0$. It is well known that (1.1) entails that α is the unique quasi-stationary distribution for X, that is the unique probability measure satisfying

$$\alpha(\cdot) = \mathbb{P}_\alpha\left(X_t \in \cdot \mid t < \tau_\partial\right), \ \forall t \geq 0.$$

We provide two independent sets of assumptions to check the exponential mixing property of the conditional distribution of X.

Our first result (Theorem 2.1) shows that a two-sided estimate for the transition density of a general absorbed Markov process at some time $t > 0$ is sufficient to ensure (1.1). Two sided estimates have been recently developed for a wide range of processes, including diffusion processes in general domains of $\mathbb{R}^d$ (we refer to the beginning of Section 2 for a bibliography). In particular, we recover the results of Knobloch and Partzsch [14], who proved that (1.1) holds for a class of diffusion processes evolving in $\mathbb{R}^d$ ($d \geq 3$), using two sided estimates combined with non-trivial spectral properties of the infinitesimal generator of X. We actually prove that the two sided estimates are sufficient for diffusion processes in $\mathbb{R}^d$ but also for general Markov processes, while some spectral properties can be recovered from our results.

Our second result (Theorem 3.1) is based on gradient estimates of the Dirichlet semi-group obtained by Wang [22] and Priola and Wang [21]. The gradient estimates of [22] hold for Brownian motions with C^1 drift evolving in bounded manifolds with C^2 boundary ∂M and killed when they hit ∂M. The gradient estimates of [21] hold for uniformly elliptic diffusion processes with Hölder diffusion coefficient and bounded drift evolving in a bounded domain M of $\mathbb{R}^d$ ($d \geq 1$) with C^2 boundary and killed when they hit ∂M. In both situations, we prove that the convergence (1.1) holds true. Up to our knowledge, this result is the first one of this kind for diffusion processes in Riemannian manifolds and our conditions for diffusion processes on $\mathbb{R}^d$ improve significantly the existing results of [14] and [10]. In particular, we

prove that the global Lipschitz and local regularity assumptions near the boundary of [10] are unnecessary. *A fortiori* the C^1 regularity assumption on the diffusion coefficient of [14] is also strongly relaxed, when the boundary of the domain is of class C^2 and the drift is bounded and continuous (in [14] as in Section 2, domains with $C^{1,1}$ boundary and drifts in a Kato class are allowed).

Our two approaches are complementary. The one based on two sided estimates is particularly adapted to irregular domains (see for instance [16]), while the gradient estimate approach applies to diffusion processes with less regular coefficients.

The usual tools to prove convergence as in (1.1) involve coupling arguments: for example, contraction in total variation norm for the non-conditioned semi-group can be obtained using mirror and parallel coupling, see [17, 22, 21]. However, the process conditioned not to be killed up to a given time $t > 0$ is a time-inhomogeneous diffusion process with a singular drift for which these methods fail. For instance, a standard d-dimensional Brownian motion $(B_t)_{t\geq 0}$ conditioned not to exit a smooth domain $D \subset \mathbb{R}^d$ up to a time $t > 0$ has the law of the solution $(X_s^{(t)})_{s\in[0,t]}$ to the stochastic differential equation

$$dX_s^{(t)} = dB_s + [\nabla \ln \mathbb{P}_\cdot(t - s < \tau_\partial)]\,(X_s^{(t)})ds.$$

Since $\mathbb{P}_x(t - s < \tau_\partial)$ vanishes when x converges to the boundary ∂D of D, the drift term in the above SDE is singular and existing coupling methods do not apply. Hence, convergence of conditioned diffusion processes have been obtained up to now using (sometimes involved) spectral theoretic arguments (see for instance, [2, 15, 18, 20] for one-dimensional diffusion processes and [3, 14] for multi-dimensional diffusion processes) which are strictly limited to diffusion processes and have limitations in terms of generality and flexibility.

The strength of our approach is that, instead of using a spectral theoretic perspective, we rely, as was done in [5] for one-dimensional diffusions, on recent probabilistic criteria for convergence of conditioned processes obtained in [4, Theorem 2.1], overcoming the difficulties pointed out in the previous paragraph. This result states that the exponential convergence (1.1) is equivalent to the following condition.

Condition (A). There exist $t_0, c_1, c_2 > 0$ such that, for all $x, y \in M$, there exists a probability measure $\nu_{x,y}$ on M satisfying

(A1)

$$\mathbb{P}_x(X_{t_0} \in \cdot \mid t_0 < \tau_\partial) \geq c_1 \nu_{x,y}(\cdot)$$

and

$$\mathbb{P}_y(X_{t_0} \in \cdot \mid t_0 < \tau_\partial) \geq c_1 \nu_{x,y}(\cdot);$$

(A2) for all $z \in M$ and all $t \geq 0$,

$$\mathbb{P}_{\nu_{x,y}}(t < \tau_\partial) \geq c_2 \mathbb{P}_z(t < \tau_\partial).$$

Condition (A) also implies that the conditional distributions of X are in fact contracting in total variation [4, Corollary 2.2], that $e^{\lambda_0 t}\mathbb{P}_x(t < \tau_\partial)$ converges when $t \to +\infty$, uniformly in x, to a positive eigenfunction η of the infinitesimal generator of $(X_t, t \geq 0)$ for the eigenvalue $-\lambda_0$ characterized by the relation $\mathbb{P}_\alpha(t < \tau_\partial) = e^{-\lambda_0 t}$, $\forall t \geq 0$ [4, Proposition 2.3], a spectral gap property [4, Corollary 2.4], and the existence and exponential ergodicity of the so-called Q-process, defined as the process X conditioned to never hit the boundary [4, Theorem 3.1].

The paper is organized as follows. In Section 2, we state and prove a sufficient criterion for (1.1) based only on two-sided estimates. We also provide several references giving two sided estimates for many kinds of processes. In Section 3, we state a sufficient criterion for (1.1) for drifted Brownian motion on Riemannian manifolds and for diffusion processes with Hölder diffusion coefficients. The proof, based on gradient estimates obtained in [22, 21], is given in Section 4.

2 Quasi-stationary behavior under two-sided estimates

In this section, we consider absorbed Markov processes satisfying two-sided estimates. More specifically, we consider a diffusion process X in some closed Riemanian manifold M (or open subset of $\mathbb{R}^d$), absorbed at the boundary of M at first hitting time τ_∂. We assume that M is equipped with a measure μ such that for some $t > 0$ and for all $x \in M$, X_t admits a density function $p(t, x, y)$ in M with respect to $\mu(dy)$ given $X_0 = x$. We say that X satisfies a two-sided estimate at time t if there exists a constant $c > 0$ and two measurable function $f_1, f_2 : M \to [0, +\infty)$ such that

$$c^{-1} f_1(x) f_2(y) \leq p(t, x, y) \leq c f_1(x) f_2(y), \ \forall x, y \in M. \tag{2.1}$$

Our main result states that this condition, together with the assumption $\mu(f_1 f_2) > 0$, guaranties the uniform exponential convergence in total variation of the conditional distributions of X_s given $s < \tau_\partial$ to a unique quasi-stationary distribution.

Two-sided estimates were already a key ingredient to prove exponential convergence of conditional distributions in [14]. However, their analysis requires additional assumptions (continuity of the transition density and existence of ground states) and they assume that the two-sided estimate holds with f_1 and f_2 the ground states of the generator and its adjoint. In addition, checking their criterion involves intricate spectral theory arguments, whereas our proof only requires to check the above two-sided estimate. Moreover, our approach leads to explicit rates of convergence in terms of c, f_1, f_2 and μ. Also note that the existence of the ground states are recovered afterwards from our results: an eigenmeasure for the adjoint generator is given by the quasi-stationary distribution α [19] and an eigenfunction for the generator by [4, Proposition 2.3]. Moreover, as will appear clearly in the proof, we don't use at all the Riemannian structure of M nor the diffusion property of X. In particular, our result applies to any sub-Markov process in any measurable state space M equipped with some measure μ, whose semi-group exhibits the property (2.1) for some $t > 0$. Finally, since our criterion implies Condition (A), we obtain many complementary results as explained in the Introduction. In particular, we obtain a contraction result in total variation (see (2.3) below), the uniform convergence of $e^{\lambda_0 t}\mathbb{P}_x(t < \tau_\partial)$ to the ground state when $t \to +\infty$ [4, Proposition 2.3], a spectral gap result [4, Corollary 2.4] and the uniform exponential ergodicity of the Q-process [4, Theorem 3.1].

Estimates of the form (2.1) have been proved in a variety of contexts. In the case of diffusion processes in $\mathbb{R}^d$, $d \geq 3$ absorbed at the boundary of a bounded domain, this goes back to the seminal paper of Davies and Simon [9]. The case of standard Brownian motion in a bounded $C^{1,1}$ domain of $\mathbb{R}^d$, $d \geq 3$ was studied in [23]. This result has then been extended in [13] to diffusions in a bounded $C^{1,1}$ domain in $\mathbb{R}^d$, $d \geq 3$ or more, with inifnitesimal generator

$$L = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}\partial_i\partial_j + \sum_{i=1}^{d} b_i\partial_i,$$

with symmetric, uniformly elliptic and C^1 diffusion matrix $(a_{ij})_{1\leq i,j\leq d}$, and with drift $(b_i)_{1\leq i\leq d}$ in the Kato class $K_{d,1}$, which contains $L^p(dx)$ functions for $p > d$. Diffusions on bounded, closed Riemannian manifolds with very

irregular boundary and with generator

$$L = \Delta + X,$$

where Δ is the Laplace-Beltrami operator and X is a smooth vector field, were also studied in [16].

Estimates of the form (2.1) are also known for quite general one-dimensional diffusions (see for instance [20], where the link with quasi-stationary distribution is studied).

Two-sided estimates are also known for other classes of processes, among which α-stable Lévy processes (generated by the fractional Laplacian) with drift in a Kato class in a $C^{1,1}$ domain of $\mathbb{R}^d$ ($d \geq 2$) [7], or without drift for more general domains in any dimension [1]. Extensions to more general jump processes similar to the α-stable Lévy process are given in [6, 12, 8]. For example, the results of the last reference covers α-stable Lévy processes with killing (either at some state-dependent rate, of with some state-dependent probability at each jump time).

Remark 1. Note that most of these results are actually obtained for Markov processes with additional killing inside the domain, and the next result also covers this case. Indeed, two sided estimates for processes with bounded killing rate are immediate consequences of the two sided estimates for the process without killing

Theorem 2.1. *Assume that $\mathbb{P}_x(\tau_\partial < \infty) = 1$ for all $x \in M$, that there exists $t_0 > 0$ such that X_{t_0} admits a density function $p(t_0, x, y)$ in M with respect to a probability measure $\mu(dy)$ on M given $X_0 = x$ and that there exist a constant $c > 0$ and two functions $f_1, f_2 : M \to [0, +\infty)$ such that*

$$c^{-1} f_1(x) f_2(y) \leq p(t_0, x, y) \leq c f_1(x) f_2(y), \ \forall x, y \in M, \tag{2.2}$$

and such that $\mu(f_1 f_2) > 0$, then Assumption (A), and hence (1.1)*, are satisfied. In addition, for all μ_1, μ_2 probability measures on M,*

$$\begin{aligned} &\|\mathbb{P}_{\mu_1}(X_t \in \cdot \mid t < \tau_\partial) - \mathbb{P}_{\mu_2}(X_t \in \cdot \mid t < \tau_\partial)\|_{TV} \\ &\qquad \leq \frac{c^3}{\mu(f_2)} \left(1 - c^{-5}\mu(f_1 f_2)\right)^{\lfloor t/t_0 \rfloor} \frac{\|\mu_1 - \mu_2\|_{TV}}{\mu_1(f_1) \wedge \mu_2(f_1)}. \end{aligned} \tag{2.3}$$

In addition, there exists a unique quasi-stationary distribution α for X, which satisfies

$$c^{-2} \frac{f_2(y)\mu(dy)}{\int_M f_2(x)\mu(dx)} \leq \alpha(dy) \leq c^2 \frac{f_2(y)\mu(dy)}{\int_M f_2(x)\mu(dx)}. \tag{2.4}$$

Proof of Theorem 2.1. Let us first check that (A1) holds with $\nu_{x,y} = \nu$ independent of x, y given by

$$\nu(dy) = \frac{f_2(y)\mu(dy)}{\mu(f_2)}.$$

Note that the assumptions of Theorem 2.1 imply that $\mu(f_2) \in (0, +\infty)$, and

$$\|f_1\|_\infty \mu(f_2) \leq c. \tag{2.5}$$

In particular, ν is a well defined probability measure. For all $x \in M$ and $A \subset M$, we have, by (2.2)

$$c^{-1} f_1(x) \int_A f_2(y)\, d\mu(y) \leq \mathbb{P}_x(X_{t_0} \in A) \leq c f_1(x) \int_A f_2(y)\, d\mu(y).$$

In particular,

$$\mathbb{P}_x(X_{t_0} \in A \mid t_0 < \tau_\partial) = \frac{\mathbb{P}_x(X_{t_0} \in A)}{\mathbb{P}_x(X_{t_0} \in E)} \geq c^{-2} \frac{\int_A f_2(y)\, \mu(dy)}{\mu(f_2)}.$$

We thus obtained (A1) with $c_1 = c^{-2}$ and $\nu_{x,y} = \nu$.

Let us now check (A2). We have, for all $z \in M$,

$$\begin{aligned}
\mathbb{P}_\nu(X_{t_0} \in A) &\geq c^{-1} \int_M f_1(x)\, d\nu(x) \int_A f_2(x)\, \mu(dx) \\
&\geq c^{-1} \frac{\mu(f_1 f_2)}{\mu(f_2)} \frac{f_1(z)}{\|f_1\|_\infty} \int_A f_2(x)\, \mu(dx) \\
&\geq c^{-3} \mu(f_1 f_2)\, \mathbb{P}_z(X_{t_0} \in A),
\end{aligned}$$

where we used (2.5) in the last inequality.

Hence, for all $t \geq t_0$ and all $z \in M$,

$$\begin{aligned}
\mathbb{P}_\nu(t < \tau_\partial) &= \mathbb{E}_\nu \left(\mathbb{P}_{X_{t_0}}(t - t_0 < \tau_\partial) \right) \\
&\geq c^{-3} \mu(f_1 f_2)\, \mathbb{E}_z \left(\mathbb{P}_{X_{t_0}}(t - t_0 < \tau_\partial) \right) \\
&= c^{-3} \mu(f_1 f_2)\, \mathbb{P}_z(t < \tau_\partial).
\end{aligned}$$

For $t < t_0$, we have $\mathbb{P}_\nu(t < \tau_\partial) \geq \mathbb{P}_\nu(t_0 < \tau_\partial) \geq \mathbb{P}_\nu(t_0 < \tau_\partial)\mathbb{P}_z(t < \tau_\partial)$. Integrating the first inequality in (2.2) with respect to $\nu(dx)\mu(dy)$, we have

$$\mu(f_1 f_2) = \nu(f_1)\mu(f_2) \leq c\mathbb{P}_\nu(t_0 < \tau_\partial).$$

Hence $\mathbb{P}_\nu(t < \tau_\partial) \geq c^{-1}\mu(f_1 f_2)\mathbb{P}_z(t < \tau_\partial)$ for $t < t_0$ and we have proved (A2) for $c_2 = (c^{-3} \wedge c^{-1})\mu(f_1 f_2) = c^{-3}\mu(f_1 f_2)$.

We now prove (2.3). We make use of the following general consequence of Condition (A), proved in [4, Corollary 2.2]: for all probability measures μ_1, μ_2 on M and for all $t > 0$,

$$\|\mathbb{P}_{\mu_1}(X_t \in \cdot \mid t < \tau_\partial) - \mathbb{P}_{\mu_2}(X_t \in \cdot \mid t < \tau_\partial)\|_{TV} \leq \frac{(1 - c_1 c_2)^{\lfloor t/t_0 \rfloor}}{c(\mu_1) \wedge c(\mu_2)} \|\mu_1 - \mu_2\|_{TV}, \tag{2.6}$$

with $c(\mu_i)$ defined by

$$c(\mu_i) = \inf_{t \geq 0} \frac{\mathbb{P}_{\mu_i}(t < \tau_\partial)}{\sup_{z \in M} \mathbb{P}_z(t < \tau_\partial)}.$$

Hence we need to prove that, for $i = 1, 2$,

$$c(\mu_i) \geq c^{-3} \mu(f_2) \mu_i(f_1). \tag{2.7}$$

We proceed as above. Assume first $t \geq t_0$. Then, using (2.2),

$$\begin{aligned}
\mathbb{P}_{\mu_i}(X_{t_0} \in A) &\geq c^{-1} \int_M f_1(x)\, d\mu_i(x) \int_A f_2(x)\, d\mu(x) \\
&\geq c^{-1} \mu_i(f_1) \frac{f_1(z)}{\|f_1\|_\infty} \int_A f_2(x)\, d\mu(x) \\
&\geq c^{-3} \mu_i(f_1) \mu(f_2)\, \mathbb{P}_z(X_{t_0} \in A),
\end{aligned}$$

where we also used (2.5). Now, for all $t \geq t_0$,

$$\begin{aligned}
\mathbb{P}_{\mu_i}(t < \tau_\partial) &= \mathbb{E}_{\mu_i}\left(\mathbb{P}_{X_{t_0}}(t - t_0 < \tau_\partial)\right) \\
&\geq c^{-3} \mu_i(f_1) \mu(f_2)\, \mathbb{E}_z\left(\mathbb{P}_{X_{t_0}}(t - t_0 < \tau_\partial)\right) \\
&= c^{-3} \mu_i(f_1) \mu(f_2)\, \mathbb{P}_z(t < \tau_\partial),
\end{aligned}$$

which implies (2.7) for $t \geq t_0$. For $t \leq t_0$, using (2.2),

$$\begin{aligned}
\mathbb{P}_{\mu_i}(t < \tau_\partial) &\geq \mathbb{P}_{\mu_i}(t_0 < \tau_\partial) \\
&\geq \mathbb{P}_{\mu_i}(t_0 < \tau_\partial) \sup_{x \in M} \mathbb{P}_x(t < \tau_\partial) \\
&\geq c^{-1} \mu_i(f_1) \mu(f_2) \sup_{x \in M} \mathbb{P}_x(t < \tau_\partial) \\
&\geq c^{-3} \mu_i(f_1) \mu(f_2) \sup_{x \in M} \mathbb{P}_x(t < \tau_\partial),
\end{aligned}$$

where the last inequality comes from the fact that $c \geq 1$. This completes the proof of (2.3). The existence and uniqueness of a quasi-stationary distribution α follows easily (see [4, Thm. 2.1]).

To conclude the proof of Theorem 2.1, it only remains to check (2.4). For this, given any bounded measurable $g : M \to \mathbb{R}_+$, since α is a quasi-stationary distribution,

$$\alpha(g) = \mathbb{E}_\alpha[g(X_t) \mid t < \tau_\partial] = \frac{\iint_{M^2} g(y)p(t,x,y)\alpha(dx)\mu(dy)}{\iint_{M^2} p(t,x,y)\alpha(dx)\mu(dy)}.$$

Using twice (2.2), we obtain

$$c^{-2}\nu(g) \leq \alpha(g) \leq c^2\nu(g).$$

The conclusion follows. □

3 Quasi-stationary behavior under gradient estimates

In this section, we consider the two following distinct situations and prove that the exponential convergence (1.1) holds for both of them. More precisely, we assume that either

S1. X is a diffusion process evolving in a bounded, connected and closed Riemannian manifold M with C^2 boundary ∂M and the infinitesimal generator of X is given by $L = \frac{1}{2}\Delta + Z$, where Δ is the Laplace-Beltrami operator and Z is a C^1 vector field. The process is absorbed when it hits the boundary ∂M.

S2. X is a diffusion process evolving in a bounded and connected domain M of $\mathbb{R}^d$ with C^2 boundary ∂M and X satisfies the SDE $dX_t = s(X_t)dB_t + b(X_t)dt$, where $(B_t, t \geq 0)$ is a r-dimensional standard Brownian motion, $b : M \to \mathbb{R}^d$ is bounded and continuous and $s : M \to \mathbb{R}^{d\times r}$ is continuous, ss^* is uniformly elliptic and for all $r > 0$,

$$\sup_{x,y\in M,\ |x-y|=r} \frac{|s(x) - s(y)|^2}{r} \leq g(r) \tag{3.1}$$

for some function g such that $\int_0^1 g(r)dr < \infty$. The process is absorbed when it hits the boundary ∂M.

The reason why we concentrate on these two examples is that we will make an extensive use of gradient estimates developed by Wang in [22] and Priola and Wang in [21] respectively for the former and the latter case. In both

cases, the authors have proved that there exists $t_1 > 0$ and $C < \infty$ such that, for all f bounded measurable,

$$\|\nabla P_{t_1} f\|_\infty \leq C\|f\|_\infty, \tag{3.2}$$

where (P_t) denotes the semi-group of X. Both [22] and [21] actually give a stronger version of (3.2):

$$\|\nabla P_t f\|_\infty \leq \frac{c}{1 \wedge \sqrt{t}}\|f\|_\infty, \quad \forall t > 0, \tag{3.3}$$

with proofs relying on a careful study of some coupling for copies of X.

Situation S1 corresponds exactly to the assumptions of [22]. In Situation S2, we need to assume that M is bounded and b is bounded on M, which is stronger than what [21] assumes. Our assumptions clearly imply (i), (ii), (iv) of [21, Hyp. 4.1] (see [21, Lemma 3.3] for the assumption on s). Since we assume that M is bounded and C^2, assumptions (iii') and (v) of [21] are also satisfied (see [21, Rk. 4.2]). Note that (3.1) is satisfied as soon as s is uniformly α-Hölder on M for some $\alpha > 0$.

We denote by ρ the Riemannian distance on M and by $\rho_{\partial M}(\cdot) := \inf_{x \in \partial M} \rho(x, \cdot)$ the distance to ∂M.

Theorem 3.1. *Assume that X is a diffusion process as in situations S1 or S2 above. Then Condition (A) and hence* (1.1) *are satisfied. Moreover, there exist two constants $C, \gamma > 0$ such that, for any initial distributions μ_1 and μ_2 on M,*

$$\begin{aligned} &\|\mathbb{P}_{\mu_1}(X_t \in \cdot \mid t < \tau_\partial) - \mathbb{P}_{\mu_2}(X_t \in \cdot \mid t < \tau_\partial)\|_{TV} \\ &\qquad\qquad \leq \frac{Ce^{-\gamma t}}{\mu_1(\rho_{\partial M}) \wedge \mu_2(\rho_{\partial M})}\|\mu_1 - \mu_2\|_{TV}. \end{aligned} \tag{3.4}$$

Before turning to the proof of this theorem, we emphasize the novelty of Theorem 3.1 in situation S1, since, up to our knowledge, no quasi-stationary distribution result has been obtain for diffusion processes evolving in such general Riemannian manifolds. As explained in the introduction, Theorem 3.1 in situation S2 is a significant improvement of [10] and [14]. In addition, compared with these two papers, the contraction inequality (3.4) is a completely new result for diffusion processes, as well as the complementary results mentioned in the introduction (uniform convergence of $e^{\lambda_0 t}\mathbb{P}_x(t < \tau_\partial)$ to the ground state when $t \to +\infty$, a spectral gap result and the uniform exponential ergodicity of the Q-process, see [4]).

Remark 2. The gradient estimates of [21] are proved for diffusion processes with space-dependent killing rate $V : M \to [0, \infty)$. More precisely, they consider infinitesimal generators of the form

$$L = \frac{1}{2} \sum_{i,j=1}^{d} [ss^*]_{ij} \partial_i \partial_j + \sum_{i=1}^{d} b_i \partial_i - V$$

with V bounded measurable. Our proof also applies to this setting, although adding a bounded killing rate would make our notations much more intricate.

4 Proof of Theorem 3.1

Since a bounded domain of $\mathbb{R}^d$ with C^2 boundary is also a Riemannian manifold, we will keep the latter terminology in the proof. It is important to keep in mind that S1 and S2 correspond to quite different situations. However, most of the proof applies to both situations simultaneously because it mainly relies on (3.2), which has been proved in both settings, respectively in [22] and [21].

4.1 Estimates of the boundary's hitting probability and return time to a compact

Since the boundary ∂M is C^2 and compact, there exists $\varepsilon_0 > 0$ such that $\rho_{\partial M}$ is C^2 on $M \setminus M_{\varepsilon_0}$, where

$$M_{\varepsilon_0} := \{x \in M : \rho_{\partial M}(x) \geq \varepsilon_0\}.$$

The goal of this subsection is to prove that there exists constants $\varepsilon_2 \in (0, \varepsilon_0)$ and $A > 0$ such that, for all $x \in M$ and all $s \geq t_1$, where t_1 comes from (3.2),

$$\mathbb{P}_x(X_s \in M_{\varepsilon_2} \mid s < \tau_\partial) \geq A. \tag{4.1}$$

First, we deduce immediately from the gradient inequality (3.2) applied to $f = 1_M$ that there exists a constant $C > 0$ such that, for all $x \in M$,

$$\mathbb{P}_x(t_1 < \tau_\partial) \leq C\, \rho_{\partial M}(x). \tag{4.2}$$

The following Lemma is proved at the end of this subsection.

Lemma 4.1. *There exist $\varepsilon_1 \in (0, \varepsilon_0)$ and a constant $C' > 0$ such that, for all $x \in M$,*

$$\mathbb{P}_x(T_{\varepsilon_1} < t_1 < \tau_\partial) \geq C' \, \rho_{\partial M}(x), \tag{4.3}$$

where $T_\varepsilon = \inf\{t \geq 0, \ X_t \in M_\varepsilon\}$.

This lemma and (4.2) imply that, for all $x \in M$,

$$\mathbb{P}_x(T_{\varepsilon_1} \leq t_1 \mid t_1 < \tau_\partial) = \frac{\mathbb{P}_x(T_{\varepsilon_1} \leq t_1 < \tau_\partial)}{\mathbb{P}_x(t_1 < \tau_\partial)} \geq \frac{C'}{C}.$$

Since the diffusion is uniformly elliptic, it is clear using local charts that

$$C_2 := \inf_{y \in M_{\varepsilon_1}} \mathbb{P}_y(\rho(X_s, y) \leq \varepsilon_1/2, \ \forall s \in [0, t_1]) > 0.$$

Hence, for all $x \in M \setminus M_{\varepsilon_1}$,

$$\begin{aligned} \mathbb{P}_x(X_{t_1} \in M_{\varepsilon_1/2}) &\geq \mathbb{P}_x(T_{\varepsilon_1} < t_1) \inf_{y \in \partial M_{\varepsilon_1}} \mathbb{P}_y(T_{\varepsilon_1/2} > t_1) \\ &\geq \frac{C' C_2 \mathbb{P}_x(t_1 < \tau_\partial)}{C} \end{aligned}$$

For $x \in M_{\varepsilon_1}$, we have

$$\mathbb{P}_x(X_{t_1} \in M_{\varepsilon_1/2}) \geq C_2 \geq C_2 \mathbb{P}_x(t_1 < \tau_\partial).$$

Finally, we deduce that there exists a constant $A > 0$ such that, for all $x \in M$,

$$\mathbb{P}_x(X_{t_1} \in M_{\varepsilon_1/2} \mid t_1 < \tau_\partial) \geq A. \tag{4.4}$$

This is (4.1) for $s = t_1$ and $\varepsilon_2 = \varepsilon_1/2$.

To conclude, for some fixed $s > t_1$ and $x \in M$, we deduce from the Markov property that

$$\begin{aligned} \mathbb{P}_x(X_s \in M_{\varepsilon_1/2}) &= \mathbb{E}_x \left[1_{s - t_1 < \tau_\partial} \mathbb{P}_{X_{s-t_1}}(X_{t_1} \in M_{\varepsilon_1/2}) \right] \\ &\geq A \, \mathbb{E}_x \left[1_{s - t_1 < \tau_\partial} \mathbb{P}_{X_{s-t_1}}(t_1 < \tau_\partial) \right] \\ &= A \, \mathbb{P}_x(s < \tau_\partial). \end{aligned}$$

Proof of Lemma 4.1. For all $t < T_{\varepsilon_0}$, we define $Y_t = \rho_{\partial M}(X_t)$. In both situations S1 and S2, we have

$$dY_t = \sigma_t dB_t + b_t dt,$$

where B is a standard Brownian motion, where $\sigma_t \in [\underline{\sigma}, \bar{\sigma}]$ and $|b_t| \leq \bar{b}$ are adapted continuous processes, with $0 < \underline{\sigma}, \bar{\sigma}, \bar{b} < \infty$. Using Dubins-Schwarz Theorem, there exists a differentiable function $c(s)$ such that $c(0) = 0$ and

$$W_s := \int_0^{c(s)} \sigma_t dB_t$$

is a Brownian motion and $c'(s) \in [\bar{\sigma}^{-2}, \underline{\sigma}^{-2}]$. In addition,

$$\int_0^{c(s)} b_t \, dt \geq -\bar{b}c(s) \geq -\bar{b}\underline{\sigma}^{-2}s.$$

As a consequence, setting $Z_s = Y_0 + W_s - \bar{b}\underline{\sigma}^{-2}s$, we have almost surely $Z_s \leq Y_{c(s)}$ for all s such that $c(s) \leq T_{\varepsilon_0}$.

Setting $a = \bar{b}\underline{\sigma}^{-2}$, the function

$$f(x) = \frac{e^{2ax} - 1}{2a}$$

is a scale function for the drifted Brownian motion Z. The diffusion process defined by $N_t = f(Z_t)$ is a martingale and its speed measure is given by $s(dv) = \frac{dv}{(1+2av)^2}$. The Green function for one-dimensional diffusion processes [11, Lemma 23.10] entails, for $\varepsilon = f(\varepsilon_0)$ and all $u \in (0, \varepsilon/2)$ (in the following lines, $\mathbb{P}_u^N$ denotes the probability with respect to N with initial position $N_0 = u$),

$$\begin{aligned}
\mathbb{P}_u^N(t < T_0^N \wedge T_{\varepsilon/2}^N) &\leq \frac{\mathbb{E}_u^N(T_0^N \wedge T_{\varepsilon/2}^N)}{t} = \frac{2}{t}\int_0^{\varepsilon/2}\left(1 - \frac{u \vee v}{\varepsilon/2}\right)(u \wedge v)s(dv) \\
&\leq \frac{2}{t}\int_0^{\varepsilon/2}(u \wedge v)\frac{dv}{(1+2av)^2} \\
&\leq u\,\frac{C_\varepsilon}{t}, \quad \text{where } C_\varepsilon = 2\int_0^{\varepsilon/2}\frac{dv}{(1+2av)^2},
\end{aligned} \tag{4.5}$$

where we set $T_\varepsilon^N = \inf\{t \geq 0,\ N_t = \varepsilon\}$. Let us fix $t_2 = \varepsilon C_\varepsilon$. Since N is a martingale, we have, for all $u \in (0, \varepsilon/2)$,

$$\begin{aligned}
u = \mathbb{E}_u^N(N_{t_2 \wedge T_{\varepsilon/2}^N \wedge T_0^N}) &\leq \frac{\varepsilon}{2}\mathbb{P}_u^N(T_{\varepsilon/2}^N < t_2 \wedge T_0^N) + \frac{\varepsilon}{2}\mathbb{P}_u^N(t_2 < T_{\varepsilon/2}^N \wedge T_0^N) \\
&\leq \frac{\varepsilon}{2}\mathbb{P}_u^N(T_{\varepsilon/2}^N < t_2 \wedge T_0^N) + \frac{u}{2},
\end{aligned}$$

where the last inequality comes from (4.5). Reducing ε_0 (and hence ε) if necessary, we can assume without loss of generality that $t_2 = \varepsilon C_\varepsilon \leq \underline{\sigma}^2 t_1$. Hence there exists a constant $C > 0$ only depending on ε_0 such that

$$\mathbb{P}_u^N(T^N_{\varepsilon/2} < \underline{\sigma}^2 t_1 \wedge T^N_0) \geq C\, u.$$

Hence, setting $\varepsilon_1 = f^{-1}(\varepsilon/2)$, for all $x \in M \setminus M_{\varepsilon_1}$,

$$\mathbb{P}_x(T^Z_{\varepsilon_1} < \underline{\sigma}^2 t_1 \wedge T^Z_0) \geq C\, f(\rho_{\partial M}(x)) \geq C\, \rho_{\partial M}(x).$$

Now, using the fact that the derivative of the time change c_s belongs to $[\underline{\sigma}, \bar{\sigma}]$, it follows that for all $x \in M \setminus M_{\varepsilon_1}$,

$$\begin{aligned}\mathbb{P}_x(T^Y_{\varepsilon_1} < t_1 \wedge T^Y_0) &\geq \mathbb{P}_x(c(T^Z_{\varepsilon_1}) \leq c(\underline{\sigma}^2 t_1) \wedge c(T^Z_0)) \\ &= \mathbb{P}_x(T^Z_{\varepsilon_1} < \underline{\sigma}^2 t_1 \wedge T^Z_0) \\ &\geq C \rho_{\partial M}(x).\end{aligned}$$

Therefore,

$$\begin{aligned}\mathbb{P}_x(T^Y_{\varepsilon_1} \leq t_1 \leq T^Y_0) &\geq \mathbb{E}_x\left[1_{T^Y_{\varepsilon_1} \leq t_1 \wedge T^Y_0}\, \mathbb{P}_{X_{T^Y_{\varepsilon_1}}}(t_1 < \tau_\partial)\right] \\ &\geq \mathbb{P}_x(T^Y_{\varepsilon_1} \leq t_1 \wedge T^Y_0) \inf_{y \in M_{\varepsilon_1}} \mathbb{P}_y(t_1 < \tau_\partial) \\ &\geq C' \rho_{\partial M}(x),\end{aligned}$$

where we used that $\inf_{y \in M_{\varepsilon_1}} \mathbb{P}_y(t_1 < \tau_\partial) > 0$. Finally, since $T_{\varepsilon_1} = 0$ under $\mathbb{P}_x$ for all $x \in M_{\varepsilon_1}$, replacing C' by $C' \wedge \mathrm{diam}(M)^{-1}$ entails (4.3) for all $x \in M$. □

4.2 Dobrushin coefficient for the fixed time horizon conditioned distribution

For all $x, y \in M$ and $t > 0$, let $\mu^t_{x,y}$ be the infimum measure of $\delta_x P_t$ and $\delta_y P_t$, i.e. for all measurable $A \subset M$,

$$\mu^t_{x,y}(A) := \inf_{A_1 \cup A_2 = A} (\delta_x P_t 1_{A_1} + \delta_y P_t 1_{A_2}).$$

For all $z \in M$ and $r_0 > 0$, we define for all $x \in M$

$$f_{z,r_0}(x) = (r_0 - \rho(x,z))_+,$$

where $a_+ = a \vee 0$. For all $t > 0$ and $r_0 > 0$, we define

$$\mathcal{E}_{t,r_0} := \left\{(x,y,z) \in M^3 : \forall r > r_0,\ \mu^t_{x,y}(f_{z,r}) > 0\right\}$$

and

$$\mathcal{E}_{r_0} := \bigcap_{t\in(1,2]} \mathcal{E}_{t,r_0}.$$

Our first goal is to prove the next lemma.

Lemma 4.2. *For all $t > 0$, $(x,y) \mapsto \mu^t_{x,y}(M)$ is continuous on M^2. In addition, $\mathcal{E}_{r_0} = M^3$ for all $r_0 > 0$.*

Proof. By (3.3), for all positive measurable f on M bounded by 1, for all $x, y \in M$ and for all $t \in (1,2]$

$$|P_t f(x) - P_t f(y)| \leq \frac{c\rho(x,y)}{1 \wedge \sqrt{t}} \leq c\rho(x,y). \tag{4.6}$$

This implies the uniform Lipschtiz-continuity of $P_t f$ for all f bounded by 1. In particular, we deduce that

$$\mu^t_{x,y}(M) = \inf_{A_1 \cup A_2 = M} (P_t 1_{A_1}(x) + P_t 1_{A_2}(y))$$

is continuous w.r.t. $(x,y) \in M^2$ (and even Lipschitz).

We now prove the second statement of Lemma 4.2. Fix $r_0 > 0$ and $x \in M$. Since the drift is uniformly bounded and the diffusion is uniformly elliptic, it is clear using local charts that

$$\mathbb{P}_x(X_t \in B(x, r_0/2),\ \forall t \in [0,2]) > 0.$$

In particular

$$\inf_{t\in(1,2]} \mathbb{E}_x(f_{x,r_0}(X_t)) > 0, \tag{4.7}$$

and thus $(x,x,x) \in \mathcal{E}_{r_0}$ for all $x \in M$.

Now, for all A_1 and A_2 measurable such that $A_1 \cup A_2 = M$, the map

$$(x,y,z) \mapsto P_t(f_{z,r_0} 1_{A_1})(x) + P_t(f_{z,r_0} 1_{A_2})(y)$$

is uniformly Lipschitz on M^3 by (4.6) and since $\|f_{z,r_0} - f_{z',r_0}\|_\infty \leq \rho(z,z')$. Therefore,

$$(x,y,z) \mapsto \inf_{t\in(1,2]} \mu^t_{x,y}(f_{z,r_0}) = \inf_{t\in(1,2]} \inf_{A_1\cup A_2 = M} (P_t(f_{z,r_0} 1_{A_1})(x) + P_t(f_{z,r_0} 1_{A_2})(y))$$

is Lipschitz on M^3. Hence, it follows from (4.7) that $\mathcal{E}_{r_0}$ contains an open ball. In particular, the interior $\mathrm{int}(\mathcal{E}_{r_0})$ is non-empty. Let us prove that $\mathrm{int}(\mathcal{E}_{r_0})$ is closed, which will imply by connexity that $\mathcal{E}_{r_0} \supset \mathrm{int}(\mathcal{E}_{r_0}) = M^3$.

Let (x_0, y_0, z_0) be an accumulation point of $\mathrm{int}(\mathcal{E}_{r_0})$. Fix $r > r_0$. Then there exist $(x, y, z) \in \mathcal{E}_{r_0}$ such that $\rho(z, z_0) \leq (r - r_0)/2$, and $\hat{r} \in (0, (r - r_0)/2)$ such that the closed ball centered at (x, y, z) and of radius $\hat{r}$ is included in $\mathcal{E}_{r_0}$. Moreover, since M is connected, the diffusion is uniformly elliptic and the drift is bounded, for all $s > 0$,

$$\mathbb{P}_{x_0}(X_s \in B(x, \hat{r})) =: c_s^{x_0} > 0 \quad \text{and} \quad \mathbb{P}_{y_0}(X_s \in B(y, \hat{r})) =: c_s^{y_0} > 0.$$

It then follows that, for all $t_0 \in (1, 2]$, defining $\eta = (t_0 - 1)/2 \in (0, \frac{1}{2}]$,

$$\begin{aligned}
\mu_{x_0,y_0}^{t_0}(f_{z_0,r}) &= \inf_{A_1 \cup A_2 = M} [\delta_{x_0} P_{1+2\eta}(f_{z_0,r} 1_{A_1}) + \delta_{y_0} P_{1+2\eta}(f_{z_0,r} 1_{A_2})] \\
&\geq c_\eta^{x_0} \wedge c_\eta^{y_0} \\
&\quad \times \inf_{A_1 \cup A_2 = M} \left(\inf_{\hat{x} \in B(x,\hat{r})} \delta_{\hat{x}} P_{1+\eta}(f_{z_0,r} 1_{A_1}) + \inf_{\hat{y} \in B(x,\hat{r})} \delta_{\hat{y}} P_{1+\eta}(f_{z_0,r} 1_{A_2}) \right) \\
&\geq c_\eta^{x_0} \wedge c_\eta^{y_0} \inf_{\hat{x} \in B(x,\hat{r}),\ \hat{y} \in B(y,\hat{r})} \mu_{\hat{x},\hat{y}}^{1+\eta}(f_{z_0,r}) \\
&\geq c_\eta^{x_0} \wedge c_\eta^{y_0} \inf_{\hat{x} \in B(x,\hat{r}),\ \hat{y} \in B(y,\hat{r})} \mu_{\hat{x},\hat{y}}^{1+\eta}(f_{z,(r+r_0)/2}),
\end{aligned}$$

where we used in the last line the inequality

$$f_{z_0,r}(x) = (r - \rho(x, z_0))_+ \geq (r - \rho(z, z_0) - \rho(x, z))_+ \geq f_{z_0, \frac{r+r_0}{2}}(x).$$

Since $(\hat{x}, \hat{y}, z) \in \mathcal{E}_{r_0}$ for all $\hat{x} \in \overline{B(x, \hat{r})}$ and $\hat{y} \in \overline{B(y, \hat{r})}$, and since $(x, y) \mapsto \mu_{x,y}^{1+\eta}(f_{z_0, \frac{r+r_0}{2}})$ is continuous and positive on $\mathcal{E}_{r_0}$ (since $\frac{r+r_0}{2} > r_0$), we deduce that $\mu_{x_0,y_0}^{t_0}(f_{z_0,r}) > 0$ for all $r > r_0$ and $t_0 \in (1, 2]$. and hence $(x_0, y_0, z_0) \in \mathcal{E}_{r_0}$. This concludes the proof of Lemma 4.2. □

Recall the constant ε_2 introduced in (4.1). Fix $s \geq t_1$ and $t \in (1, 2]$. We have, for all $x, z' \in E$,

$$\begin{aligned}
\delta_x P_{t+s} f &\geq \int_M 1_{z \in M_{\varepsilon_2}} P_t f(z)\, \delta_x P_s(dz) \\
&\geq \int_M 1_{z \in M_{\varepsilon_2}} \mu_{z,z'}^t(f)\, \delta_x P_s(dz).
\end{aligned}$$

Integrating both sides over $z' \in M_{\varepsilon_2}$ with respect to $\delta_y P_s(dz')/\delta_y P_s 1_M$, we deduce that

$$\delta_x P_{t+s} f \geq \int_{M \times M} 1_{z,z' \in M_{\varepsilon_2}} \mu_{z,z'}^t(f)\, \delta_x P_s(dz)\, \frac{\delta_y P_s(dz')}{\delta_y P_s 1_M}.$$

But $\delta_x P_{t+s} 1_M \leq \delta_x P_s 1_M$, hence

$$\frac{\delta_x P_{t+s} f}{\delta_x P_{t+s} 1_M} \geq \int_{M_{\varepsilon_2} \times M_{\varepsilon_2}} \mu^t_{z,z'}(f) \, \frac{\delta_x P_s(dz)}{\delta_x P_s 1_M} \, \frac{\delta_y P_s(dz')}{\delta_y P_s 1_M}.$$

Since $t \in (1, 2]$, it follows from Lemma 4.2 that

$$m_t = \inf_{z,z' \in M_{\varepsilon_2}} \mu^t_{z,z'}(M) > 0,$$

and therefore, since $s \geq t_1$,

$$\begin{aligned} m_t^{x,y} &:= \int_{M_{\varepsilon_2} \times M_{\varepsilon_2}} \mu^t_{z,z'}(M) \, \frac{\delta_x P_s(dz)}{\delta_x P_s 1_M} \, \frac{\delta_y P_s(dz')}{\delta_y P_s 1_M} \\ &\geq m_t \int_{M_{\varepsilon_2} \times M_{\varepsilon_2}} \frac{\delta_x P_s(dz)}{\delta_x P_s 1_M} \, \frac{\delta_y P_s(dz')}{\delta_y P_s 1_M} \\ &\geq A^2 m_t, \end{aligned} \tag{4.8}$$

by (4.1). Therefore, defining the probability measure

$$\nu^{s,t}_{x,y}(\cdot) = \frac{1}{m_t^{x,y}} \int_{M_{\varepsilon_2} \times M_{\varepsilon_2}} \mu^t_{z,z'}(\cdot) \, \frac{\delta_x P_s(dz)}{\delta_x P_s 1_M} \, \frac{\delta_y P_s(dz')}{\delta_y P_s 1_M}, \tag{4.9}$$

we obtain that

$$\frac{\delta_x P_{t+s} f}{\delta_x P_{t+s} 1_M} \geq m_t^{x,y} \nu^{s,t}_{x,y}(f) \geq A^2 m_t \nu^{s,t}_{x,y}(f).$$

Hence we have proved Condition (A1) for $t_0 = t_1 + 2$ (choosing $s = t_1$ and $t = 2$ and defining $\nu_{x,y} = \nu^{t_1,2}_{x,y}$).

4.3 Balance condition on the absorption probabilities

Our goal is now to prove Condition (A2). We first show that the first step of our proof (where we show (4.1)) implies the following stronger version of the gradient estimates (3.3), but only for $f = 1_M$.

Proposition 4.3. *There exists a constant $c' > 0$ such that, for all $t > 0$,*

$$\|\nabla P_t 1_M\|_\infty \leq \frac{c'}{\sqrt{t} \wedge 1} \|P_t 1_M\|_\infty. \tag{4.10}$$

Proof. This follows directly from the following computation: for all $t \geq t_0 = t_1 + 2$,

$$\begin{aligned} P_t 1_M(x) &\geq \mathbb{P}_x(X_{t-2} \in M_{\varepsilon_2}) \inf_{y \in M_{\varepsilon_2}} \mathbb{P}_y(2 < \tau_\partial) \\ &\geq A \mathbb{P}_x(t-2 < \tau_\partial)\, c_{\varepsilon_2}, \end{aligned}$$

by (4.1) and where $c_{\varepsilon_2} := \inf_{y \in M_{\varepsilon_2}} \mathbb{P}_y(2 < \tau_\partial)$. Hence,

$$\|P_t 1_M\|_\infty \geq c_{\varepsilon_2} A \|P_{t-2} 1_M\|_\infty.$$

Hence it follows from (3.3) that, for all $t > t_0$,

$$\begin{aligned} \|\nabla P_t 1_M\|_\infty &= \|\nabla P_2(P_{t-2} 1_M)\|_\infty \\ &\leq \frac{c}{\sqrt{2}} \|P_{t-2}(1_M)\|_\infty \\ &\leq \frac{c}{\sqrt{2} c_\varepsilon A} \|P_t 1_M\|_\infty. \end{aligned}$$

Since in addition $\|P_t 1_M\|_\infty \geq \|P_{t_0} 1_M\|_\infty$ for all $t \leq t_0$, we also deduce from (3.3) that

$$\|\nabla P_t 1_M\|_\infty \leq \frac{c}{\sqrt{t} \wedge 1} \leq \frac{c}{(\sqrt{t} \wedge 1) \|P_{t_0} 1_M\|_\infty} \|P_t 1_M\|_\infty. \qquad \square$$

This proposition implies that the function

$$x \mapsto \frac{P_t 1_M(x)}{\|P_t 1_M\|_\infty} \tag{4.11}$$

is c'-Lipschitz for all $t \geq 1$. Since this functions vanishes on ∂M and its maximum over M is 1, we deduce that, for all $t \geq 1$, the argmax of this function, denoted by z_t, exists in M. Moreover,

$$\frac{P_t 1_M(x)}{\|P_t 1_M\|_\infty} \geq \frac{f_{z_t, r_0}(x)}{r_0} \tag{4.12}$$

for all $x \in M$, with $r_0 := 1/c'$. In particular, this implies that there exists $\varepsilon' > 0$ such that $z_t \in M_{\varepsilon'}$ for all $t \geq 1$. Without loss of generality, we assume that $\varepsilon' \in (0, \varepsilon_2)$, where the constant ε_2 is taken from (4.1).

Therefore, for all $x, y \in M$ and for all $t \geq 1$,

$$\begin{aligned} \mathbb{P}_{\nu_{x,y}^{t_1,2}}(t < \tau_\partial) &\geq \frac{\|P_t 1_M\|_\infty}{r_0} \nu_{x,y}^{t_1,2}(f_{z_t, r_0}) \\ &= \frac{\|P_t 1_M\|_\infty}{r_0\, m_2^{x,y}} \int_{M_{\varepsilon'} \times M_{\varepsilon'}} \mu_{z,z'}^2(f_{z_t, r_0}) \frac{\delta_x P_{t_1}(dz)}{\delta_x P_{t_1} 1_M} \frac{\delta_y P_{t_1}(dz')}{\delta_y P_{t_1} 1_M}, \end{aligned}$$

where we used the definition (4.9) of the measure $\nu_{x,y}^{t_1,2}$. Now, it follows from the proof of Lemma 4.2 that $(x,y,z) \mapsto \mu_{x,y}^2(f_{z,r_0})$ is positive and continuous on M^3. Hence, using that $m_2^{x,y} \leq 1$ and defining $c'' := \inf_{(x,y,z)M_{\varepsilon'}^3} \mu_{x,y}^2(f_{z,r_0}) > 0$, we obtain

$$\begin{aligned}\mathbb{P}_{\nu_{x,y}^{t_1,2}}(t < \tau_\partial) &\geq \frac{c''\|P_t 1_M\|_\infty}{r_0} \int_{M_{\varepsilon_2} \times M_{\varepsilon_2}} \frac{\delta_x P_{t_1}(dz)}{\delta_x P_{t_1} 1_M} \frac{\delta_y P_{t_1}(dz')}{\delta_y P_{t_1} 1_M} \\ &\geq \frac{c'' A^2}{r_0} \|P_t 1_M\|_\infty,\end{aligned}$$

where the last inequality follows from (4.1).

This entails Condition (A2) for all $t \geq 1$. For $t \leq 1$, we make use of the following direct consequence of the proof Lemma 4.2:

$$c''' := \inf_{(x,y) \in M_\varepsilon^2} \mu_{x,y}^2(M_\varepsilon) > 0.$$

We deduce from (4.9) that $\nu_{x,y}^{t_1,2}(M_{\varepsilon_2}) \geq c'''/m_2^{x,y}$ and hence, for all $t \leq 1$,

$$\begin{aligned}\mathbb{P}_{\nu_{x,y}^{t_1,2}}(t < \tau_\partial) &\geq \frac{c'''}{m_2^{x,y}} \inf_{z \in M_{\varepsilon_2}} \mathbb{P}_z(t < \tau_\partial) \geq \frac{c'''}{m_2^{x,y}} \inf_{z \in M_{\varepsilon_2}} \mathbb{P}_z(1 < \tau_\partial) \\ &\geq \frac{c'''}{m_2^{x,y}} \inf_{z \in M_{\varepsilon_2}} \mathbb{P}_z(1 < \tau_\partial) \sup_{z \in M} \mathbb{P}_z(t < \tau_\partial) > 0.\end{aligned}$$

This ends the proof of (A2) and hence of (1.1).

4.4 Contraction in total variation norm

Our aim is now to conclude the proof of Theorem 3.1 by proving (3.4). In order to complete this step, we make use of (2.6). Hence, we need to prove that for all probability measure μ on M,

$$c(\mu) = \inf_{t \geq 0} \frac{\mathbb{P}_\mu(t < \tau_\partial)}{\mathbb{P}_{z_t}(t < \tau_\partial)} \geq c\mu(\rho_{\partial M}) \tag{4.13}$$

for some constant $c > 0$, where we recall that z_t is defined as the argmax of (4.11).

Enlarging t_0 and reducing c_1 and c_2, one can assume without loss of generality that $\nu = \nu_{x,y}$ does not depend on x, y in (A) (this follows from the equivalence between (A), (A') and (A") proved in [4, Theorem 2.1]).

Then, using (4.12) and (A1), we deduce that for all $t \geq t_0 + 1$,

$$\begin{aligned}
\mathbb{P}_\mu(t < \tau_\partial) &= \mu(P_{t_0} P_{t-t_0} 1_M) \\
&\geq \frac{\|P_{t-t_0} 1_M\|_\infty}{r_0} \mu(P_{t_0} f_{z_{t-t_0}, r_0}) \\
&\geq \frac{\|P_t 1_M\|_\infty}{r_0} \mu(P_{t_0} f_{z_{t-t_0}, r_0}) \\
&\geq \frac{\mathbb{P}_{z_t}(t < \tau_\partial)}{r_0} \mu(P_{t_0} f_{z_{t-t_0}, r_0}) \\
&\geq c_1 \frac{\mathbb{P}_{z_t}(t < \tau_\partial) \nu(f_{z_{t-t_0}, r_0})}{r_0} \mathbb{P}_\mu(t_0 < \tau_\partial). \qquad (4.14)
\end{aligned}$$

Now, Lemma 4.1 implies the existence of a constants $C'(t_0), \varepsilon_1(t_0) > 0$ such that, for all $x \in M \setminus M_{\varepsilon_1(t_0)}$,

$$\mathbb{P}_x(t_0 \wedge T_{\varepsilon_1(t_0)} < \tau_\partial) \geq \mathbb{P}_x(T_{\varepsilon_1(t_0)} < t_0 < \tau_\partial) \geq C'(t_0) \rho_{\partial M}(x).$$

Now, by Markov's property,

$$\mathbb{P}_x(t_0 < \tau_\partial) \geq \mathbb{P}_x(t_0 \wedge T_{\varepsilon_1(t_0)} < \tau_\partial) \inf_{y \in \partial M_{\varepsilon_1(t_0)}} \mathbb{P}_y(t_0 < \tau_\partial).$$

Since $\inf_{y \in \overline{M_{\varepsilon_2(t_0)}}} \mathbb{P}_y(t_0 < \tau_\partial) > 0$, we deduce that there exists a constant $c > 0$ such that, for all $x \in M$,

$$\mathbb{P}_x(t_0 < \tau_\partial) \geq c \rho_{\partial M}(x).$$

Integrating the last inequality with respect to μ, we deduce from (4.14) that, for all $t \geq t_0 + 1$,

$$\mathbb{P}_\mu(t < \tau_\partial) \geq c \frac{\nu(f_{z_{t-t_0}, r_0})}{r_0} \mu(\rho_{\partial M}) \mathbb{P}_{z_t}(t < \tau_\partial).$$

Since we have when $t < t_0 + 1$

$$\begin{aligned}
\mathbb{P}_\mu(t < \tau_\partial) \geq \mathbb{P}_\mu(t_0 + 1 < \tau_\partial) &\geq c \frac{\mathbb{P}_{z_{t_0+1}}(t_0 + 1 < \tau_\partial) \nu(f_{z_1, r_0})}{r_0} \mu(\rho_{\partial M}) \\
&\geq c \frac{\mathbb{P}_{z_{t_0+1}}(t_0 + 1 < \tau_\partial) \nu(f_{z_1, r_0})}{r_0} \mu(\rho_{\partial M}) \mathbb{P}_{z_t}(t < \tau_\partial),
\end{aligned}$$

the proof of (4.13) is completed. This ends the proof of Theorem 3.1.

References

[1] K. Bogdan, T. Grzywny, and M. Ryznar. Heat kernel estimates for the fractional Laplacian with Dirichlet conditions. *Ann. Probab.*, 38(5):1901–1923, 2010.

[2] P. Cattiaux, P. Collet, A. Lambert, S. Martínez, S. Méléard, and J. San Martín. Quasi-stationary distributions and diffusion models in population dynamics. *Ann. Probab.*, 37(5):1926–1969, 2009.

[3] P. Cattiaux and S. Méléard. Competitive or weak cooperative stochastic lotka-volterra systems conditioned to non-extinction. *J. Math. Biology*, 60(6):797–829, 2010.

[4] N. Champagnat and D. Villemonais. Exponential convergence to quasi-stationary distribution and Q-process. *Probability Theory and Related Fields*, pages 1–41, 2015. Online first.

[5] N. Champagnat and D. Villemonais. Exponential convergence to quasi-stationary distribution for one-dimensional diffusions. Arxiv preprint 1506.02385, 2015.

[6] Z.-Q. Chen, P. Kim, and R. Song. Heat kernel estimates for $\Delta + \Delta^{\alpha/2}$ in $C^{1,1}$ open sets. *J. Lond. Math. Soc. (2)*, 84(1):58–80, 2011.

[7] Z.-Q. Chen, P. Kim, and R. Song. Dirichlet heat kernel estimates for fractional Laplacian with gradient perturbation. *Ann. Probab.*, 40(6):2483–2538, 2012.

[8] Z.-Q. Chen, P. Kim, and R. Song. Stability of Dirichlet heat kernel estimates for non-local operators under Feynman-Kac perturbation. *Trans. Amer. Math. Soc.*, 367(7):5237–5270, 2015.

[9] E. B. Davies and B. Simon. Ultracontractivity and the heat kernel for Schrödinger operators and Dirichlet Laplacians. *J. Funct. Anal.*, 59(2):335–395, 1984.

[10] P. Del Moral and D. Villemonais. Exponential mixing properties for time inhomogeneous diffusion processes with killing. *ArXiv e-prints*, Dec. 2015.

[11] O. Kallenberg. *Foundations of modern probability.* Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.

[12] K.-Y. Kim and P. Kim. Two-sided estimates for the transition densities of symmetric Markov processes dominated by stable-like processes in $C^{1,\eta}$ open sets. *Stochastic Process. Appl.*, 124(9):3055–3083, 2014.

[13] P. Kim and R. Song. Estimates on Green functions and Schrödinger-type equations for non-symmetric diffusions with measure-valued drifts. *J. Math. Anal. Appl.*, 332(1):57–80, 2007.

[14] R. Knobloch and L. Partzsch. Uniform conditional ergodicity and intrinsic ultracontractivity. *Potential Analysis*, 33:107–136, 2010.

[15] M. Kolb and A. Wübker. Spectral analysis of diffusions with jump boundary. *J. Funct. Anal.*, 261(7):1992–2012, 2011.

[16] J. Lierl and L. Saloff-Coste. The Dirichlet heat kernel in inner uniform domains: local results, compact domains and non-symmetric forms. *J. Funct. Anal.*, 266(7):4189–4235, 2014.

[17] T. Lindvall and L. C. G. Rogers. Coupling of multidimensional diffusions by reflection. *Ann. Probab.*, 14(3):860–872, 1986.

[18] J. Littin C. Uniqueness of quasistationary distributions and discrete spectra when ∞ is an entrance boundary and 0 is singular. *J. Appl. Probab.*, 49(3):719–730, 2012.

[19] S. Méléard and D. Villemonais. Quasi-stationary distributions and population processes. *Probability Surveys*, 2012. To appear.

[20] Y. Miura. Ultracontractivity for Markov semigroups and quasi-stationary distributions. *Stoch. Anal. Appl.*, 32(4):591–601, 2014.

[21] E. Priola and F.-Y. Wang. Gradient estimates for diffusion semigroups with singular coefficients. *J. Funct. Anal.*, 236(1):244–264, 2006.

[22] F.-Y. Wang. Gradient estimates of dirichlet heat semigroups and application to isoperimetric inequalities. *Ann. Probab.*, 32(1A):424–440, 01 2004.

[23] Q. S. Zhang. The boundary behavior of heat kernels of Dirichlet Laplacians. *J. Differential Equations*, 182(2):416–430, 2002.