

Unique quasi-stationary distribution, with a possibly stabilizing extinction

Aurélien Velleret

▶ **To cite this version:**

Aurélien Velleret. Unique quasi-stationary distribution, with a possibly stabilizing extinction. Stochastic Processes and their Applications, 2022, 148, pp.98-138. 10.1016/j.spa.2022.02.004 . hal-03781505

HAL Id: hal-03781505
https://univ-eiffel.hal.science/hal-03781505

Submitted on 20 Sep 2022

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.

Unique quasi-stationary distribution, with a possibly stabilizing extinction

Aurélien Velleret

Institut für Mathematik, Goethe Universität, Fachbereich 12, 60054 Frankfurt am Main, Germany,
email: velleret@math.uni-frankfurt.de

arXiv:1802.02409v6 [math.PR] 6 Sep 2021

Abstract

We establish sufficient conditions for exponential convergence to a unique quasi-stationary distribution in the total variation norm. These conditions also ensure the existence and exponential ergodicity of the Q-process, the process conditionned upon never being absorbed. The technique relies on a coupling procedure that is related to Harris recurrence (for Markov Chains). It applies to general continuous-time and continuous-space Markov processes. The main novelty is that we modulate each coupling step depending both on a final horizon of time (for survival) and on the initial distribution. By this way, we could notably include in the convergence a dependency on the initial condition. As an illustration, we consider a continuous-time birth-death process with catastrophes and a diffusion process describing a (localized) population adapting to its environment.

Keywords: quasi-stationary distribution, survival capacity, Q-process, Harris recurrence, birth-and-death process, diffusion

1. Introduction

1.1. Presentation

Given a continuous-time and continuous-space Markov process with an absorbing state, we are interested in this work in the long time behavior of the process conditionally on not being absorbed (not being "extinct").

More precisely, our first concern is on the marginal —at time t— conditioned on not being extinct —also at time t— (the **MCNE** in short). We wish to highlight key conditions on the process such that these MCNE converge as $t \to \infty$ to a unique distribution α. This limiting distribution is called the **quasi-stationary distribution** (the **QSD**) —cf Subsections 1.3 and 2.2, or chapter 2 in [19] for more details on this notion. The techniques we use allow us to establish not only the existence and uniqueness of the QSD, but also the exponential convergence in total variation norm, cf Theorem 2.1.

In addition, we deduce, under the same conditions, the existence of a specific eigenfunction h of the infinitesimal generator, with the same eigenvalue as the QSD. As time goes to infinity, the renormalizing factor at time t behaves asymptotically as $h \exp[-\lambda t]$, cf (2.4) and Theorem 2.2. This convergence motivates the name **survival capacity** that we give to h (sometimes described as the "reproductive value" in ecological models). Again, the

convergence is exponential, but not uniform over the state space in our case. Moreover, we deduce the existence of the **Q-process**. Its marginal at time t is given by the limit (as $T \to \infty$) of the marginal of the original process at time t conditioned on not being extinct at time T, cf Theorem 2.3. Thus, it is often described as the process conditioned to never be absorbed. Finally, we deduce for the Q-process the existence and uniqueness of its stationary distribution β together with a property related to exponential ergodicity.

To deduce these results, our aim is to combine a large degree of generality with conditions as easy to verify as possible. A specificity of our approach is that it allows to deduce a coupling procedure depending on the initial condition that ensures a contraction in total variation towards the limiting distribution. It is only for commodity that we have restricted the analysis to cases where there is a unique QSD. One can find in [48] an application to group selection models where our procedure of proof is exploited to deduce the convergence to some QSD in a specific basin of attraction. Also, the proof can be adapted quasi verbatim to discrete time processes.

We exploit the idea, first exploited in [13], to rely on a more constructive method in the form of a strong regeneration condition, analogous to Harris' recurrence (what we can see maybe a bit more clearly in the present work). At the foundation of our proof is clearly the characterization given in [13] of the uniform exponential convergence to a unique QSD. As we can see in the applications we present (cf Section 4) lack of reversibility is not at all an issue for our proofs. The hope with these techniques is also to include easily more complexity on the stochastic models, (for instance time inhomogeneity) while relying on the same method with uniform in time estimates (cf [14], [2], [21]).

The remainder of Section 1 is organized as follows. Subsection 1.2 describes our general notations ; Subsection 1.3 presents our specific setup of a Markov process with extinction ; and Subsection 1.4 the decomposition of the state space on which we base our assumptions. Subsection 2.1 presents the main set of conditions which we show to be sufficient for the exponential convergence to the QSD. Subsection 2.2 states the three main theorems of the present paper, dealing respectively with the QSD, the survival capacity and the Q-process. The conditions that we present are then certainly numerous ; yet we believe that they are quite convenient to deal with in practice, except maybe for $(A3)$, for which we can only give a few hints in the present work (cf Subsection 2.3 and 4.2). Other issues on the assumptions are discussed in Subsection 2.4. Subsection 2.5 is devoted to the comparison with the literature. We turn in Section 3 to elementary properties that relate our assumptions. Before we deal with the main proofs of the general theorems in Section 5, we present in Section 4 two applications of these.

Theses results seem to be new, but concern toy-models. We hope that they will help the reader get insight on our approach. The application of our theorems to more significant biological models is intended for following work (cf. already [47], [48], [31]).

1.2. Elementary notations

In the following, the notation $k \geq 1$ has generally to be understood as $k \in \mathbb{N}$ while $t \geq 0$ (resp. $c > 0$) should be understood as $t \in \mathbb{R}_+ := [0, \infty)$ (resp. $c \in \mathbb{R}_+^* := (0, \infty)$). In this

context (with $m \leq n$), we denote classical sets of integers by: $\mathbb{Z}_+ := \{0, 1, 2...\}$, $\mathbb{N} := \{1, 2, 3...\}$, $[\![m, n]\!] := \{m,\, m+1, ..., n-1,\, n\}$, where the notation $:=$ makes explicit that we define some notation by this equality. For maxima and minima, we usually denote: $s \vee t := \max\{s, t\}$, $s \wedge t := \min\{s, t\}$. Accordingly, for a function ψ, $\psi^\wedge$ (resp. $\psi^\vee$) will usually be used for a lower-bound (resp. for an upper-bound) of ψ.

Let $\big(\Omega; (\mathcal{F}_t)_{t\geq 0}; (X_t)_{t\geq 0}; (P_t)_{t\geq 0}; (\mathbb{P}_x)_{x\in\mathcal{X}\cup\partial}\big)$ be a time homogeneous strong Markov process with cadlag paths on some Polish space $\mathcal{X} \cup \{\partial\}$ [[40], Definition III.1.1], where $(\mathcal{X}; \mathcal{B})$ is a measurable space and $\partial \notin \mathcal{X}$. We also assume that the filtration $(\mathcal{F}_t)_{t\geq 0}$ is right-continuous and complete. We recall that $\mathbb{P}_x(X_0 = x) = 1$, P_t is the transition function of the process satisfying the usual measurability assumptions and Chapman-Kolmogorov equation. The first entry time (resp. the first exit time) of $\mathcal{D}$, for some domain $\mathcal{D} \subset \mathcal{X}$, will generally be denoted by $\tau_{\mathcal{D}}$ (resp. by $T_{\mathcal{D}}$). While dealing with the Markov property between different stopping times, we wish to clearly indicate with our notation that we introduce a copy of X (ie with the same semigroup (P_t)) whose dependency upon X is limited to its initial condition. This copy (and the associated stopping times) is then denoted with a tilde ($\widetilde{X}$, $\widetilde{\tau}_\partial$, $\widetilde{T}_{\mathcal{D}}$ etc.). In the notation $\mathbb{P}_{X_{\tau_E}}(t - \tau_E < \widetilde{\tau}_\partial)$ for instance, τ_E and X_{τ_E} refer to the initial process X while $\widetilde{\tau}_\partial$ refers to the copy $\widetilde{X}$.

1.3. The stochastic process with absorption

We consider a strong Markov processes absorbed at ∂: the cemetery. More precisely, we assume that $X_s = \partial$ implies $X_t = \partial$ for all $t \geq s$ and that the extinction epoch: $\tau_\partial := \inf\{t \geq 0;\ X_t = \partial\}$ is a stopping time. Thus, the family $(P_t)_{t\geq 0}$ defines a non-conservative semigroup of operators on the set $\mathcal{B}_+(\mathcal{X})$ (resp. $\mathcal{B}_b(\mathcal{X})$) of positive (resp. bounded) $(\mathcal{X}, \mathcal{B})$-measurable functions. For any probability measure μ on $\mathcal{X}$, that is $\mu \in \mathcal{M}_1(\mathcal{X})$, and $f \in \mathcal{B}_+(\mathcal{X})$ (or $f \in \mathcal{B}_b(\mathcal{X})$) we use the notations:

$$\mathbb{P}_\mu(.) := \int_{\mathcal{X}} \mathbb{P}_x(.)\ \mu(dx), \quad \langle \mu \,|\, f\rangle := \int_{\mathcal{X}} f(x)\ \mu(dx).$$

We denote by $\mathbb{E}_x$ (resp. $\mathbb{E}_\mu$) the expectation corresponding to $\mathbb{P}_x$ (resp. $\mathbb{P}_\mu$).

$$\begin{aligned} &\mu P_t(dy) := \mathbb{P}_\mu(X_t \in dy), && \langle \mu P_t \,|\, f\rangle \;=\; \langle \mu \,|\, P_t f\rangle \;= \mathbb{E}_\mu[f(X_t)], \\ &\mu A_t(dy) := \mathbb{P}_\mu(X_t \in dy \,|\, t < \tau_\partial), && \langle \mu A_t \,|\, f\rangle \;= \mathbb{E}_\mu[f(X_t) \,|\, t < \tau_\partial], \end{aligned}$$

μA_t is what we called the MCNE (at time t, with initial distribution μ). In this setting, the family $(P_t)_{t\geq 0}$ (resp. $(A_t)_{t\geq 0}$) defines a linear but non-conservative semigroup (resp. a conservative but non-linear semigroup) of operators on $\mathcal{M}_1(\mathcal{X})$ endowed with the total variation norm: $\|\mu\|_{TV} := \sup\{|\mu(A)|;\ A \in \mathcal{B}\}$ for $\mu \in \mathcal{M}(\mathcal{X})$. A probability measure α is said to be the *quasi-limiting distribution* of an initial condition μ if:

$$\forall\, B \in \mathcal{B}, \quad \lim_{t\to\infty} \mathbb{P}_\mu(X_t \in B \,|\, t < \tau_\partial) := \lim_{t\to\infty} \mu A_t(B) = \alpha(B).$$

It is now classical (cf e.g. Proposition 1 in [MV12]) that α is then a quasi-stationary distribution or QSD, in the sense that: $\forall\, t \geq 0, \quad \alpha A_t(dy) = \alpha(dy)$.

Our first purpose will be to prove that the assumptions in Subsection 2.1 provide sufficient conditions for the existence of a unique quasi-limiting distribution α, independent of the initial condition.

1.4. Specification on the state space

In the following, we will always assume the following decomposition of $\mathcal{X}$:

Assumption 0. *: "**Exhaustion of** $\mathcal{X}$" There exists a sequence $(\mathcal{D}_\ell)_{\ell\geq 1}$ of closed subsets of $\mathcal{X}$ such that:*

$$\forall\, n \geq 1,\ \mathcal{D}_\ell \subset \mathcal{D}_{\ell+1}^{\circ} \quad \text{and} \quad \bigcup_{\ell\geq 1} \mathcal{D}_\ell = \mathcal{X}. \tag{A0}$$

This sequence will serve as a reference for the following statements. For instance, we will have control on the process through the fact that the initial distribution belongs to some set of the form:

$$\mathcal{M}_{\ell,\,\xi} := \{\mu \in \mathcal{M}_1(\mathcal{X});\ \mu(\mathcal{D}_\ell) \geq \xi\}, \qquad \text{with } \xi \in (0,1). \tag{1.1}$$

Note that for any $\xi > 0$: $\mathcal{M}_1(\mathcal{X}) = \bigcup_{\ell\geq 1} \mathcal{M}_{\ell,\,\xi}$. Let also:

$$\mathbf{D} := \{\mathcal{D};\ \mathcal{D} \text{ is closed and there exists } \ell \geq 1 \text{ such that } \mathcal{D} \subset \mathcal{D}_\ell\}. \tag{1.2}$$

2. Exponential convergence to the *QSD*

2.1. Hypotheses

We recall that for any set $\mathcal{D}$, we defined the first exit and entry times as:

$$T_{\mathcal{D}} := \inf\{t \geq 0;\ X_t \notin \mathcal{D}\}, \quad \tau_{\mathcal{D}} := \inf\{t \geq 0;\ X_t \in \mathcal{D}\}.$$

Assumption 1. *: "**Mixing property**"*
There exists some probability measure $\zeta \in \mathcal{M}_1(\mathcal{X})$ such that, for any $\ell \geq 1$, there exists $L \geq \ell$, $c, t > 0$ such that:

$$\forall\, x \in \mathcal{D}_\ell, \qquad \mathbb{P}_x\left[X_t \in dx;\ t < \tau_\partial \wedge T_{\mathcal{D}_L}\right] \geq c\,\zeta(dx). \tag{A1}$$

Assumption 2. *: "**Escape from the Transitory domain**"*
For given $\rho > 0$ and $E \in \mathbf{D}$:

$$e_{\mathcal{T}} := \sup_{x\in\mathcal{X}} \mathbb{E}_x\left(\exp\left[\rho\left(\tau_\partial \wedge \tau_E\right)\right]\right) < \infty. \tag{A2}$$

The order ρ in the previous exponential moment is required to be larger than the following **"survival estimate"** that involves the measure ζ in (*A*1):

$$\rho_S := \sup\left\{\rho \in \mathbb{R};\ \sup_{\ell\geq 1} \inf_{t>0}\ e^{\rho t}\,\mathbb{P}_\zeta(t < \tau_\partial \wedge T_{\mathcal{D}_\ell}) = 0\right\}. \tag{2.1}$$

Assumption 3. *: "**Asymptotic comparison of survival**"*
For a given $E \in \mathbf{D}$ and $\zeta \in \mathcal{M}_1(\mathcal{X})$:

$$\limsup_{t\to\infty} \sup_{x\in E} \frac{\mathbb{P}_x(t < \tau_\partial)}{\mathbb{P}_\zeta(t < \tau_\partial)} < \infty. \tag{A3}$$

We say that Assumption $(\mathbf{A})$ holds, whenever:
"$(A1)$ holds for some $\zeta \in \mathcal{M}_1(\mathcal{X})$ and a sequence $(\mathcal{D}_\ell)$ that satisfies $(A0)$. Moreover, there exist $E \in \mathbf{D}$ such that $(A2)$, holds with some $\rho > \rho_S$ as well as $(A3)$."

As we shall see in Subsection 3, $(A1)$ implies that $\rho_S < \infty$. In order to ensure Assumption $(\mathbf{A})$, we may not need to estimate precisely ρ_S: it is possible (depending on the process) that $(A2)$ is satisfied for any potential value for $\rho > 0$ (where E is likely to depend on ρ). Moreover, ρ_S as well as assumption $(A3)$ actually do not depend of the choice of ζ satisfying $(A1)$.

2.2. Main Theorems: the simplest set of assumptions

Theorem 2.1. *Assume that Assumption* $(\mathbf{A})$ *holds. Then, there exists a unique QSD* α*. Moreover, we have exponential convergence to* α *of the MCNE's at a given rate* $\gamma > 0$*. More precisely, for any pair* $\ell \geq 1$ *and* $\xi \in (0,1)$*, there exists* $C = C(\ell, \xi) > 0$ *such that:*

$$\forall\, t > 0,\ \forall\, \mu \in \mathcal{M}_{\ell,\xi}, \quad \| \mathbb{P}_\mu [\, X_t \in dx \mid t < \tau_\partial] - \alpha(dx) \|_{TV} \leq C\, e^{-\gamma\, t}. \tag{2.2}$$

It is classical (cf e.g. Theorem 2.2 in [19]) that, as a QSD, α is associated to some extinction rate λ:

$$\forall\, t \geq 0, \quad \mathbb{P}_\alpha(t < \tau_\partial) = e^{-\lambda\, t}, \text{ so that } \alpha P_t = e^{-\lambda\, t}\, \alpha. \tag{2.3}$$

$$\text{Let:} \quad h_t(x) := e^{\lambda\, t} \mathbb{P}_x(t < \tau_\partial). \tag{2.4}$$

Theorem 2.2. *Again under Assumption* $(\mathbf{A})$*, we have exponential convergence in the supremum norm of* $(h_t)_{t\geq 0}$ *to a limit* h*, with the rate* γ *deduced from* (2.2)*. The function* h*, which describes the "survival capacity" of the initial condition* μ*, has a positive lower-bound on any* $\mathcal{D}_\ell$*, an upper-bound on* $\mathcal{X}$ *and vanishes on* ∂*. It also belongs to the domain of the infinitesimal generator* $\mathcal{L}$*, associated with the semi-group* $(P_t)_{t\geq 0}$ *on* $(B(\mathcal{X} \cup \{\partial\}); \|.\|_\infty)$*, and:*

$$\mathcal{L}\, h = -\lambda\, h, \qquad so \quad \forall\, t \geq 0,\ P_t\, h = e^{-\lambda\, t} h. \tag{2.5}$$

Remark: Like in [13], it is also not difficult to show that there is no eigenvalue of $\mathcal{L}$ between 0 and $-\lambda$, and that h is the unique eigenvector associated to $-\lambda$.

Theorem 2.3. *Under again Assumption* $(\mathbf{A})$*, we have:*
(i) **Existence of the** Q**-process:**
There exists a family $(\mathbb{Q}_x)_{x\in\mathcal{X}}$ *of probability measures on* Ω *defined by:*

$$\lim_{t\to\infty} \mathbb{P}_x(\Lambda_s \,\big|\, t < \tau_\partial) = \mathbb{Q}_x(\Lambda_s), \tag{2.6}$$

for all $\mathcal{F}_s$*-measurable set* Λ_s*. The process* $(\Omega; (\mathcal{F}_t)_{t\geq 0}; (X_t)_{t\geq 0}; (\mathbb{Q}_x)_{x\in\mathcal{X}})$ *is an* $\mathcal{X}$*-valued homogeneous strong Markov process.*

(ii) Transition kernel:
The transition kernel of the Markov process X under $(Q_x)_{x\in\mathcal{X}}$ is given by:

$$q(x;t;dy) = e^{\lambda\, t}\, \frac{h(y)}{h(x)}\, p(x;t;dy), \tag{2.7}$$

where $p(x;t;dy)$ is the transition kernel of the Markov process X under $(P_x)_{x\in\mathcal{X}}$. In other words, for all $\psi \in \mathcal{B}_b(\mathcal{X})$ and $t \geq 0$, $\quad \langle \delta_x\, Q_t \,\big|\, \psi\rangle = e^{\lambda\, t}\, \langle \delta_x\, P_t \,\big|\, h\times\psi\rangle \,/\, h(x), \quad$ where $(Q_t)_{t\geq 0}$ is the semi-group of X under $\mathbb{Q}$.

(iii) Exponential ergodicity:
There is a unique invariant distribution of X under $\mathbb{Q}$, defined by:

$$\beta(dx) := h(x)\, \alpha(dx).$$

Moreover, there exists $\gamma > 0$ and $C = C(\ell, \xi)$ such that:

$$\forall\, t > 0,\ \forall\, \mu \in \mathcal{M}_{\ell,\,\xi}, \quad \|\mathbb{Q}_{\mu B[h]}(X_t \in dx) - \beta(dx)\|_{\frac{1}{h}} \leq C\, e^{-\gamma\, t}. \tag{2.8}$$

$$\text{where } \|\mu\|_{\frac{1}{h}} := \left\| \frac{\mu(dx)}{h(x)} \right\|_{TV} \geq \frac{\|\mu(dx)\|_{TV}}{\|h\|_\infty}, \quad \mu B[h](dx) := h(x)\, \mu(dx) \,/ \langle \mu \,\big|\, h\rangle,$$
$$\mathbb{Q}_\mu(dw) := \textstyle\int_{\mathcal{X}} \mu(dx)\, \mathbb{Q}_x(dw).$$

2.3. How to verify $(A3)$?

For discrete space, it is quite natural to deduce $(A3)$ from the fact that there exists t such that: $\inf_{x\in E} \mathbb{P}_\zeta(X_t = x) > 0$. We can thus couple some trajectories starting from ζ and passing in x at time t to the set of all trajectories starting from x. From this we can infer a lower-bound of the asymptotic survival ability of the former (starting from ζ) compared to the latter (starting from x). For an illustration, this coupling is exploited in the birth and death process in Subsection 4.1.

For continuous space however, the process starting from ζ will never hit precisely x. We need to wait a bit for the process starting from x to diffuse before the association we expect can be ensured. Although it appears quite more complicated, our argument is very similar. In cases where the Harnack inequality holds (notably pure diffusive processes, cf Subsection 4.2.2), one is usually able to prove:

$$\forall\, x \in E, \qquad \mathbb{P}_x\big(X_t \in dx;\ t < \tau_\partial\big) \leq c\, \mathbb{P}_\zeta\big(X_{t_\alpha} \in dx;\ t_\alpha < \tau_\partial\big),$$

where $t, t_\alpha, c > 0$ are independent of x. Like for the discrete space case, we then deduce $(A3)$ from the Markov property and an additional control of the survival on finite time-interval. Note that when the Harnack inequality holds, it is natural to exploit it already in the proof of $(A1)$, cf Subsection 4.2.2.

In a much general setting, and especially when jumps are involved in the process, the situations might get much tricky when one wishes to look for a similar coupling of trajectories. The issue is notably on exceptional behavior along which we have poor controls (no jump for a long time, too many jumps, too large etc.). In [46], we provide a very efficient and more easily verified condition which ensures $(A3)$, given the other assumptions. Since

this condition is technical and may appear too abstract without the illustration of various examples, we encourage any interested reader to look at [47] and [33], besides the simple illustrations given in [46].

2.4. Remarks on the Assumptions and the results

Remark 2.4.1. *Since X is right-continuous and the filtration is both right-continuous and complete, the first entry time of any Borel set is a stopping time, cf. Theorem 52 in* [20]*, or more recently Theorem 2.4 in* [7]*. It means in particular that the first exit time $T_{\mathcal{D}_\ell}$ and the first entry time $\tau_{\mathcal{D}_\ell}$ are stopping times (for any $\ell \geq 1$ and any initial condition). The result extends in fact to any iterated combination of the kind "next entry time of $\mathcal{D}_\ell$ after the first exit time of $\mathcal{D}_L$ following the first entry time of $\mathcal{D}_\ell$". For this, we shall use that there is a positive gap between each of the three random times (say $\tau_0 < T_1 < \tau_1$) involved, and that for any t, $: (s, \omega) \mapsto \mathbf{1}_{\{\tau_0(\omega) < s \leq t\}}$ has left continuous paths (and similarly with T_1 instead of τ_0 and possibly so on by induction).*

Remark 2.4.2. *This property on first entry times is the main reason for us to assume $\mathcal{X}$ Polish. The space topology is not much exploited. It means notably that one could treat càdlàg processes X that are known to satisfy the strong Markov property only w.r.t. the basic filtration (that is a priori not right-continuous), as long as this strong Markov property can be ensured for sequences of entry and exit times of sets from the family $(\mathcal{D}_\ell)_{\ell \geq 1}$. It is for simplicity that we assume that the strong Markov property is fulfilled for a right-continuous filtration, noting that it holds for our examples. At least, as stated in Theorem 7.7 of [43], a Feller semigroup on a locally compact and separable space generates a process that is strong Markov for the augmented filtration that we consider. Yet, the required condition on the family $(\mathcal{D}_\ell)_{\ell \geq 1}$ covers a broader range of processes.*

For an example of càdlàg process for which one may easily find suitable sets $(\mathcal{D}_\ell)_{\ell \geq 1}$ although the strong Markov property is not satisfied for the augmented filtration, one can adapt the counter-example provided in page 90 of [43]: such counter-examples are easily produced by specifying a partially absorbing set whose exit time is non-predictible. We can think for instance that the process stays where it hits this sets (in [43], the set is $\{0\}$ for a process on $\mathbb{R}$) for an exponential time, before leaving.

Remark 2.4.3. *As we can see in the illustrations of [46] (exploiting this result), it is not required for the process to be strong Feller: for jump processes, there may exist bounded measurable function f such that $P_t f$ is discontinuous. h itself might not be continuous, notably for discontinuous jump rate.*

Remark 2.4.4. $(A1)$ *imposes a weak form of irreducibility condition, with this reference measure ζ, and a coherence in time to prevent periodicity.*

It may happen that there exists absorbing domains $\mathcal{D}_A$ (whose escape can only happen at τ_∂). Any MCNE with initial condition $x \in \mathcal{D}_A$ is necessarily supported in $\mathcal{D}_A$. Any ζ that satisfies $(A1)$ is thus also supported in $\mathcal{D}_A$. Moreover, if these MCNE converge to a unique QSD as in our result, this QSD is necessarily supported on $\mathcal{D}_A$ as well.

Remark 2.4.5. *Assumption* $(A1)$ *is a stronger version of Doeblin's condition that appears for the convergence of Markov Chains without extinction. It also implies that any border of*

extinction shall be approached by the sequence $\mathcal{D}_\ell$ while $\ell \to \infty$, but never from inside any $\mathcal{D}_\ell$, since by Lemma 3.0.3:

$$\forall\, \ell \geq 1,\ \forall\, t > 0, \quad \inf\{\mathbb{P}_x(t < \tau_\partial);\ x \in \mathcal{D}_\ell\}\ > 0.$$

Remark 2.4.6. *When it concerns pure jump processes, one can generally choose $\mathcal{D}_L := \mathcal{D}_\ell$. For other processes, one often needs "a bit of space" between $\mathcal{D}_\ell$ and $\mathcal{D}_L^c$ to obtain a lower bound uniform in $x \in \mathcal{D}_\ell$ over trajectories from x to ζ staying inside $\mathcal{D}_L$ (as in our second application with a diffusion).*

Remark 2.4.7. *To understand* (2.8)*, it is worth noticing that, considering some general initial condition in the left-hand side of* (2.6)*, we obtain for the Q-process a biased initial condition:*

$$\forall\, \mu \in \mathcal{M}_1(\mathcal{X}), \quad \lim_{t\to\infty} \mathbb{P}_\mu(\Lambda_s \,\big|\, t < \tau_\partial) = \mathbb{Q}_{\mu B[h]}(\Lambda_s). \tag{2.9}$$

To deduce (2.8) *from* (2.2)*, we reformulate* (2.7) *in terms of $B[h]$, P_t, A_t and Q_t:*

$$\forall\, t \geq 0,\ \forall\, \mu \in \mathcal{M}_1(\mathcal{X}), \quad (\mu B[h])Q_t = (\mu P_t)\, B[h] = (\mu A_t)\, B[h]. \tag{2.10}$$

Originally, we intended to adapt the proof of Theorem 2.1 on the marginal at time t conditioned on survival at time $t + T$ to deduce a control uniform in T. This approach is effective but leads to a weaker result where $\|.\|_{\frac{1}{h}}$ is replaced by $\|.\|_{TV}$. The convergence of the MCNE is here more informative, because h is bounded.

Remark 2.4.8 (On the indices)**.** *Throughout the proof, the constants and sets that we consider will be indexed by a capital letter referring to the property they are involved in. The indexes S, E, M, C, A, R, D, P and L stand respectively for "Survival" (notably in* (2.1)*,* (3.1) *and* (5.15)*), "first Entry" (in* (A2)*, where it can also refer to "Escape", and in Lemma 5.1.2), "Mixing" (in* (A1)*), "Containment" (throughout Section 5.1), "Absorption" (in* (5.17)*), "Renewal" (in* (5.22)*), "Doeblin" (also in* (5.22)*), "Persistence" (in* (5.16)*) and finally "Last exit" (in Lemma 5.1.4). $\circ$ as an index plays a similar role to indicate parameters referring to the core of convergence as specified in Theorem 5.1.*

2.5. Comparison with the literature

General perspective. Although there is already a vast literature on QSDs (see notably the impressive bibliography collected by Pollett [39]), the approach we follow seems to have been explored only in the very recent years. For a review on the results that were previously obtained, we refer notably to general surveys as in [19], in [23] or more specifically for population dynamics in [MV12]. We see already in these surveys how essential is the role played by the spectral theory. The spectral theory is very effective both to relate the QSD and the survival capacity to the first eigenvector of a diagonalizable operator and to identify the convergence rate as the gap between the first and the second eigenvalues (cf e.g. [11]). The principal drawback of the spectral theory is that it usually relies on reversibility. Certainly, for 1 dimensional processes, this condition of time-symmetry is quite easily satisfied ; while, more generally, it can be deduced from conditions easy to verify (detailed balance notably). This may explain why reversibility is so extensively

studied. Yet, it is a very restrictive condition for higher dimensions, as it is well explained in the appendix A of [12].

Alternative methods are usually much less effective. In [18], the authors prove the existence of the QSD via a Tychonov fixed point theorem. Another proof for the existence of the QSD is presented in [26] for Markov Chains on $\mathbb{Z}_+$, based on compactness arguments and renewal techniques. In [5], the authors prove, under quite stringent conditions, the existence and uniqueness of the QSD and propose estimations of this QSD up to some computable time, again with renewal arguments. The authors of [22] relate the speed of convergence to QSD to the one of a related Doob's transform towards its stationary distribution. Yet the conditions of the last two papers seem to apply essentially to discrete-space processes, or at least when the extinction is in some sense uniformly bounded. The existence of the QSD and the survival capacity has also been related, at least for discrete time and discrete space, to the notion of R-positivity (cf e.g. [44], [45] or [42]). This is especially useful when the process is easily described by generating functions (in particular for Galton-Watson processes) but seems quite an abstract criterion otherwise. Still, it provides the main principle of focusing on the exponential rate of extinction, which is at the core of our study. Our proof can reasonably be judged as an extension of the one presented in [27] with a focus on general practical assumptions for R-positivity, noting that their analysis is restricted to discrete-time and discrete-space Markov processes.

The dependency on the initial condition. Upper-bounds of the form:

$$\| \mathbb{P}_\mu [X_t \in dx \mid t < \tau_\partial] - \alpha(dx) \|_{TV} \leq C(\mu)\, e^{-\gamma\, t}. \tag{2.11}$$

assume generally $C(\mu) = \langle \mu \,|\, W \rangle$ in the case of α being a stationary distribution (i.e. for processes without extinction). The use of such a reference function W has been thoroughly studied in [36] in the case of Markov Chains, or in e.g. [24], [10], [9] for continuous time processes. The condition on W is what characterizes it as a Lyapunov function and relate a priori to a control of the first entry time τ_E. Different probabilistic bounds have generally been proposed, although, including extinction, exponential moments appear compulsory (all the more since λ, the limiting rate of extinction, is not precisely known). In a loose version, and still for P_t as a conservative semi-group, such exponential control may take the form:

$$P_{t_0} W \leq e^{-\rho_W t_0} W + C_W, \qquad \text{where } \rho_W, C_W > 0.$$

E is then generally chosen as the set $\{W \leq d_W\}$ for d_W sufficiently large for

$$P_{t_0} W \leq e^{-\rho'_W t_0} W + C'_W \mathbf{1}_E$$

to hold (for a smaller value $\rho'_W < \rho_W$ and some $C'_W > 0$). For E to be convenient with respect to the other criteria (mixing or comparison of survival) and especially when ρ_S is not estimated precisely, it is usually assumed that W is proper (i.e. $W(x)$ converges to infinity as $\|x\|$ tends to infinity).

An extension of this assumption for the non-conservative case has been recently proposed in [3] through their Assumption **A**. Another usual version involves the infinitesimal

generator $\mathcal{L}$ and assumes the following:

$$\mathcal{L}W \leq -\rho_W W + C_W, \qquad \text{where } \rho_W, C_W > 0.$$

The inequality is stronger than the previous one for proper W. A related estimate for non-conservative cases is also proposed in Proposition 2.10 of [3]. When considering extinction, we lose also the property of linearity over the initial condition. This explains why upper-bounds like $\langle \mu \,|\, W \rangle$ are not so general and why we focus on general initial distributions and not only Dirac Masses.

The conclusion that we present is quite natural for the models we have in mind, where extinction plays a stabilizing role, preventing transient dynamics. In the perspective of natural selection, we expect to observe the prevalence of trajectories leading to and gravitating around some basin of attraction, notably compared to those dragged away in deadlier regions. Although the burden of mal-adaptation may seem light in the short run, if it is too hard for the process to escape from less adapted areas, one can presume that the process cannot have been there for long. In particular, the trajectories starting from favorable initial conditions may outcompete what remains of the distribution, so that it becomes the leading part in the convergence to the QSD.

Other expressions of $C(\mu)$ have been presented in [16], [28] and [3], with similar interpretation. Note that the proofs in [16] and [3] concern the convergence towards a unique Yaglom distribution, which may not be unique as a QSD, for any initial conditions with a light enough tail. In [16], (2.11) is obtained with a non-linear dependency of the form $C(\mu) = C \langle \mu \,|\, \psi_1 \rangle / \langle \mu \,|\, \psi_2 \rangle$. As can be seen in our following paper[46], the dependency we introduce implies (2.11) with $C(\mu) = C/\langle \mu \,|\, h \rangle$, for some $C > 0$. So the former extends our result by including the more classical dependency through a Lyapunov function. In [28] and [3], the convergence is stated in a weighted norm involving a weight function W (resp.V) related to the previous ψ_1. The dependency $C(\mu)$ stated in their analog of (2.11) is implicitly related to both $\langle \mu \,|\, h \rangle$ and $\langle \mu \,|\, W \rangle$ (their function h plays the same role as our). A dependency on ψ_1 (or on W) is neglected in our article: $(A2)$ ensures in a way that we can find some upper-bounded ψ_1 (we refer e.g. to Lemma 3.6 in [16]).

A hint to connect the current techniques to their setting is to adjust the probabilities of transition of the Markov process according to such weight function W. The set of positive measures μ such that $\int W(x)\mu(dx) \leq 1$ for instance takes the place of the set of probability measures (for more details, we refer to Subsection 2.3.2 of [46]). In general practice, it does not seem so clear to us how to find such Lyapunov functions especially when one wishes to combine simple bounds on different parts of the space. So we believe that our assumption $(A2)$ is more natural to verify in many examples (cf e.g. our second application), while easier to interpret.

The assumptions. If one can relate our set of assumptions to the ones proposed in [27], the similarity is clearly greater with [16], [28] and [3], because of the introduction of a continuous-space setting. It is proved in [3] that their conditions are not only sufficient, but also necessary, and similarly in Theorem 2.3 of [17] with regards to [16]. Similar reciprocal results are obtained in [46] for the current setting.

Due to our trajectorial approach, we require additional confinement properties (with restrictions of the probabilities upon the events $\{t < T_{\mathcal{D}_L}\}$). As explained above, Lyapunov functions are also not directly exploited in our approach. The survival estimates presented in [16], [28] and [3] appears less natural to interpret than the condition on survival that we propose ($\rho > \rho_S$): they require the introduction of a function, say ψ_2, for which the time behavior of $\psi_2(X_t)$ must satisfy a certain minimization property. Given Lemma 3.2 in [16], their condition appears more general and can certainly be convenient for specific models. Nonetheless Theorem 2.3 in [46] proves that ρ_S takes the optimal value λ provided that the convergence results (2.2) and of the survival capacity hold.

The approach. The techniques exploited in [16], [28] and [3] are quite different from ours. In the steps of the R-theory, the study of the h-transformed process is at the core of [28], with a weighted norm. Contraction estimates under similar weighted norms are exploited in [16] and [3]. Our proofs are much more constructive and rely on a control on entry times of core sets thanks to the competition between different behaviors. It extends to models where the uniqueness of the QSD does not hold due to transitivity conditions, as one can observe in the applications we have in [48]. In particular, our work offers a new constructive perspective even for the results in [13] (cf Subsection 5.3) since the coupling steps which we introduce apply directly to the MCNE (and not to their linearized versions).

3. Several implications of $(A1)$

Lemma 3.0.1. *Assume that* $(A1)$ *holds for two probability measures* ζ^1 *and* ζ^2*. Then, the associated values for* ρ_S *coincide. Moreover, the sets* E *for which assumption* $(A3)$ *holds are the same for both measures.*

"From mixing to regeneration, then survival": $(A1)$ trivially implies, for any ℓ such that $\zeta(\mathcal{D}_\ell) > 0$ the following **regeneration estimate**:

There exists $t, c > 0$, $L \geq \ell$ such that, with $\mathcal{D}_S := \mathcal{D}_\ell \subset \mathcal{D}_L$:

$$\forall\, x \in \mathcal{D}_S, \qquad \mathbb{P}_x(X_t \in \mathcal{D}_S;\ t < \tau_\partial \wedge T_{\mathcal{D}_L}) \geq c. \tag{3.1}$$

Lemma 3.0.2. *Assume that* (3.1) *holds (for* t, $c > 0$, $\mathcal{D}_S \subset \mathcal{D}_L$ *and* $\zeta(\mathcal{D}_S) > 0$*). Then,* $\rho_S \leq -\frac{1}{t}\ln(c)$.

In particular, we deduce that $(A1)$ implies $\rho_S < \infty$.

Lemma 3.0.3. $(A1)$ *is equivalent to the apparently stronger version (with the same* ζ*): For any* $\ell \geq 1$*, and* $t_\vee > 0$*, there exists* $L \geq \ell$*,* $t \geq t_\vee$ *and* $c > 0$ *such that:*

$$\forall\, x \in \mathcal{D}_\ell, \qquad \mathbb{P}_x\left[X_t \in dx;\ t < \tau_\partial \wedge T_{\mathcal{D}_L}\right] \geq c\,\zeta(dx). \tag{$\overline{A1}$}$$

3.1. Proof of Lemma 3.0.1

Assume that ρ_S^1 is associated to a first choice of ζ^1 satisfying $(A1)$, and consider another choice ζ^2. By $(A0)$, there exists $\ell \geq 1$ such that $\zeta^2(\mathcal{D}_\ell) \geq 1/2$. By $(A1)$ applied to ζ^1, for some $c_J, t_J > 0$ and $L > \ell$:

$$\mathbb{P}_{\zeta^2}(X_{t_J} \in dx;\ t_J < \tau_\partial \wedge T_{\mathcal{D}_L}) \geq c_J \zeta^1(dx). \tag{3.2}$$

By definition of ρ_S^1, for any $\rho > \rho_S^1$, there exists $c_S, t_S > 0$ and $L' \geq L$ such that:

$$\forall\, t \geq t_S,\ \mathbb{P}_{\zeta^1}(t < \tau_\partial \wedge T_{\mathcal{D}_{L'}}) \geq c_S \exp[-\rho t]. \tag{3.3}$$

By combining (3.2), (3.3) and the Markov property, we deduce:

$$\limsup_{t>0} \frac{\exp[-\rho(t+t_J)]}{\mathbb{P}_{\zeta^2}(t+t_J < \tau_\partial \wedge T_{\mathcal{D}'_L})} \leq \left(c_J\, c_S\, \exp[\rho t_J]\right)^{-1} < \infty.$$

By optimizing in ρ, we deduce $\rho_S^2 \leq \rho_S^1$ and the equality by symmetry.

Concerning assumption $(A3)$, (3.2) and the Markov property imply that for any $t \geq 0$ and $x \in \mathcal{X}$:

$$\begin{aligned} &\mathbb{P}_{\zeta^2}(t+t_J < \tau_\partial) \geq c_J \mathbb{P}_{\zeta^1}(t < \tau_\partial) \\ \text{Thus} \qquad &\frac{\mathbb{P}_x(t+t_J < \tau_\partial)}{\mathbb{P}_{\zeta^2}(t+t_J < \tau_\partial)} \leq (c_J)^{-1} \frac{\mathbb{P}_x(t < \tau_\partial)}{\mathbb{P}_{\zeta^1}(t < \tau_\partial)}. \end{aligned}$$

If assumption $(A3)$ holds for ζ^1 and E, it thus holds also for ζ^2 and the same E. □

3.2. Proof of Lemma 3.0.2

Assume (3.1). Let $x \in \mathcal{D}_S$, $\rho := -\frac{1}{t_{RG}} \ln(c_{RG})$, $T_L := \inf\{t \geq 0,\ X_t \notin \mathcal{D}_L\}$. By induction over $k \in \mathbb{N}$ and the Markov property:

$$\forall\, k \geq 1, \qquad \inf_{x \in \mathcal{D}_S} \mathbb{P}_x(k\, t_{RG} < T_L) \geq \exp(-\rho\, k\, t_{RG}).$$

Thus, for a general value of $t > 0$:

$$\begin{aligned} \inf_{x \in \mathcal{D}_S} \mathbb{P}_x(t < T_L) &\geq \inf_{x \in \mathcal{D}_S} \mathbb{P}_x\left(\left\lceil \tfrac{t}{t_{RG}} \right\rceil t_{RG} < T_L\right) \geq \exp\left(-\rho \left\lceil \frac{t}{t_{RG}} \right\rceil t_{RG}\right) \\ &\geq \exp(-\rho\,(t + t_{RG})) = c_S\, e^{-\rho\, t} \quad \text{ with } c_S := \exp(-\rho\, t_{RG}) = c_{RG}. \end{aligned}$$ □

3.3. Proof of Lemma 3.0.3:

Let $\ell \geq L_S$ for which we apply $(A1)$. By induction with the Markov property, it is quite straightforward to extend the property $(A1)$ on $\mathcal{D}_\ell$ with the same L_M, $t^{(k)} := k \times t$, $c^{(k)} := c \times (c\,\zeta(\mathcal{D}_S))^{k-1}$. Then, for any $t_{\underline{\vee}} > 0$, we only need to apply this extension for some $k \geq 1$ such that $t^{(k)} \geq t_{\underline{\vee}}$. On the other hand, $(\overline{A1})$ clearly implies $(A1)$ (take $t_{\underline{\vee}} = 0$), so that we have indeed proved $(A1) \Leftrightarrow (\overline{A1})$. □

4. Two models to which our results apply

4.1. Birth-and-death process with catastrophes

We choose to illustrate our result with this example for its clear simplicity. In this birth and death process, the population can get extinct punctually at any time during what we call a catastrophe. These events happen at a rate depending on the current number of alive individuals. Otherwise, the process gets extinct when there is only a unique individual that

ends up dying. To ensure uniqueness of a QSD, we will impose that the catastrophe rate is large enough when the population size is large. Biologically, we could imagine that the population is under the threat of some voracious predators, but can stay hidden as long as the population size is not too large.

In fact, one has now quite a complete description of quasi-stationarity for birth-and-death processes. It is proved in [34] that there exists a unique QSD for one dimensional birth and death processes if and only if (2.11) holds with a uniform constant $C(\mu) = C > 0$. This equivalence is probably due to the fact that in these models, extinction can only occur once the process is inside some given compact set (i.e. once it has descended from infinity), as suggested in Theorem 19 in [23]. Like in [13] and as we will do, the authors of [23] include direct extinction from any state of the birth-and-death process (what is called a "catastrophe"). Theorem 19 in [23] states that the behavior of the process is the same if catastrophe only happens in a compact set. In Theorem 4.1 of [13], the authors prove that, for a bounded catastrophe rate, there is descent from infinity (see notably [4]) iff (2.11) holds with a uniform constant $C(\mu) = C$. This does not exclude however that (2.11) could hold without descent from infinity, which we prove with our technique.

4.1.1. Description of the process

X, the population size, is a time-homogeneous Markov Chain on $\mathbb{Z}_+$ where $\partial = 0$ is the absorbing state and $\mathcal{X} = \mathbb{N}$. Given $X_0 = n \geq 1$, there is a death with rate $d_n > 0$ (leading to $X = n-1$), a birth with rate $b_n > 0$ (leading to $X = n+1$) and a catastrophe with rate $c_n \geq 0$ (leading to $X = 0$). Since c_1 and d_1 play the same role (the transition is from $X = 1$ to $X = \partial$), we assume w.l.o.g. $c_1 = 0$. Actually, $d_1 > 0$ is not required in the following statements.

Theorem 4.1. *Assume that:* *for some $n \geq 1$ (thus for all n) $\mathbb{P}_n(\tau_\partial < \infty) = 1$*

$$\text{and} \qquad \liminf_{n\to\infty} c_n > \inf_{k\geq 1}(b_k + d_k + c_k). \tag{4.1}$$

Then, the conclusions of Theorem 2.1, 2.2 and 2.3 hold.

At least for some of the models, the speed of convergence towards the QSD cannot be uniformly bounded over all initial conditions, since:

Proposition 4.1.1. *We can define some positive values for $(b_n, d_n, c_n)_{n\geq 1}$ such that (4.1) holds and for which, whatever large the time $t > 0$, and whatever small the similarity threshold $\epsilon \in (0,1)$, we can still find some initial condition $x \in \mathcal{X}$ such that:*

$$\left\|\mathbb{P}_x\left(X_t \in dy \,\middle|\, t < \tau_\partial\right) - \mathbb{P}_1\left(X_t \in dy \,\middle|\, t < \tau_\partial\right)\right\|_{TV} \geq 1 - \epsilon.$$

The proof of Theorem 4.1 and Proposition 4.1.1 are achieved resp. in Subsection 4.1.2-3.

Remark 4.1.2. *Explicit values for ρ_S can hardly be obtained except for very specific models. Yet, it might be of interest to find, depending on the specific model under consideration, more precise estimates as our value $\inf_k(b_k + d_k + c_k)$. This upper-bound comes from a survival estimate of the simplest form: the process reaches some position k (as optimal as we need) on which to stay up to (large) time t. We are a priori very far from a necessary and sufficient condition: it seems hardly possible to infer generically the level of catastrophe rate that affects the process as it evolves at large values.*

Remark 4.1.3. *By the condition $\mathbb{P}_n(\tau_\partial < \infty) = 1$, we mean that the process is non-explosive. With: $T_\infty := \lim_{n\to\infty} \inf\{t \geq 0;\ X_t \geq n\}$, our condition means that "for some $n \geq 1$ (thus for all), $\mathbb{P}_n(T_\infty = \infty) = 1$". Clearly, this property holds provided it holds in the associated birth-and-death process without catastrophe (i.e. imposing $c_n \equiv 0$). The simplest case is then when the sequence b_n/n is upper-bounded. We refer to Theorem 5.5.2 in [32] for a more general condition (still deduced from the case without catastrophe).*

We do not exclude that it could hold more generally. Yet, for catastrophe to play a role in this condition would require the family of catastrophe rate (c_n) to quickly reach very large values as n goes to infinity. It does not seem likely in practical applications.

Remark 4.1.4. *Considering $\bar{\tau}_\partial := \inf\{t \geq 0;\ X_t = 0\} \wedge T_\infty$ as the extinction epoch, our theorem extends to the case $T_\infty < \infty$. It also extends to models where catastrophes do not entirely exterminate the population. Assume for instance that after a catastrophe, from a population of size larger than some $K \geq 1$, only K individuals are to survive. We can keep the extinction for population of size initially lower than K, but it's not very significant here. Then $(A2)$ can easily be adapted with $K \in E = [\![1, \ell_E]\!]$. The proof of the other assumptions remains the same.*

Remark 4.1.5. *The alternative conditions given in [16] seem also very efficient to obtain Theorem 4.1. Since ρ_S is finite, this will certainly not be the case for the ones in [28].*

4.1.2. Proof of Theorem 4.1

By (4.1), let $k, \ell_E \geq 1$ and $\rho_E > 0$ be such that:

$$0 < \widetilde{\rho}_S := b_k + d_k + c_k < \rho_E < \inf_{\{n \geq \ell_E + 1\}} c_n := \widetilde{\rho}_E. \tag{4.2}$$
$$\zeta := \delta_k, \quad \mathcal{D}_S := \{k\}, \quad E := [\![1, \ell_E]\!].$$

Let $\mathcal{D}_\ell = [\![1, \ell \vee k]\!]$, for $\ell \geq 1$. In the following, we ensure $(\mathbf{A})$ (where $\zeta(\mathcal{D}_S) > 0$ is obvious). First, $(A0)$ is obvious.

Proof of $(A1)$ and $(A3)$.

Let $n \geq k$. Consider

$$\partial^n := \{0\} \cup [\![n+1, \infty[\![, \qquad \tau^n_\partial := \inf\{t \geq 0;\ X_t \in \partial^n\}.$$

Then the process $Y_t := X_t \mathbf{1}_{\{t < \tau^n_\partial\}}$ is a Markov Chain on the finite space $[\![0, n]\!]$, absorbed at $\partial = 0$. Since $\forall\, \ell \geq 1,\ d_\ell > 0, b_\ell > 0$, this Chain (Y_t) is irreducible and it is elementary to prove that:

$$\forall\, t_Y > 0,\ \exists\, c_Y > 0, \quad \forall\, i, j \in [\![1, n]\!], \quad \mathbb{P}_i(Y_{t_Y} = j) \geq c_Y. \tag{4.3}$$

With $j := k$ and $n := \ell_M \vee k$, (4.3) clearly implies $(A1)$ (with parameters $\ell = \ell_M, L = L_M, c = c_M, t = t_M$). We can indeed choose $\zeta := \delta_k$, $L_M := n$, $t_M = 1$ (arbitrary), and c_M the value of c_Y associated to the choice of $t_Y = t_M$.

With $i := k$ and $\ell = \ell_E$, (4.3) and the Markov property imply $(A3)$ for any E, because:

$$\begin{aligned} \forall\, t > 0,\ \forall\, j \in [\![1, \ell_E]\!], \quad \mathbb{P}_j(t < \tau_\partial) &\leq (1/c_Y) \times \mathbb{P}_k(X_{t_Y} = j;\ t + t_Y < \tau_\partial) \\ &\leq (1/c_Y) \times \mathbb{P}_k(t < \tau_\partial). \end{aligned}$$

Proof of $(A2)$.

By (4.2), the catastrophe rate is larger than $\widetilde{\rho}_E$ as long as the process remains outside E. It implies that we can upper-bound $\tau_\partial \wedge \tau_E$ by an exponential variable with rate $\widetilde{\rho}_E$. Thus:

$$\forall\, t > 0,\ \forall\, n \geq \ell_E + 1, \quad \mathbb{P}_n(t < \tau_E \wedge \tau_\partial) \leq \exp(-\widetilde{\rho}_E\, t). \tag{4.4}$$

It is classical -by Fubini Theorem, and the integral expression of the exponential- to relate the exponential moment with the repartition function by:

$$\mathbb{E}_n\left(\exp\left[\rho_E\left(\tau_E \wedge \tau_\partial\right)\right]\right) = 1 + \rho_E \int_0^\infty \exp[\rho_E\, t]\, \mathbb{P}_n(t < \tau_E \wedge \tau_\partial)\, dt. \tag{4.5}$$

By (4.4) and (4.5), we conclude:

$$\begin{aligned}\forall\, n \geq \ell_E + 1, \quad \mathbb{E}_n\left(\exp\left[\rho_E\left(\tau_E \wedge \tau_\partial\right)\right]\right) &\leq 1 + \rho_E \int_0^\infty \exp[-(\widetilde{\rho}_E - \rho_E)\, t]\, dt \\ &= 1 + \ \{\rho_E \ / \ (\widetilde{\rho}_E - \rho_E)\} < \infty.\end{aligned}$$

Proof that for any $k \geq 1$*:* $b_k + d_k + c_k \geq \rho_S$.

Immediately, by (4.2):

$$\forall\, t \geq 0, \quad \mathbb{P}_k(X_t = k;\ t < \tau_\partial \wedge T_{\mathcal{D}_k}) \geq \mathbb{P}_k(\forall\, s \leq t,\ X_s = k) = \exp(-\widetilde{\rho}_S\, t).$$

4.1.3. Proof of Proposition 4.1.1

We consider one of the simplest choice, which is to take b_n, d_n linear in n (the classical Malthus' growth model, without competition) and c_n constant for $n \geq 2$. We can then choose arbitrarily:

$$\begin{aligned} b_1, d_1, \bar{b}, \bar{d} \in (0,\infty)^5, \quad c_2 > (b_1 + d_1), \\ \text{with } c_1 = 0, \qquad \forall\, n \geq 2, \quad b_n := \bar{b}\, n,\ d_n := \bar{d}\, n,\ c_n := c_2. \end{aligned} \tag{4.6}$$

(4.1) is clearly satisfied. There is no explosion for this model, so that extinction happens a.s. (note Remark 4.1.3 of Subsection 4.1.1 on this aspect).

We shall only need to consider transitions between values of the form $2^n, n \geq 2$. Let:

$$T_n := \inf\left\{t \geq 0;\ X_t \leq 2^{n-1} \text{ or } X_t \geq 2^{n+1}\right\}, \tag{4.7}$$

$$\tau_n := \inf\left\{t \geq 0;\ X_t \leq 2^n\right\}. \tag{4.8}$$

We use the following lemma, whose proof is deferred after the one of Proposition 4.1.1:

Lemma 4.1.6. *For some* $u > 0$*, it holds:* $\lim_{n\to\infty} \mathbb{P}_{2^n}(T_n \leq u) = 0.$

For given $t, \epsilon > 0$, let $K := \lfloor t/u \rfloor + 1$ and $N \geq 1$ (by Lemma 4.1.6) such that:

$$\mathbb{P}_1(X_t \leq 2^N \,\big|\, t < \tau_\partial) \geq 1 - \epsilon/2, \tag{4.9}$$

$$\forall\, n \geq N, \quad \mathbb{P}_{2^n}(T_n \leq u) \leq \epsilon \times e^{-(c_2 - d_1)t}/(4\,K). \tag{4.10}$$

With initial condition $x := 2^{N+K+1}$, in order that X reaches 2^N before time $t \le K\,u$, it must at least once have got from 2^{N+k} to 2^{N+k-1} during a time-interval less than u, for some $1 \le k \le K+1$. With the Markov property, this implies, with (4.7), (4.8), (4.10) and the fact that the extinction rate is always lower-bounded by d_1:

$$\mathbb{P}_x(\tau_N \le t) \le \textstyle\sum_{\{k \le K+1\}} \mathbb{P}_{2^{N+k}}(T_{N+k} \le u) \le \epsilon\, e^{-(c_2-d_1)t}/2.$$
$$\mathbb{P}_x(\tau_N \le t;\ t < \tau_\partial) \le e^{-d_1\,t}\,\mathbb{P}_x(\tau_N \le t) \le \epsilon\, e^{-c_2\,t}/2.$$

Since the extinction rate is upper-bounded by c_2: $\mathbb{P}_x(t < \tau_\partial) \ge e^{-c_2\,t}$.
This implies $\mathbb{P}_x(\tau_N \le t \,\big|\, t < \tau_\partial) \le \epsilon/2$. Therefore, with also (4.9):

$$\begin{aligned}\|\delta_1 A_t - \delta_x A_t\|_{TV} &\ge \mathbb{P}_1(X_t \le 2^N \,\big|\, t < \tau_\partial) - \mathbb{P}_x(X_t \le 2^N \,\big|\, t < \tau_\partial)\\ &\ge 1 - \epsilon/2 - \epsilon/2 \ge 1 - \epsilon.\end{aligned}$$ □

Proof of Lemma 4.1.6. :
With initial condition 2^n, we can decompose X as a semi-martingale, up to time $t \wedge T_n$:

$$\forall\, t > 0, \quad X_{t\wedge T_n} := 2^n + \textstyle\int_0^{t\wedge T_n} (\bar b - \bar d)\, X_s\, ds + M_{t\wedge T_n}, \tag{4.11}$$

where $(M_{t\wedge T_n})_t$ is a martingale with bounded quadratic variation, with (4.7):

$$< M >_{t\wedge T_n} = \textstyle\int_0^{t\wedge T_n} (\bar b + \bar d)\, X_s\, ds \le (\bar b + \bar d)\, 2^{n+1}\, t. \tag{4.12}$$

Let $u := (8\,|\bar b - \bar d| \vee 1)^{-1}$ so that, by (4.7), a.s.:

$$\forall\, t \le u, \quad \Big|\textstyle\int_0^{t\wedge T_n} (\bar b - \bar d)\, X_s\, ds\Big| \le |\bar b - \bar d|\, 2^{n+1}\, u \le 2^{n-2}. \tag{4.13}$$

$$\begin{aligned}\mathbb{P}_{2^n}(T_n \le u) &\le \mathbb{P}_{2^n}\left(\sup_{\{t\le u\}} M_{t\wedge T_n} \ge 2^{n-2}\right) \qquad \text{by (4.11) and (4.13)}\\ &\le 2^{-(2\,n-4)}\ \mathbb{E}_{2^n}(< M >_{u\wedge T_n}) \text{ by Doob's inequality}\\ &\le 2^{-(2\,n-4)}\ (\bar b + \bar d)\, 2^{n+1}\, u \qquad \text{by (4.12)}\\ &= \frac{4\,(\bar b + \bar d)}{|\bar b - \bar d| \vee 1}\, 2^{-n} \underset{n\to\infty}{\longrightarrow} 0 \quad \text{with the definition of } u.\end{aligned}$$ □

4.2. Adaptation of a population to its environment: application to a diffusion process

In this illustration, the notion of being in a mal-adapted region is quite intuitive and the criteria for the exponential convergence to a unique QSD rather natural. Again, the general proof for this illustrative example is unclear without our techniques, except maybe with those of [16]. Yet, in this case, it is presumably quite technical to find a proper Lyapunov function (although our argument proves in fact that they exist). In fact, our control is deduced from local bounds ensuring both a rapid escape from several specific local domains together with sufficiently low transition rates between these domains.

4.2.1. Presentation of the model

We consider a simple coupled process describing the eco-evolutive dynamics of a population. We model the population size by a logistic Feller diffusion $(N_t)_{t\geq 0}$ where the growth rate $(r(X_t))_{t\geq 0}$ is changing randomly. Namely, the adaptation of the population and the change of the environment are assumed to act on a hidden process (X_t) in $\mathbb{R}^d$, from which the growth rate is deduced. For simplicity, we will assume that X_t evolves as a continuous Markov process driven by some Brownian Motion and a drift (possibly depending on N and X). For very low values of $r(X_t)$, it is expected that the population shall vanish very quickly. It would thus not change much of the result to introduce an absorbing boundary at some threshold of mal-adaptation. Yet, we want our result to be independent of any such truncation of the trait space and say that this large extinction is sufficient in itself to bound the mal-adaptation, while highlighting that the initial condition indeed matters here.

In a general setting, the process can be described as:

$$(S)\begin{cases} dN_t = (r(X_t) - c\ N_t)\ N_t\ dt + \sigma\ \sqrt{N_t}\ dB_t^N \\ dX_t = b(X_t, N_t)\ dt + \theta(X_t, N_t)\ dB_t^X \end{cases}$$

with initial conditions (n, x), B^N and B^X two independent Brownian Motions, $c, \sigma > 0$, and r, b, θ being locally Hölder continuous functions. We also require that θ is locally elliptic, in the following sense: for any compact set K of $\mathbb{R}^{d+1}$, there exists $\bar{\theta} > 0$ such that for any $(n, x) \in K$ and $\xi \in \mathbb{R}^d$: $\sum_{i,j} \theta_{i,j}(n, x)\xi_i\xi_j \geq \bar{\theta}\,|\xi|^2$.

Theorem 4.2. *Consider the process (X, N) with the notations specified above and the assumption that $\limsup_{\|x\|\to\infty} r(x) = -\infty$. Then, all the results of Subsection 2.2 hold. In particular, there is exponential convergence in total variation of the MCNEs to the unique QSD.*

It is also not much more costly to introduce catastrophes, arising at rate $\rho_c(x, n)$, leading to the complete extinction of the population. Partial extinction of the population (with jumps on the population size), are however quite more technical to deal with (because the Harnack inequality is not as obvious). In [46], where the focus is on processes with jumps, we shall present techniques that makes it much more manageable.

The main issue for this model is to specify the conditions for $(A2)$ to hold. We provide in Subsection 4.2.4 a way to prove it in a strong case where it holds for any ρ, i.e. for the following Theorem 4.2.1. For diffusions like this, $(A1)$ and $(A3)$ may be deduced quite roughly thanks to the Harnack inequalities, as presented in the next subsection. In order to satisfy assumption $(A3)$, there is no additional restriction on the set E, so that the only requirement on E is for $(A2)$.

4.2.2. Harnack inequalities for $(A1)$ and $(A3)$

In the following, we say that a process (Y_t) on $\mathcal{Y} \subset \mathbb{R}^d$ with generator $\mathcal{L}$ (including possibly an extinction rate ρ_c) satisfies Assumption (H) if the following property holds:

Consider any path-connected open relatively compact sets $\mathcal{D}, \mathcal{D}' \subset \mathcal{Y}$, such that $\overline{\mathcal{D}} \subset \mathcal{D}'$, with C^∞ boundaries, and such that for any point $x \in \partial\mathcal{D}'$, there exists a closed ball $C \in \mathbb{R}^d$ (of non-empty interior) such that $C \cap \overline{\mathcal{D}'} = \{y\}$. For any $0 < t_1 < t_2$ and non-negative

C^2 constraints: $u_{\partial\mathcal{D}'} : (\{0\} \times \mathcal{D}') \cup ([0, t_2] \times \partial\mathcal{D}') \to [0, \infty)$, there exists a unique solution $u \in C^{1,2}((0, t_2) \times \mathcal{D}') \cap C^0([0, t_2] \times \overline{\mathcal{D}'})$ to the problem:

$$\begin{aligned} &\partial_t u(t, y) = \mathcal{L}u(t, y) && \text{on } [0, t_2] \times \mathcal{D}'; \\ &u(t, y) = u_{\partial\mathcal{D}'}(y) && \text{on } (\{0\} \times \mathcal{D}') \cup ([0, t_2] \times \partial\mathcal{D}'). \end{aligned}$$

It is non-negative on $int(\mathcal{D}')$ and it satisfies, for some $C = C(\mathcal{L}, t_1, t_2, \mathcal{D}, \mathcal{D}') > 0$ independent of $u_{\partial\mathcal{D}'}$:

$$\inf_{y\in\mathcal{D}} u(t_2, y) \geq C \sup_{y\in\mathcal{D}} u(t_1, y).$$

Proposition 4.2.1. *Assume that Assumption (H) holds, and $\limsup_{\|x\|\to\infty} r(x) = -\infty$. Then, all the results of Subsection 2.2 hold, and we have in particular exponential convergence in total variation of the MCNE to the unique QSD.*

Assumption (H) is crucial for the proofs of both $(A1)$ and $(A3)$, yet not at all for the one of $(A2)$. As stated in the next proposition, the form of the equation for N is the main ingredient.

Proposition 4.2.2. *Assume that (X, N) is a càdlàg process on $\mathbb{R}^d \times \mathbb{R}_+$ such that N is solution to:*

$$dN_t = (r(X_t) - c\ N_t)\ N_t\ dt + \sigma\ \sqrt{N_t}\ dB_t^N,$$

where B^N is a Brownian motion. Assume that τ_∂ is upper-bounded by $\inf\{t \geq 0; N_t = 0\}$. Provided that $\limsup_{\|x\|\to\infty} r(x) = -\infty$, it holds that for any $\rho > 0$, there exist $n > 0$ such that:

$$\sup_{x\in\mathcal{X}} \mathbb{E}_x \left(\exp\left[\rho\left(\tau_\partial \wedge \tau_E\right)\right]\right) < \infty,$$

where $E := \bar{B}(0, n) \times [1/n, n]$, with $\bar{B}(0, n)$ the closed ball of R^d centered in 0 and of radius n.

The main elements of the proof are given in Subsection 4.2.4, with the most elementary arguments deferred to the Appendix.

From Proposition 4.2.1 to Theorem 4.2. Looking at the system (S) of equations, Assumption (H) holds for the generator:

$$\begin{aligned} \mathcal{L}f(x, n) := [r(x) - c\, n]\, \partial_n f(x, n) + b(x, n)\, \partial_x f(x, n) \\ + n\ \sigma^2/2 \times \Delta_n f(x, n) + \theta^2(x, n)/2 \times \Delta_x f(x, n). \end{aligned}$$

The proof for existence and uniqueness of the solution u for such second-order partial operator with Hölder coefficients and elliptic diffusion coefficient can be found for instance in Corollary 2, Section 4, Chapter 3 of [29]. It also ensures that the solution has two continuous x-derivatives and one continuous t-derivative. The fact it is non-negative is then a consequence of the Maximum principle (cf. e.g. Theorem 1, Chapter 2 in [29]). Finally, the comparison comes from the parabolic Harnack inequality, exploiting the regularity of u and the fact it is non-negative. For its proof, we refer to Theorem 1.1 in [30]. The Harnack

inequality on the open and path-connected set $\mathcal{D}$ is not too difficult to deduce from the local Harnack inequalities these authors provide. One essentially needs to cover $\bar{\mathcal{D}}$ by a finite number of balls included in $\mathcal{D}'$ on which the local inequalities can be applied. For any points $x, y \in \bar{\mathcal{D}}$, we can then construct a path between them for which the number of visited such balls is uniformly bounded. The interval $[t_1, t_2]$ shall then be split into as many time-intervals and the local Harnack inequalities applied recursively to conclude that Assumption (H) holds.

Assumption (H) shall hold more generally, notably under the Hörmander condition instead of the condition of ellipticity (cf e.g. [38]). Lots of articles are dedicated to prove such estimates under various conditions.

4.2.3. Proof of Theorem 4.2.1

In order to conveniently exploit Assumption (H), we choose to consider $(\mathcal{D}_\ell)$ as some sequence (can be anyone) of strictly increasing compact and path-connected sets with C^∞ boundaries whose union is $\mathcal{Y} := \mathbb{R}^d \times \mathbb{R}_+^*$.

Such a sequence clearly satisfies $(A0)$ and any set of the form $\bar{B}(0,n) \times [1/n, n]$ is a subset of $\mathcal{D}_\ell$ for ℓ sufficiently large. $(A2)$ thus also hold for this process.

Assumption (H) with $Y_t = (X_t, N_t)$ implies $(A1)$.

For some non-negative C^∞ function f with support in $\mathcal{D}_\ell$, we apply Assumption (H) with $\mathcal{D} := \mathcal{D}_\ell$, $\mathcal{D}' := \mathcal{D}_{\ell+1}$ with $u_{\partial\mathcal{D}'}(t,y) := f(x)$ on $\{0\} \times \mathcal{D}_\ell$ and $u_{\partial\mathcal{D}'}(t,y) := 0$ on $\mathbb{R}_+^* \times \partial\mathcal{D}_{\ell+1}$. The solution u we obtain is identified thanks to Itô formula as: $u(t,y) := \mathbb{E}_y\left(f(Y_t);\ t < \tau_\partial^{\ell+1}\right)$, with an additional extinction when the process exits $\mathcal{D}_{\ell+1}$. Applying Harnack inequalities implies thus that for any $y \in \mathcal{D}_\ell$, and some reference $y_0 \in \mathcal{D}_1$:

$$\mathbb{E}_y\left(f(Y_{t_2});\ t_2 < \tau_\partial^{\ell+1}\right) \geq c\, \mathbb{E}_{y_0}\left(f(Y_{t_1});\ t_1 < \tau_\partial^1\right).$$

Since it is classical that $\mathbb{P}_{y_0}\left(Y_t \in \mathcal{D}_1;\ t < \tau_\partial^1\right) > 0$, we can obtain a probability measure ζ, independent of ℓ, such that (since c does not depend on f):

$$\forall\, y \in \mathcal{D}_\ell, \quad \mathbb{E}_y\left(Y_t \in dy;\ t < \tau_\partial^{\ell+1}\right) \geq c\, \zeta(dy).$$

Assumption (H) with $Y_t = (X_t, N_t)$ implies $(A3)$.

The proof of $(A3)$ is a bit similar but much more technical because the reference measure is now in the upper-bound, so that we can no longer neglect trajectories exiting $\mathcal{D}'$. W.l.o.g., we consider E to be of the form $\mathcal{D}_\ell$ for ℓ sufficiently large. Since the support of ζ is included in $\mathcal{D}_1$, we wish to prove that there exists $c > 0$ such that for any $y_1 \in \mathcal{D}_1$ and $y_E \in E$:

$$\mathbb{P}_{y_E}\left(Y_t \in dy;\ t < \tau_\partial\right) \leq c\, \mathbb{P}_{y_1}\left(Y_{t_\alpha} \in dy;\ t_\alpha < \tau_\partial\right), \tag{4.14}$$

where we can choose here $0 < t_\alpha < t$ arbitrary (c depending on this choice). (4.14) directly implies $(A3)$ with the functions $f_s(y) = \mathbb{P}_y(s - t_\alpha < \tau_\partial)$, and the Markov property.

In the step 4 of the proof given in Section 4 of [15], N. Champagnat and D. Villemonais used a trick to obtain results such as (4.14). Their idea is to apply the parabolic Harnack

inequality on some regular and compact domain $\mathcal{R}$ such that $E \subset \mathcal{R} \subset \mathcal{X}$ and $d(E, \partial\mathcal{R}) > 0$ while approximating the function:

$$u(t,y) := \mathbb{E}_y \left(f(Y_t);\ t < \tau_\partial \right), \qquad \text{with } \ t \geq t_\mathcal{R},\ y \in \mathcal{R},$$

defined for some non-negative $f \in C^\infty(\mathcal{X})$ and any choice of $0 < t_\mathcal{R} < t_\alpha$. Although we can prove (as they do) that u is continuous, it is a priori not regular enough to apply Harnack inequality directly. Thus, we approximate it on the parabolic boundary $[t_\mathcal{R}, \infty) \times \partial\mathcal{R} \bigcup \{t_\mathcal{R}\} \times \mathcal{R}$ by the family $(U_k)_{k\geq 1}$ of smooth non-negative functions ($\mathcal{C}^\infty_+$ w.l.o.g.). We then deduce approximations of u in $[t_\mathcal{R}, \infty) \times \mathcal{R}$ by (smooth) solutions of:

$$\begin{aligned} &\partial_t u_k(t,y) - \mathcal{L} u_k(t,y) = 0, && t \geq t_\mathcal{R},\ y \in \mathcal{R}^\circ \\ &u_k(t,y) = U_k(t,y), && t \geq t_\mathcal{R},\ y \in \partial\mathcal{R}, \quad \text{or } t = t_\mathcal{R},\ y \in \mathcal{R}. \end{aligned}$$

By Assumption (H), the constant involved in the Harnack inequality does not depend on the values on the boundary. Thus, it applies with the same constant for the whole family of approximations u_k. We refer to the proof in [15] to state that the Harnack inequality then extends to the approximated function u, where the regularity of $u \in C^{1,2}$ is required to apply the Itô formula on the process $u(t-s, Y_s)$. Thus, (4.14) indeed holds (where we could have chosen any $y \in E$).

Now we have concluded that Assumption $(\mathbf{A})$ holds, so that the conclusions of Theorems 2.1 to 2.3 hold. □

4.2.4. Proof of Proposition 4.2.2: escape from the transitory domain

The purpose of this section is to demonstrate how to prove $(A2)$ when, depending on the position in the transitory domain, there are various reasons for a quick escape. To combine several local estimates, dealing with suprema in the initial condition of exponential moments appear much more convenient than Lyapunov estimates, see Appendix A. Moreover, these exponential moments can be naturally deduced from probabilities of retention (or transfer) in the transitory domain for a finite given time, see Appendix B, C or D.

Decomposition of the transitory domain

In our choice for E, with three parameters $n_0 < n_\infty < n_E$ to be fixed, its complementary $\mathcal{T} = \mathcal{X} \setminus E$ is made up of 3 subdomains: "$y = \infty$", "$y = 0$", and "$\|x\| = \infty$", according to figure 1. They are formally defined as follow:

- $\mathcal{T}^N_\infty := \left\{\mathbb{R}^d \setminus B(0, n_E)\right\} \times (n_\infty, \infty) \bigcup\ B(0, n_E) \times (n_E, \infty)$ ("$y = \infty$"),
- $\mathcal{T}_0 := B(0, n_E) \times [0, n_0]$ ("$y = 0$"),
- $\mathcal{T}^X_\infty := \left\{\mathbb{R}^d \setminus B(0, n_E)\right\} \times (n_0, n_E]$ ("$\|x\| = \infty$").

Essentially, we will need to choose n_∞ sufficiently large to have the property of descent from infinity for $\mathcal{T}^N_\infty$; $n_E > n_\infty$ sufficiently large to have a growth rate so low that the population cannot maintain itself in $\mathcal{T}^X_\infty$; n_0 sufficiently small to prove that the population can hardly survive after entering $\mathcal{T}_0$. Thus, the process will escape each region with an exponential moment. Yet, we also need to prove that the process will not circulate between

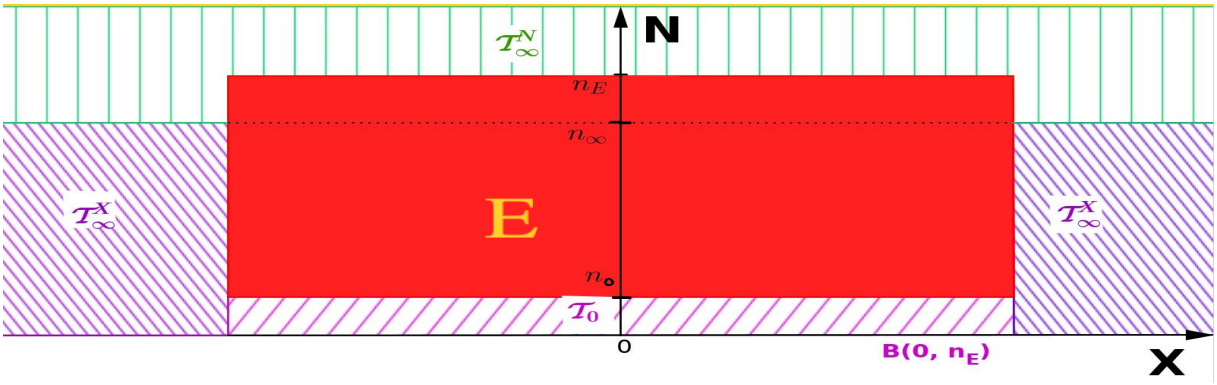


Figure 1: subdomains for $(A2)$

the different transitory areas. Therefore, we will set these areas such that at least some of the transitions (those associated with an increase of the population size) happens with so low probability that Theorem 1 holds true.

For each of these domains, we define the following exponential moments that we shall relate by specific inequalities. Let $t_h > 0$ (a threshold needed to ensure the boundedness) and $\widehat{\tau}_E := \tau_E \wedge \tau_\partial \wedge t_h$ (remember that τ_E is the hitting time of E):

- $\mathcal{E}_\infty^N := \sup\{\mathbb{E}_{(x,n)}[\exp(\rho\, \widehat{\tau}_E)];\ (x,n) \in \mathcal{T}_\infty^N\}$,
- $\mathcal{E}_\infty^X := \sup\{\mathbb{E}_{(x,n)}[\exp(\rho\, \widehat{\tau}_E)];\ (x,n) \in \mathcal{T}_\infty^X\}$,
- $\mathcal{E}_0 := \sup\{\mathbb{E}_{(x,n)}[\exp(\rho\, \widehat{\tau}_E)];\ (x,n) \in \mathcal{T}_0\}$.

Implicitly, we assume ρ to be given. Then, $\mathcal{E}_\infty^N$, $\mathcal{E}_\infty^X$ and $\mathcal{E}_0$ can be seen as functions of n_0, n_∞ and n_E. These values are to be specified depending on ρ for the proof of $(A2)$ to hold. The dependency in t_h shall be negligible as $t_h \to \infty$.

A set of inequalities associating the local bounds

The local exponential moments that we introduce are related thanks to the three following propositions, obtained from local bounds mentioned in the following three lemmas. We refer to Appendices A, B, C and D to see first how to deduce $(A2)$ from the three propositions that follow, and then respectively for the (technical) proofs of the propositions (including the lemmas):

Proposition 4.2.3. *Given any $\rho > 0$, we can define $n_\infty > 0$ and $C \geq 1$ such that, whatever $n_E > n_\infty$ and $t_h > 0$:* $\mathcal{E}_\infty^N \leq C\ \left(1 + \mathcal{E}_\infty^X\right)$.

Proposition 4.2.4. *Given any ρ, ϵ, $n_\infty > 0$, we have, for some $C \geq 1$ (in fact independent of any parameters), and any n_E sufficiently large and $t_h > 0$:*
$\mathcal{E}_\infty^X \leq C\ (1 + \mathcal{E}_0) + \epsilon\, \mathcal{E}_\infty^N$.

Proposition 4.2.5. *Given any ρ, ϵ, $n_\infty > 0$, we have, for some $C \geq 1$, any n_0 sufficiently small, any $n_E \geq n_\infty$ and any $t_h > 0$:* $\mathcal{E}_0 \leq C + \epsilon \left(\mathcal{E}_\infty^N + \mathcal{E}_\infty^X\right)$.

The associated elementary bounds on finite time

The main ingredient for these propositions are simple comparison properties that are specific to each of the transitory domain. By focusing on each of the domains separately (with the transitions between them), we can highly simplify our control on the dependency of the processes. Specific autonomous one-dimensional processes indeed act as upper-bounds for each of the domains. The values of (X_t) do not affect these auxilliary processes, but only the regions on which they act as upper-bounds.

Propositions 4.2.3 and 4.2.4 are deduced from the estimates given in Lemmas 4.2.6 and 4.2.7 on autonomous processes of the form:

$$N_t^D := n + \int_0^t (r - c\, N_s^D)\, N_s^D\, ds + \int_0^t \sigma\, \sqrt{N_s^D}\, dB_s. \tag{4.15}$$

Propositions 4.2.3 relies on the property of descent from infinity valid for any value of r:

Lemma 4.2.6. *Let N^D be the solution of (4.15), for some $r \in \mathbb{R}$ and $c > 0$, with n the initial condition. Then, for any $t, \epsilon > 0$ there exists $n_\infty > 0$ such that:*

$$\sup_{n>0} \mathbb{P}_n(t < \tau_\downarrow^D) \leq \epsilon \qquad \text{with } \tau_\downarrow^D := \inf\left\{s \geq 0,\ N_s^D \leq n_\infty\right\}.$$

Proposition 4.2.4 relies on the strong negativity on the drift term:

Lemma 4.2.7. *Considering any $c, t > 0$, with $\tau_\partial^D := \inf\left\{t \geq 0,\ N_t^D = 0\right\}$:*

$$\sup_{n>0} \mathbb{P}_n\left(t < \tau_\partial^D\right) \underset{r\to-\infty}{\longrightarrow} 0.$$

Moreover, for any $n, \epsilon > 0$, there exists n_c such that, for any r sufficiently low, with $T_\infty^D := \inf\left\{t \geq 0,\ N_t^D \geq n_c\right\}$: $\mathbb{P}_n\left(T_\infty^D \leq t\right) + \mathbb{P}_n\left(N_t^D \geq n\right) \leq \epsilon$.

Finally, Proposition 4.2.5 relies on an upper-bound given as a Continuous State Branching Process, for which the extinction rate is much more explicit. It is clearly as strong as needed for sufficiently small initial condition.

We recall that the complete proofs (of Proposition 4.2.2 from the propositions and of the propositions themselves) are deferred to respectively Appendices A, B, C and D. With this, we have concluded the proof of Proposition 4.2.2.

5. Proof of Theorems 2.1-3

In Subsection 5.3, we present the general principles of our coupling that concludes the proof of Theorem 2.1. These principles would alone end the proof in the context of the Assumption (A) in [13]. Yet, with our more general assumptions, these principles require the results of the two previous subsections. First, we prove in Subsection 5.1 that the MCNE will keep in the long run some mass on some specific set $\mathcal{D}_\circ$ (which is weaker but related in some sense to the tension of the laws); then we prove in Subsection 5.2 that $(A3)$ holds in fact for $\mathcal{X}$ instead of just E. At the end of Subsection 5.3, the proof of Theorem 2.1 is then complete. The following Subsection 5.4 and 5.5 then prove respectively Theorem 2.2 and 2.3.

5.1. Stabilization of the process in the long run

The main purpose of this section is to prove:

Theorem 5.1. *Assume that* (**A**) *holds. Then, there exists* $\mathcal{M}_\circ = \mathcal{M}_{\ell_\circ,\xi_\circ}$ *(with* $\ell_\circ \geq 1$, $\xi_\circ > 0$*) such that for any* $\ell \geq 1$ *and* $\xi \in (0,1)$*, there exists* $t_\circ = t_\circ(\ell,\xi) > 0$ *such that:*

$$\forall\, \mu \in \mathcal{M}_{\ell,\xi},\ \forall\, t \geq t_\circ, \qquad \mu A_t \in \mathcal{M}_\circ. \tag{5.1}$$

Remark 5.1.1. *Assumption* $(A3)$ *is not involved in the proof of Theorem 5.1. This will be important in [46] since we will exploit Theorem 5.1 to provide an alternative criterion to* $(A3)$*.*

Proof of Theorem 5.1:

According to $(A2)$, let $\ell_E \geq 1$ and $\rho_E > \rho_S$ be such that:

$$\begin{aligned} &\text{with } E := \mathcal{D}_{\ell_{_E}},\ \tau_E^1 := \inf\{t \geq 0;\ X_t \in E\}, \\ &e_{\mathcal{T}} := \sup_x \mathbb{E}_x\left(\exp\left[\rho_E\,(\tau_E^1 \wedge \tau_\partial)\right]\right) < \infty. \end{aligned} \tag{5.2}$$

From (2.1), i.e. the definition of ρ_S, there exists $\widetilde{\rho}_S \in (\rho_S, \rho_E)$, $c_S > 0$ and $\ell_S \geq 1$, such that:

$$\forall\, t \geq 0,\ \mathbb{P}_\zeta(t < \tau_\partial \wedge T_{\mathcal{D}_{\ell_S}}) \geq c_S\ \exp(-\widetilde{\rho}_S\, t). \tag{5.3}$$

We then apply $(A1)$ with $\ell = \ell_{_E}$ to state that there exists $L_C \geq \ell_S \vee \ell_E$, $t_C, c_C > 0$ such that, with $\mathcal{D}_C := \mathcal{D}_{L_C}$:

$$\forall\, x \in E, \quad \mathbb{P}_x\Big[X_{t_C} \in dy,\ t_C < \tau_\partial \wedge T_{\mathcal{D}_C}\Big] \geq c_C\, \zeta(dy). \tag{5.4}$$

W.l.o.g., we are allowed to replace, in the following usage of (5.3), $\mathcal{D}_{\ell_S}$ by $\mathcal{D}_C$.

In order to conclude the proof of Theorem 5.1, we need the following three Lemmas, for which we define by induction over $i \in \mathbb{N}$:

$$\begin{aligned} &T_C^i := \inf\left\{t \geq \tau_E^i;\ X_t \notin \mathcal{D}_C\right\},\ T_C^0 := 0, \\ &\tau_E^{i+1} := \inf\left\{t \geq T_C^i;\ X_t \in E\right\}. \end{aligned}$$

Lemma 5.1.2. ***First entry in*** E***:*** *Assume that* (5.2), $(A1)$ *and* (5.3) *hold. Then, for any* ℓ, ξ*, there exists* $C_E = C_E(\ell,\,\xi) > 0$ *such that:*

$$\forall\, t_h > 0,\ \forall\, \mu \in \mathcal{M}_{\ell,\xi}, \quad \mathbb{P}_\mu(t_h \leq \tau_E^1 \,|\, t_h < \tau_\partial) \leq C_E\, e^{-(\rho_E - \widetilde{\rho}_S)\, t_h}.$$

Lemma 5.1.3. ***Containment of the process after*** T_C^i***:***
Suppose that (5.3) *and* $(A1)$ *hold. Then, there exists* $\ell_\circ \geq L_C$*, and* $c_\circ > 0$ *such that:*

$$\forall\, x \in \mathcal{D}_C,\ \forall\, t > 0, \qquad \mathbb{P}_x\left(t < T_C^1 \wedge T_{\mathcal{D}_{\ell_\circ}} \wedge \tau_\partial\right) \geq c_\circ\, \exp[-\widetilde{\rho}_S\, t].$$

Lemma 5.1.4. ***Last exit from*** $\mathcal{D}_C$***:***
Suppose that $(A0)$, $(A1)$, (5.2), (5.3) *hold (with* $E \subset \mathcal{D}_C$*) and* $\rho_E > \widetilde{\rho}_S$*. Then, there exists* $C_L > 0$*, such that for any* $\mu \in \mathcal{M}_1(\mathcal{X})$ *with* $t_h > t > 0$*:*

$$\begin{aligned} &\mathbb{P}_\mu\left(T_C^{I(t_h)} \leq t_h - t,\ t_h \leq \tau_E^{I(t_h)+1},\ \tau_E^1 < t_h \,\middle|\, t_h < \tau_\partial\right) \leq C_L\, e^{-(\rho_E - \widetilde{\rho}_S)\, t}, \\ &\text{with } I(t_h) := \max\left\{i \geq 0;\ T_C^i \leq t_h\right\} (< \infty \ a.s.). \end{aligned}$$

The proofs of these Lemmas are deferred, in the order of occurrence, after the proof that they imply Theorem 5.1.

5.1.1. Proof that Lemmas 5.1.2-3 imply Theorem 5.1

With Lemma 5.1.2 and 5.1.4, we obtain an upper-bound (with high probability) on how much time the process may have spent outside $\mathcal{D}_C$. Thus, we can associate most of trajectories ending outside $\mathcal{D}_C$ to others ending inside $\mathcal{D}_C$. From this association, we deduce a lower-bound on the probability to see the process in $\mathcal{D}_C$.

Let us first define $\mathcal{D}_\circ$ according to Lemma 5.1.3. In the following, we will define: $\mathcal{M}_\circ := \{\mu \in \mathcal{M}_1(\mathcal{X});\ \mu(\mathcal{D}_\circ) \geq \xi_\circ\}$ for a well-chosen $\xi_\circ$. Thanks to Lemma 5.1.4, we choose some $t > 0$ sufficiently large to ensure: $\forall\, t_h > t,\ \forall\, \mu \in \mathcal{M}_1(\mathcal{X})$,

$$\mathbb{P}_\mu\left(T_C^{I(t_h)} \leq t_h - t,\ t_h \leq \tau_E^{I(t_h)+1},\ \tau_E^1 < t_h \,\middle|\, t_h < \tau_\partial\right) \leq {}^1\!/_4\ . \tag{5.5}$$

Let $\ell \geq 1$, $\xi \in (0,1)$. Thanks to Lemmas 5.1.2, we know that for some $t_\circ \geq t > 0$:

$$\forall\, t_h \geq t_\circ,\ \forall\, \mu \in \mathcal{M}_{\ell,\xi}, \quad \mathbb{P}_\mu\left(t_h \leq \tau_E^1 \,\middle|\, t_h < \tau_\partial\right) \leq {}^1\!/_4\ . \tag{5.6}$$

Let $\mu \in \mathcal{M}_{\ell,\xi}$. Let us first assume that:

$$\mathbb{P}_\mu\left(\tau_E^{I(t_h)+1} \leq t_h \,\middle|\, t_h < \tau_\partial\right) \geq {}^1\!/_4\ . \tag{5.7}$$

By definition of $I(t_h)$, on the event $\left\{\tau_E^{I(t_h)+1} \leq t_h\right\} \bigcap \{t_h < \tau_\partial\}$, we know that the process stays in $\mathcal{D}_C$ in the time-interval $[\tau_E^{I(t_h)+1}, t_h]$. In particular:

$$\mu A_{t_h}(\mathcal{D}_\circ) \geq \mu A_{t_h}(\mathcal{D}_C) \geq \mathbb{P}_\mu\left(\tau_E^{I(t_h)+1} \leq t_h \,\middle|\, t_h < \tau_\partial\right) \geq {}^1\!/_4\ . \tag{5.8}$$

where we recall that $\ell_\circ \geq L_C$ by Lemma 5.1.3.

Now that this case has been easily treated, we consider the complementary:

$$\mathbb{P}_\mu\left(\tau_E^{I(t_h)+1} \leq t_h \,\middle|\, t_h < \tau_\partial\right) < {}^1\!/_4\ .$$

Thus, by (5.5) and (5.6): $\mathbb{P}_\mu\left(t_h - t < T_C^{I(t_h)},\ \tau_E^1 \leq t_h \,\middle|\, t_h < \tau_\partial\right) \geq {}^1\!/_4$. By defining the stopping time: $\tau_C := \inf\{s \geq t_h - t;\ X_s \in \mathcal{D}_C\}$, we deduce:

$$\mathbb{P}_\mu\left(t_h - t \leq \tau_C < t_h \,\middle|\, t_h < \tau_\partial\right) \geq {}^1\!/_4\ . \tag{5.9}$$

By the Markov property, then Lemma 5.1.3:

$$\begin{aligned}
&\mathbb{P}_\mu\left(X_{t_h} \in \mathcal{D}_\circ,\ t_h - t \leq \tau_C < t_h,\ t_h < \tau_\partial\right)\\
&\quad \geq \mathbb{E}_\mu\left[\mathbb{P}_{X_{\tau_C}}\left(\widetilde{X}_{t_h-\tau_C} \in \mathcal{D}_\circ,\ t_h - \tau_C < \widetilde{\tau}_\partial\right);\ t_h - t \leq \tau_C < t_h \wedge \tau_\partial\right]\\
&\quad \geq c_\circ \exp[-\widetilde{\rho}_S t]\, \mathbb{P}_\mu\left[t_h - t \leq \tau_C < t_h \wedge \tau_\partial\right]\\
&\quad \geq c_\circ \exp[-\widetilde{\rho}_S t]\, \mathbb{P}_\mu\left[t_h - t \leq \tau_C < t_h \,\middle|\, t_h < \tau_\partial\right] \times \mathbb{P}_\mu\left[t_h < \tau_\partial\right].
\end{aligned}$$

So (5.9) indeed implies $\mu A_{t_h}(\mathcal{D}_\circ) \geq \xi_\circ$ with $\xi_\circ := c_\circ\, e^{-\widetilde{\rho}_S\, t}/4$. With $\mathcal{M}_\circ := \{\mu \in \mathcal{M}_1(\mathcal{X});\ \mu(\mathcal{D}_\circ) \geq \xi_\circ\}$ ($\xi_\circ$ given by the previous formula does not depend on ℓ, ξ or μ), we indeed prove (5.1) for the case where (5.7) does hold false. Recall that the first case where (5.7) does hold true is directly concluded in (5.8). □

5.1.2. Proof of Lemma 5.1.2

By (5.2) and the Markov inequality:

$$\forall\, \mu,\ \forall\, t_h > 0, \quad \mathbb{P}_\mu\left(t_h \le \tau_E^1 \wedge \tau_\partial\right) \le e_{\mathcal{T}}\, e^{-\rho_E\, t_h}. \tag{5.10}$$

Let $\ell \ge 1$, $\xi \in (0,1)$. We apply $(A1)$, (5.3) and the Markov property to obtain that there exists $c > 0$ (that we can decompose as $c = \xi\, c_S\, c_M(\ell)$ for some constant $c_M(\ell) > 0$ only depending on ℓ) such that: $\forall\, \mu \in \mathcal{M}_{\ell,\xi},\ \forall\, t_h > 0$,

$$\mathbb{P}_\mu\left(t_h < \tau_\partial\right) \ge c\, e^{-\widetilde{\rho}_S\, t_h}. \tag{5.11}$$

Thus, by (5.10), (5.11), with: $C_E := e_{\mathcal{T}}/c > 0$,

$$\forall\, \mu \in \mathcal{M}_{\ell,\xi},\ \forall\, t_h > 0, \quad \mathbb{P}_\mu\left(t_h \le \tau_E^1 \,\middle|\, t_h < \tau_\partial\right) \le C_E\, \exp[-(\rho_E - \widetilde{\rho}_S)\, t_h]. \qquad \square$$

5.1.3. Proof of Lemma 5.1.3

Thanks to $(A1)$, applied with $\ell = L_C$, there exists some $\mathcal{D}_\circ$, $t_E, c_E > 0$ such that:

$$\forall\, x \in \mathcal{D}_C, \quad \mathbb{P}_x\left(\tau_E^1 \le t_E \wedge T_{\mathcal{D}_\circ} \wedge \tau_\partial\right) \ge c_E. \tag{5.12}$$

Recalling (5.4), we deduce that conditionally on $\mathcal{F}_{\tau_E^1}$ and on the event $\left\{\tau_E^1 \le t_E \wedge T_{\mathcal{D}_\circ}\right\}$:

$$\mathbb{P}_{X_{\tau_E^1}}\left(\widetilde{X}_{t_C} \in dx,\ t_C < \widetilde{T_C^1} \wedge \widetilde{T}_{\mathcal{D}_\circ} \wedge \widetilde{\tau}_\partial\right) \ge c_C\, \zeta(dx). \tag{5.13}$$

By combining (5.12), (5.13), (5.3) and the Markov property, we deduce that:

$$\mathbb{P}_x\left(t_h < T_C^1 \wedge T_{\mathcal{D}_\circ} \wedge \tau_\partial\right) \ge c_\circ\, \exp[-\widetilde{\rho}_S\, t_h],$$
$$\text{with } c_\circ := c_E\, c_C\, c_S\, \exp[\widetilde{\rho}_S\, (t_E + t_C)] > 0. \qquad \square$$

5.1.4. Proof of Lemma 5.1.4

The idea is to use that it is very unlikely for the process to still be alive after experiencing an excursion outside E for a long time (and still be there). Indeed, compared to trajectories that stay inside E (in particular those reaching quickly $\mathcal{D}_S$, for which (3.1) holds, and not leaving $\mathcal{D}_L$) the probabilities of the associated events vanish with a larger rate: $\rho_E > \widetilde{\rho}_S$. It would have been convenient if we could initiate the comparison just before $T_C^{I(t_h)}$, where the process exits E for the last time before t_h. Yet, it is not a stopping time, so that the Markov property is not directly applicable and the proof gets more technical.

Let us first prove that $I(t_h) < \infty$. Since X has càdlàg paths, we would have on the event $\{I(t_h) = \infty\}$: $\sup_j T_C^i = \sup_j \tau_E^i = T < t_h$ with $X_{T-} \in E \cap \overline{\mathcal{X} \setminus \mathcal{D}_C}$. Yet, by $(A0)$, this set is empty, so that a.s. $I(t_h) < \infty$. Then, exploiting a discretization of time in time-intervals of length t_L to be fixed later:

$$\begin{aligned}
P &:= \mathbb{P}_\mu\left(T_C^{I(t_h)} \le t_h - t,\ t_h \le \tau_E^{I(t_h)+1} \wedge \tau_\partial,\ \tau_E^1 < t_h\right) \\
&= \textstyle\sum_{\{i \ge 1\}} \mathbb{P}_\mu\left(T_C^i \le t_h - t,\ t_h \le \tau_E^{i+1} \wedge \tau_\partial\right) \\
&\le \textstyle\sum_{\{i \ge 1\}} \sum_{\{k \ge 0\}} \mathbf{1}_{\{k\, t_L \le t_h - t\}} \mathbb{P}_\mu\left(T_C^i \in (k\, t_L, (k+1)t_L],\ t_h \le \tau_E^{i+1} \wedge \tau_\partial\right) \\
&= \textstyle\sum_{\{i \ge 1\}} \sum_{\{k \ge 0\}} \mathbf{1}_{\{k\, t_L \le t_h - t\}} \mathbb{E}_\mu\big[\mathbb{P}_{X_{(k+1)t_L}}\left(t_h - (k+1)t_L \le \widetilde{\tau}_E^1 \wedge \widetilde{\tau}_\partial\right); \\
&\qquad\qquad T_C^i \in (k\, t_L, (k+1)t_L],\ (k+1)t_L \le \tau_E^{i+1} \wedge \tau_\partial\big],
\end{aligned}$$

where we used the Markov property. Exploiting (5.2):

$$P \leq e_T \sum_{\{k\geq 0\}} \mathbf{1}_{\{k\, t_L \leq t_h - t\}} \exp[-\rho_E\,(t_h - (k+1)t_L)] \\ \times \sum_{\{i\geq 1\}} \mathbb{P}_\mu\left[T_C^i \in (k\, t_L, (k+1)t_L],\ (k+1)t_L \leq \tau_E^{i+1} \wedge \tau_\partial\right]. \qquad (5.14)$$

The trick is to observe that, by definitions of $\tau_E^i < T_C^i$, one shall have $X_s \in \mathcal{D}_C$ for any $s \in [\tau_E^i, T_C^i)$, in particular on some vicinity to the left of T_C^i. Defining for $k \geq 0$:

$$\tau_C^k := \inf\{s \geq k t_L : X_s \in \mathcal{D}_C\},$$

we see that the events $\left\{T_C^i \in (k\, t_L, (k+1)t_L]\right\} \cap \left\{(k+1)t_L \leq \tau_E^{i+1} \wedge \tau_\partial\right\}$ are disjoint (for k fixed) and included in the event $\left\{\tau_C^k < (k+1)t_L \wedge \tau_\partial\right\}$. On the other hand, exploiting the Markov property together with Lemma 5.1.3:

$$\mathbb{P}_\mu\left[t_h < \tau_\partial\right] \geq c_\circ\, \exp[-\widetilde{\rho}_S(t_h - k\, t_L)]\, \mathbb{P}_\mu\left[\tau_C^k < (k+1)t_L \wedge \tau_\partial\right].$$

Coming back to (5.14), we deduce:

$$P \leq \frac{e_T\, e^{\rho_E\, t_L}}{c_\circ} \mathbb{P}_\mu\left[t_h < \tau_\partial\right] \times \sum_{\{k\geq 0\}} \mathbf{1}_{\{k\, t_L \leq t_h - t\}} \exp[-(\rho_E - \widetilde{\rho}_S) \times (t_h - k t_L)].$$

The sum over k is upper-bounded by:

$$\exp[-(\rho_E - \widetilde{\rho}_S)t] \times \sum_{\{\ell\geq 0\}} \exp[-\ell(\rho_E - \widetilde{\rho}_S)t_L] \leq \frac{e^{-(\rho_E - \widetilde{\rho}_S)t}}{1 - e^{-(\rho_E - \widetilde{\rho}_S)t_L}}.$$

This concludes the proof of the Lemma, with: $C_L := \dfrac{e_T\, e^{\rho_E\, t_L}}{c_\circ(1 - e^{-(\rho_E - \widetilde{\rho}_S)t_L})}$. The choice of t_L is free, so that we can fix it to optimize this constant. □

5.2. Persistence

5.2.1. Theorem 5.2

For the proof of the following Theorem 5.2, we need the following Corollary of Theorem 5.1:

Corollary 5.2.1. ***"Stability"****:*
Under Assumption (**A**)*, there exists* $t_S, c'_S > 0$ *and* $\widetilde{\rho}_S \in (\rho_S, \rho_E)$ *such that:*

$$\forall\, u \geq 0,\ \forall\, t \geq u + t_S, \qquad \mathbb{P}_\zeta(t - u < \tau_\partial) \leq c'_S\, e^{\widetilde{\rho}_S u}\, \mathbb{P}_\zeta\left(t < \tau_\partial\right). \qquad (5.15)$$

Theorem 5.2. *Assume that there exists* $\rho_E > \widetilde{\rho_S}$*,* $\mathcal{D}_S \subset \mathcal{X}$*,* $E \subset \mathcal{X}$ *and* $\zeta \in \mathcal{M}_1(\mathcal{X})$ *such that* (*A*3)*,* (*A*2)*,* (5.3) *and* (5.15) *hold. Then, there exists* $t_P, c_P > 0$ *such that:*

$$\forall\, x \in \mathcal{X},\ \forall\, t \geq t_P, \quad \mathbb{P}_x\left(t < \tau_\partial\right) \leq c_P\, \mathbb{P}_\zeta\left(t < \tau_\partial\right). \qquad (5.16)$$

5.2.2. Proof of Corollary 5.2.1:

By (5.1) and $(A1)$, there exists $c > 0$ such that for any v sufficiently large: $\zeta A_v \geq c\,\zeta$. with the Markov property, it implies for any $u \geq 0$:

$$\mathbb{P}_\zeta(v + u < \tau_\partial) \geq c\,\mathbb{P}_\zeta(v < \tau_\partial)\,\mathbb{P}_\zeta(u < \tau_\partial).$$

Exploiting (5.3) with $t = u$, we deduce Corollary 5.2.1 with $v = t - u$ and $c'_S = (c\,c_S)^{-1}$. □

5.2.3. Proof of Theorem 5.2

From $(A3)$, there exists $t_A, c_A > 0$ such that:

$$\forall\, t \geq t_A,\ \forall\, x \in E, \quad \mathbb{P}_x(t < \tau_\partial) \leq c_A\,\mathbb{P}_\zeta(t < \tau_\partial). \tag{5.17}$$

This proof is very close to the one in [15] (p13:"Step 2: Proof of $(A1)$"), except that, in (5.15), $t - u$ shall be larger than some value, and similarly for t in (5.17). To compare the notations, our $e_{\mathcal{T}}$, c_A and c_S refer resp. to their M, C_m and $4/c_1\,D_m\,D_{n_1}$. Thus, we won't detail it much and refer to [15].

Let $\zeta \in \mathcal{M}_1(\mathcal{X})$, $t \geq t_P := t_S \vee t_A$ and $x \in \mathcal{X}$.

$$\mathbb{P}_x(t < \tau_\partial) \leq c_A\,\mathbb{E}_x\left[\mathbb{P}_\zeta(t - \tau_E < \widetilde{\tau}_\partial);\ \tau_E < (t - t_P) \wedge \tau_\partial\right] + \mathbb{P}_x(t - t_P \leq \tau_E \wedge \tau_\partial) \tag{5.18}$$

thanks to property $(A3)$, since $t - \tau_E \geq t_P \geq t_A$ on $\{\tau_E < (t - t_P) \wedge \tau_\partial\}$. By $(A2)$ (with the Markov inequality) and Corollary 5.2.1, with $u = t_A$ for the first term of (5.18) and $u = t - t_S$ for the second: $\forall t \geq 0$, $\forall x \in \mathcal{X}$,

$$\mathbb{P}_x(t < \tau_\partial) \leq \left(c_A + e^{\widetilde{\rho}_S(t_P - t_S)} \,/\, \mathbb{P}_\zeta\left(t_S < \tau_\partial\right)\right) \times c'_S\,e_{\mathcal{T}}\ \mathbb{P}_\zeta(t < \tau_\partial) \qquad \square$$

5.3. Coupling procedure: proof of Theorem 2.1

5.3.1. Definition of the uncoupled part

With a given set of parameters t_D, c_D, t_P, $c_P > 0$ (cf following subsection) we define for $t_h > t_P$:

$$J(t_h) := \lfloor (t_h - t_P)/t_D \rfloor\,. \tag{5.19}$$

For $t \geq 0$, $\mu \in \mathcal{M}_1\left(\mathcal{X}\right)$, $t_h > t_P$, and $k \in \mathbb{N}$, let:

$$\begin{aligned} a(k,t) = a_\mu^{t_h}(k,t) := \mathbf{1}_{\{k \leq J(t_h),\, k\,t_D \leq t\}} \times {}^{c_D}\!/_{c_P}\ \left(1 - {}^{c_D}\!/_{c_P}\right)^{k-1} \\ \times \frac{\mathbb{P}_\mu(t_h < \tau_\partial)}{\mathbb{P}_\mu(t < \tau_\partial)} \times \frac{\mathbb{P}_\zeta(t - k\,t_D < \tau_\partial)}{\mathbb{P}_\zeta(t_h - k\,t_D < \tau_\partial)}. \end{aligned} \tag{5.20}$$

Remark 5.3.1. *As we can see in the proof of Fact 5.3.8, $a(k,t)$ corresponds to the mass associated with the k-th step of coupling, considered at time t with the constraint that it must represent a fixed proportion of μA_{t_h} (at time t_h). We refer to Figure 2 for a presentation of the coupling architecture.*

Let $r_j := 1 - \sum_{\{k \le j\}} a(k, j\, t_D)$. Under the condition $r_j > 0$, that we will prove to be true by induction over $j \le J(t_h)$, we define:

$$\nu_j(dx) := ({}^1/_{r_j}) \times \left[\mu A_{j\, t_D}(dx) - \sum_{\{k \le j\}} a(k, j\, t_D)\, \zeta A_{(j-k)\, t_D}(dx) \right], \tag{5.21}$$

with the convention $\nu_0 := \mu$. Remark that this definition ensures $\nu_j(\mathcal{X}) = 1$.

Remark 5.3.2. *ν_j shall correspond to the marginal of the process conditioned of not being already coupled at time $j\, t_D$. We normalize what remains of $\mu A_{j\, t_D}$ when we subtract the contribution of each coupling step (only those up to the j-th will contribute to the sum). The main difficulty will be to prove that, under suitable conditions, ν_j is indeed a positive measure, thus a probability measure. In Figure 2, the associated coupling procedure is presented for the more general case where we compare two initial conditions in some $\mathcal{M}_{\ell,\xi}$ rather than already in $\mathcal{M}_R$.*

Remark 5.3.3. *The procedure extends to cases where the QSD is not unique provided that $\mu^{(1)}$ and $\mu^{(2)}$ are in the same basin of attraction, as we can see in [48]. This procedure already deals with an inhomogeneity in time due to the conditioning for survival at time t_h, so that adaptations to time-inhomogeneous processes are likely to be easy.*

Explanation of the procedure as presented in Figure 5.3.1: Since both initial conditions belong to the same $\mathcal{M}_{\ell,\xi}$, the time $t_\circ = t_\circ^{\ell,\xi}$ needed to reach $\mathcal{M}_R$ can be chosen uniformly. Then, after every time-interval of size t_D and as long as time $t_h - t_P$ is not reached, we shall exploit property (5.22). We split the "remaining MCNE" (the ν_j after j splitting steps) in order to extract a component whose contribution to the MCNE at time t_h is explicit. This contribution (in the expression of $\zeta[t_h - t_\circ^{\ell,\xi}]$) is proportional to $\zeta A_{t_h - t}$ for a splitting at time t. For this contribution at time t_h to be fixed, note that the contribution to the MCNE at time t has to depend both on the remaining time $t_h - t$ and the specific value of ν_j.

5.3.2. Definition of the constants involved

For clarity, we denote by t_h (for horizon of time) the time t that appears in Theorem 2.1. During this coupling procedure, it will stay fixed, and won't appear in the other sections. The constants $c_P, t_P > 0$ come from Theorem 5.2, while $c_D, t_D > 0$ come from this corollary of Theorem 5.1:

Proposition 5.3.4. ***"Coupling and Renewal"***
Suppose that (A1) holds and (5.1) also for some $\mathcal{M}_\circ := \mathcal{M}_{\ell_\circ, \xi_\circ}$. Then, with $\ell_R := \ell_\circ$, $\xi_R := \xi_\circ/2$, $\mathcal{M}_R := \mathcal{M}_{\ell_R, \xi_R}$, $\mathcal{D}_R := \mathcal{D}_{\ell_R}$, there exits $c_D \in (0,1)$ and $t_D \ge t_\circ(\ell_R, \xi_R)$ such that:

$$\forall\, \mu \in \mathcal{M}_R, \qquad \mu A_{t_D}(dx) \ge c_D\, \zeta(dx) \text{ and } \frac{\mu A_{t_D}(\mathcal{D}_R) - c_D}{1 - c_D} \ge \xi_R. \tag{5.22}$$

Remark 5.3.5. *The subscript D refers to "Doeblin's" condition, since we will likewise iteratively couple a proportion at most c_D of the distribution. The properties (5.22) and (5.16) make us able to prove the induction: $\nu_j \in \mathcal{M}_R \Rightarrow \nu_{j+1} \in \mathcal{M}_R$.*

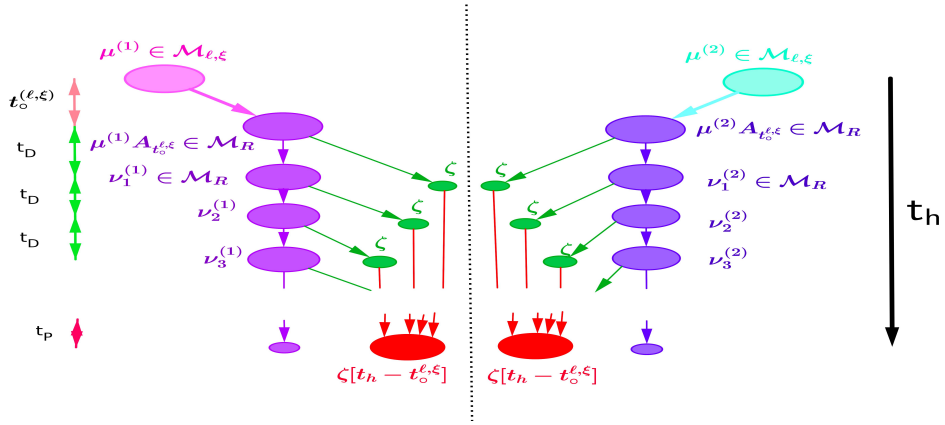


Figure 2: Illustration of the coupling procedure

The figure illustrates the coupling procedure on two initial conditions $\mu^{(1)}$ and $\mu^{(2)}$. We can observe by symmetry how the MCNE are progressively decomposed with time descending along the vertical axis. By construction, the middle red parts (at time t_h) are common for both initial conditions (both its distribution $\zeta[t_h - t_\circ^{\ell,\xi}]$ and the amount of mass).

Proof of Proposition 5.3.4.

We apply (5.1) with $\ell = \ell_R$ and $\xi = \xi_R$. Thus, with $t_R := t_\circ(\ell_R, \xi_R)$:

$$\forall\, \mu \in \mathcal{M}_R,\ \forall\, t \geq t_R, \qquad \mu A_t \in \mathcal{M}_\circ,\ \text{i.e. } \mu A_t(\mathcal{D}_R) \geq 2\,\xi_R \tag{5.23}$$

We can then define $c_M \in (0,\,1)$, $t_M \geq t_R$ thanks to $(A1)$, cf Subsection 3.3, such that:

$$\forall\, x \in \mathcal{D}_R,\ \mathbb{P}_x\left[X_{t_M} \in dx;\ t_M < \tau_\partial\right] \geq c_M\, \zeta(dx). \tag{5.24}$$

We can then choose $c_D := c_M\ \xi_R \in (0,1)$, $t_D := t_M \geq t_R$ (for the statement of the proposition), and observe that:

$$\begin{aligned}
\forall\, \mu \in \mathcal{M}_R,\ \mu A_{t_M}(dx) &\geq \mu(\mathcal{D}_R)\, c_M\, \zeta(dx) \qquad \text{by (5.24)} \\
&\geq \xi_R\, c_M\, \zeta(dx) = c_D\, \zeta(dx) \qquad \text{because } \mu \in \mathcal{M}_R. \\
\frac{\mu A_{t_D}(\mathcal{D}_R) - c_D}{1 - c_D} &\geq \frac{2\xi_R - c_D}{1 - c_D} = \frac{2 - c_M}{1 - \xi_R\, c_M}\xi_R \geq \xi_R,
\end{aligned}$$

by (5.23), where we used $1 \geq (1 - \xi_R)\, c_M$ (of course $c_M \in (0,1)$ and $\xi_R > 0$). □

Remark: *Our choice for ξ_R and c_D is done for simplicity and can certainly be improved regarding the convergence rate γ. What we require is rather: $\xi_R \leq \frac{c_M}{c_D} \wedge \frac{\xi_\circ - c_D}{1 - c_D}$.*

5.3.3. Lower-bound on the marginals

At time t_h, for any initial condition $\mu \in \mathcal{M}_R$, the MCNE shall be lower-bounded by:

$$\zeta[t_h](dx) := \textstyle\sum_{\{j \leq J(t_h)\}} \left({}^{c_D}\!/_{c_P} \right) \times \left(1 - {}^{c_D}\!/_{c_P}\right)^{j-1} \zeta A_{t_h - j\,t_D}(dx) \geq 0. \tag{5.25}$$

Remark: *The definition of* $(\zeta[t])_{t\geq 0}$ *implicitly depends on* c_D, t_D, c_P *and* t_P, *but not on* μ, ℓ *or* ξ.

The proof of Theorem 2.1 will be completed thanks to Theorems 5.1, 5.2 and the following proposition:

Proposition 5.3.6. *Suppose* (5.1), (*A*1) *and* (5.16) *hold, with* c_D *and* t_D *chosen according to Proposition 5.3.4,* c_P, t_P *according to* (5.16). *Then, to any pair* $\ell \geq 0$ *and* $\xi \in (0,1)$, *we can associate a time* $t_\circ = t_\circ(n,\xi) > 0$ *such that:*

$$\forall\, \mu \in \mathcal{M}_{\ell,\,\xi},\ \forall\, t_{h,2} \geq t_{h,1} \geq t_\circ, \qquad \mu A_{t_{h,2}} \geq \zeta[t_{h,1} - t_\circ]. \tag{5.26}$$

But as a first step to conclude the proof of Theorem 5.1, we prove the following lemma:

Lemma 5.3.7. *Assume that for some* $j \leq J(t_h) - 1$, $r_j > 0$, $\nu_j \in \mathcal{M}_1(\mathcal{X})$ *and* (5.27) *holds. Then:*

$$r_{j+1} > 0 \quad \text{and } \exists\, 0 < c_j \leq c_D, \quad \nu_{j+1}(dx) = \left(\nu_j A_{t_D}(dx) - c_j\ \zeta(dx)\right) / (1 - c_j).$$

Then, in order to achieve the induction "$\nu_j \in \mathcal{M}_R$ implies $\nu_{j+1} \in \mathcal{M}_R$" we ensure iteratively:

$$\mathbb{P}_{\nu_j}\left(t_h - j\, t_D < \tau_\partial\right) \leq c_P\, \mathbb{P}_{\nu_j}\left(t_D < \tau_\partial\right)\, \mathbb{P}_\zeta\left(t_h - [j+1]\, t_D < \tau_\partial\right), \tag{5.27}$$

$$\text{and } \nu_j A_{t_D} \geq c_D\, \zeta, \tag{5.28}$$

that come respectively from (5.16) and (5.22). The proof of Proposition 5.3.6 is achieved in this second step, while the third one concludes the proof of Theorem 2.1.

Step 1: proof of Lemma 5.3.7.

First of all, we need to relate $1 - \sum_{k\geq 1} a(k, j\, t_D)$ to the repartition of mass at time t_h, which is done in the proof of the following lemma, whose proof (similar to the next paragraph, yet much simpler) is reported in Appendix E:

Fact 5.3.8. *Assume that for some* $j \leq J(t_h) - 1$*:* $\quad r_j > 0,\ \nu_j \in \mathcal{M}_1(\mathcal{X})$.

$$\textit{Then } r_j = \left[1 - \frac{c_D}{c_P}\right]^j \times \frac{\mathbb{P}_\mu(t_h < \tau_\partial)}{\mathbb{P}_\mu(j\, t_D < \tau_\partial)} \times \frac{1}{\mathbb{P}_{\nu_j}(t_h - j\, t_D < \tau_\partial)}.$$

By the definition of ν_j, cf (5.21): $\mu A_{j\, t_D} = \sum_{k=1}^{j} a(k, j\, t_D)\, \zeta A_{(j-k)\, t_D} + r_j\, \nu_j$.

$$\begin{aligned}
\mu A_{[j+1]\, t_D} &= \frac{\mathbb{P}_\mu(j\, t_D < \tau_\partial)}{\mathbb{P}_\mu([j+1]\, t_D < \tau_\partial)}\ \mu A_{j\, t_D} \cdot P_{t_D} \\
&= \textstyle\sum_{\{k\leq j\}} a(k, j\, t_D) \times \dfrac{\mathbb{P}_\mu(j\, t_D < \tau_\partial) \times \mathbb{P}_\zeta([j+1-k]\, t_D < \tau_\partial)}{\mathbb{P}_\mu([j+1]\, t_D < \tau_\partial) \times \mathbb{P}_\zeta([j-k]\, t_D < \tau_\partial)} \zeta A_{[j+1-k]\, t_D} \\
&\qquad + \ell_j\ \nu_j A_{t_D}, \qquad (5.29) \\
&\text{where } \ell_j := r_j \times \frac{\mathbb{P}_\mu(j\, t_D < \tau_\partial)}{\mathbb{P}_\mu([j+1]\, t_D < \tau_\partial)} \times \mathbb{P}_{\nu_j}(t_D < \tau_\partial). \qquad (5.30)
\end{aligned}$$

By (5.20), i.e. the definition of $a(k, j\,t_D)$:

$$a(k, j\,t_D) \times \frac{\mathbb{P}_\mu(j\,t_D < \tau_\partial)}{\mathbb{P}_\mu([j+1]\,t_D < \tau_\partial)} \times \frac{\mathbb{P}_\zeta([j+1-k]\,t_D < \tau_\partial)}{\mathbb{P}_\zeta([j-k]\,t_D < \tau_\partial)} = a(k, [j+1]\,t_D).$$

Thus $\quad 1 = \sum_{\{k \le j\}} a(k, [j+1]\,t_D) + \ell_j \quad$ i.e. $\quad r_{j+1} = \ell_j - a(j+1, [j+1]\,t_D), \qquad (5.31)$

by evaluating (5.29) on $\mathcal{X}$ and the definition of r_{j+1}, cf.(5.21).

By (5.30), (5.20) and by Fact 5.3.8:

$$\begin{aligned} c_j :=& a(j+1, [j+1]\,t_D)/\ell_j \qquad (5.32)\\ =& \left[1 - \frac{c_D}{c_P}\right]^{-j} \times \frac{\mathbb{P}_\mu(j\,t_D < \tau_\partial)}{\mathbb{P}_\mu(t_h < \tau_\partial)} \times \frac{\mathbb{P}_{\nu_j}(t_h - j\,t_D < \tau_\partial)}{\mathbb{P}_{\nu_j}(t_D < \tau_\partial)} \times \frac{\mathbb{P}_\mu([j+1]\,t_D < \tau_\partial)}{\mathbb{P}_\mu(j\,t_D < \tau_\partial)}\\ & \times \frac{c_D}{c_P}\left(1 - \frac{c_D}{c_P}\right)^{j} \times \frac{\mathbb{P}_\mu(t_h < \tau_\partial)}{\mathbb{P}_\mu([j+1]\,t_D < \tau_\partial)} \times \frac{1}{\mathbb{P}_\zeta(t_h - [j+1]\,t_D < \tau_\partial)}\\ =& \frac{c_D}{c_P} \frac{\mathbb{P}_{\nu_j}(t_h - j\,t_D < \tau_\partial)}{\mathbb{P}_{\nu_j}(t_D < \tau_\partial) \times \mathbb{P}_\zeta(t_h - [j+1]\,t_D < \tau_\partial)}. \end{aligned}$$

Thanks to (5.27): • $0 < c_j \le c_D$.

Since $c_D < 1$, using (5.31) and (5.32): • $r_{j+1} = \ell_j\,(1 - c_j) > 0$.

Finally, by (5.21), i.e. the definition of ν_{j+1}, (5.32) and (5.29):

$$\begin{aligned} \nu_{j+1} = (1/r_{j+1}) \times \left[\mu A_{[j+1]\,t_D} - \textstyle\sum_{\{k \le j+1\}} a(k, (j+1)\,t_D)\,\zeta A_{(j+1-k)\,t_D}\right]\\ = (\nu_j A_{t_D} - c_j\,\zeta) \times \ell_j\,/r_{j+1}\\ \bullet \quad \nu_{j+1} = (\nu_j A_{t_D} - c_j\,\zeta)\,/(1 - c_j). \end{aligned}$$

□

Step 2: proof of Proposition 5.3.6 with Lemma 5.3.7.

We first define $\mathcal{M}_R$ thanks to Proposition 5.3.4 together with $(A1)$ and (5.1) such that (5.22) holds.

Step 2.1: under the assumption that $\mu \in \mathcal{M}_R$. Then, by induction over $j \le J(t_h)$, we state (I_j): $r_j > 0$ and $\nu_j \in \mathcal{M}_R$. We initialize at $j = 0$, with $r_0 = 1$ and $\nu_0 := \mu \in \mathcal{M}_R$ by hypothesis.

Assume (I_j) for some $j \le J(t_h) - 1$. Then, by (I_j) and (5.22), (5.28) holds. $j \le J(t_h) - 1$ means notably $t_h - [j+1]\,t_D \ge t_P$, thus thanks to (5.16):

$$\begin{aligned} \mathbb{P}_{\nu_j}\left(t_h - j\,t_D < \tau_\partial\right) = \mathbb{E}_{\nu_j}\left[\mathbb{P}_{X_{t_D}}\left(t_h - [j+1]\,t_D < \tau_\partial\right);\ t_D < \tau_\partial\right]\\ \le c_P\,\mathbb{P}_\zeta\left(t_h - [j+1]\,t_D < \tau_\partial\right) \times \mathbb{P}_{\nu_j}\left[t_D < \tau_\partial\right]. \end{aligned}$$

Thanks to Lemma 5.3.7 together with (5.28): $\nu_j \ge 0$ thus $\nu_j \in \mathcal{M}_1(\mathcal{X})$. Moreover, for any measurable set $\mathcal{D}$:

$$\begin{aligned} \nu_{j+1}(\mathcal{D}) \ge (\nu_j A_{t_D}(\mathcal{D}) - c_j)\ /(1 - c_j) = 1 - (1 - \nu_j A_{t_D}(\mathcal{D}))\,/\,(1 - c_j)\\ \ge (\nu_j A_{t_D}(\mathcal{D}) - c_D)\ /\,(1 - c_D), \end{aligned}$$

since $\zeta(\mathcal{D}) \vee \nu_j A_{t_D}(\mathcal{D}) \leq 1$, $c_j \leq c_D$ and $: c \to 1 - (1 - \nu_j A_{t_D}(\mathcal{D}))\,/\,(1-c)$ is decreasing. In particular, $\nu_{j+1} \in \mathcal{M}_1(\mathcal{X})$ holds true and thanks again to (5.22) we prove finally:

$$\nu_{j+1}(\mathcal{D}_R) \geq \frac{1}{1-c_D}\,(\nu_j A_{t_D}(\mathcal{D}_R) - c_D) \geq \xi_R.$$

Therefore, (I_{j+1}) holds.

By induction, we get $(I_{J(t_h)})$ thus $r_{J(t_h)} > 0$ and $\nu_{J(t_h)} \in \mathcal{M}_R \subset \mathcal{M}_1(\mathcal{X})$. By (5.21), i.e. the definition of $\nu_{J(t_h)}$, and since (A_t) is a semigroup:

$$\begin{aligned}
\mu A_{t_h} &= \frac{\mathbb{P}_\mu(J(t_h)\,t_D < \tau_\partial)}{\mathbb{P}_\mu(t_h < \tau_\partial)}\,\mu A_{J(t_h)\,t_D} P_{t_h - J(t_h)\,t_D} \\
&\geq \frac{\mathbb{P}_\mu(J(t_h)\,t_D < \tau_\partial)}{\mathbb{P}_\mu(t_h < \tau_\partial)}\,\left[\textstyle\sum_{\{k \leq J(t_h)\}} a(k, J(t_h)\,t_D)\,\zeta A_{[J(t_h)-k]\,t_D}\,P_{t_h - J(t_h)\,t_D}\right] \\
&\geq \sum_{k \leq J(t_h)} \frac{\mathbb{P}_\mu(J(t_h)\,t_D < \tau_\partial)}{\mathbb{P}_\mu(t_h < \tau_\partial)}\,a(k, J(t_h)\,t_D)\,\frac{\mathbb{P}_\zeta(t_h - k\,t_D < \tau_\partial)}{\mathbb{P}_\zeta([J(t_h)-k]\,t_D < \tau_\partial)}\zeta A_{t_h - k\,t_D}.
\end{aligned}$$

Finally, thanks to (5.20) and (5.25), we conclude: $\mu A_{t_h} \geq \zeta[t_h]$.

Step 2.2: $\mu \in \mathcal{M}_1(\mathcal{X})$. For general initial conditions, recall that in Proposition 5.3.4, we constructed $\mathcal{M}_R$ such that $\mathcal{M}_\circ \subset \mathcal{M}_R$. Thus, (5.1) holds with $\mathcal{M}_R$ instead of $\mathcal{M}_\circ$. Since $t_{h,2} \geq t_{h,1}$, we obtain $\mu A_{t_\circ + t_{h,2} - t_{h,1}} \in \mathcal{M}_R$. As $\mu A_{t_{h,2}} = \mu A_{t_\circ + t_{h,2} - t_{h,1}} A_{t_{h,1} - t_\circ}$, we finally deduce from (5.29):

$$\mu A_{t_{h,2}} \geq \zeta[t_{h,1} - t_\circ].$$

□

Step 3: conclusion of the proof of Theorem 2.1.

Thanks to Proposition 5.3.6:

$$\begin{aligned}
&\forall\,\ell \in \mathbb{N},\ \forall\,\xi \in (0,1),\ \forall\,(\mu_1,\,\mu_2) \in (\mathcal{M}_{\ell,\xi})^2,\ \forall\,t_{h,2} \geq t_{h,1} \geq t_\circ, \\
&\qquad \mu_1 A_{t_{h,1}} \geq \zeta[t_{h,1} - t_\circ] \quad \text{and} \quad \mu_2 A_{t_{h,2}} \geq \zeta[t_{h,1} - t_\circ] \\
&\text{thus } \|\mu_2 A_{t_{h,2}} - \mu_1 A_{t_{h,1}}\|_{TV} \leq \|\mu_2 A_{t_{h,2}} - \zeta[t_{h,1} - t_\circ]\|_{TV} + \|\mu_1 A_{t_{h,1}} - \zeta[t_{h,1} - t_\circ]\|_{TV} \\
&\qquad\qquad \leq 2 \times [1 - \zeta[t_{h,1} - t_\circ](\mathcal{X})]. \qquad (5.33)
\end{aligned}$$

$$\begin{aligned}
1 - \zeta[t_h](\mathcal{X}) &= 1 - \textstyle\sum_{\{j \leq J(t_h)\}} \binom{c_D}{c_P} \left[1 - {}^{c_D}\!/_{c_P}\right]^{j-1} \qquad \text{by (5.25)}, \\
&= \left[1 - {}^{c_D}\!/_{c_P}\right]^{J(t_h)} \leq \exp[-\gamma\,(t_h - t_P - t_D)] \qquad \text{by (5.19)}, \\
&\qquad \text{with } \gamma := -\frac{1}{t_D}\,\ln\left[1 - \frac{c_D}{c_P}\right] \qquad (5.34)
\end{aligned}$$

Finally, with (5.33), (5.34) and $C = C(\ell,\,\xi) := 2\,\exp[\gamma\,(t_P + t_D + t_\circ(\ell,\xi))]$:

$$\|\mu_2 A_{t_{h,2}} - \mu_1 A_{t_{h,1}}\|_{TV} \leq C\,e^{-\gamma\,t_{h,1}}. \qquad (5.35)$$

This states that for any $\mu \in \mathcal{M}_{\ell,\xi}$, $(\mu A_{t_h})_{\{t_h \geq 0\}}$ is a Cauchy-sequence for the total variation distance. Thus, it converges for this distance to some distribution $\alpha^{\ell,\xi}$. Since for any $\ell \leq \ell'$ and $\xi \geq \xi' > 0$, it is clear by definition that $\mathcal{M}_{\ell,\xi} \subset \mathcal{M}_{\ell',\xi'}$, we deduce $\alpha^{\ell,\xi} = \alpha^{\ell \vee \ell', \xi \wedge \xi'} = \alpha^{\ell',\xi'}$. This means (since $\mathcal{M}_1(\mathcal{X}) = \bigcup_{(\ell,\xi)} \mathcal{M}_{\ell,\xi}$) that a unique distribution α is the attractor. In particular, there cannot be a QSD different from α.

For any initial condition μ: $\lim_{t\to\infty} \mathbb{P}_\mu(X_t \in dy \,|\, t < \tau_\partial) = \alpha(dy)$, where the convergence holds in the weak topology (ie α is a *quasi-limiting distribution*). One can then easily adapt the proof of Lemma 7.2 in [8] to deduce that α is effectively a QSD and $\forall\, t \geq 0$, $\mathbb{P}_\alpha(t < \tau_\partial) = e^{-\lambda\, t}$. By letting $t_{h,2} \to \infty$ in (5.35), with $\mu_2 = \mu_1 = \mu \in \mathcal{M}_{\ell,\xi}$:

$$\| \mathbb{P}_\mu [X_t \in dx \mid t < \tau_\partial] - \alpha(dx) \|_{TV} \leq C(\ell,\xi)\, e^{-\gamma\, t}$$

This ends the proof of Theorem 2.1 (up to Appendix E). □

5.4. Proof of Theorem 2.2:

Step 1: proof of the uniform convergence to h

Considering the arguments in the proof of Theorem 5.2, it is easily seen that for any probability measure μ, there exists c'_P, t'_P such that:

$$\forall\, x \in \mathcal{X},\ \forall\, t \geq t'_P, \quad \mathbb{P}_x(t < \tau_\partial) \leq c'_P\, \mathbb{P}_\mu(t < \tau_\partial)\,.$$

Here, we need this estimate for $\mu := \alpha$. To achieve this, we only need to apply $(A1)$ and adjust the value for c_P: $c'_P := c_P\, e^{-\lambda\, t_M} / (\alpha(\mathcal{D}_{\ell_M})\, c_M)$, where t_M, c_M are given by $(A1)$ for initial condition in $\mathcal{D}_{\ell_M}$. This can be translated in term of a uniform bound on h by:

$$\|h_\bullet\|_\infty := \sup_{\{t\geq 0\}} \|h_t\|_\infty \leq c'_P \vee e^{\lambda\, t_P} < \infty. \tag{5.36}$$

Like in the proof of Proposition 2.3 in [13], we deduce that, for any $s, t > 0$, $\mu \in \mathcal{M}_{\ell,\xi}$:

$$|\langle \mu | h_t \rangle - \langle \mu | h_{t+s} \rangle| = \langle \mu | h_t \rangle\, |\langle \alpha - \mu | h_s \rangle| \leq \|h_\bullet\|^2_\infty\, C(\ell,\xi)\, e^{-\gamma\, t}. \tag{5.37}$$

The constant C can actually be taken independently of ℓ, ξ. Indeed, because the previous expression is linear in μ and $\langle \alpha \,|\, h_t \rangle \equiv 1$:

$$|\langle \mu \,|\, h_t - h_{t+s} \rangle| = 2|\langle \bar{\mu} \,|\, h_t - h_{t+s} \rangle|, \quad \text{where } \bar{\mu} := (\mu + \alpha)/2.$$

By choosing ℓ sufficiently large to ensure $\xi := \alpha(\mathcal{D}_{\ell_M})/2 > 0$, we deduce that for any $\mu \in \mathcal{M}_1(\mathcal{X})$, $\bar{\mu} \in \mathcal{M}_{\ell,\xi}$. The inequality (5.37) is thus uniform in $\mu \in \mathcal{M}_1(\mathcal{X})$, so that (h_t) defines a Cauchy sequence for the uniform norm. We deduce that h_t converges to some unique function h, whose norm is also bounded by $\|h_\bullet\|_\infty$. □

Step 2: Characterization of the survival capacity h

The rest of the proof is directly taken from [13]. As the punctual limit of (h_t), and since for any $t \geq 0$, h_t vanishes on ∂, this also hold for h. With the uniform bound (5.36), we deduce that h is also bounded. As stated in the beginning of this Subsection 5.4, we

can replace ζ by any probability measure μ in (5.16) (with specific values for $c_P(\mu) = c_P(\ell)/\xi, t_P(\mu) = t_P(\ell) > 0$). In particular, for $\mu = \delta_x$, with $x \in \mathcal{D}_\ell$:

$$\forall\, t \geq t_P(\ell), \quad \mathbb{P}_\alpha(t < \tau_\partial) \leq c_P(\ell)\, \mathbb{P}_x(t < \tau_\partial) \quad \text{thus } \forall\, t \geq t_P(\ell), \quad h_t(x) \geq c_P(\ell) > 0.$$

This proves that h has a positive lower-bound on any $\mathcal{D}_\ell$. By the Markov property and (2.3):

$$\forall\, u > 0, \quad P_u h(x) = \lim_{t\to\infty} \frac{\mathbb{E}_x\left[\mathbb{P}_{X_u}(t < \tau_\partial)\right]}{\mathbb{P}_\alpha(t < \tau_\partial)} = e^{-\lambda\, u} \lim_{t\to\infty} \frac{\mathbb{P}_x(t + u < \tau_\partial)}{\mathbb{P}_\alpha(t + u < \tau_\partial)} = e^{-\lambda\, u}\, h(x).$$

From this and (5.36), we immediately deduce that h is in the domain of $\mathcal{L}$ and $\mathcal{L}\, h = -\lambda\, h$. □

5.5. Proof of Theorem 2.3:

Except for (*iii*), for which we will prove (2.8), and for the uniqueness of the stationary distribution, the proof is almost the same as in [13].

Step 1: Proof that the Q-process is well-defined and characterization

Let Λ_s be a $\mathcal{F}_s$-measurable set and $\mu \in \mathcal{M}_1(\mathcal{X})$. By the Markov property:

$$\mathbb{P}_\mu(\Lambda_s \,|\, t < \tau_\partial) = \mathbb{E}_\mu\left[e^{\lambda s}\, h_{t-s}(X_s)/\langle \mu \,|\, h_t\rangle;\ s < \tau_\partial,\ \Lambda_s\right].$$

By Theorem 2.2, the random variable $M_s^t := \mathbf{1}_{\{s<\tau_\partial\}}\, e^{\lambda s}\, h_{t-s}(X_s)\ /\ \langle \mu \,|\, h_t\rangle$, (where $t \geq s$) converges a.s. to:

$$M_s := \mathbf{1}_{\{s<\tau_\partial\}}\, e^{\lambda s}\, h(X_s)\ /\ \langle \mu \,|\, h\rangle,$$

where $\langle \mu \,|\, h\rangle > 0$ because h is positive on $\mathcal{X}$. For t sufficiently large (a priori depending on μ), we deduce from (5.36) and the convergence of $\langle \mu \,|\, \eta_t\rangle$ to $\langle \mu \,|\, \eta\rangle$:

$$0 \leq M_s^t \leq 2\, e^{\lambda s}\, \|h_\bullet\|_\infty\ /\ \langle \mu \,|\, h\rangle. \tag{5.38}$$

Thus, by the dominated convergence Theorem, we obtain that $E_\mu(M_s) = 1$.

By the penalisation's theorem of Roynette, Vallois and Yor (cf Theorem 2.1 in [41]) these two conditions imply that M is a martingale under P_μ and that $\mathbb{P}_\mu(\Lambda_s \,|\, t < \tau_\partial)$ converges to $E_\mu(M_s;\ \Lambda_s)$ for all $\Lambda_s \in \mathcal{F}_s$ when $t \to \infty$. In particular for $\mu = \delta_x$, this means that $\mathbb{Q}_x$ is well defined and:

$$\frac{d\mathbb{Q}_x}{d\mathbb{P}_x}_{\,|\,\mathcal{F}_s} = \mathbf{1}_{\{s<\tau_\partial\}}\, e^{\lambda s}\, \frac{h(X_s)}{h(x)}. \tag{5.39}$$

(5.39) implies directly (2.7). Concerning (2.10):

$$\begin{aligned}
\mu B[h] Q_t(dy) &= \int \frac{h(x)}{\langle \mu \,|\, h\rangle} \mu(dx)\, \frac{h(y)}{h(x)}\, e^{\lambda\, t}\, p(x;t;dy) \\
&= \frac{h(y)}{\langle \mu \,|\, P_t\, h\rangle}\, \mu P_t(dy) = \mu P_t B[h] \qquad \text{by (2.5)}, \\
&= \frac{h(y)\, \mathbb{P}_\mu(t < \tau_\partial)}{\langle \mu\, P_t \,|\, h\rangle} \times \frac{\mu P_t(dy)}{\mathbb{P}_\mu(t < \tau_\partial)} = \frac{h(y)}{\langle \mu\, A_t \,|\, h\rangle}\, \mu A_t(dy) = \mu A_t B[h].
\end{aligned}$$

For a more general convergence, with μ as initial condition and $\Lambda_s \in \mathcal{F}_s$, we deduce:

$$\begin{aligned}\lim_{t\to\infty} \mathbb{P}_\mu(\Lambda_s \,|\, t < \tau_\partial) &= E_\mu\left(e^{\lambda s}\, h(X_s) \,/\, \langle \mu \,|\, h\rangle;\ s < \tau_\partial, \Lambda_s\right)\\ &= \int_{\mathcal{X}} \mu(dx) \frac{h(x)}{\langle \mu \,|\, h\rangle} E_x\left(e^{\lambda s}\, \frac{h(X_s)}{h(x)};\ s < \tau_\partial, \Lambda_s\right)\\ &= \int_{\mathcal{X}} \mu(dx) \frac{h(x)}{\langle \mu \,|\, h\rangle} \mathbb{Q}_x(\Lambda_s) = \mathbb{Q}_{\mu B[h]}(\Lambda_s).\end{aligned}$$

by (5.39) and the definition of $\mu B[h]$ in (2.9).

Moreover, the convergence holds in fact in total variation over $\mathcal{F}_s$, as we prove it now. By the previous calculations, (5.38) and (5.36), for any $\epsilon > 0$:

$$\begin{aligned}\left\|\mathbb{P}_\mu(dw \,|\, t < \tau_\partial) - \mathbb{Q}_{\mu B[h]}(dw)\right\|_{TV, \mathcal{F}_s} &\le \mathbb{E}_\mu |M_s^t - M_s|\\ &\le 4\, e^{\lambda s}\, \frac{\|h_\bullet\|_\infty}{\langle \mu \,|\, h\rangle}\, \mathbb{P}_\mu(|M_s^t - M_s| \ge \epsilon) + \epsilon,\\ \text{so} \quad \limsup_{t\to\infty} \left\|\mathbb{P}_\mu(dw \,|\, t < \tau_\partial) - \mathbb{Q}_{\mu B[h]}(dw)\right\|_{TV, \mathcal{F}_s} &\le \epsilon.\end{aligned}$$

By letting $\epsilon \to 0$, we conclude:

$$\forall\, s \in \mathbb{R}_+, \quad \left\|\mathbb{P}_\mu(dw \,|\, t < \tau_\partial) - \mathbb{Q}_{\mu B[h]}(dw)\right\|_{TV, \mathcal{F}_s} \underset{t\to\infty}{\longrightarrow} 0.$$

For the proof that X defines a strong Markov process under $(\mathbb{Q}_x)_{x\in\mathcal{X}}$, we refer again to the proof in [13].

Step 2: The invariant distribution for X under $\mathbb{Q}$

For all $t \ge 0$ and $f \in \mathcal{B}_b(\mathcal{X})$, with (5.39):

$$\begin{aligned}\langle \beta \,|\, Q_t\, f\rangle &= \langle \alpha \,|\, h \times Q_t\, f\rangle = e^{\lambda\, t}\, \langle \alpha \,|\, P_t\, (h \times f)\rangle\\ &= \langle \alpha \,|\, h \times f\rangle = \langle \beta \,|\, f\rangle,\end{aligned} \tag{5.40}$$

where we used (2.3). We prove the uniqueness with the next subsection.

Step 3: Proof of (2.8)

Exploiting (2.10), we deduce from our definitions:

$$\begin{aligned}\|(\mu B[h])Q_t - \beta\|_{\frac{1}{h}} &= \left\|\frac{\mu A_t}{\langle \mu A_t \,|\, h\rangle} - \alpha\right\|_{TV}\\ &\le \langle \mu A_t \,|\, h\rangle^{-1} \times [\|\mu A_t - \alpha\|_{TV} + |\langle \mu A_t \,|\, h\rangle - 1|].\end{aligned} \tag{5.41}$$

To ensure a lower-bound on $\langle \mu A_t \,|\, h\rangle$, we exploit (2.5) and write:

$$\langle \mu A_t \,|\, h\rangle = \frac{e^{\lambda t} \langle \mu P_t \,|\, h\rangle}{\langle \mu \,|\, h_t\rangle} = \frac{\langle \mu \,|\, h\rangle}{\langle \mu \,|\, h_t\rangle}.$$

We already know that h_t is uniformly upper-bounded and h has a lower-bound on any $\mathcal{D}_\ell$. Since $|\langle \mu A_t \,|\, h\rangle - 1| = |\langle \mu A_t - \alpha \,|\, h\rangle| \le \|\mu A_t - \alpha\|_{TV}\ \|h_\bullet\|_\infty$, and exploiting (5.41) and (2.2), we conclude that there exists $C' = C'(\ell, \xi) > 0$ such that:

$$\forall\, t > 0,\ \forall\, \mu \in \mathcal{M}_{\ell,\xi}, \quad \|\mathbb{Q}_{\mu B[h]}(X_t \in dx) - \beta(dx)\|_{\frac{1}{h}} \le C'\, e^{-\gamma\, t}. \qquad \square$$

Step 4: Convergence with initial condition for the Q-process

When μ_Q is the initial condition of the Q-process, it is in general not possible to interpret it as $\mu B[h]$. Indeed, we should expect in this case $\mu(dx)$ to be proportional to $h(x)^{-1}\,\mu_Q(dx)$, which may not be integrable. Thus, the convergence to β might in general not be exponential.

However, it is exponential for measures with support in any of the $\mathcal{D}_\ell$, in particular Dirac masses. Indeed, we have a lower-bound of h: $\quad h^{(\ell)} := \inf\{h_x;\ x \in \mathcal{D}_\ell\}$, $\quad$ which is positive because of $(A1)$ and (5.16). Thus, if $\mu_Q \in \mathcal{M}_1(\mathcal{X})$ has support on $\mathcal{D}_\ell$, $\langle \mu_Q \,|\, 1/h\rangle \leq 1/h^{(\ell)} < \infty$, so:

$$\mu_Q = \mu B[h], \quad \text{with } \mu(dx) := \mu_Q B(1/h) := \mu_Q(dx) \,/\, (h(x) \times \langle \mu_Q \,|\, 1/h\rangle).$$

Now, μ has the same support as μ_Q, thus $\mu(\mathcal{D}_\ell) = 1$, i.e. $\mu \in \mathcal{M}_{\ell,1}$. By (2.8):

$$\|\mu_Q Q_t - \beta\|_{TV} = \|\mu B[h] Q_t - \beta\|_{TV} \leq C(\ell,1)\, e^{-\gamma t}.$$

More generally, since the Q-process is linear with its initial condition, and by $(A0)$, the property of uniqueness of the stationary distribution β holds.

Besides, to have exponential convergence, it suffices that: $\langle \mu_Q \,|\, 1/h\rangle < \infty$. It can be deduced from $\sum_{\ell \geq 1} \mu_Q(\mathcal{D}_\ell \setminus \mathcal{D}_{\ell-1})\,/h^{(\ell)} < \infty$ (note that one has lower-bounds of $h^{(\ell)}$). In any case, the convergence still holds in total variation. □

Appendices:

Appendix A: *Combine all the inequalities to prove Proposition 4.2.2*

We shall first prove that an upper-bound of the global supremum can be deduced from the upper-bounds in Propositions 4.2.3-5. So we start by assuming that the inequalities derived in these propositions hold for some parameters ϵ^X, ϵ^0, C^N_∞, C^X_∞ and C_0 (C^N_∞ coming from Proposition 4.2.3; ϵ^X and C^X_∞ from Proposition 4.2.4; ϵ^0 and C_0 from Proposition 4.2.5) and explain how these inequalities can imply the global supremum in $(A2)$. This implication shall hold at least for ϵ^X and ϵ^0 sufficiently small, which is obtained with n_E sufficiently large. The constraints on ϵ^X and ϵ^0 are mentioned while we handle the inequalities. We prove next that we can indeed find suitable choices of ϵ^X, ϵ^0, C^X_∞, C^N_∞ and C_0 for the upper-bounds in Propositions 4.2.3-5 to hold with these constraints.

t_h is introduced to make sure that $\mathcal{E}^X_\infty \vee \mathcal{E}^N_\infty \vee \mathcal{E}_0 < \infty$ ($\leq \exp[\rho\, t_h]$). It is needed to justify the following inequalities, but this specific upper-bound plays no role. By the upper-bounds in Propositions 4.2.4 and 4.2.3:

$$\begin{aligned} \mathcal{E}^X_\infty &\leq C^X_\infty\,(1 + \mathcal{E}_0) + \epsilon^X\, C^N_\infty (1 + \mathcal{E}^X_\infty) \\ (1 - \epsilon^X\, C^N_\infty)\, \mathcal{E}^X_\infty &\leq C^X_\infty + \epsilon^X\, C^N_\infty + C^X_\infty\, \mathcal{E}_0. \end{aligned}$$

Provided that: $\epsilon^X \leq (2\, C^N_\infty)^{-1}$, recalling that $C^N_\infty \wedge C^X_\infty \geq 1$, and combining it with the upper-bounds of Proposition 4.2.5, it yields:

$$\begin{aligned} \mathcal{E}^X_\infty \leq 3\, C^X_\infty + 2\, C^X_\infty\, \mathcal{E}_0, \qquad \mathcal{E}^N_\infty &\leq 4\, C^N_\infty\, C^X_\infty + 2\, C^N_\infty\, C^X_\infty\, \mathcal{E}_0 \\ \mathcal{E}_0 \leq C_0 + 7\, \epsilon^0\, C^N_\infty\, C^X_\infty + 4\, \epsilon^0\, C^N_\infty\, C^X_\infty \mathcal{E}_0 \\ \text{thus } \left(1 - 4\, \epsilon^0 C^N_\infty\, C^X_\infty\right)\, \mathcal{E}_0 &\leq C_0 + 7\, \epsilon^0 C^N_\infty\, C^X_\infty. \end{aligned}$$

Provided: $\epsilon^0 \le \left(8\, C_\infty^N\, C_\infty^X\right)^{-1}$, and recalling that $C_\infty^N \wedge C_0 \ge 1$, we deduce:

$$\mathcal{E}_0 \le 4\, C_0, \qquad \mathcal{E}_\infty^X \le 11\, C_\infty^X\, C_0, \qquad \mathcal{E}_\infty^N \le 12\, C_\infty^N\, C_\infty^X\, C_0.$$

Finally, provided: $\epsilon^X \le (2\, C_\infty^N)^{-1}$, $\epsilon^0 \le \left(8\, C_\infty^N\, C_\infty^X\right)^{-1}$, conditions which we can satisfy and restrict the choices of n_∞ and $n_E > n_\infty$, we deduce:

$$sup_{(x,n)} \left\{\mathbb{E}_{(x,n)}[\exp(\rho\, \widehat{\tau}_E)]\right\} \le 12\, C_\infty^N\, C_\infty^X\, C_0 < \infty. \tag{6.1}$$

More precisely, for any ρ, we obtain from Proposition 4.2.3 the constants n_∞ and C_∞^N, so that we can set $\epsilon^X := (2\, C_\infty^N)^{-1}$. We then deduce, thanks to Proposition 4.2.4, some value for n_E and C_∞^X. Setting $\epsilon_0 = (8 C_\infty^N C_\infty^X)^{-1}$, we can choose, according to Proposition 4.2.5, some value $n_0 > 0$ and C_0. Taking the limit in (6.1) as $t_h \to \infty$ (recall that $\widehat{\tau}_E := \tau_E \wedge \tau_\partial \wedge t_h$) and choosing $n := n_\infty \vee n_E \vee n_0$ conclude the proof of Proposition 4.2.2.

□

Appendix B: *Descent from infinity, proof of Proposition 4.2.3*
Lemma 4.2.6 implies Proposition 4.2.3.

We obtain by induction and the Markov property: $\forall\, n > 0, \quad \mathbb{P}_n(k\, t < \tau_\downarrow^D) \le \epsilon^k$. Thus, by choosing ϵ sufficiently small (for any given value of $t > 0$), we ensure:

$$C_\infty^N := \sup_{\{n>0\}} \left\{\mathbb{E}_n[\exp(\rho \tau_\downarrow^D)]\right\} < +\infty.$$

A fortiori with $T_\downarrow := \inf\{t, N_t \le n_\infty\} \wedge \tau_E \le \tau_\downarrow^D$,

$$\sup_{(x,n)} \left\{\mathbb{E}_{(x,n)}[\exp(\rho\, T_\downarrow)]\right\} \le C_\infty^N < \infty.$$

At time $T_\downarrow$, the process is either in E or in $\mathcal{T}_\infty^X$. Thus:

$$\begin{aligned}\mathbb{E}_{(x,n)}[\exp(\rho\, \widehat{\tau}_E)] &\le \mathbb{E}_{(x,n)}[\exp(\rho\, T_\downarrow)\,;\ (x,n)_{T_\downarrow} \in E] \\ &\quad + \mathbb{E}_{(x,n)}\Big[\exp(\rho\, T_\downarrow)\mathbb{E}_{(X,N)_{T_\downarrow}}[\exp(\rho\, \widehat{\tau}_E)]\,;\ (X,N)_{T_\downarrow} \in \mathcal{T}_\infty^X\Big],\end{aligned}$$

with the Markov property and the fact that $(\widehat{\tau}_E - \mathcal{T}_\infty^X)_+ \le t_h$ on the event $\{(X,N)_{T_\downarrow} \in \mathcal{T}_\infty^X\}$. Therefore:

$$\mathcal{E}_\infty^N \le C_\infty^N\, (1 + \mathcal{E}_\infty^X).$$

□

Proof of Lemma 4.2.6.

The proof of this Lemma relies mainly on the same arguments as in [1], part 6, related to the descent from infinity. Let $Z_t := \sigma/2 \times \sqrt{N_t^D}$. It is solution to the following EDS:

$$Z_t := z + \int_0^t \psi\,(Z_s)\ ds + B_t, \text{ where } \psi(z) := -\frac{1}{2\, z} + \frac{r\, z}{2} - c\, z^3. \tag{6.2}$$

As long as Z is very large and $|B|$ not exceptionally large, the leading term $-c\,Z_t^3$ indeed makes the process come down in finite time. Let $V := Z - B$. It is the solution of the ODE:

$$\frac{dV_t}{dt} = -\frac{1}{2(V_t+B_t)} + \frac{r\,(V_t+B_t)}{2} - c\,(V_t+B_t)^3, \tag{6.3}$$

$$\text{Let } z_2 \geq z_1 := \sup\left\{z>0,\ \left|-\frac{1}{2z}+\frac{rz}{2}\right| \geq \frac{cz^3}{2}\right\}, \tag{6.4}$$

$$T_B := \inf\{t>0, B_t \notin [-z_2,\ 2\,z_2]\}, \qquad T_V := \inf\{t>0, V_t < 2z_2\}, \tag{6.5}$$

where we consider w.l.o.g. an initial condition z strictly bigger than $2\,z_2$, so that T_V is positive a.s. Then, as in [1], we get on the time interval $[0, T_B \wedge T_V]$:

$$B_t \geq -z_2 \geq -V_t/2, \text{ implying } V_t + B_t \geq V_t/2 \text{ and } V_t + B_t \geq z_2,$$

$$\left|-\frac{1}{2(V_t+B_t)} + \frac{r\,(V_t+B_t)}{2}\right| \leq \frac{c}{2}(V_t+B_t)^3,$$

$$\frac{d}{dt}\left[(V_t)^{-2}\right] = \frac{-2}{(V_t)^3}\frac{dV}{dt} \geq 2\times\left(c-\frac{c}{2}\right)\times\left(\frac{V_t+B_t}{V_t}\right)^3 \geq \frac{c}{8},$$

$$\text{thus } V_t^{(-2)} - z^{(-2)} \geq c\,t/8 \quad \text{and in particular} \qquad V_t \leq \sqrt{8/(c\times t)}.$$

Thus, $\{t \leq T_B\} \subset \{T_V \leq t\} \cup \left\{V_t \leq \sqrt{8/(c\times t)}\right\}$.

By (6.5), let z_2 be sufficiently big to ensure: $\mathbb{P}(T_B < t) \leq \epsilon$. Then, denote: $z_\infty := \left(\sqrt{8/(c\times t)} + 2\,z_2\right) \vee (4\,z_2)$. We deduce that, on the event $\{t \leq T_B\}$, either $Z_t \leq z_\infty$ or $T_V \leq t$ while $Z_{T_V} \leq 4z_2 \leq z_\infty$. In any case, $\tau_\downarrow^D \leq t$. Hence: $\forall\, z>0,\ \mathbb{P}_z(t < \tau_\downarrow^D) \leq \epsilon$. $\square$

Appendix C: *Mal-adaptation too large, proof of Proposition 4.2.4:*
Lemma 4.2.7 implies Proposition 4.2.4.

Let ρ, ϵ, $n_\infty > 0$ ($c > 0$ is the same as for the definition of Z). For simplicity, we choose $t := \log(2)/\rho > 0$ (i.e. $\exp[\rho\,t] = 2$), and assume w.l.o.g. $t < t_h$. We choose $r_\vee \in \mathbb{R}$ according to Lemma 4.2.7 such that:

$$\forall\, n>0,\ \forall\, r \leq r_\vee, \quad \mathbb{P}_n\left(t < \tau_\partial^D\right) \leq e^{-\rho\,t}/2 = 1/4,$$
$$\forall\, r \leq r_\vee, \quad \mathbb{P}_{n_\infty}\left(T_\infty^D \leq t\right) + \mathbb{P}_{n_\infty}\left(N_t^D \geq n_\infty\right) \leq \epsilon/4.$$

Since $\limsup_{\|x\|\to\infty} r(x) = -\infty$, with n_E chosen sufficiently large:

$$\forall\, x \notin B(0,n_E),\ r(x) \leq\ r_\vee.$$

Let (X,N) with initial condition $(x,n) \in \mathcal{T}_\infty^X$. In the following, we denote:

$$T_\infty^N := \inf\{t \geq 0,\ N_t \geq n_c\}, \quad \tau_0 := \inf\{t>0,\ (X,\ N)_t \in \mathcal{T}_0\},$$
$$T := t \wedge T_\infty^N \wedge \tau_0 \wedge \tau_E \wedge \tau_\partial. \tag{6.6}$$

Since, on the event $\{T = t\}$, either $N_t \geq n_\infty$ or $(X, Y)_t \in \mathcal{T}_\infty^X$:

$$\begin{aligned}\mathbb{E}_{(x,n)}[\exp(\rho\,\widehat{\tau}_E)] &= \mathbb{E}_{(x,n)}[\exp(\rho\,T)\,;\ T = \widehat{\tau}_E] + \mathbb{E}_{(x,n)}[\exp(\rho\,\widehat{\tau}_E)\,;\ T = \tau_0] \\ &\quad + \mathbb{E}_{(x,n)}[\exp(\rho\,\widehat{\tau}_E)\,;\ T = t] + \mathbb{E}_{(x,n)}[\exp(\rho\,\widehat{\tau}_E)\,;\ T = T_\infty^N] \\ &\leq \exp(\rho\,t)\ (1 + \mathcal{E}_0) + \exp(\rho\,t)\,\mathbb{P}_{(x,n)}[T = t]\mathcal{E}_\infty^X \\ &\quad + \exp(\rho\,t)\ \left(\mathbb{P}_{(x,n)}[T = T_\infty^N] + \mathbb{P}_{(x,n)}[N_t \geq n_\infty,\, T = t]\right)\ \mathcal{E}_\infty^N,\end{aligned}$$

by the Markov property. Now, by (6.6), N^D is an upper-bound of N before T. Thus, by our definitions of $t, n_E, r_\vee$:

$$\mathbb{E}_{(x,n)}[\exp(\rho\,\widehat{\tau}_E)] \leq 2\ (1 + \mathcal{E}_0) + \mathcal{E}_\infty^X/2 + \epsilon\,\mathcal{E}_\infty^N/2.$$

Taking the supremum over $(x, n) \in \mathcal{T}_\infty^X$ in the last inequality yields:
$\mathcal{E}_\infty^X \leq 4\ (1 + \mathcal{E}_0) + \epsilon\,\mathcal{E}_\infty^N$. □

Proof of Lemma 4.2.7.

We recall our definition of Z and ψ in (6.2). In the following, we condider ψ as a function of r, thus the notation $\psi_r(z) := -1/(2\,z) + (rz)/2 - c\,z^3$.

Step 1: $\sup_{z>0}\ \psi_r(z) \underset{r\to-\infty}{\longrightarrow} -\infty$.

Let $A > 0$, $z_A := \frac{2}{A}$ and $r_\vee := -A^2$. Then:

$$\begin{aligned}&\forall\, z \leq z_A,\ \forall\, r \leq 0, && \psi_r(z) \leq -1/(2\,z_A) = -A, \\ &\forall\, z \geq z_A,\ \forall\, r \leq r_\vee \leq 0, && \psi_r(z) \leq r_\vee\, z_A = -2A.\end{aligned}$$ □

Step 2: bound on $Z_t^A := z - A\,t + B_t$ for A large:

Let $\epsilon, t_D > 0$. We can choose $\Delta z > 0$ such that, with $N \sim \mathcal{N}(0, 1)$:

$$\mathbb{P}\left(\sup_{\{t\leq t_D\}} B_t \geq \Delta z\right) = 2\,\mathbb{P}\left(N \geq \Delta z\,/\,\sqrt{t_D}\right) \leq \epsilon. \tag{6.7}$$

Then, we can choose $A > 0$ (sufficiently big) such that:

$$\mathbb{P}\left(B_{t_D} \geq A\,t_D\right) = \mathbb{P}\left(N \geq A\,\sqrt{t_D}\right) \leq\ \epsilon.$$

We also choose $r_\vee$ thanks to step 1 such that:

$$\forall\, r \geq r_\vee, \quad \sup_{z>0}\ \psi_r(z) \leq -A \leq 0.$$

We now assume that the initial condition of Z satisfies $Z \leq z_\infty$ ($z_\infty = \sigma\sqrt{n_\infty}/2$).

For any $z_E \geq z_\infty + \Delta z$ and $r \leq r_\vee$, we deduce:

$$\begin{aligned}&\sup_{\{t\leq t_D\}} Z_t \leq z_\infty + \sup_{\{t\leq t_D\}} B_t, \\ &\mathbb{P}\left(\sup_{\{t\leq t_D\}} Z_t \geq z_E\right) \leq \mathbb{P}\left(\sup_{\{t\leq t_D\}} B_t \geq \Delta z\right) \leq \epsilon \quad \text{by (6.7)}, \\ &\mathbb{P}\left(Z_{t_D} \geq z_\infty\right) \leq \mathbb{P}\left(B_{t_D} \geq A\,t_D\right) \leq \epsilon \quad \text{by our choice of A and } r_\vee.\end{aligned}$$

Thus

$$\mathbb{P}\left(T^D_\infty \le t_D\right) \le \mathbb{P}\left(Z_{t_D} \ge z_\infty\right) + \mathbb{P}\left(\sup_{\{t\le t_D\}} Z_t \ge z_E\right) \le 2\,\epsilon,$$

with $n_E = (2\,z_E/\sigma)^2$. It proves the second claim of the Lemma (up to a change of ϵ by $\epsilon/2$).

Step 3: descent from infinity and extinction

Now, we need to assume $c > 0$. Let again $\epsilon, t_D > 0$. Thanks to Lemma 4.2.6 (for $r = 0$ since $\mathbb{P}(t_D < \tau_\downarrow)$ is decreasing with r) we choose $z_\downarrow > 0$ such that, with $\tau_\downarrow := \inf\{t \ge 0;\ Z_t \le z_\downarrow\}$:

$$\forall\, r \le 0,\ \forall\, z > 0, \quad \mathbb{P}_{z_\infty}(t_D < \tau_\downarrow) \le \epsilon \tag{6.8}$$

Like in the previous step, we choose $A > 0$ such that:

$$\mathbb{P}\left(B_{t_D} \ge A\,t_D - z_\downarrow\right) \le \epsilon.$$

Again, we choose $r_\vee$ thanks to step 1 such that: $\forall\, r \le r_\vee, \quad \sup_{z>0} \psi_r(z) \le -A \le 0$.

Then, with $r \le r_\vee$, on the event $\{\tau_\downarrow \le t_D\}$, conditionally on $Z_{\tau_\downarrow}$:

$$\mathbb{P}_{N^D_{\tau_\downarrow}}\left(2\,t_D - \tau_\downarrow < \widetilde{\tau_\partial}^{\,D}\right) \le \mathbb{P}_{Z_{\tau_\downarrow}}\left(\widetilde{Z}_{t_D} > 0\right) \le \mathbb{P}\left(z_\downarrow - A\,t_D + B_{t_D} > 0\right) \le \epsilon, \tag{6.9}$$

by our choices of A and $r_\vee$. Finally, by the Markov property, for any $z > 0$:

$$\begin{aligned}\mathbb{P}_{z_\infty}\left(2\,t_D < \tau^D_\partial\right) &\le \mathbb{P}_{z_\infty}\left(t_D < \tau_\downarrow\right) + \mathbb{E}_{z_\infty}\left[\mathbb{P}_{Z_{\tau_\downarrow}}\left(2\,t_D - \tau_\downarrow < \widetilde{\tau^D_\partial}\right);\ \tau_\downarrow \le t_D\right] \\ &\le 2\,\epsilon \quad \text{with (6.8), (6.9)}\end{aligned}$$

which proves the first claim of the Lemma (replace ϵ by $\epsilon/2$ in the proof and take $t_D = t/2$). □

Appendix D: *Too few individuals, proof of Proposition 4.2.5:*

For $(x, n) \in \mathcal{T}_0$, with n_0 sufficiently small, we would like to say that mortality is so strong in this area that it overcomes an exponential growth at rate ρ. In order to get an estimate of mortality in $\mathcal{T}_0$, we will use some coupling with branching processes and consider the process after a time $t_D = 1$ (arbitrary). In practice, we prove that for any $\rho, \epsilon' > 0$, there exists $C' \ge 1$ such that for any n_E sufficiently large:

$$\mathcal{E}_0 \le C' + \epsilon'\,\left(\mathcal{E}^N_\infty + \mathcal{E}^X_\infty + \mathcal{E}_0\right).$$

By taking $\epsilon' = (\epsilon \wedge 1)/2$, $C = 2\,C'$, it clearly implies Proposition 4.2.5.

The equation $N^U_t = n_0 + \int_0^t r_+ N^U_s ds + \sigma \int_0^t \sqrt{N^U_s}\, dB^N_s$ defines an upper-bound of N on $[0, t_D]$ provided $n \le n_0$, while N^U is a classical branching process. The survival of (X, N) beyond t_D clearly implies the survival of N^U beyond t_D. Let us define ρ_0 by the relation: $\mathbb{P}_{n_0}\left(t_D < \tau^U_\partial\right) =: \exp(-\rho_0\, t_D)$. For a branching process like N^U, it is classical that: $\rho_0 \to \infty$ as $n_0 \to 0$. Indeed, with $u(t, \lambda)$ the Laplace exponent of N^U (cf e.g. [37]

Subsection 4.2, notably Lemma 5): $\mathbb{P}_{n_0}\left(\tau_\partial^U \le t_D\right) = \exp[-n_0 \lim_{\lambda\to\infty} u(t_D,\lambda)] \to 1$, as $n_0 \to 0$.

So we can impose that $\rho_0 > \rho$, and even that $\exp(-(\rho_0-\rho)\, t_D)$ is sufficiently small to make transitions from $\mathcal{T}_0$ to $\mathcal{T}_0$, $\mathcal{T}_\infty^N$ or $\mathcal{T}_\infty^X$ of little incidence.

$$\begin{aligned}
\mathbb{E}_{(x,n)}[\exp(\rho\widehat{\tau}_E)] &\le \mathbb{E}_{(x,n)}\Big[\exp(\rho\widehat{\tau}_E);\ \widehat{\tau}_E < t_D\Big] \\
&\qquad + \mathbb{E}_{(x,n)}\Big[\exp(\rho\widehat{\tau}_E); (x,n)_{t_D} \in \mathcal{T}_0 \cup \mathcal{T}_\infty^N \cup \mathcal{T}_\infty^X\Big] \\
&\le \exp[\rho\, t_D] + \exp(\rho\, t_D)\ (\mathcal{E}_0 + \mathcal{E}_\infty^N + \mathcal{E}_\infty^X)\ \mathbb{P}_{(x,n)}(t_D < \tau_\partial) \\
&\le C' + \epsilon'\ (\mathcal{E}_0 + \mathcal{E}_\infty^N + \mathcal{E}_\infty^X), \qquad \text{where } C' := \exp[\rho\, t_D] \\
\text{and } \epsilon' &:= \exp(-(\rho_0-\rho)\, t_D) \to 0 \text{ as } n_0 \to 0.
\end{aligned}$$

□

Appendix E: *Proof of Fact 5.3.8*

Like in the proof of Lemma 5.3.7 with Fact 5.3.8:

$$\begin{aligned}
\mu A_{t_h} &= \frac{\mathbb{P}_\mu(j\, t_D < \tau_\partial)}{\mathbb{P}_\mu(t_h < \tau_\partial)}\ \mu A_{j\, t_D} \cdot \mathbb{P}_{t_h - j\, t_D} \\
&= \sum_{k=1}^{j} a(k, j\, t_D) \times \frac{\mathbb{P}_\mu(j\, t_D < \tau_\partial)}{\mathbb{P}_\mu(t_h < \tau_\partial)} \times \frac{\mathbb{P}_\zeta(t_h - k\, t_D < \tau_\partial)}{\mathbb{P}_\zeta([j-k]\, t_D < \tau_\partial)} \zeta A_{t_h - k\, t_D} \\
&\qquad + r_j \times \frac{\mathbb{P}_\mu(j\, t_D < \tau_\partial)}{\mathbb{P}_\mu(t_h < \tau_\partial)} \times \mathbb{P}_{\nu_j}(t_h - j\, t_D < \tau_\partial)\ \nu_j A_{t_h - j\, t_D} \qquad (6.10)
\end{aligned}$$

Yet, by (5.20): $a(k, j\, t_D) \times \dfrac{\mathbb{P}_\mu(j\, t_D < \tau_\partial)}{\mathbb{P}_\mu(t_h < \tau_\partial)} \times \dfrac{\mathbb{P}_\zeta(t_h - k\, t_D < \tau_\partial)}{\mathbb{P}_\zeta([j-k]\, t_D < \tau_\partial)} = \dfrac{c_D}{c_P}\left(1 - \dfrac{c_D}{c_P}\right)^{k-1}$,
so that we obtain, by evaluating the measures in (6.10) on $\mathcal{X}$:

$$r_j \times \frac{\mathbb{P}_\mu(j\, t_D < \tau_\partial)}{\mathbb{P}_\mu(t_h < \tau_\partial)} \times \mathbb{P}_{\nu_j}(t_h - j\, t_D < \tau_\partial) = 1 - \sum_{k=1}^{j} \frac{c_D}{c_P}\left(1 - \frac{c_D}{c_P}\right)^{k-1} = \left(1 - {}^{c_D}/_{c_P}\right)^j .$$ □

Aknowledgment

I am very grateful to Etienne Pardoux, my PhD supervisor, for his great support all along the redaction of this article. The comments of my reviewer have also been particularly detailed and helpful. I express my sincere thanks to him. I wish to thank also Nicolas Champagnat for their advices and support.

I would like finally to thank the very inspiring meetings and discussions brought about by the Chair "Modélisation Mathématique et Biodiversité" of VEOLIA-Ecole Polytechnique-MnHn-FX.

References

[1] Ba M and Pardoux E, The effect of competition on the height and length of the forest of genealogical trees of a large population, Malliavin Calculus and Related Topics, Springer Proceedings in Mathematics and Statistics Vol 34, 445-467 (2012).

[2] Bansaye, V., Cloez, B., Gabriel, P.; Ergodic behavior of non-conservative semigroups via generalized Doeblin's conditions, P. Acta Appl Math, pp 1–44, (2019)
[3] Bansaye, V., Cloez, B., Gabriel, P., Marguet, A.; A non-conservative Harris' ergodic theorem, arXiv:1903.03946 (2019)
[4] Bansaye, V., Collet,P., Martinez, S., Méléard, S., San Martin, J.; Diffusions from Infinity, Trans. Amer. Math. Soc., V372, pp 5781-5823 (2019)
[5] Barbour, A.D., Pollett, P.K.; Total variation approximation for quasi-stationary distributions, J. of Appl. Probab., V. 47, pp. 934–946 (2010)
[6] Bass, R. F.; Diffusions and Elliptic Operators, Probab. and Its Applications, Springer, New York (1998)
[7] Bass, R. F.; The measurability of hitting times, Elec. Comm. in Probab., v15, pp99-105 (2010)
[8] Cattiaux, P., and all; Quasi-Stationary Distributions and Diffusion Models in Population Dynamics, The Annals of Probab., V. 37, No. 5, pp. 1926–1969 (2009)
[9] Cattiaux, P., Guilin, A.; Hitting times, functional inequalities, Lyapunov conditions and uniform ergodicity, J. of Funct. Analysis, V. 272, Issue 6, pp. 2361-2391 (2016)
[10] Cattiaux, P., Guilin, A., Zitt, P.A.; Poincaré inequalities and hitting times, Annales de l'institut Henri Poincare (B) Probab. and Stat., V. 49, pp. 95-118 (2010)
[11] Chazottes, R., Collet, P., Méléard, S.; Sharp asymptotics for the quasi-stationary distribution of birth-and-death processes, Probab. Theory Relat. Fields, V. 164, Issue 1–2, pp. 285–332 (2016)
[12] Chazottes, R., Collet, P., Méléard, S.; On time scales and quasi-stationary distributions for multitype birth-and-death processes, Annales de l'Institut H. Poincaré, 55-4 (2017)
[13] Champagnat, N., Villemonais, D.; Exponential convergence to quasi-stationary distribution and Q-process, Probab. Theory Relat. Fields, V. 164, pp. 243–283 (2016)
[14] Champagnat, N., Villemonais, D.; Uniform convergence of time-inhomogeneous penalized Markov processes, ESAIM: Probab. and Stat., vol. 22, pp. 129-162 (2018)
[15] Champagnat, N., Villemonais, D.; Lyapunov criteria for uniform convergence of conditional distributions of absorbed Markov processes, preprint on ArXiv: 1704.01928 (2017)
[16] Champagnat, N., Villemonais, D.; General criteria for the study of quasi-stationarity, preprint on ArXiv: 1712.08092 (2017)
[17] Champagnat, N., Villemonais, D.; Practical criteria for R-positive recurrence of unbounded semigroups, Electron. Commun. Probab., V.25, N.6, pp.1–11 (2020)
[18] Collet, P., Martínez, S., Méléard, S., San Martín, J.; Quasi-stationary distributions for structured birth and death processes with mutations, Probab. Theory Relat. Fields, V. 151, pp. 191–231 (2011)
[19] Collet, P., Martínez, S., San Martín, J.; Quasi-Stationary Distributions, Probab. and Its Appl., Springer, Berlin Heidelberg (2013)
[20] Dellacherie, C., Meyer, P.A.; Probabilities and potential, North Holland (2011)
[21] Del Moral, P., Villemonais, D.; Exponential mixing properties for time inhomogeneous diffusion processes with killing; Bernoulli J., Vol. 24, Nbr. 2, 1010-1032. (2018)
[22] Diaconis, P., Miclo, L.; On quantitative convergence to quasi-stationarity; Ann. Fac. Sci. Toulouse Math., Sér. 6, V. 24, no. 4, pp. 973-1016 (2015)
[23] van Doorn, E.A., Pollett, P.K.; Quasi-stationary distributions for discrete-state models, Eur. J. of Operational Research, V. 230, pp. 1-14 (2013)
[24] Down, D., Meyn, S.P., Tweedie, R. L.; Exponential and Uniform Ergodicity of Markov Processes, The Annals of Probab., V. 23, pp. 1671-1691 (1995)
[25] C. Evans, L.; Partial Differential Equations; Graduate Studies in Mathematics, V. 19, Am. Math. Society (1998)
[26] Ferrari, P. A., Kesten, H., Martinez S., Picco, P.; Existence of quasi-stationary distributions. A renewal dynamical approach; Ann. of Probab., V. 23, No. 2, pp. 501–521 (1995)
[27] Ferrari, P. A., Kesten, H., Martinez S.; R-positivity, quasi-stationary distributions and ratio

limit theorems for a class of probabilistic automata; Ann. Appl. Probab. V.6, Nbr. 2, 577-616 (1996)
[28] Ferré, G., Rousset, M., Stoltz, G.; More on the long time stability of Feynman-Kac semigroups; Stoch PDE: Anal Comp (2020)
[29] Friedman, A.; Partial differential equations of parabolic type, Dover publications, New York (2008)
[30] Krylov, N. V., Safonov,M. V.; A property of the solutions of parabolic equations with measurable coefficients. Izv. Akad. Nauk SSSR Ser. Mat., 44(1):161–175, 239 (1980)
[31] Mariani, M., Pardoux, E., Velleret, A.; Metastability between the clicks of the Müller ratchet; preprint on ArXiv : 2007.14715 (2021)
[32] Méléard, S.; Modèles aléatoires en Ecologie et Evolution, Math. et appl., V. 77, Springer, Berlin Heidelberg (2016)
[33] Mariani, M., Pardoux, E., Velleret, A.; Metastability between the clicks of the Muller ratchet; preprint on ArXiv: 2007.14715 (2020)
[34] Martínez, S., San Martin, J., Villemonais, D.; Existence and uniqueness of a quasi-stationary distribution for Markov processes with fast return from infinity, J. of Appl. Probab., V. 51, Nbr. 3, pp. 756-768 (2014)
[MV12] Méléard, S., Villemonais, D.; Quasi-stationary distributions and population processes, Probab. Surveys, V. 9, pp. 340-410 (2012)
[36] Meyn, S.P., Tweedie, R. L.; Markov Chains and Stochastic Stability, Commun. and Control Eng. Series, Springer, London (1993)
[37] Pardoux, E; Probabilistic Models of Population Evolution: Scaling Limits, Genealogies and Interactions; Springer (2016)
[38] Pascucci, A., and Polidoro, S.; On the Harnack inequality for a class of hypoelliptic evolution equations, V. 356, N. 11, Transactions of the American Mathematical Society, pp4383-4394
[39] Pollett, P. K.; Quasi-stationary distributions: A bibliography., available at `people.smp.uq.edu.au/PhilipPollett/papers/qsds/qsds.html` (2015)
[40] Rogers, L. C. G., Williams, D.; Diffusions, Markov processes, and martingales; V. 1; Cambridge Math. Library (2000)
[41] Roynette, B., Vallois, P., Yor, M.; Some penalisations of the Wiener measure; Jpn. J. Math., V. 1, Issue 1, pp. 263–290 (2006)
[42] Seneta, E., Vere-Jones, D.; On quasi stationary distributions in discrete-time Markov chains with a denumerable infinity of states; J. Appl. Prob. 3, 403-434 (1996)
[43] Swart, J., Winter, A.; Markov processes: theory and examples, Mimeo, `https://www.uni-due.de/hm0110/Markovprocesses/sw20.pdf`, (2013)
[44] Tweedie, R.L.; R-Theory for Markov Chains on a General State Space I: Solidarity Properties and R-Recurrent Chains, Ann. Probab., V.2, N.5, pp. 840-864 (1974)
[45] Tweedie, R.L.; Quasi-Stationary Distributions for Markov Chains on a General State Space, J. of Appl. Probab., V. 11, N. 4, pp. 726-741 (1974)
[46] Velleret, A.; Exponential quasi-ergodicity for processes with discontinuous trajectories, preprint on ArXiv: 1902.01441 (2019)
[47] Velleret, A.; Adaptation of a population to a changing environment under the light of quasi-stationarity; preprint on ArXiv: 1903.10165 (2019)
[48] Velleret, A.; Two level natural selection with a quasi-stationarity approach; preprint on ArXiv: 1903.10161 (2019)