




# Transcritical Bifurcation for the Conditional Distribution of a Diffusion Process

Michel Benaïm<sup>1</sup> · Nicolas Champagnat<sup>2</sup> · William Oçafrain<sup>2</sup> · Denis Villemonais<sup>2</sup> 

Received: 20 December 2021 / Revised: 31 October 2022 / Accepted: 14 November 2022 /  
Published online: 28 November 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

## Abstract

In this article, we describe a simple class of models of absorbed diffusion processes with parameter, whose conditional law exhibits a transcritical bifurcation. Our proofs are based on the description of the set of quasi-stationary distributions for general two-clusters reducible processes.

**Keywords** Absorbed Markov processes · Quasi-stationary distributions · Exponential mixing · Stochastic differential equations · Bifurcation

**Mathematics Subject Classification (2020)** 60J25 · 60J60 · 60B10 · 37A25

## 1 Model, Motivation and Main Result

Let  $D = (0, 5)$  and consider Lipschitz functions  $\varphi^1, \varphi^2, \psi^1, \psi^2 : D \rightarrow [0, 1]$  such that  $\psi^1 + \psi^2 = 1$  and (see Fig. 1)

---

✉ Denis Villemonais  
denis.villemonais@univ-lorraine.fr

Michel Benaïm  
michel.benaïm@unine.ch

Nicolas Champagnat  
Nicolas.Champagnat@inria.fr

William Oçafrain  
william.oçafrain@univ-lorraine.fr

<sup>1</sup> Institut de Mathématiques, Université de Neuchâtel, Neuchâtel, Switzerland

<sup>2</sup> Université de Lorraine, CNRS, Inria, IECL, 54000 Nancy, France

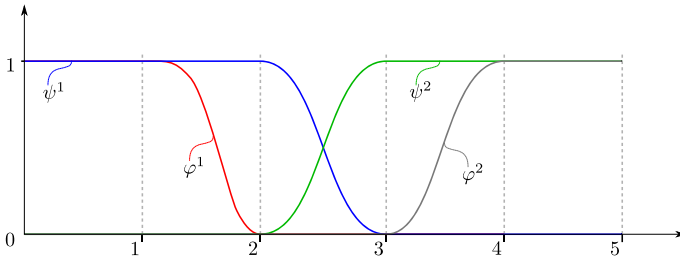


Fig. 1 Representation of the graphs of  $\varphi^1, \varphi^2, \psi^1$  and  $\psi^2$

$$\begin{cases} \varphi^1_{|(0,1]} \equiv 1, \varphi^1_{|[1,2]} \leq 1, & \varphi^1_{|[2,5]} \equiv 0, \\ \varphi^2_{|(0,3]} \equiv 0, \varphi^2_{|[3,4]} \leq 1, & \varphi^2_{|[4,5]} \equiv 1, \\ \psi^1_{|(0,2]} \equiv 1, 0 < \psi^1_{|(2,3)} \leq 1, & \psi^1_{|[3,5]} \equiv 0, \\ \psi^2_{|(0,2]} \equiv 0, 0 < \psi^2_{|(2,3)} \leq 1, & \psi^2_{|[3,5]} \equiv 1. \end{cases}$$

For all  $\alpha > 0$ , we consider the absorbed diffusion process  $X^\alpha$  evolving according to the Itô’s stochastic differential equation

$$dX_t^\alpha = \left( \varphi^1(X_t^\alpha) + \sqrt{\alpha} \varphi^2(X_t^\alpha) \right) dB_t + (\psi^1(X_t) + \alpha \psi^2(X_t)) dt, \tag{1}$$

stopped when it reaches  $\partial D = \{0, 5\}$  and where  $B$  is a standard one dimensional Brownian motion. This defines a sub-Markov semi-group  $(P_t^\alpha)_{t \in \mathbb{R}_+}$  on  $D$  by

$$\delta_x P_t^\alpha f = \mathbb{E}_x(f(X_t^\alpha) 1_{X_t^\alpha \in D}), \quad \forall f \in L^\infty(D)$$

and a semi-flow  $(\Phi_t^\alpha)_{t \in \mathbb{R}_+}$  on the set  $\mathcal{P}(D)$  of probability measures on  $D$  by

$$\Phi_t^\alpha(\mu) = \mathbb{P}_\mu(X_t \in \cdot \mid X_t \in D).$$

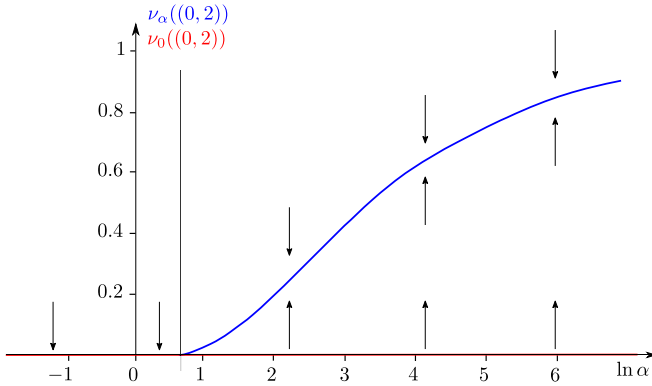
Observe that the diffusion coefficient in (1) vanishes on the interval  $[2, 3]$ . Since in addition the drift coefficient is positive on  $[2, 3]$  for any  $\alpha > 0$ , the set  $[3, 5]$  is absorbing and  $\mathbb{P}_2(\forall t \geq 0, X_t \geq 2, \exists t \geq 0, X_t = 3) = 1$ .

Hence, the family of diffusion processes  $(X^\alpha)_{\alpha \geq 0}$  has some similarities with the family of discrete-time Markov chains  $(X^{a,b})_{a,b \in (0,1)}$ , where  $X^{a,b}$  is defined on  $\{1, 2, \partial\}$ , absorbed at  $\partial$ , with transition submatrix on  $\{1, 2\}$  given by

$$\begin{pmatrix} a & 1-a \\ 0 & b \end{pmatrix}.$$

It is pointed in [1, Example 3.5] that the probability measures  $\nu_2 := \delta_2$  and  $\nu := \frac{a-b}{1-b} \delta_1 + \frac{1-a}{1-b} \delta_2$  (when  $a > b$ ) are such that

- If  $a > b$ ,  $\lim_{n \rightarrow \infty} \mathbb{P}_\mu(X_n^{a,b} = i \mid X_n^{a,b} \neq \partial) = \nu(\{i\})$  for all  $i = 1, 2$  and  $\mu \neq \delta_2$ .



**Fig. 2** Domain of attraction of the attractors, with quadratic  $\varphi^1, \varphi^2, \psi^1, \psi^2$ , namely:  $\varphi_1(x) = x(2 - x)$  for all  $x \in (1, 2)$ ,  $\varphi_2(x) = (x - 3)^2$  for all  $x \in (3, 4)$ ,  $\psi_1(x) = (3 - x)(x - 1)$  for all  $x \in (2, 3)$ ,  $\psi_2(x) = 1 - \psi_1(x)$  for all  $x \in (2, 3)$ .  $\nu_\alpha((0, 2))$  has been computed numerically using Fleming–Viot type approximation techniques (see for instance [2])

- Otherwise,  $\lim_{n \rightarrow \infty} \mathbb{P}_\mu(X_n^{a,b} = i \mid X_n^{a,b} \neq \partial) = \nu_2(\{i\})$  for all  $i = 1, 2$  and  $\mu \in \mathcal{P}(\{1, 2\})$ .

Our goal in this note is to prove a similar property for the family  $(X^\alpha)_{\alpha \geq 0}$ , which can be formulated as a bifurcation of the dynamical system on  $\mathcal{P}(D)$  generated by  $\Phi^\alpha$ . Our motivation is to quantify precisely the speed of convergence and the basin of attraction of the fixed points of this dynamical systems (see Fig. 2). Another motivation is to provide an illustration to the fact that, for an absorbed diffusion process satisfying the weak Hörmander condition and regularity properties at the absorbing set, uniqueness of a quasi-stationary distribution does not necessary hold true unless some accessibility properties are satisfied (see Theorem 1.8 of [3]). The definition of a quasi-stationary distribution is recalled in Sect. 2.

To state our main result, we define the absorption parameters of  $X^\alpha$  for  $\alpha = 1$ :

$$\lambda_1 := \inf \left\{ \lambda \in \mathbb{R}, \liminf_{t \rightarrow +\infty} e^{\lambda t} \mathbb{P}_{x_1}(X_t^1 \in (0, 2)) > 0 \right\}, \text{ for some } x_1 \in (0, 2).$$

and

$$\lambda_2 := \inf \left\{ \lambda \in \mathbb{R}, \liminf_{t \rightarrow +\infty} e^{\lambda t} \mathbb{P}_{x_2}(X_t^1 \in [3, 5]) > 0 \right\}, \text{ for some } x_2 \in [3, 5],$$

As expected, we will see that the parameters  $\lambda_1$  and  $\lambda_2$  are positive and do not depend on  $x_1$  nor  $x_2$ .

**Theorem 1.1** *The dynamical system generated by  $\Phi^\alpha$  parametrized by  $\alpha > 0$  admits a transcritical bifurcation at  $\lambda_1/\lambda_2$ . More precisely, there exist a family of probability measure  $(\mu_\alpha)_{\alpha > \lambda_1/\lambda_2}$  on  $(0, 5)$  and a probability measure  $\mu_0$  on  $[3, 5)$  such that:*

- for  $\alpha \leq \lambda_1/\lambda_2$ ,  $\mu_0$  is a global attractor for  $\Phi^\alpha$  for the total variation distance,

- for  $\alpha > \lambda_1/\lambda_2$ ,  $\mu_0$  is a saddle point whose stable manifold for the total variation distance is the set  $W^s(\mu_0) := \{\mu \in \mathcal{P}((0, 5)) : \mu((0, 2)) = 0\}$ , and  $\mu_\alpha$  is a stable point whose basin of attraction for the total variation distance is  $W^u(\mu_0) = \{\mu \in \mathcal{P}((0, 5)) : \mu((0, 2)) > 0\}$ .

Our method also provides an estimate for the speed of convergence of the dynamical system generated by  $\Phi^\alpha$  to its limit fixed point. One can check from the proof that it is exponential when  $\alpha \neq \lambda_1/\lambda_2$  and polynomial in  $O(1/t)$  when  $\alpha = \lambda_1/\lambda_2$ .

Let us now give heuristic arguments explaining why the system admits a bifurcation at  $\lambda_1/\lambda_2$ . Informally,  $\alpha$  is a time-scale parameter for the process  $X^\alpha$  restricted to  $[3, 5)$ , so that it has the same law as  $(X_{\alpha t}^1)_{t \geq 0}$ . In particular, the absorption parameter of  $X^\alpha$  restricted to  $[3, 5)$  is  $\alpha\lambda_2$ , while its absorption parameter restricted to  $(0, 2)$  (and hence  $(0, 3)$ ) is  $\lambda_1$ . When  $\alpha\lambda_2$  is smaller than  $\lambda_1$ , the process  $X^\alpha$  escapes  $(0, 3)$  at a higher pace than it exits  $[3, 5)$ , so that the equilibrium distribution is concentrated on  $[3, 5)$ . When  $\alpha\lambda_2$  is strictly larger than  $\lambda_1$ , then the process escapes  $(0, 3)$  at a lower pace than it exits  $[3, 5)$ , so that the equilibrium distribution puts mass on  $(0, 3)$  (and on  $[3, 5)$  since the process can travel from  $(0, 3)$  to  $[3, 5)$ , while the converse is not true). This heuristic explanation also entails that larger values of  $\alpha$  lead to higher values of  $\nu_\alpha((0, 2))$ , as illustrated by Fig. 2.

The proof of this result relies on the theory of quasi-stationary distributions. We show in particular that the process  $X^\alpha$  admits either one or two quasi-stationary distributions, depending on the value of  $\alpha$ , which correspond to  $\mu_0$  and  $\mu_\alpha$  (when  $\alpha > \lambda_0/\lambda_1$ ). A central feature allowing this property is that  $X^\alpha$  is reducible. Indeed, it is proved in [3] that irreducibility for such diffusions entails the uniqueness of a quasi-stationary distribution. In order to prove Theorem 1.1, we start with considerations on quasi-stationary distributions for Markov processes in reducible state spaces. Since they apply to general Markov processes and may have independent interest, they are stated independently in Sect. 2. We conclude the proof of the theorem in Sect. 3.

## 2 Quasi-Stationarity for Two-Clusters Reducible Processes

Let  $X$  be a Markov process with state space  $M = D \cup \{\partial\}$  with  $\partial \notin D$ , in discrete or continuous time, such that  $M$  admits a measurable partition  $D_1 \cup D_2 \cup \{\partial\}$ . We assume that  $\{\partial\}$  and  $D_2 \cup \{\partial\}$  are absorbing sets (see Fig. 3), which means that

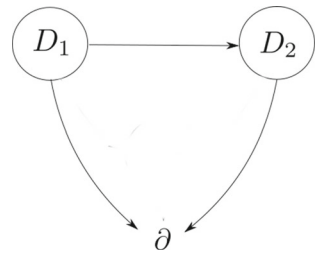
$$\mathbb{P}_\partial(X_t = \partial) = 1 \quad \text{and} \quad \mathbb{P}_x(X_t \in D_2 \cup \{\partial\}) = 1, \quad \forall x \in D_2, \forall t \geq 0.$$

A probability measure  $\nu$  is said to be a *quasi-stationary distribution* if, for all  $t \geq 0$ ,

$$\mathbb{P}_\nu(X_t \in \cdot \mid X_t \neq \partial) = \nu(\cdot). \quad (2)$$

It is well known (see [4, Proposition 1]) that the notion of quasi-stationary distribution is equivalent to the one of *quasi-limiting distribution*, defined as a probability measure  $\nu$  such that there exist some initial distributions  $\mu$  such that, for all measurable subset

**Fig. 3** Transition graph displaying the relation between the sets  $D_1$ ,  $D_2$  and  $\partial$



$A \subset D$ ,

$$\lim_{t \rightarrow \infty} \mathbb{P}_\mu(X_t \in A \mid X_t \neq \partial) = \nu(A).$$

It is also well known that, to any quasi-stationary distribution  $\nu$  is associated the so-called *exponential absorption rate*  $\lambda_0 > 0$  such that  $\mathbb{P}_\nu(X_t \in \cdot, X_t \neq \partial) = e^{-\lambda_0 t} \nu$ . We refer the reader to [4–6] for a general overview on the theory of quasi-stationarity, and to [2, 7–19] for the study of the quasi-stationary distribution of diffusion processes.

The aim of this section is to provide conditions on  $X$  allowing to obtain the existence of a quasi-stationary distribution  $\nu$  for the process  $X$ , as well as the so-called *Malthusian behavior* (see [20] for the terminology), that is to say the existence of a positive function  $\eta$  on  $D$  such that

$$\lim_{t \rightarrow \infty} e^{\lambda_0 t} \mathbb{P}_x(X_t \in \cdot, X_t \neq \partial) = \eta(x) \nu(\cdot),$$

where  $\lambda_0$  is the exponential absorption rate associated to  $\nu$  (in particular, this convergence entails the convergence of the conditional probability measure  $\mathbb{P}_x(X_t \in \cdot \mid X_t \neq \partial)$  towards  $\nu$ ). In what follows, we will present three different sets of assumptions, each discussed in three different subsections, and each entailing different Malthusian behavior for the process  $X$ .

### 2.1 Exponential Convergence on $D_2$ and Faster Exit from $D_1$

Let us introduce our first set of assumptions.

**Assumption QSD2** a. There exist a positive function  $\eta_2$  on  $D_2$ , a probability measure  $\nu_2$  on  $D_2$ , a constant  $\lambda_2 > 0$ , and positive constants  $C_2, \gamma_2 > 0$  such that

$$\|e^{\lambda_2 t} \mathbb{P}_x(X_t \in \cdot, X_t \in D_2) - \eta_2(x) \nu_2(\cdot)\|_{TV} \leq C_2 e^{-\gamma_2 t}, \quad \forall x \in D_2, t \geq 0. \tag{3}$$

b. In addition,  $\sup_{x \in D_1} e^{\lambda_2 t} \mathbb{P}_x(X_t \in D_1) \leq f(t)$ , where  $f$  is non-increasing and  $L^1$  on  $\mathbb{R}_+$ .

The assumption a. refers to the Malthusian behavior, as described before, of the restriction of  $X$  on the subset  $D_2$ , holding uniformly in  $x$  in total variation and expo-

nentially fast. We refer the reader to [3, 17, 21–28] for general criteria entailing such behavior.

Also, remark that the inequality (3) entails that  $\eta_2$  is bounded (take for example  $t = 0$ ). In addition, it implies that  $\eta_2(x) = \lim_{t \rightarrow +\infty} e^{\lambda_2 t} \mathbb{P}_x(X_t \neq \partial)$  for all  $x \in D_2$ . In addition, as noticed above, it implies that  $\nu_2$  is a quasi-stationary distribution for  $X$ .

Then Assumption QSD2 entails the following result on quasi-stationarity.

**Theorem 2.1** *Under QSD2, there exists  $\eta : D \rightarrow \mathbb{R}_+$  positive on  $D_2$  such that*

$$\sup_{x \in D} \left\| e^{\lambda_2 t} \mathbb{P}_x(X_t \in \cdot, X_t \neq \partial) - \eta(x) \nu_2(\cdot \cap D_2) \right\|_{TV} \xrightarrow{t \rightarrow +\infty} 0. \tag{4}$$

*If in addition  $\mathbb{P}_x(\exists n \geq 0, X_n \in D_2) > 0$  for all  $x \in D_1$ , then  $\eta$  is positive on  $D$  and  $\nu_2(\cdot \cap D_2)$  is the unique quasi-stationary distribution for  $X$  on  $D$ .*

In other terms, the Malthusian behavior of  $X$ , only assumed for  $x \in D_2$  in Assumption QSD2, holds for all  $x \in D$  and uniformly in  $x$  in total variation. Speed of convergence for (4) is discussed after the proof of Theorem 2.1.

**Proof of Theorem 2.1** For the rest, we first prove this theorem in the discrete time setting, and then consider the continuous-time setting.

*Step 1: Proof in the discrete-time setting* We define the stopping time  $\tau_1^c := \min\{n \geq 0, X_n \notin D_1\}$ . For all  $x \in D$  and all measurable set  $A \subset D_1 \cup D_2$ , for all  $n \in \mathbb{Z}_+$ , we have, using the strong Markov property at time  $\tau_1^c$ ,

$$\begin{aligned} \mathbb{P}_x(X_n \in A) &= \mathbb{P}_x(X_n \in A \cap D_1) + \mathbb{P}_x(X_n \in A \cap D_2) \\ &= \mathbb{P}_x(X_n \in A \cap D_1) + \sum_{k=0}^n \mathbb{E}_x \left( 1_{\tau_1^c=k} \mathbb{P}_{X_k}(X_{n-k} \in A \cap D_2) \right). \end{aligned} \tag{5}$$

If  $x \in D_2$ , then the result is an immediate consequence of (3) with  $\eta(x) = \eta_2(x)$ . It only remains to consider the case  $x \in D_1$ . On the one hand, we have by assumption

$$e^{\lambda_2 n} \mathbb{P}_x(X_n \in A \cap D_1) \leq e^{\lambda_2 n} \mathbb{P}_x(X_n \in D_1) \xrightarrow{n \rightarrow +\infty} 0. \tag{6}$$

On the other hand, using (3) and extending  $\eta_2$  to  $\{\partial\}$  by  $\eta_2(\partial) = 0$ , we obtain, for all  $k \geq 0, t \geq 0$  and measurable set  $A \subset D_2$ ,

$$\begin{aligned} &\left| \mathbb{E}_x \left( 1_{\tau_1^c=k} \mathbb{P}_{X_k}(X_{n-k} \in A) \right) - \mathbb{E}_x \left( 1_{\tau_1^c=k} \eta_2(X_k) \nu_2(A) e^{-\lambda_2(n-k)} \right) \right| \\ &\leq \mathbb{E}_x \left( 1_{\tau_1^c=k} C_2 e^{-(\lambda_2+\gamma_2)(n-k)} 1_{X_k \in D_2} \right) \\ &\leq \begin{cases} C_2 \mathbb{P}_x(X_{k-1} \in D_1, X_k \in D_2) e^{-(\lambda_2+\gamma_2)(n-k)} & \text{if } k \geq 1, \\ 0 & \text{if } k = 0. \end{cases} \end{aligned}$$

Summing over  $k$ , we deduce that, for all  $n \geq 0$  and all measurable set  $A \subset D_2$ ,

$$\begin{aligned} & \left| \sum_{k=0}^n \mathbb{E}_x \left( 1_{\tau_1^c=k} \mathbb{P}_{X_k} (X_{n-k} \in A) \right) - \sum_{k=0}^n \mathbb{E}_x \left( 1_{\tau_1^c=k} \eta_2(X_k) v_2(A) e^{-\lambda_2(n-k)} \right) \right| \\ & \leq C_2 e^{-\lambda_2 n} e^{-\gamma_2 n} \sum_{k=1}^n e^{\lambda_2 k} e^{\gamma_2 k} \mathbb{P}_x (X_{k-1} \in D_1, X_k \in D_2) \\ & \leq C_2 e^{-\lambda_2(n-1)} e^{-\gamma_2 n} \sum_{k=1}^n e^{\gamma_2 k} f(k-1). \end{aligned} \tag{7}$$

Moreover

$$\begin{aligned} \sum_{k \geq n+1} \mathbb{E}_x \left( 1_{\tau_1^c=k} \eta_2(X_k) e^{\lambda_2 k} \right) & \leq \|\eta_2\|_\infty \sum_{k \geq n+1} e^{\lambda_2 k} \mathbb{P}_x (X_{k-1} \in D_1, X_k \in D_2) \\ & \leq \|\eta_2\|_\infty \sum_{k \geq n+1} e^{\lambda_2 k} (\mathbb{P}_x (X_{k-1} \in D_1) - \mathbb{P}_x (X_k \in D_1)) \\ & \leq \|\eta_2\|_\infty (e^{\lambda_2} - 1) \sum_{k \geq n} e^{\lambda_2 k} \mathbb{P}_x (X_k \in D_1) \\ & \leq \|\eta_2\|_\infty (e^{\lambda_2} - 1) \int_{n-1}^{+\infty} f(t) dt. \end{aligned}$$

This, (5), (6) and (7) entail that, for all  $x \in D_1$  and all measurable  $A \subset D$ ,

$$\sup_{x \in D_1} \left| e^{\lambda_2 n} \mathbb{P}_x (X_n \in A) - \sum_{k=0}^{+\infty} \mathbb{E}_x \left( 1_{\tau_1^c=k} \eta_2(X_k) e^{\lambda_2 k} \right) v_2(A \cap D_2) \right| \xrightarrow{n \rightarrow +\infty} 0.$$

Setting

$$\eta(x) := \begin{cases} \sum_{k=0}^{+\infty} \mathbb{E}_x \left( 1_{\tau_1^c=k} \eta_2(X_k) e^{\lambda_2 k} \right) & \text{if } x \in D_1 \\ \eta_2(x) & \text{if } x \in D_2, \end{cases}$$

this concludes the proof of (4).

If in addition  $\mathbb{P}_x(\exists n \geq 0, X_n \in D_2) > 0$  for all  $x \in D_1$ , then  $\eta$  is positive on  $D$  and hence  $v_2(\cdot \cap D_2)$  is the unique quasi-limiting distribution of the process and hence its unique quasi-stationary distribution.

*Step 2: Proof in the continuous-time setting* The proof is done by applying the discrete time result to the Markov chain  $(X_n)_{n \in \mathbb{Z}_+}$ .

Let  $X$  satisfy Assumption QSD2. Then the discrete time process  $(X_n)_{n \in \mathbb{N}}$  also satisfies Assumption QSD2 and hence, by Theorem 2.1 in the discrete time setting, we have

$$\sup_{x \in D} \left\| e^{\lambda_2 n} \mathbb{P}_x (X_n \in \cdot, X_n \neq \partial) - \eta(x) v_2(\cdot \cap D_2) \right\|_{TV} \xrightarrow{n \rightarrow +\infty} 0. \tag{8}$$

For any  $x \in D$  and  $h > 0$ , integrating the above convergence result with respect to  $\mathbb{P}_x(X_h \in \cdot)$  entails that

$$\|e^{\lambda 2n} \mathbb{P}_x(X_{n+h} \in \cdot, X_{n+h} \neq \partial) - \mathbb{E}_x(\eta(X_h))v_2(\cdot \cap D_2)\|_{TV} \xrightarrow{n \rightarrow +\infty} 0$$

Similarly, for any  $x \in D$  and  $h > 0$ , applying (8) to the test function  $y \in D \mapsto \mathbb{P}_y(X_h \in \cdot, X_h \neq \partial)$ , gives

$$\|e^{\lambda 2n} \mathbb{P}_x(X_{n+h} \in \cdot, X_{n+h} \neq \partial) - \eta(x)\mathbb{P}_{v_2}(X_h \in \cdot, X_h \neq \partial)\|_{TV} \xrightarrow{n \rightarrow +\infty} 0.$$

But  $D_2 \cup \{\partial\}$  being absorbing, we have  $\mathbb{P}_{v_2}(X_h \in \cdot, X_h \neq \partial) = \mathbb{P}_{v_2}(X_h \in \cdot, X_h \in D_2)$ , which implies that

$$\mathbb{E}_x(\eta(X_h))v_2(\cdot \cap D_2) = \eta(x)\mathbb{P}_{v_2}(X_h \in \cdot \cap D_2) = e^{-\lambda 2h}\eta(x)v_2(\cdot \cap D_2),$$

so that  $\mathbb{E}_x(\eta(X_h)) = e^{-\lambda 2h}\eta(x)$ , and hence

$$\|e^{\lambda 2n} \mathbb{P}_x(X_{n+h} \in \cdot, X_{n+h} \neq \partial) - e^{-\lambda 2h}\eta(x)v_2(\cdot \cap D_2)\|_{TV} \xrightarrow{n \rightarrow +\infty} 0$$

Since the convergence holds uniformly in  $h \in [0, 1]$ , this concludes the proof of (4) in the continuous time setting. The uniqueness of the quasi-stationary distribution is immediate, since any quasi-stationary distribution for  $(X_t)_{t \in \mathbb{R}_+}$  is also a quasi-stationary distribution for  $(X_n)_{n \in \mathbb{Z}_+}$ .  $\square$

Let us do some remarks before passing to the second set of assumptions.

**Remark 2.1** In the proof, the speed of convergence in the above theorem is explicit in terms of the quantities appearing in the assumptions. In particular, if Assumption QSD2 holds true with

$$e^{(\lambda_2 + \varepsilon)t} \sup_{x \in D_1} \mathbb{P}_x(X_t \in D_1) \xrightarrow{t \rightarrow +\infty} 0,$$

for some  $\varepsilon > 0$ , then the convergence in (4) is exponential.

**Remark 2.2** Non-uniform speed of convergence can also be proved under weaker form of Assumption QSD2. For instance if (3) holds true and if, for some  $x \in D_1$ ,  $e^{\lambda 2t} \mathbb{P}_x(X_t \in D_1) \leq f(t)$  with  $f$  non-increasing and  $L^1$  on  $\mathbb{R}_+$  (but non-uniformly in  $x \in D_1$ ), then,

$$\|e^{\lambda 2n} \mathbb{P}_x(X_n \in \cdot) - \eta(x)v_2(\cdot \cap D_2)\|_{TV} \xrightarrow{n \rightarrow +\infty} 0.$$

**Remark 2.3** The use of the absorbed Markov process setting is for convenience only: the above result applies more generally to semi-groups on  $L_\infty(D)$  admitting an isolated simple leading eigenvalue  $\lambda \in \mathbb{C}$ .

### 2.2 Exponential Convergence on $D_1$ and Faster Absorbtion in $D_2$

Let us now state our second set of assumptions.

**Assumption QSD1** There exist a positive function  $\eta_1$  on  $D_1$ , a probability measure  $\nu_1$  on  $D_1$ , a constant  $\lambda_1 > 0$ , and positive constants  $C_1, \gamma_1 > 0$  such that

$$\|e^{\lambda_1 t} \mathbb{P}_x(X_t \in \cdot, X_t \in D_1) - \eta_1(x) \nu_1(\cdot)\|_{TV} \leq C_1 e^{-\gamma_1 t}, \quad \forall x \in D_1, t \geq 0. \tag{9}$$

In addition, for all  $t \geq 0$ ,

$$e^{\lambda_1 t} \sup_{x \in D_2} \mathbb{P}_x(X_t \in D_2) \leq f(t) \tag{10}$$

where  $f$  is non-decreasing and  $L^1$  on  $\mathbb{R}_+$ .

Assumption QSD1 is very similar to Assumption QSD2, except that this assumption now deals with the quasi-stationarity for the restriction of the process  $X$  considered as absorbed by  $D_2 \cup \{\partial\}$ . Similarly to Assumption QSD2, Assumption QSD1 entails that  $\eta_1$  is bounded,  $\eta_1(x) = \lim_{t \rightarrow \infty} e^{\lambda_1 t} \mathbb{P}_x(X_t \in D_1)$  for all  $x \in D_1$ , and that  $\nu_1$  is a quasi-stationary distribution for the process  $X$  considered as absorbed by  $D_2 \cup \{\partial\}$ . The condition (10) tells that  $\sup_{x \in D_2} \mathbb{P}_x(X_t \neq \partial)$  vanishes faster than the probability of survival starting from the quasi-stationary distribution  $\nu_1$ .

**Theorem 2.2** *Under Assumption QSD1, there exists a positive finite measure  $\nu$  on  $D$  such that*

$$\sup_{x \in D} \|e^{\lambda_1 t} \mathbb{P}_x(X_t \in \cdot, X_t \neq \partial) - \eta(x) \nu\|_{TV} \xrightarrow{t \rightarrow +\infty} 0, \tag{11}$$

where  $\eta(x) = \eta_1(x)$  for all  $x \in D_1$  and  $\eta(x) = 0$  for all  $x \in D_2$ . In addition,  $\nu/\nu(D)$  is the unique quasi-stationary distribution of  $X$  such that  $\nu(D_1)/\nu(D) > 0$ .

Hence, (11) entails that  $\mathbb{P}_x(X_t \in \cdot | X_t \neq \partial)$  converges in total variation towards  $\nu/\nu(D)$ . Note also that Assumption QSD1 does not tell anything on the convergence of  $\mathbb{P}_x(X_t \in \cdot | X_t \neq \partial)$  when  $x \in D_2$ .

**Remark 2.4** Similarly as in Remark 2.1, if

$$e^{(\lambda_1 + \epsilon)t} \sup_{x \in D_2} P_x(X_k \in D_2) \rightarrow 0$$

for some  $\epsilon > 0$ , then the convergence in (11) is exponential.

**Remark 2.5** Similarly as in Remark 2.2, if (9) holds true and if

$$\sum_{k=0}^{\infty} e^{\lambda_1 k} \sup_{x \in D_2} \mathbb{P}_x(X_k \in A) < +\infty.$$

for some measurable  $A \subset D$ , then

$$\sup_{x \in D_1} \left| e^{\lambda_1 n} \mathbb{P}_x(X_n \in A) - \eta_1(x) \nu(A) \right| \xrightarrow{n \rightarrow +\infty} 0.$$

This last case is particularly interesting, since it applies to situations where  $\nu$  is not necessarily a finite measure (one may have  $\nu(D_2) = +\infty$ ).

**Proof of Theorem 2.2** We only prove the result in the discrete-time setting. The adaptation to the continuous-time setting follows from the same argument as in the proof of Theorem 2.1.

Fix  $x \in D_1$ . The result is an immediate consequence of (9) if  $A \subset D_1$  with  $\nu(\cdot \cap D_1) = \nu_1$ . It remains to consider the case  $A \subset D_2$ , the general case being obtained by linearity. For all measurable  $A \subset D_2$  and all  $x \in D_1$ , we have, for all  $k \in \{0, \dots, n - 1\}$ ,

$$\begin{aligned} & \left| \mathbb{E}_x \left( 1_{\tau_1^c = n-k} \mathbb{P}_{X_{n-k}}(X_k \in A) \right) - \eta_1(x) e^{-\lambda_1(n-k-1)} \mathbb{E}_{\nu_1} \left( 1_{\tau_1^c = 1} 1_{X_{k+1} \in A} \right) \right| \\ &= \left| \mathbb{E}_x \left( f_A(X_{n-k-1}) 1_{n-k-1 < \tau_1^c} \right) - \eta_1(x) e^{-\lambda_1(n-k-1)} \nu_1(f_A) \right| \end{aligned}$$

where, for all  $y \in D_1$ ,  $f_A(y) = \mathbb{P}_y(X_1 \notin D_1, X_{k+1} \in A)$ , and where  $\tau_1^c$  was defined in the proof of Theorem 2.1. By Markov’s property, we have  $f_A(y) \leq \sup_{z \in D_2} \mathbb{P}_z(X_k \in A)$ . Hence, according to (9), we have

$$\begin{aligned} & \left| \mathbb{E}_x \left( 1_{\tau_1^c = n-k} \mathbb{P}_{X_{n-k}}(X_k \in A) \right) - \eta_1(x) e^{-\lambda_1(n-k-1)} \mathbb{E}_{\nu_1} \left( 1_{\tau_1^c = 1} 1_{X_{k+1} \in A} \right) \right| \\ & \leq C_1 e^{-\lambda_1 n} e^{-\gamma_1(n-k-1)} \left( e^{\lambda_1(k+1)} \sup_{z \in D_2} \mathbb{P}_z(X_k \in A) \right) \\ & \leq C_1 e^{-\lambda_1 n} e^{-\gamma_1(n-k-1)} e^{\lambda_1} f(k). \end{aligned}$$

Summing over  $k \in \{0, \dots, n - 1\}$  and multiplying by  $e^{\lambda_1 n}$ , we get

$$\begin{aligned} & \left| e^{\lambda_1 n} \mathbb{P}_x(X_n \in A) - \eta_1(x) \sum_{k=0}^{n-1} e^{\lambda_1(k+1)} \mathbb{E}_{\nu_1} \left( 1_{\tau_1^c = 1} 1_{X_{k+1} \in A} \right) \right| \\ & \leq C_1 e^{-\gamma_1 n} \sum_{k=0}^{n-1} e^{\gamma_1(k+1)} e^{\lambda_1(k+1)} \sup_{z \in D_2} \mathbb{P}_z(X_k \in A) \\ & \leq C_1 e^{\lambda_1} \sum_{k=0}^{n-1} e^{-\gamma_1(n-k-1)} f(k), \end{aligned}$$

where the right hand term goes to 0 when  $n \rightarrow +\infty$ . Finally, we observe that

$$\begin{aligned} \sum_{k=n}^{+\infty} e^{\lambda_1(k+1)} \mathbb{E}_{\nu_1} \left( \mathbf{1}_{\tau_1^c=1} \mathbf{1}_{X_{k+1} \in A} \right) &\leq \sum_{k=n}^{+\infty} e^{\lambda_1(k+1)} \sup_{z \in D_2} \mathbb{P}_x(X_k \in A) \\ &\leq e^{\lambda_1} \int_{n-1}^{+\infty} f(t) dt, \end{aligned}$$

which also goes to 0 when  $n \rightarrow +\infty$  under the assumption (10). This concludes the proof of (11) with

$$\nu(A) := \nu_2(A \cap D_1) + \sum_{k=0}^{+\infty} e^{\lambda_1(k+1)} \mathbb{E}_{\nu_1} \left( \mathbf{1}_{\tau_1^c=1} \mathbf{1}_{X_{k+1} \in A \cap D_2} \right)$$

In particular,  $\nu/\nu(D)$  is a quasi-limiting distribution of  $X$  and is thus a quasi-stationary distribution.

To conclude, let  $\nu'$  be a quasi-stationary distribution for  $X$  such that  $\nu'(D_1) > 0$ . Integrating (11) with respect to  $\nu'$  and noting that  $\nu'(\eta_1) > 0$ , we deduce that the exponential absorption rate of  $\nu'$  is  $\lambda_1$  and that  $\nu/\nu(D)$  is a quasi-limiting distribution for  $X$  starting from  $\nu'$ , and thus  $\nu = \nu'$ . □

### 2.3 Exponential Convergence in $D_1$ and $D_2$ with the Same Rate

Let us now present our last set of assumptions.

**Assumption QSD1–2** There exist two positive functions  $\eta_1$  on  $D_1$  and  $\eta_2$  on  $D_2$ , two probability measures  $\nu_1$  on  $D_1$  and  $\nu_2$  on  $D_2$ , a positive constant  $\lambda_0 > 0$ , and positive constants  $C_1, \gamma_1, C_2, \gamma_2 > 0$  such that (9) and (3) hold true with  $\lambda_1 = \lambda_2 = \lambda_0$ .

In other terms,  $\nu_2$  and  $\nu_1$  are the quasi-stationary distributions for  $X$ , respectively, started from  $D_2$  and absorbed at  $\partial$ , and started from  $D_1$  and absorbed at  $D_2 \cup \{\partial\}$ , associated to the same absorption rate  $\lambda_0 > 0$ . Under this assumption, we have the following result.

**Theorem 2.3** *Under Assumption QSD1–2, the process admits  $\nu_2(\cdot \cap D_2)$  as unique quasi-stationary distribution and*

$$\sup_{x \in D_1} \left\| \frac{e^{\lambda_0 t}}{t} \mathbb{P}_x(X_t \in \cdot, X_t \in D_2) - \eta(x) \nu_2(\cdot \cap D_2) \right\|_{TV} \leq \frac{C}{t+1}, \tag{12}$$

where  $\eta$  is a positive function on  $D_1$  and  $C$  is a positive constant.

In particular, Malthusian behavior (3) does not hold for  $x \in D_1$ . However, (12) still entails that the probability measure  $\mathbb{P}_x(X_t \in \cdot \mid X_t \neq \partial)$  converges in total variation towards  $\nu_2$  for all  $x \in D_1$  (and also for all  $x \in D_2$  by Hypothesis (3)).

**Proof of Theorem 2.3** As for the proof of Theorem 2.1, we only deal with the discrete-time setting.

Fix  $x \in D_1$  and measurable set  $A \subset D_2$ . For all  $k \geq 0$ , we have, using (3) (recall that  $\lambda_1 = \lambda_2 = \lambda_0$ ),

$$\left| e^{\lambda_0(k+1)} \mathbb{P}_{v_1}(\tau_1^c = 1, X_{k+1} \in A) - e^{\lambda_0} \mathbb{E}_{v_1} \left( 1_{\tau_1^c=1} \eta_2(X_1) \right) v_2(A) \right| \leq e^{\lambda_0} C_2 e^{-\gamma_2 k}, \quad (13)$$

where  $\tau_1^c$  was defined in the proof of Theorem 2.1. Moreover, for all  $n \geq 1$  and  $k \in \{0, \dots, n-1\}$ ,

$$\begin{aligned} & \left| e^{\lambda_0(k+1)} \eta_1(x) \mathbb{P}_{v_1}(\tau_1^c = 1, X_{k+1} \in A) - e^{\lambda_0 n} \mathbb{E}_x \left( 1_{\tau_1^c=n-k} 1_{X_n \in A} \right) \right| \\ &= e^{\lambda_0(k+1)} \left| \eta_1(x) v_1(f_A) - e^{\lambda_0(n-k-1)} \mathbb{E}_x (f_A(X_{n-k-1})) \right|, \end{aligned}$$

where  $f_A(y) = \mathbb{P}_y(\tau_1^c = 1, X_{k+1} \in A) \leq \sup_{z \in D_2} \mathbb{P}_z(X_k \in D_2)$ . Hence using (9), we deduce that

$$\begin{aligned} & \left| e^{\lambda_0(k+1)} \eta_1(x) \mathbb{P}_{v_1}(\tau_1^c = 1, X_{k+1} \in A) - e^{\lambda_0 n} \mathbb{E}_x \left( 1_{\tau_1^c=n-k} 1_{X_n \in A} \right) \right| \\ & \leq e^{\lambda_0(k+1)} C_1 e^{-\gamma_1(n-k-1)} \sup_{z \in D_2} \mathbb{P}_z(X_k \in D_2) \\ & = C_1 e^{\lambda_0} e^{-\gamma_1(n-k-1)} \sup_{z \in D_2} e^{\lambda_0 k} \mathbb{P}_z(X_k \in D_2), \end{aligned}$$

where, according to (3),  $\sup_{z \in D_2} e^{\lambda_0 k} \mathbb{P}_z(X_k \in D_2)$  is uniformly bounded in  $k$  by  $C_2 + \|\eta_2\|_\infty$ . This, (13) and summing over  $k \in \{0, \dots, n-1\}$  imply that

$$\begin{aligned} & \left| e^{\lambda_0 n} \mathbb{P}_x(X_n \in A) - n e^{\lambda_0} \eta_1(x) \mathbb{E}_{v_1} \left( 1_{\tau_1^c=1} \eta_2(X_1) \right) v_2(A) \right| \\ & \leq C_2 \eta_1(x) e^{\lambda_0} \sum_{k=0}^{\infty} e^{-\gamma_2 k} + \sum_{k=0}^{+\infty} e^{-\gamma_1 k} C_1 e^{\lambda_0} (C_2 + \|\eta_2\|_\infty). \end{aligned}$$

This concludes the proof of (12) with  $\eta(x) := e^{\lambda_0} \eta_1(x) \mathbb{E}_{v_1} \left( 1_{\tau_1^c=1} \eta_2(X_1) \right)$ .  $\square$

### 3 Proof of Theorem 1.1

We start with a proposition related to the theory of quasi-stationary distributions for diffusion processes.

**Proposition 3.1** *There exist a positive function  $\eta_2 : [3, 5) \rightarrow (0, +\infty)$ , a probability measure  $v_2$  on  $[3, 5)$  and positive constants  $C_2, \lambda_2, \gamma_2 > 0$  such that, for all  $\alpha > 0$*

and for all probability measure  $\mu$  on  $[3, 5)$ ,

$$\|e^{\alpha\lambda_2 t} \mathbb{P}_\mu (X_t^\alpha \in \cdot \cap [3, 5)) - \mu(\eta_2)v_2(\cdot)\|_{TV} \leq C_2 e^{-\alpha\gamma_2 t}, \quad \forall t \geq 0. \tag{14}$$

There exist a positive function  $\eta_1 : (0, 2) \rightarrow (0, +\infty)$ , a probability measure  $v_1$  on  $(0, 2)$  and positive constants  $C_1, \lambda_1, \gamma_1 > 0$  such that, for all probability measure  $\mu$  on  $(0, 2)$  and all  $\alpha > 0$ ,

$$\|e^{\lambda_1 t} \mathbb{P}_\mu (X_t^\alpha \in \cdot \cap (0, 2)) - \mu(\eta_1)v_1(\cdot)\|_{TV} \leq C_1 e^{-\gamma_1 t}, \quad \forall t \geq 0. \tag{15}$$

**Proof of Proposition 3.1** For any  $\alpha > 0$ , on the event  $X_0^\alpha \in [3, 5)$ , the process remains almost surely in  $[3, 5)$  until it reaches  $\partial D$  at the end point 5. It is known (this can be proved for instance using Section 4.5 of [14]) that, considering the process  $X^\alpha$  restricted to  $[3, 5)$  absorbed when it reaches 5, there exists a probability measure  $\mu_\alpha$  on  $[3, 5)$  (the quasi-stationary distribution of  $X^\alpha$  restricted to  $[3, 5)$ ), a positive function  $\zeta_\alpha : [3, 5) \rightarrow (0, +\infty)$ , and positive constants  $c_\alpha, \delta_\alpha, \delta'_\alpha > 0$  such that, for all probability measure  $\mu$  on  $[3, 5)$ ,

$$\|e^{\delta_\alpha t} \mathbb{P}_\mu (X_t^\alpha \in \cdot) - \mu(\zeta_\alpha)\mu_\alpha(\cdot)\|_{TV} \leq c_\alpha e^{-\delta'_\alpha t}, \quad \forall t \geq 0. \tag{16}$$

Also, since  $\text{Law}((X_t^\alpha)_{t \geq 0}) = \text{Law}((X_{\alpha t}^1)_{t \geq 0})$ , we deduce that  $\zeta_\alpha = \zeta_1, \mu_\alpha = \mu_1$  and  $\delta_\alpha = \alpha\delta_1$  for all  $\alpha > 0$ . Moreover, one can take  $c_\alpha = c_1$  and  $\delta'_\alpha = \alpha\delta'_1$ . Setting  $\eta_2 = \zeta_2, v_2 = \mu_1, C_2 = c_1, \lambda_2 = \delta_1$  and  $\gamma_2 = \delta'_1$ , this proves (14).

Similarly, the law of the process  $X^\alpha$  restricted to  $(0, 2)$  and absorbed when it reaches  $\{0, 2\}$  does not depend on  $\alpha$ , and there exists a probability measure  $v_1$  on  $(0, 2)$  (the quasi-stationary distribution of the process  $X^\alpha$  conditioned to remain in  $(0, 2)$ ), a positive function  $\eta_1 : (0, 2) \rightarrow (0, +\infty)$  and positive constants  $C_1, \lambda_1, \gamma_1 > 0$  such that, for all probability measure  $\mu$  on  $(0, 2)$  and all  $\alpha > 0$ ,

$$\|e^{\lambda_1 t} \mathbb{P}_\mu (X_t^\alpha \in \cdot, t < T_0 \wedge T_2) - \mu(\eta_1)\mu_1(\cdot)\|_{TV} \leq C_1 e^{-\gamma_1 t}, \quad \forall t \geq 0, \tag{17}$$

where  $T_a$  denotes the first hitting time of  $\{a\}$  (see again Section 4.5 in [14]). Since the process cannot enter  $(0, 2)$  after time  $T_0 \wedge T_2$ , we deduce that  $\mathbb{P}_\mu (X_t^\alpha \in \cdot, t < T_0 \wedge T_2) = \mathbb{P}_\mu (X_t^\alpha \in \cdot \cap (0, 2))$ .

This concludes the proof of Proposition 3.1. □

We consider now the behavior of the process with initial position in  $[2, 3)$ . For any  $\alpha > 0$  and  $x \in [2, 3)$ , denote by  $(f_t^\alpha(x))_{t \geq 0}$  the solution of the ODE  $\partial f_t^\alpha(x)/\partial t = \psi^1(f_t^\alpha(x)) + \alpha\psi^2(f_t^\alpha(x))$  and  $f_0^\alpha(x) = x$ . For all  $x \in [2, 3)$ , denote by  $t_3(x) = \inf\{t \geq 0, f_t^\alpha(x) = 3\}$ , then

$$\mathbb{P}_x(X_t^\alpha \in A) = \begin{cases} \mathbf{1}_{f_t^\alpha(x) \in A} & \text{if } f_t^\alpha(x) < 3, \\ \mathbb{P}_3(X_{t-t_3(x)}^\alpha \in A) & \text{if } f_t^\alpha(x) \geq 3, \end{cases}$$

where we used the strong Markov property at time  $T_3$  and the fact that  $X^\alpha$  is deterministic with drift equal to  $\psi^1 + \alpha\psi^2$  on  $[2, 3]$  (so that  $T_3 = t_3(x)$   $\mathbb{P}_x$ -almost surely). In particular, we deduce from Proposition 3.1 that, for all  $x \in [2, 3)$  and  $t \geq t_3(x)$ ,

$$\left\| e^{\alpha\lambda_2(t-t_3(x))} \mathbb{P}_x (X_t^\alpha \in \cdot) - \eta_2(3)v_2(\cdot) \right\|_{TV} \leq C_2 e^{-\alpha\gamma_2(t-t_3(x))}$$

and hence

$$\left\| e^{\alpha\lambda_2 t} \mathbb{P}_x (X_t^\alpha \in \cdot) - \eta_\alpha(x)v_2(\cdot) \right\|_{TV} \leq e^{\alpha(\lambda_2+\gamma_2)t_3(x)} C_2 e^{-\alpha\gamma_2 t} \leq C'_2 e^{-\alpha\gamma_2 t},$$

where  $\eta_\alpha(x) := e^{\alpha\lambda_2 t_3(x)} \eta_2(3)$  and  $C'_2 = e^{\alpha(\lambda_2+\gamma_2)t_3(2)} C_2$ . By integration of the last inequality and by (14), we thus proved that, for any initial distribution in  $[2, 5)$ ,

$$\left\| e^{\alpha\lambda_2 t} \mathbb{P}_\mu (X_t^\alpha \in \cdot) - \mu(\eta_\alpha)v_2(\cdot) \right\|_{TV} \leq C'_2 e^{-\alpha\gamma_2 t}, \quad \forall t \geq 0, \tag{18}$$

where

$$\eta_\alpha(x) = \begin{cases} e^{\alpha\lambda_2 t_3(x)} \eta_2(3) & \text{if } x \in [2, 3] \\ \eta_2(x) & \text{if } x \in [3, 5). \end{cases}$$

We are now in a position to apply the results of Sect. 2, with  $D_1 = (0, 2)$  and  $D_2 = [2, 5)$ .

**Case  $\alpha < \lambda_1/\lambda_2$ .**

In this case we observe that Assumption QSD2 is satisfied (of course with  $\alpha\lambda_2$  instead of  $\lambda_2$ ). Indeed, on the one hand (3) is immediate from (18), while, for all  $x \in D_1$ ,

$$\sup_{x \in D_1} e^{\alpha\lambda_2 t} \mathbb{P}_x (X_t \in D_1) \leq e^{(\alpha\lambda_2 - \lambda_1)t} (C_1 + \|\eta_1\|_\infty) \xrightarrow{t \rightarrow +\infty} 0$$

and is  $L^1(\mathbb{R}_+)$ , where we used (15). Moreover, we have  $\mathbb{P}_x(X_1 \in D_2) > 0$  for all  $x \in D_1$  and hence, according to Theorem 2.1,  $v_2(\cdot \cap D_2)$  is the unique quasi-stationary distribution  $\nu$  for the process  $X^\alpha$  absorbed at  $\{0, 5\}$ , and (4) implies that, for all probability measure  $\mu$  on  $D$ ,

$$\Phi_t^\alpha(\mu) \xrightarrow{t \rightarrow +\infty} v_2(\cdot \cap D_2).$$

**Case  $\alpha = \lambda_1/\lambda_2$ .**

In this case, we observe that Assumption QSD1–2 is satisfied. To deduce the convergence of  $\Phi_t^\alpha(\mu)$ , we need to distinguish two cases: either  $\mu(D_1) = 0$  and then the fact that

$$\Phi_t^\alpha(\mu) \xrightarrow{t \rightarrow +\infty} v_2(\cdot \cap D_2).$$

follows from (3), or  $\mu(D_1) > 0$  and then it follows from (9), (3) and (12) that, on the one hand,

$$\begin{aligned} \frac{e^{\lambda_0 t}}{t} \mathbb{P}_\mu(X_t \neq \partial) &= \frac{e^{\lambda_0 t}}{t} \mathbb{P}_\mu(X_t \in D_1) + \frac{e^{\lambda_0 t}}{t} \mathbb{P}_\mu(X_t \in D_2, X_t \neq \partial) \\ &\xrightarrow{t \rightarrow +\infty} \mu(\eta) > 0 \end{aligned}$$

and on the other hand, that for all  $A \subset D_1 \cup D_2$  measurable,

$$\begin{aligned} \frac{e^{\lambda_0 t}}{t} \mathbb{P}_\mu(X_t \in A) &= \frac{e^{\lambda_0 t}}{t} \mathbb{P}_\mu(X_t \in A \cap D_1) + \frac{e^{\lambda_0 t}}{t} \mathbb{P}_\mu(X_t \in A \cap D_2) \\ &\xrightarrow{t \rightarrow +\infty} \mu(\eta) \nu_2(A \cap D_2), \end{aligned}$$

where the convergence is uniform with respect to  $A$ . Hence, in all cases,

$$\Phi_t^\alpha(\mu) \xrightarrow{TV, t \rightarrow +\infty} \nu_2(\cdot \cap D_2).$$

**Case  $\alpha > \lambda_1/\lambda_2$ .**

In this case, we observe that Assumption QSD1 is satisfied. Indeed, on the one hand (9) holds true, while, for all  $x \in D_2$ ,

$$\sum_{k \in \mathbb{Z}_+} \sup_{x \in D_2} e^{\lambda_1 k} \mathbb{P}_x(X_k \in D_2) \leq \sum_{k \in \mathbb{Z}_+} e^{(\lambda_1 - \alpha \lambda_2)k} (C'_2 + \|\eta_\alpha\|_\infty) < +\infty,$$

where we used (18) and  $\alpha \lambda_2 > \lambda_1$ . Then, (11) in Theorem 2.2 implies that there exists a probability measure  $\nu$  on  $D$  such that  $\nu(D_1) > 0$  and such that, for all probability measure  $\mu$  on  $D$  such that  $\mu(D_1) > 0$ ,

$$\Phi_t^\alpha(\mu) \xrightarrow{TV, t \rightarrow +\infty} \nu.$$

In addition, (18) entails that, for all probability measure  $\mu$  on  $D$  such that  $\mu(D_1) = 0$ ,

$$\Phi_t^\alpha(\mu) \xrightarrow{TV, t \rightarrow +\infty} \nu_2(\cdot \cap D_2).$$

This concludes the proof of Theorem 1.1.

**Data Availability** This manuscript has no associated data. Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

**Declarations**

**Conflict of interest** The authors have no relevant financial or non-financial interests to disclose.

## References

1. Benaïm, M., Cloez, B., Panloup, F.: Stochastic approximation of quasi-stationary distributions on compact spaces and applications. *Ann. Appl. Probab.* **28**(4), 2370–2416 (2018)
2. Del Moral, P., Villemonais, D.: Exponential mixing properties for time inhomogeneous diffusion processes with killing. *Bernoulli* **24**(2), 1010–1032 (2018)
3. Benaïm, M., Champagnat, N., Oçafrain, W., Villemonais, D.: Degenerate processes killed at the boundary of a domain. arXiv e-prints [arXiv:2103.08534](https://arxiv.org/abs/2103.08534) [math.PR] (2021)
4. Méléard, S., Villemonais, D.: Quasi-stationary distributions and population processes. *Probab. Surv.* **9**, 340–410 (2012)
5. Collet, P., Martínez, S., San Martín, J.: Markov chains, diffusions and dynamical systems. In: *Quasi-stationary Distributions. Probability and its Applications (New York)*, p. 280. Springer, Berlin (2013). <https://doi.org/10.1007/978-3-642-33131-2>
6. van Doorn, E.A., Pollett, P.K.: Quasi-stationary distributions for discrete-state models. *Eur. J. Oper. Res.* **230**(1), 1–14 (2013). <https://doi.org/10.1016/j.ejor.2013.01.032>
7. Pinsky, R.G.: On the convergence of diffusion processes conditioned to remain in a bounded region for large time to limiting positive recurrent diffusion processes. *Ann. Probab.* **13**(2), 363–378 (1985)
8. Gong, G.L., Qian, M.P., Zhao, Z.X.: Killed diffusions and their conditioning. *Probab. Theory Related Fields* **80**(1), 151–167 (1988)
9. Cattiaux, P., Collet, P., Lambert, A., Martínez, S., Méléard, S., San Martín, J.: Quasi-stationary distributions and diffusion models in population dynamics. *Ann. Probab.* **37**(5), 1926–1969 (2009). <https://doi.org/10.1214/09-AOP451>
10. Cattiaux, P., Méléard, S.: Competitive or weak cooperative stochastic Lotka–Volterra systems conditioned to non-extinction. *J. Math. Biol.* **60**(6), 797–829 (2010). <https://doi.org/10.1007/s00285-009-0285-4>
11. Knobloch, R., Partzsch, L.: Uniform conditional ergodicity and intrinsic ultracontractivity. *Potential Anal.* **33**, 107–136 (2010). <https://doi.org/10.1007/s11118-009-9161-5>
12. Kolb, M., Steinsaltz, D.: Quasilimiting behavior for one-dimensional diffusions with killing. *Ann. Probab.* **40**(1), 162–212 (2012). <https://doi.org/10.1214/10-AOP623>
13. Littin C., J.: Uniqueness of quasistationary distributions and discrete spectra when  $\infty$  is an entrance boundary and 0 is singular. *J. Appl. Probab.* **49**(3), 719–730 (2012). <https://doi.org/10.1239/jap/1346955329>
14. Champagnat, N., Villemonais, D.: General criteria for the study of quasi-stationarity. arXiv e-prints [arXiv:1712.08092](https://arxiv.org/abs/1712.08092) [math.PR] (2017)
15. Champagnat, N., Coulibaly-Pasquier, K.A., Villemonais, D.: Criteria for exponential convergence to quasi-stationary distributions and applications to multi-dimensional diffusions, pp. 165–182. Springer, Cham (2018). [https://doi.org/10.1007/978-3-319-92420-5\\_5](https://doi.org/10.1007/978-3-319-92420-5_5)
16. Hening, A., Kolb, M.: Quasistationary distributions for one-dimensional diffusions with singular boundary points. *Stoch. Process. Appl.* **129**(5), 1659–1696 (2019). <https://doi.org/10.1016/j.spa.2018.05.012>
17. Guillin, A., Nectoux, B., Wu, L.: Quasi-stationary distribution for strongly Feller Markov processes by Lyapunov functions and applications to hypoelliptic Hamiltonian systems. Working paper or preprint (2020). <https://hal.archives-ouvertes.fr/hal-03068461>
18. Lelièvre, T., Ramil, M., Reygner, J.: Quasi-stationary distribution for the langevin process in cylindrical domains, part I: existence, uniqueness and long-time convergence. arXiv preprint [arXiv:2101.11999](https://arxiv.org/abs/2101.11999) (2021)
19. Ramil, M.: Quasi-stationary distribution for the Langevin process in cylindrical domains, part II: overdamped limit. arXiv preprint [arXiv:2103.00338](https://arxiv.org/abs/2103.00338) (2021)
20. Bertoin, J., Watson, A.R.: The strong Malthusian behavior of growth-fragmentation processes. *Annales Henri Lebesgue* **3**, 795–823 (2020)
21. Del Moral, P., Miclo, L.: On the stability of nonlinear Feynman–Kac semigroups. *Ann. Fac. Sci. Toulouse Math.* (6) **11**(2), 135–175 (2002)
22. Del Moral, P.: Genealogical and interacting particle systems with applications. In: *Feynman–Kac Formulae. Probability and its Applications (New York)*, p. 555. Springer (2004). <https://doi.org/10.1007/978-1-4684-9393-1>
23. Champagnat, N., Villemonais, D.: Uniform convergence to the  $Q$ -process. *Electron. Commun. Probab.* **22**, 7 (2017). <https://doi.org/10.1214/17-ECP63>

24. Ferré, G., Rousset, M., Stoltz, G.: More on the long time stability of Feynman–Kac semigroups. arXiv e-prints [arXiv:1807.00390](https://arxiv.org/abs/1807.00390) [math.PR] (2018)
25. Velleret, A.: Unique quasi-stationary distribution, with a possibly stabilizing extinction. arXiv e-prints [arXiv:1802.02409](https://arxiv.org/abs/1802.02409) [math.PR] (2018)
26. Bansaye, V., Cloez, B., Gabriel, P., Marguet, A.: A non-conservative Harris’ ergodic theorem. arXiv preprint [arXiv:1903.03946](https://arxiv.org/abs/1903.03946) (2019)
27. Champagnat, N., Villemonais, D.: Practical criteria for r-positive recurrence of unbounded semigroups. arXiv preprint [arXiv:1904.08619](https://arxiv.org/abs/1904.08619) (2019)
28. Champagnat, N., Villemonais, D.: Lyapunov criteria for uniform convergence of conditional distributions of absorbed Markov processes. *Stoch. Process. Appl.* **135**, 51–74 (2021)

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.