

Quasi-stationary distributions and population processes

Sylvie Méléard, Denis Villemonais

▶ **To cite this version:**

Sylvie Méléard, Denis Villemonais. Quasi-stationary distributions and population processes. Probability Surveys, 2012, 9, pp.340-410. 10.1214/11-PS191 . hal-00653834

HAL Id: hal-00653834
https://hal.science/hal-00653834

Submitted on 20 Dec 2011

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.

Quasi-stationary distributions and population processes

Sylvie Méléard and Denis Villemonais

December 20, 2011

Abstract

This survey concerns the study of quasi-stationary distributions with a specific focus on models derived from ecology and population dynamics. We are concerned with the long time behavior of different stochastic population size processes when 0 is an absorbing point almost surely attained by the process. The hitting time of this point, namely the extinction time, can be large compared to the physical time and the population size can fluctuate for large amount of time before extinction actually occurs. This phenomenon can be understood by the study of quasi-limiting distributions. In this paper, general results on quasi-stationarity are given and examples developed in detail. One shows in particular how this notion is related to the spectral properties of the semi-group of the process killed at 0. Then we study different stochastic population models including nonlinear terms modeling the regulation of the population. These models will take values in countable sets (as birth and death processes) or in continuous spaces (as logistic Feller diffusion processes or stochastic Lotka-Volterra processes). In all these situations we study in detail the quasi-stationarity properties. We also develop an algorithm based on Fleming-Viot particle systems and show a lot of numerical pictures.

Keywords: population dynamics, quasi-stationarity, Yaglom limit, birth and death process, logistic Feller diffusion, Fleming-Viot particle system.

1 Introduction

We are interested in the long time behavior of isolated biological populations with a regulated (density-dependent) reproduction. Competition for limited resources impedes these natural populations without immigration to grow indefinitely and leads them to become extinct. When the population's size attains zero, nothing happens anymore and this population's size process stays at zero. This point 0 is thus an absorbing point for the process. Nevertheless, the time of extinction can be large compared to the individual time scale and it is common that population sizes fluctuate for large amount of time before extinction actually occurs. For example, it has been observed in populations of endangered species, as the Arizona ridge-nose rattlesnakes studied in Renault-Ferrière-Porter [52], that the statistics of some biological traits seem to stabilize. Another stabilization phenomenon is given by the mortality plateau. While demographers thought for a long time that

the rate of mortality of individuals grows as an exponential function of the age, it has been observed more recently that the rate of mortality slows at advanced ages, or even stabilizes. To capture these phenomena, we will study the long time behavior of the process conditioned on non extinction and the related notion of quasi-stationarity. In particular, we will see that a Markov process with extinction which possesses a quasi-stationary distribution has a mortality plateau.

In all the following, the population's size process $(Z_t, t \geq 0)$ will be a Markov process going almost surely to extinction. We are interested in looking for characteristics of this process giving more information on its long time behavior. One way to approach this problem is to study the "quasi-limiting distribution" (QLD) of the process (if it exists), that is the limit, as $t \to +\infty$, of the distribution of Z_t conditioned on non-absorption up to time t. This distribution, which is also called Yaglom limit, provides particularly useful information if the time scale of absorption is substantially larger than the one of the quasi-limiting distribution. In that case, the process relaxes to the quasi-limiting regime after a relatively short time, and then, after a much longer period, absorption will eventually occur. Thus the quasi-limiting distribution bridges the gap between the known behavior (extinction) and the unknown time-dependent behavior of the process.
There is another point of view concerning quasi-stationarity. A quasi-stationary distribution for the process $(Z_t, t \geq 0)$ denotes any proper initial distribution on the non-absorbing states such that the distribution of Z_t conditioned on non-extinction up to time t is independent of $t, t \geq 0$.

There is a large literature on quasi-stationary distributions and Yaglom limits (see for example the large bibliography updated by Pollett [50]) and a lot of references will be given during the exposition. The present paper is by no means exhaustive, but is a survey presenting a collection of tools for the study of QSD concerning specific population's size models. More than the originality of the proofs, we emphasize some general patterns for qualitative and quantitative results on QSD. We also provide a lot of numerical illustrations of the different notions.

In Section 2 of this survey, we will introduce the different notions of QSD and review theoretical properties on QSD and QLD. We will also highlight the relations between QSDs and mortality plateaus. In Section 3, we will study the simple case of QSD for processes in continuous time with finite state space. We develop a simple example to make things more concrete. Thus we will concentrate on QSD for several stochastic population models corresponding to different space and time scales. We will underline the importance of spectral theory as mathematical tool for the research of QSD, in these different contexts. In Section 4, we will consider birth and death processes. We will state results giving explicit conditions on the coefficients ensuring the almost sure extinction of the process, and the existence and uniqueness (or not) of a QSD. We will especially focus on the density-dependence case, when the death rate of each individual is proportional to the population's size (called logistic birth and death process). We will show that in that case, the process goes almost surely to extinction, and that there is a unique QSD, coinciding with the unique QLD. In Section 5, the birth and death process is rescaled by a growing initial population and by small individual weights (small biomass). This process is proved to converge, as the initial population's size tends to infinity, to the unique solution of the deterministic logistic equation, whose unique stable equilibrium is

given by the carrying capacity. If the individual birth and death rates are proportional to the population's size while preserving the ecological balance, more numerous are the individuals, smaller are their weights and faster are their birth and death events, reflecting allometric demographies. In that case, the rescaled birth and death process converges, as the initial population size increases, to the solution of a stochastic differential equation with a 1/2-Hölder diffusion coefficient and a quadratic drift, called logistic Feller equation. The existence of the QSD is proved in this case and uniqueness is characterized by a condition meaning the return of the process from infinity in finite time. The proof relies QSD investigation to spectral theory tools developed in a functional $\mathbb{L}^2-$space. The logistic Feller equation describes the size of a mono-type population where individuals are indistinguable. Motivated by ecological and biodiversity problems, we generalize the model to multi-type populations with intra-specific and inter-specific competition. That leads us to consider stochastic Lotka-Volterra processes. We give conditions ensuring mono-type transient states or coexistence, preserving one or several dominant types in a longer scale. A large place in the survey is given to illustrations obtained by simulations. Some of them derive from an approximation method based on Fleming-Viot interacting particle systems, which is carefully described in Section 6.

A brief bibliography on quasi-stationary distributions

The study of quasi-stationary distributions began with the work of Yaglom on sub-critical Galton-Watson processes [67]. Since then, the existence, uniqueness and other properties of quasi-stationary distributions for various processes have been studied.

In particular, the case of Markov processes on finite state spaces has been studied by Darroch and Seneta, who proved under some irreducibility conditions the existence and uniqueness of the QSD, for both discrete [18] and continuous time settings [19] (detailed proofs and results are reproduced in Section 3 of the present paper). We also refer the reader to the works of van Doorn and Pollett [61] for a relaxation of the irreducibility condition.

The case of discrete time birth and death processes has been treated by Seneta and Vere-Jones [54] and Ferrari, Martínez and Picco [24]. For continuous time birth and death processes, we refer to van Doorn [59]. This last case is quite enlightening, since it leads to examples of processes with no QSD, of processes with a Yaglom limit and an infinite number of QSD and to processes with a Yaglom limit which is the unique QSD (detailed proofs and results are also developed in Section 4 of this survey). For further developments, we may refer to Pakes and Pollett [48] (where results on continuous-time birth and death processes with catastrophic events are obtained), to Bansaye [5] (where a discrete time branching process in random environment is studied), to Coolen-Schrijner [17] (where general discrete time processes are studied) and references therein.

Diffusion processes have also been extensively studied in the past decades, beginning with the seminal work of Mandl [45] for the one-dimensional case and of Pinsky [49] and Gong, Qian and Zhao [29] in the multi-dimensional situation. Martínez, Picco and San Martín [46] and Lladser and San Martín [44] highlighted cases of diffusions with infinitely many quasi-limiting distributions, with a non-trivial dependence on the initial distribution of the process. For recent development of the theory of QSDs for diffusion processes, we refer to Steinsaltz and Evans [57] and Kolb and Steinsaltz [41] where the case of one

dimensional diffusions with different boundary conditions is studied. We also emphasize that in the case of Wright-Fisher diffusions and some of its relatives, Huillet [35] derived explicit values of QSDs. Other diffusion processes related to demographic models have been studied in Cattiaux, Collet, Lambert, Martínez, Méléard and San Martín [13], where the case of the Feller logistic diffusion is developed (proofs and results are also written in detail in Section 5 of this paper), and in Cattiaux and Méléard [14], where the case of a two dimensional stochastic Lotka-Volterra system is developed (k types stochastic Lotka-Volterra systems are also studied in Section 5.4 of this survey).

Let us mention the original renewal approach of Ferrari, Kesten, Martínez and Picco [22], also studied recently by Barbour and Pollett [6] in order to provide an approximation method for the QSD of discrete state space Markov processes in continuous time. Other approximation methods have been proposed by Pollett and Stewart [51] and by Hart and Pollett [34]. In this survey, we describe the approximation method based on Fleming-Viot type interacting particle systems, introduced by Burdzy, Holyst, Ingerman et March [10] in 1996 and studied later by Burdzy, Holyst and March [11], Grigorescu and Kang [32], Ferrari and Marič [23], Villemonais [65] [66] and Asselah, Ferrari and Groisman [3].

For studies on the so-called Q-process, which is the process distributed as the original process conditioned to never extinct, we refer the reader to the above cited articles [45], [49], [29], [18], [19] and, for further developments, to the works of Collet, Martínez and San Martín [16] and of Lambert [43] and references therein.

The framework

Let us now introduce our framework in more details. The population's size $(Z_t, t \geq 0)$ is a Markov process taking values in a subset E of $\mathbb{N}$ or $\mathbb{R}_+$, in a discrete or continuous time setting. If the population is isolated, namely without immigration, then the state 0, which describes the extinction of the population, is a trap. Indeed, if there are no more individuals, no reproduction can occur and the population disappears. Thus if the system reaches 0, it stays there forever, that is, if $Z_t = 0$ for some t, then $Z_s = 0$ for any $s \geq t$.

We denote by T_0 the extinction time, i.e. the stopping time

$$T_0 = \inf\{t > 0, Z_t = 0\}. \tag{1}$$

We will consider cases for which the process goes almost surely to zero, whatever the initial state is, namely, for all $z \in E$,

$$\mathbb{P}_z(T_0 < \infty) = 1. \tag{2}$$

Before extinction, the process takes its values in the space $E^* = E\backslash\{0\}$. Any long time distribution of the process conditioned on non-extinction will be supported by E^*.

Notations For any probability measure μ on E^*, we denote by $\mathbb{P}_\mu$ (resp. $\mathbb{E}_\mu$) the probability (resp. the expectation) associated with the process Z initially distributed with respect to μ. For any $x \in E^*$, we set $\mathbb{P}_x = \mathbb{P}_{\delta_x}$ and $\mathbb{E}_x = \mathbb{E}_{\delta_x}$. We denote by $(P_t)_{t\geq 0}$ the semi-group of the process Z killed at 0. More precisely, for any $z > 0$ and f measurable and bounded on E^*, one defines

$$P_t f(z) = \mathbb{E}_z(f(Z_t)\mathbf{1}_{t<T_0}). \tag{3}$$

For any finite measure μ and any bounded measurable function f, we set

$$\mu(f) = \int_{E^*} f(x)\mu(dx),$$

and we also define the finite measure μP_t by

$$\mu P_t(f) = \mu(P_t f) = \mathbb{E}_\mu(f(Z_t)\mathbf{1}_{t<T_0}).$$

2 Definitions, general properties and first examples

There are several natural questions associated with this situation.

Question 1 What is the distribution of the size of a non-extinct population at a large time t ? The mathematical quantity of interest is thus the conditional distribution of Z_t defined, for any Borel subset $A \subset E^*$, by

$$\mathbb{P}_\nu(Z_t \in A | T_0 > t) = \frac{\mathbb{P}_\nu(Z_t \in A; T_0 > t)}{\mathbb{P}_\nu(T_0 > t)} = \frac{\nu P_t(\mathbf{1}_A)}{\nu P_t(\mathbf{1}_{E^*})}, \tag{4}$$

where ν is the initial distribution of the population's size Z_0. We will study the asymptotic behavior of this conditional probability when t tends to infinity. The first definition that we introduce concerns the existence of a limiting conditional distribution.

Definition 1. *Let α be a probability measure on E^*. We say that α is a **quasi-limiting distribution** (QLD) for Z, if there exists a probability measure ν on E^* such that, for any measurable set $A \subset E^*$,*

$$\lim_{t\to\infty} \mathbb{P}_\nu \left(Z_t \in A | T_0 > t\right) = \alpha(A).$$

In some cases the long time behavior of the conditioned distribution can be proved to be initial state independent. This leads to the following definition.

Definition 2. *We say that Z has a **Yaglom limit** if there exists a probability measure α on E^* such that, for any $x \in E^*$ and any measurable set $A \subset E^*$,*

$$\lim_{t\to\infty} \mathbb{P}_x \left(Z_t \in A | T_0 > t\right) = \alpha(A). \tag{5}$$

When it exists, the Yaglom limit is a QLD. The reverse isn't true in general and (5) will actually not imply the same property for any initial distribution.

Question 2 As in the ergodic case, we can ask if this Yaglom limit has the conditional stationarity property given by the following definition.

Definition 3. *Let α be a probability measure on E^*. We say that α is a **quasi-stationary distribution** (QSD) if, for all $t \geq 0$ and any measurable set $A \subset E^*$,*

$$\alpha(A) = \mathbb{P}_\alpha \left(Z_t \in A | T_0 > t\right).$$

The main questions are: Does a QSD exists? Is there a unique QSD for the process? We will study examples where QSDs do not exist, or with an infinity of QSDs, or with a unique QSD. The relation between the existence of QSD, QLD and Yaglom limit is clarified in Proposition 1 below. Namely, we will prove that

$$\text{Yaglom limit } \Rightarrow \text{QSD } \Leftrightarrow \text{ QLD.}$$

Question 3 Since the processes we are interested in become extinct in finite time almost surely, the event $t < T_0$ becomes a rare event when t becomes large. An important question is then to know whether the convergence to the Yaglom limit happens before the typical time of extinction, or if it happens only after very large time periods, in which case the populations whose size are distributed with respect to the Yaglom limit are very rare. Both situations can appear, as illustrated by the simple example of Section 2.3.

Question 4 While most of theoretical results on QLDs, QSDs and Yaglom limits are concerned with existence and uniqueness problems, it would be useful in practice to have qualitative information on the Yaglom limit. We present here particle approximation results and numerical computations of the Yaglom limit for some population's size models, providing some enlightenment on Question 3 above.

Question 5 Another mathematical quantity related to this conditioning is based on a pathwise point of view. In the finite state space case of Section 3 and the logistic Feller diffusion case of Section 5, we will describe the distribution of the trajectories who never attain the trap. This will allow us to define a process, commonly referred to as the Q process for Z. We will prove that the new process defined by this distribution is ergodic, and that its stationary distribution is absolutely continuous with respect to the QSD (but not equal).

The present section is organized as follows. In Subsection 2.1, we state general properties of QLDs, QSDs and Yaglom limits. In Subsection 2.2, we develop the case of the Galton-Watson process. This discrete time process is of historical importance, since the notion of Yaglom limit has originally been developed for this process by Yaglom itself (see [67]). In Subsection 2.3, we develop a very simple example of a process evolving in a finite subset of $\mathbb{N}$. For this process, one can easily prove the existence of the Yaglom limit, the uniqueness of the QSD, and compare the speed of extinction to the speed of convergence to the Yaglom limit. We also provide numerical computation of the relevant quantities.

2.1 General properties

Most of the following results are already known by the QSD community. In this section, we emphasize their generality.

2.1.1 QSD, QLD and Yaglom limit

It is clear that any Yaglom limit and any QSD is also a QLD. The reverse implication has been proved by Vere-Jones [63] for continuous time Markov chains evolving in a countable state space. The following proposition extends this result to the general setting.

Proposition 1. *Let α be a probability measure on E^*. The distribution α is a QLD for Z if and only if it is a QSD for Z.*

Remark 1. When it exists, the Yaglom limit is uniquely defined, while there are processes with an infinity of QSDs (see the birth and death process case of Section 4). We immediately deduce that there exist QSDs which aren't a Yaglom limit.

Proof. (1) If α is a QSD then it is a QLD for Z starting with distribution α.

(2) Assume now that α is a QLD for Z and for an initial probability measure μ on E^*. Thus, for any measurable and bounded function f on E^*,

$$\alpha(f) = \lim_{t\to\infty} \mathbb{E}_\mu(f(Z_t)|T_0 > t) = \lim_{t\to\infty} \frac{\mathbb{E}_\mu(f(Z_t); T_0 > t)}{\mathbb{P}_\mu(T_0 > t)}.$$

Applying the latter with $f(z) = \mathbb{P}_z(T_0 > s)$, we get by the Markov property

$$\mathbb{P}_\alpha(T_0 > s) = \lim_{t\to\infty} \frac{\mathbb{P}_\mu(T_0 > t+s)}{\mathbb{P}_\mu(T_0 > t)}.$$

Let us now consider $f(z) = \mathbb{P}_z(Z_s \in A, T_0 > s)$, with $A \subset E^*$. Applying the Markov property again, it yields

$$\begin{aligned}\mathbb{P}_\alpha(Z_s \in A; T_0 > s) &= \lim_{t\to\infty} \frac{\mathbb{P}_\mu(Z_{t+s} \in A; T_0 > t+s)}{\mathbb{P}_\mu(T_0 > t)} \\ &= \lim_{t\to\infty} \frac{\mathbb{P}_\mu(Z_{t+s} \in A; T_0 > t+s)}{\mathbb{P}_\mu(T_0 > t+s)} \frac{\mathbb{P}_\mu(T_0 > t+s)}{\mathbb{P}_\mu(T_0 > t)}.\end{aligned}$$

By definition of the QLD α, $\frac{\mathbb{P}_\mu(Z_{t+s}\in A; T_0>t+s)}{\mathbb{P}_\mu(T_0>t+s)}$ converges to $\alpha(A)$ and $\frac{\mathbb{P}_\mu(T_0>t+s)}{\mathbb{P}_\mu(T_0>t)}$ converges to $\mathbb{P}_\alpha(T_0 > s)$, when t goes to infinity. We deduce that, for any Borel set A of E^* and any $s > 0$,

$$\alpha(A) = \mathbb{P}_\alpha(Z_s \in A|T_0 > s).$$

The probability measure α is then a QSD. □

2.1.2 Exponential extinction rate

Proposition 2. *Let us consider a Markov process Z with absorbing point 0 satisfying (2). Assume that α is a QSD for the process. Then there exists a positive real number $\theta(\alpha)$ depending on the QSD such that*

$$\mathbb{P}_\alpha(T_0 > t) = e^{-\theta(\alpha)t}. \tag{6}$$

This theorem shows us that starting from a QSD, the extinction time has an exponential distribution with parameter $\theta(\alpha)$ independent of $t > 0$, given by

$$\theta(\alpha) = -\frac{\ln \mathbb{P}_\alpha(T_0 > t)}{t}.$$

Proof. By the Markov property,

$$\begin{aligned}\mathbb{P}_\alpha\left(T_0 > t+s\right) &= \mathbb{E}_\alpha\left(\mathbb{P}_{Z_t}(T_0 > s)\mathbf{1}_{T_0>t}\right)\\ &= \mathbb{P}_\alpha(T_0>t)\mathbb{E}_\alpha\left(\mathbb{P}_{Z_t}(T_0>s)|T_0>t\right),\end{aligned}$$

since $T_0 \leq t$ implies $Z_t = 0$, and $\mathbb{P}_0(T_0 > s) = 0$. By definition of a QSD, we get

$$\mathbb{E}_\alpha\left(\mathbb{P}_{Z_t}(T_0>s)|T_0>t\right) = \mathbb{P}_\alpha(T_0>s).$$

Hence we obtain that for all s,$t > 0$, $\mathbb{P}_\alpha(T_0 > t+s) = \mathbb{P}_\alpha(T_0>s)\mathbb{P}_\alpha(T_0>t)$. Let us denote $g(t) = \mathbb{P}_\alpha(T_0>t)$. We have $g(0) = 1$ and, because of (2), $g(t)$ tends to 0 as t tends to infinity. An elementary proof allows us to conclude that there exists a real number $\theta(\alpha) > 0$ such that

$$\mathbb{P}_\alpha(T_0>t) = e^{-\theta(\alpha)t}.$$

□

2.1.3 QSD and exponential moments

The following statement gives a necessary condition for the existence of QSDs in terms of existence of exponential moments of the hitting time T_0.

Proposition 3. *Assume that α is a QSD. Then, for any $0 < \gamma < \theta(\alpha)$,*

$$\mathbb{E}_\alpha(e^{\gamma T_0}) < +\infty. \tag{7}$$

In particular, there exists a positive number z such that $\mathbb{E}_z(e^{\gamma T_0}) < +\infty$.

Proposition 3 suggests that if the population can escape extinction for too long times with positive probability, then the process has no QSD. This is the case for the critical Galton-Watson process: its extinction time is finite almost surely, but its expectation isn't finite.

Proof. We compute the exponential moment in continuous and discrete time settings. In both cases, it is finite if and only if $\theta(\alpha) > \gamma$.

In the continuous time setting, (6) says that, under $\mathbb{P}_\alpha$, T_0 has an exponential distribution with parameter $\theta(\alpha)$. We deduce that, for any $\theta(\alpha) > \gamma$,

$$\mathbb{E}_\alpha\left(e^{\gamma T_0}\right) = \frac{\theta(\alpha)}{\theta(\alpha)-\gamma}.$$

In the discrete time setting, (6) says that under $\mathbb{P}_\alpha$, T_0 has a geometric distribution with parameter $e^{-\theta(\alpha)}$. We deduce that

$$\mathbb{E}_\alpha\left(e^{\gamma T_0}\right) = \frac{1-e^{-\theta(\alpha)}}{e^{-\gamma}-e^{-\theta(\alpha)}}.$$

Since $\mathbb{E}_\alpha(e^{\gamma T_0})$ is equal to $\int_{E^*}\mathbb{E}_z(e^{\gamma T_0})\alpha(dz)$, the finiteness of the integral implies the last assertion.

□

Remark 2. In the particular case of an irreducible continuous time Markov chain with state space $\mathbb{N}$ such that $\lim_{z\to+\infty} \mathbb{P}_z(T_0 \leq t) = 0$, $\forall t \geq 0$, Ferrari, Kesten, Martínez and Picco [22] proved that the existence of the moment (7) for some $z \in \mathbb{N}$ and some $\gamma > 0$ is equivalent to the the existence of a quasi-stationary distribution.

It is actually not true in any case, as shown by the following counter-example. Let Z be a continuous time random walk on $\mathbb{N}$ reflected on 1 and killed at rate 1. Thus, for any $\lambda \in [0,1[$ and any probability measure μ on $\mathbb{N}$, $\mathbb{E}_\mu(e^{\lambda T_0})$ is finite. Nevertheless the conditional distribution $\mathbb{P}_z(Z_t \in \cdot | t < T_0)$ is the distribution of a standard continuous time random walk reflected on 1, which converges to 0 as t tends to infinity. In particular Z has no QLD and thus no QSD.

2.1.4 A spectral point of view

In this section, the results are stated in the continuous time setting. The operator L with domain $\mathcal{D}(L)$ denotes the infinitesimal generator of the sub-Markovian semi-group (P_t) associated with the killed process Z. The next proposition links the existence of QSDs for Z and the spectral properties of the dual of the operator L. It is one of the main tools used in a large literature studying QSDs.

Proposition 4. *Let α be a probability measure on E^*. We assume that there exists a set $D \subset \mathcal{D}(L)$ such that, for any measurable subset $A \subset E^*$, there exists a uniformly bounded sequence (f_n) in D converging point-wisely to $\mathbf{1}_A$.*

Then α is a quasi-stationary distribution if and only if there exists $\theta(\alpha) > 0$ such that

$$\alpha(Lf) = -\theta(\alpha)\alpha(f),\ \forall f \in D.$$

We emphasize that the existence of D is always true if the state space E^* is discrete. It is also fulfilled if E^* is an open subset of $\mathbb{R}^d$ and if Z is a diffusion with locally bounded coefficients.

Proof. (1) Let α be a QSD for Z. By definition of a QSD, we have, for every Borel set $A \subseteq E^*$,

$$\alpha(A) = \frac{\alpha P_t(\mathbf{1}_A)}{\alpha P_t(\mathbf{1}_{E^*})}.$$

By Theorem 6, there exists $\theta(\alpha) > 0$ such that for each $t > 0$,

$$\alpha P_t(\mathbf{1}_{E^*}) = \mathbb{P}_\alpha(T_0 > t) = e^{-\theta(\alpha)t}.$$

We deduce that, for any measurable set $A \subseteq E^*$, $\alpha P_t(\mathbf{1}_A) = e^{-\theta(\alpha)t}\alpha(A)$, which is equivalent to $\alpha P_t = e^{-t\theta(\alpha)}\alpha$. By Kolmogorov's forward equation and by assumption on D, we have

$$\left|\frac{\partial P_t f}{\partial t}(x)\right| = |P_t L f(x)| \leq \|Lf\|_\infty < +\infty,\ \forall f \in D.$$

In particular, one can differentiate $\alpha P_t f = \int_{E^*} P_t f(x)\alpha(dx)$ under the integral sign, which implies that

$$\alpha(Lf) = -\theta(\alpha)\alpha(f),\ \forall f \in D.$$

(2) Assume know that $\alpha(Lf) = -\theta(\alpha)\alpha(f)$ for all $f \in D$. By Kolmogorov's backward equation and the same "derivation under the integral sign" argument, we have

$$\frac{\partial \alpha(P_t f)}{\partial t} = \alpha(LP_t f) = -\theta(\alpha)\alpha P_t(f), \, \forall f \in D.$$

We deduce that

$$\alpha P_t(f) = e^{-t\theta(\alpha)}\alpha(f), \, \forall f \in D.$$

By assumption, there exists, for any measurable subset $A \subset E^*$, a uniformly bounded sequence (f_n) in D which converges point-wisely to $\mathbf{1}_A$. Finally, we deduce by dominated convergence that

$$\alpha P_t(\mathbf{1}_A) = e^{-t\theta(\alpha)}\alpha(A).$$

This implies immediately that α is a quasi-stationary distribution for Z. □

2.1.5 Long time limit of the extinction rate

Another quantity of interest in the demography and population's dynamics is given by the long time behavior of the killing or extinction rate. In the demography setting, the process Z models the vitality of some individual and t its physical age. Thus T_0 is the death time of this individual. The long time behavior of the extinction rate has been studied in detail by Steinsaltz-Evans [56] for specific cases.

The definition of the extinction rate depends on the time setting:

- In the discrete time setting, the extinction rate of Z starting from μ at time $t \geq 0$ is defined by
$$r_\mu(t) = \mathbb{P}_\mu(T_0 = t + 1 | T_0 > t).$$
- In the continuous time setting, the extinction rate of Z starting from μ at time $t \geq 0$ is defined by
$$r_\mu(t) = -\frac{\frac{\partial}{\partial t}\mathbb{P}_\mu(T_0 > t)}{\mathbb{P}_\mu(T_0 > t)},$$
when the derivative exists and is integrable with respect to μ.

Historically (*cf.* [28]), demographers applied the Gompertz law meaning that this extinction rate was exponentially increasing with time. However in 1932, Greenwood and Irwin [31] observed that in some cases, this behavior was not true. In particular there exist cases where the extinction rate converges to a constant when time increases, leading to the notion of mortality plateau. This behavior of the extinction rate has been observed in experimental situations (see for instance [12]).

The QSDs play a main role in this framework. Indeed, by Proposition 2, if α is a QSD, then the extinction rate $r_\alpha(t)$ is constant and given by

$$r_\alpha(t) = \begin{cases} 1 - e^{-\theta(\alpha)} \text{ in the discrete time setting} \\ \theta(\alpha) \text{ in the continuous time setting} \end{cases}, \, \forall t \geq 0.$$

We refer to the introduction of Steinsaltz-Evans [56] for a nice discussion of the notion of QSD in relationship with mortality plateaus.

In the next proposition, we prove that the existence of a QLD for Z started from μ implies the existence of a long term mortality plateau.

Proposition 5. *Let α be a QLD for Z, initially distributed with respect to a probability measure μ on E^*. In the continuous time setting, we assume moreover that there exists $h > 0$ such that $L(P_h \mathbf{1}_{E^*})$ is well defined and bounded. In both time settings, the rate of extinction converges in the long term:*

$$\lim_{t\to\infty} r_\mu(t) = r_\alpha(0). \tag{8}$$

Proof. In the discrete time setting, by the semi-group property and the definition of a QLD, we have

$$r_\mu(t) = 1 - \frac{\mu P_t(P_1 \mathbf{1}_{E^*})}{\mu P_t(\mathbf{1}_{E^*})} \xrightarrow[t\to+\infty]{} 1 - \alpha(P_1 \mathbf{1}_{E^*}) = r_\alpha(0).$$

The limit is by definition the extinction rate at time 0 of Z starting from α.
In the continuous time setting, by the Kolmogorov's backward equation, we have

$$\frac{\partial}{\partial t} P_{t+h} \mathbf{1}_{E^*}(x) = P_t L(P_h \mathbf{1}_{E^*})(x),\ \forall x \in E^*.$$

Since $L(P_h \mathbf{1}_{E^*})$ is assumed to be bounded, we deduce that

$$\frac{\partial}{\partial t} \mu P_{t+h}(\mathbf{1}_{E^*}) = \mu P_t L(P_h \mathbf{1}_{E^*}).$$

Then

$$\frac{\frac{\partial}{\partial t}\mu P_{t+h}(\mathbf{1}_{E^*})}{\mu P_t(\mathbf{1}_{E^*})} = \frac{\mu P_t L(P_h \mathbf{1}_{E^*})}{\mu P_t(\mathbf{1}_{E^*})} \xrightarrow[t\to\infty]{} \alpha(L P_h \mathbf{1}_{E^*}) = -\theta(\alpha)\alpha(P_h \mathbf{1}_{E^*}),$$

by the definition of a QLD and by Proposition 4. We also have

$$\frac{\mu(P_{t+h}\mathbf{1}_{E^*})}{\mu(P_t \mathbf{1}_{E^*})} \xrightarrow[t\to\infty]{} \alpha(P_h \mathbf{1}_{E^*}).$$

Finally, we get

$$r_\mu(t+h) = -\frac{\frac{\partial}{\partial t}\mu(P_{t+h}\mathbf{1}_{E^*})}{\mu(P_{t+h}\mathbf{1}_{E^*})} \xrightarrow[t\to\infty]{} \theta(\alpha),$$

which allows us to conclude the proof of Proposition 5. □

2.2 A historical example in discrete time: the Galton-Watson process

The Galton-Watson process is a population's dynamics model in discrete time, whose size $(Z_n)_{n\geq 0}$ evolves according to the recurrence formula Z_0 and

$$Z_{n+1} = \sum_{i=1}^{Z_n} \xi_i^{(n)},$$

where $(\xi_i^{(n)})_{i,n}$ is a family of independent random variables, identically distributed following the probability measure μ on $\mathbb{N}$ with generating function g. As defined, Z_n is the

size of the n^{th} generation of a population where each individual has a random number of children, chosen following μ and independently of the rest of the population. This process has been introduced by Galton and Watson (see [26]) in order to study the extinction of aristocratic surnames.
We will assume in the whole section that $0 < \mu(\{0\}) + \mu(\{1\}) < 1$. We denote by $m = E(\xi_1^{(0)})$ the average number of children by individual in our Galton-Watson process. The independence of descendants implies that starting from Z_0, the process Z is equal to the sum of Z_0 independent Galton-Watson processes issued from a single individual. By this branching property, the probability of extinction for the population starting from one individual is obtained as follows:

$$\mathbb{P}_1(\exists n \in \mathbb{N},\, Z_n = 0) = \lim_{n\to+\infty} \mathbb{E}_1(0^{Z_n}) = \lim_{n\to\infty} g \circ \cdots \circ g(0) \text{ } (n \text{ times}).$$

There are three different situations (see for instance Athreya-Ney [4]):

- The sub-critical case $m < 1$: the process becomes extinct in finite time almost surely and the average extinction time $\mathbb{E}(T_0)$ is finite.

- The critical case $m = 1$: the process becomes extinct in finite time almost surely, but $\mathbb{E}(T_0) = +\infty$.

- The super-critical case $m > 1$: the process is never extinct with a positive probability, and it yields immediately that $\mathbb{E}(T_0) = +\infty$.

Theorem 6 (Yaglom [67], 1947)**.** *Let $(Z_n)_{n\geq 0}$ be a Galton-Watson process with the reproduction generating function g. There is no quasi-stationary distribution in the critical and the super-critical case. In the sub-critical case, the Yaglom limit exists and is the unique QSD of Z. Moreover, its generating function $\hat{g}$ fulfills*

$$\hat{g}(g(s)) = m\hat{g}(s) + 1 - m,\ \forall s \in [0,1]. \tag{9}$$

Proof. The proof is adapted from Athreya-Ney [4] p. 13-14. In the critical or the super-critical case, we have $\mathbb{E}_1(T_0) = +\infty$, which implies that $\mathbb{E}_\alpha(T_0) = +\infty$ for all probability measure α on $\mathbb{N}^*$. We deduce from Proposition 3 that there is no QSD.

Assume now that $m < 1$. Let us fix an arbitrary probability measure ν on $\mathbb{N}^*$ and prove that there exists a QLD α for Z starting with distribution ν. For each $n \geq 0$, we denote by g_n the generating function of Z_n, $g_n(s) = \mathbb{E}_\nu\left(s^{Z_n}\right)\ \forall s \in [0,1]$. Recall that $g_{n+1} = g_1 \circ g_n$. Let us also denote by $\hat{g}_n$ the generating function of Z_n conditioned to $\{Z_n > 0\} = \{T_0 > n\}$:

$$\begin{aligned}
\hat{g}_n(s) &= \mathbb{E}_\nu(s^{Z_n} |\, Z_n > 0) = \frac{\mathbb{E}_\nu(s^{Z_n}\mathbf{1}_{Z_n>0})}{\mathbb{P}_\nu(Z_n > 0)} \\
&= \frac{\mathbb{E}_\nu(s^{Z_n}) - \mathbb{P}_\nu(Z_n = 0)}{1 - \mathbb{P}_\nu(Z_n = 0)} \\
&= \frac{g_n(s) - g_n(0)}{1 - g_n(0)} = 1 - \frac{1 - g_n(s)}{1 - g_n(0)} \in [0,1].
\end{aligned}$$

Note that $\hat{g}_n(0) = 0$, which is quite natural since the conditional law doesn't charge 0. For a fixed $s \in [0,1)$, we set $\Gamma(s) = \frac{1-g_1(s)}{1-s}$. Then we have, for all $n \geq 0$,

$$1 - \hat{g}_{n+1}(s) = \frac{\Gamma(g_n(s))}{\Gamma(g_n(0))} \left(1 - \hat{g}_n(s)\right).$$

Since g_1 is convex, Γ is non-decreasing. Moreover $m < 1$ implies that $g_n(x) \geq x$, so that $g_n(s)$ and $1 - \hat{g}_n(s)$ are non-decreasing in n. In particular, $\lim_{n\to\infty} \hat{g}_n(s)$ exists. Let us denote by $\hat{g}(s)$ its limit and by α the corresponding finite measure (whose mass is smaller than one). In order to prove that α is a probability measure on $\mathbb{N}^*$, it is sufficient to prove that $\hat{g}(s) \to 1$ when s goes to 1. We have

$$\Gamma(g_n(0)) \left(1 - \hat{g}_{n+1}(s)\right) = \left(1 - \hat{g}_n(g_1(s))\right).$$

Taking the limit on each size, where $\lim_{n\to\infty} \Gamma(g_n(0)) = \Gamma(1) = m$, we deduce that

$$m(1 - \hat{g}(s)) = 1 - \hat{g}(g_1(s)),$$

which implies Equation (9). Since $\lim_{s\to 1} g_1(s) = 1$ and $m < 1$, then $\hat{g}(1) = 1$. Finally, α is a QLD for Z starting with distribution ν.

One could think *a priori* that the function $\hat{g}$ depends on the starting distribution ν. We prove now that it isn't the case, so that there is a unique QLD, and then a unique QSD, which is also the Yaglom limit of the process (indeed, one could choose $\nu = \delta_x$, $x \in \mathbb{N}^*$). Assume that there exist two generating functions $\hat{g}$ and $\hat{h}$ which fulfill Equation (9). By induction, we have, for all $n \geq 1$ and all $s \in [0,1]$,

$$\begin{aligned}
\hat{g}(g_n(s)) &= m^n \hat{g}(s) + \left(m^{n-1} + \cdots + m + 1\right)(m-1), \\
\hat{h}(g_n(s)) &= m^n \hat{h}(s) + \left(m^{n-1} + \cdots + m + 1\right)(m-1).
\end{aligned}$$

We deduce that for $s \in [0,1[$

$$\hat{g}'(g_n(s))\; g_n'(s) = m^n\, \hat{g}'(s)\;;\; \hat{h}'(g_n(s))\; g_n'(s) = m^n\, \hat{h}'(s).$$

Since for the sub-critical case $g_n(0) \uparrow 1$ when $n \to \infty$, for any $s \in [0,1[$ there will be a k such that

$$g_k(0) \leq s \leq g_{k+1}(0).$$

Hence,

$$\frac{\hat{g}'(s)}{\hat{h}'(s)} = \frac{\hat{g}'(g_n(s))}{\hat{h}'(g_n(s))} \leq \frac{\hat{g}'(g_{n+k+1}(0))}{\hat{h}'(g_{n+k}(0))} = \frac{\hat{g}'(0)}{\hat{h}'(0)} \frac{m\, g_{n+k}'(0)}{g_{n+k+1}'(0)} = \frac{\hat{g}'(0)}{\hat{h}'(0)} \frac{m}{g'(g_{n+k}(0))}.$$

When n goes to infinity, we obtain $\frac{\hat{g}'(s)}{\hat{h}'(s)} \leq \frac{\hat{g}'(0)}{\hat{h}'(0)}$. The converse inequality is established similarly. Since $\hat{g}$ and $\hat{h}$ are generating functions of probability measures on $\mathbb{N}^*$, we have $\hat{g}(0) = \hat{h}(0) = 0$ and $\hat{g}(1) = \hat{h}(1) = 1$. Finally, the two functions $\hat{g}$ and $\hat{h}$ are equal, which concludes the proof of Theorem 6 . □

2.3 The simple example of an ergodic process with uniform killing in a finite state space

We present a very simple Markov process with extinction whose quasi-stationary distribution, Yaglom limit, speed of extinction and speed of convergence to the Yaglom limit are very easy to obtain.
Let $(X_t)_{t\geq 0}$ be an exponentially ergodic Markov process which evolves in the state space $E^* = \{1, \cdots, N\}$, $N \geq 1$. By exponentially ergodic, we mean that there exist a probability measure α on E^* and two positive constants C,$\lambda > 0$ such that, for all $z \in \{1, \cdots, N\}$ and all $t \geq 0$,

$$\sup_{i\in E^*} |\mathbb{P}_z(X_t = i) - \alpha(\{i\})| \leq Ce^{-\lambda t}.$$

There is no possible extinction for (X_t). Let $d > 0$ be a positive constant and let τ_d be an exponential random time of parameter d independent of the process (X_t). We define the process (Z_t) by setting

$$Z_t = \begin{cases} X_t, \text{ if } t < \tau_d \\ 0, \text{ if } t \geq \tau_d. \end{cases}$$

This model is a model for the size of a population which cannot be extinct, except at a catastrophic event which happens with rate d. Thus we have

$$\mathbb{P}_z(t < T_0) = e^{-dt}, \forall t \geq 0.$$

The conditional distribution of Z_t is simply given by the distribution of X_t:

$$\mathbb{P}_z(Z_t = i | Z_t \neq 0) = \mathbb{P}_z(X_t = i), \forall z \in E^*.$$

We deduce that the unique QSD is the Yaglom limit α and that for all $z \in E^*$ and all $t \geq 0$,

$$\sup_{i\in E^*} |\mathbb{P}_z(Z_t = i | T_0 > t) - \alpha(\{i\})| \leq Ce^{-\lambda t}.$$

Thus in this case, the conditional distribution of Z converges exponentially fast to the Yaglom limit α, with rate $\lambda > 0$ and the process becomes extinct exponentially fast, with rate $d > 0$.

Hence the comparison between the speed of convergence to the Yaglom limit and the speed of extinction will impact the observables of the process before extinction:

(a) If $\lambda \gg d$, then the convergence to the Yaglom limit happens before the typical time of extinction of the population and the quasi-stationary regime will be observable.

(b) If $\lambda \ll d$, then the extinction of the population occurs before the quasi-stationary regime is reached. As a consequence, we are very unlikely to observe the Yaglom limit.

(c) If $\lambda \sim d$, the answer is not so immediate and depends on other parameters, as in particular the initial distribution.

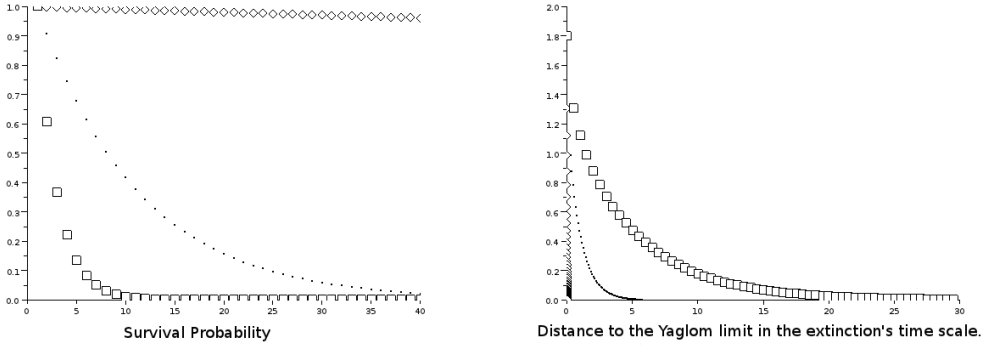


Figure 1: Example 1. A numerical computation leads to $\lambda = 0.098$. Three different situations are observed, which lead to three very different patterns for the speed of convergence to the Yaglom limit in the extinction's time scale: ($\diamond$) $\lambda \gg d = 0.001$; ($\square$) $\lambda \ll d = 0.500$; ($\cdot$) $\lambda = d = 0.098$.

Example 1. The population size Z is described by a random walk in continuous time evolving in $E = \{0,1,2,\cdots,N\}$ with transition rates given by

$$\begin{aligned} &i \to i+1 \text{ with rate 1, for all } i \in \{1,2,\cdots,N-1\},\\ &i \to i-1 \text{ with rate 1, for all } i \in \{2,3,\cdots,N\},\\ &i \to 0 \text{ with rate } d > 0, \text{ for all } i \in \{1,2,\cdots,N\}. \end{aligned}$$

The boundedness of the population size models a constraint of fixed resources which acts on the growth of the population. We will see more realistic fixed resources models including logistic death rate in the next sections. One can check that the quasi-stationary probability measure of Z is given by $\alpha_i = 1/N$ for all $i \in E^*$.

Numerical simulations. We fix $N = 100$. In that finite case, one can obtain by numerical computation the whole set of eigenvalues and eigenvectors of the infinitesimal generator L (we use here the software SCILAB). Numerical computation using the fact that λ is the spectral gap of the generator of X gives $\lambda = 0.098$. For different values $d = 0.001$, $d = 0.500$ and $d = 0.098$, we compute numerically the mathematical quantities of interest: the extinction probability $\mathbb{P}_z(T_0 > t) = e^{-dt}$ as a function of t (cf. Figure 1 left picture) and the distance $\sup_{i \in E^*} |\mathbb{P}_z(Z_t = i | T_0 > t) - \alpha(\{i\})|$ between the conditional distribution of Z_t and α as a function of $-\log \mathbb{P}_z(T_0 > t)$, which gives the extinction's time scale. (cf. Figure 1 right picture).

We observe that the convergence to the Yaglom limit happens rapidly in the case ($\diamond$) $\lambda = 0.098 \gg d = 0.001$. Indeed the distance to the Yaglom limit is equal to 0.05, while the survival probability can't be graphically distinguished from 1. On the contrary, we observe that the convergence happens very slowly in the case ($\square$) $\lambda = 0.098 \ll d = 0.500$. Indeed, the distance to the Yaglom limit is equal to 0.05 when the survival probability appears to be smaller than $e^{-15} \simeq 3 \times 10^{-7}$. The case ($\cdot$) $\lambda = 0.98 = d$ is an intermediate case, where the distance to the Yaglom limit is equal to 0.05 when the survival probability appears to be equal to $e^{-3} \simeq 0.05$.

3 The finite case, with general killing's rate

3.1 The quasi-stationary distributions

The Markov process $(Z_t)_{t\geq 0}$ evolves in continuous time in $E = \{0,1,...,N\}$, $N \geq 1$ and we still assume that 0 is its unique absorbing state. The semi-group $(P_t)_{t\geq 0}$ is the sub-Markovian semi-group of the killed process and we still denote by L the associated infinitesimal generator. In this finite state space case, the operators L and P_t are matrices, and a probability measure on the finite space E^* is a vector of non-negative entries whose sum is equal to 1. The results of this section have been originally proved by Darroch and Seneta ([18] and [19]).

Theorem 7. *Assume that Z is an irreducible and aperiodic process before extinction, which means that there exists $t_0 > 0$ such that the matrix P_{t_0} has only positive entries (in particular, it implies that P_t has positive entries for $t > t_0$). Then the Yaglom limit α exists and is the unique QSD of the process Z_t.*

Moreover, denoting by $\theta(\alpha)$ the extinction rate associated to α (see Proposition 2), there exists a probability measure π on E^* such that, for any $i,j \in E^*$,

$$\lim_{t\to\infty} e^{\theta(\alpha)t}\,\mathbb{P}_i(Z_t = j) = \pi_i\,\alpha_j$$

and

$$\lim_{t\to\infty} \frac{\mathbb{P}_i(T_0 > t+s)}{\mathbb{P}_j(T_0 > t)} = \frac{\pi_i}{\pi_j} e^{-\theta(\alpha)s}.$$

The main tool of the proof of Theorem 7 is the Perron-Frobenius Theorem, which gives us a complete description of the spectral properties of P_t and L. The main point is that the matrix P_1 has positive entries. For the proof of the Perron-Frobenius Theorem, we refer to Gantmacher [27] or Serre [55].

Theorem 8 (Perron-Frobenius Theorem)**.** *Let (P_t) be a submarkovian semi-group on $\{1,\cdots,N\}$ such that the entries of P_{t_0} are positive for $t_0 > 0$. Thus, there exists a unique positive eigenvalue ρ, which is the maximum of the modulus of the eigenvalues, and there exists a unique left-eigenvector α such that $\alpha_i > 0$ and $\sum_{i=1}^N \alpha_i = 1$, and there exists a unique right-eigenvector π such that $\pi_i > 0$ and $\sum_{i=1}^N \alpha_i\pi_i = 1$, satisfying*

$$\alpha P_{t_0} = \rho\,\alpha \;;\; P_{t_0}\pi = \rho\,\pi. \tag{10}$$

In addition, since (P_t) is a sub-Markovian semi-group, $\rho < 1$ and there exists $\theta > 0$ such that $\rho = e^{-\theta}$. Therefore

$$P_t = e^{-\theta t}A + \vartheta(e^{-\chi t}), \tag{11}$$

where A is the matrix defined by $A_{ij} = \pi_i\alpha_j$, and $\chi > \theta$ and $\vartheta(e^{-\chi t})$ denotes a matrix such that none of the entries exceeds $Ce^{-\chi t}$, for some constant $C > 0$.

Proof of Theorem 7. Applying Perron-Frobenius Theorem to the submarkovian semi-group $(P_t)_{t\geq 0}$, it is immediate from (11) that there exists $\theta > 0$ and a probability measure α on E^* such that, for any $i,j \in E^*$,

$$e^{\theta t}\mathbb{P}_i(Z_t = j) = e^{\theta t}[P_t]_{ij} = \pi_i\alpha_j + \vartheta(e^{-(\chi-\theta)t}). \tag{12}$$

Summing over $j \in E^*$, we deduce that

$$e^{\theta t}\mathbb{P}_i(T_0 > t) = \pi_i + \vartheta(e^{-(\chi-\theta)t}). \tag{13}$$

It follows that, for any $i,j \in E^*$,

$$\mathbb{P}_i(Z_t = j | T_0 > t) = \frac{\mathbb{P}_i(Z_t = j)}{\mathbb{P}_i(T_0 > t)} \xrightarrow[t\to\infty]{} \alpha_j.$$

Thus the Yaglom limit exists and is equal to α. Since E is finite, we have for any initial distribution ν on E^*

$$\lim_{t\to\infty} \mathbb{P}_\nu(Z_t = j | T_0 > t) = \sum_{i\in E^*} \nu_i \lim_{t\to\infty} \mathbb{P}_i(Z_t = j | T_0 > t) = \sum_{i\in E^*} \nu_i \alpha_j = \alpha_j.$$

We deduce that the Yaglom limit α is the unique QLD of Z, and thus it is its unique QSD. By Proposition 2, we have $\alpha P_1(\mathbf{1}_{E^*}) = e^{-\theta(\alpha)}$. By (10), this quantity is also equal to $e^{-\theta}$, so that $\theta = \theta(\alpha)$. The end of Theorem 7 is thus a straightforward consequence of (12) and (13) □

Remark 3. One can deduce from (12) and (13) that there exists a positive constant C_L such that

$$\sup_{j\in E^*, i\in E^*} |P_i(Z_t = j | Z_t > 0) - \alpha_j| \le C_L e^{-(\chi-\theta(\alpha))},$$

where the quantity $\chi - \theta(\alpha)$ is the spectral gap of L, *i.e.* the distance between the first and the second eigenvalue of L. Thus if the time-scale $\chi - \theta(\alpha)$ of the convergence to the quasi-limiting distribution is substantially bigger than the time scale of absorption ($\chi - \theta(\alpha) \gg \theta(\alpha)$), the process will relax to the QSD after a relatively short time, and after a much longer period, extinction will occur. On the contrary, if $\chi - \theta(\alpha) \ll \theta(\alpha)$, then the extinction happens before the process had time to relax to the quasi-limiting distribution.
In intermediate cases, where $\lambda - \theta(\alpha) \approx \theta(\alpha)$, the constant C_L, which depends on the whole set of eigenvalues and eigenfunctions of L, plays a main role which needs further investigations.

We generalize Example 1 to a more realistic case where the killing's rate can depend on the size of the population. For instance, it can be higher for a small population than for a big one.

Example 2. Let Z be a Markov process which models a population whose individuals reproduce and die independently, with individual birth rate $\lambda > 0$ and individual death rate $\mu = 1$. In order to take into account the finiteness of the resources, the process is reflected when it attains a given value N, that we choose here arbitrarily equal to 100. Thus the process Z evolves in the finite state space $\{0,1,\cdots,100\}$ and its transition rates are given by

$$i \to i+1 \text{ with rate } \lambda i, \text{ for all } i \in \{1,2,\cdots,99\},$$
$$i \to i-1 \text{ with rate } \mu i, \text{ for all } i \in \{1,2,3,\cdots,100\}.$$

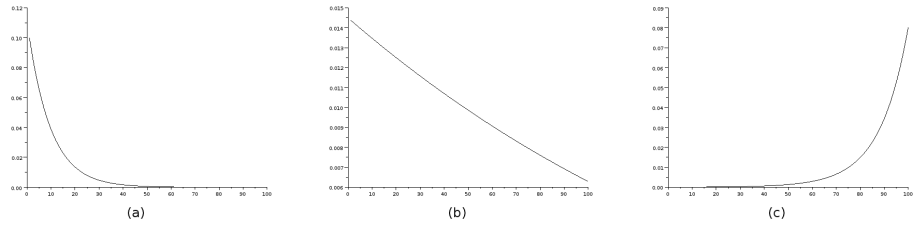


Figure 2: Example 2. Yaglom limits for different values of λ. The following values of $\theta(\alpha)$ are obtained by numerical computation. (a) $\lambda = 0.9$, $\theta(\alpha) = 0.100$; (b) $\lambda = 1.0$, $\theta(\alpha) = 0.014$; (c) $\lambda = 1.1$, $\theta(\alpha) = 5.84 \times 10^{-5}$.

The infinitesimal generator of Z is given by

$$\begin{aligned}
&L_{1,1} = -1 - \lambda \text{ and } L_{1,2} = \lambda,\\
&L_{i,i-1} = i, \quad L_{i,i} = -(1+\lambda)i \text{ and } L_{i,i+1} = \lambda\, i, \quad \forall i \in \{2, \cdots, 99\},\\
&L_{100,99} = 100 \text{ and } L_{100,100} = -100,\\
&L_{i,j} = 0, \, \forall i,j \in \{1, \cdots, 100\} \text{ such that } |j - i| > 1.
\end{aligned}$$

The process Z clearly fulfills the conditions of Theorem 7. As a consequence, it has a Yaglom limit α, which is its unique QSD. Moreover, the probability measure α is the unique normalized and positive left eigenvector of L. Since L is a finite matrix of size 100×100, one can numerically compute the whole set of eigenvectors and eigenvalues of the matrix (L_{ij}). This will allow to obtain numerically the Yaglom limit α, its associated extinction rate $\theta(\alpha)$, and the speed of convergence $\chi - \theta(\alpha)$. Moreover, for any $t \geq 0$, one can compute the value of e^{tL}, which is equal to P_t (the semi-group of Z at time t). Hence, we may obtain the numerical value of the conditioned distribution $\mathbb{P}_{Z_0}(Z_t \in .|t < T_0)$, for any initial size Z_0. Finally, we are also able to compute numerically the distance between α and the conditioned distribution $\mathbb{P}_{Z_0}(Z_t \in .|t < T_0)$, for any value of $\lambda > 0$ and $Z_0 \in \{1, \cdots, 100\}$.

In Figure 2, we represent the Yaglom limit α for different values of λ, namely $\lambda = 0.9$, $\lambda = 1.0$ and $\lambda = 1.1$. Let us comment the numerical results.

(a) In the first case ($\lambda = 0.9$), an individual is more likely to die than to reproduce and we observe that the Yaglom limit is concentrated near the absorbing point 0. The rate of extinction $\theta(\alpha)$ is the highest in this case, equal to 0.100. In fact, the process reaches the upper bound 100 very rarely, so that the behavior of the process is very similar to the one of a linear birth and death process with birth and death rates equal to λ and μ respectively. In Section 4, we study such linear birth and death processes. We show that the Yaglom limit (which exists if and only if $\lambda < \mu$) is given by a geometric law and $\theta(\alpha) = \mu - \lambda$.

(b) In the second case ($\lambda = \mu = 1$), we observe that α decreases almost linearly from α_1 to α_{100} and the upper bound $N = 100$ plays a crucial role. In fact, letting N tend to $+\infty$, one would observe that for any $i \geq 1$, α_i decreases to 0. The extinction rate $\theta(\alpha)$ which is equal to 0.014 for $N = 100$ would also go to 0. The counterpart of this phenomenon for the linear birth and death process studied in Section 4 is that the Yaglom limit will not exist when $\mu = \lambda$.

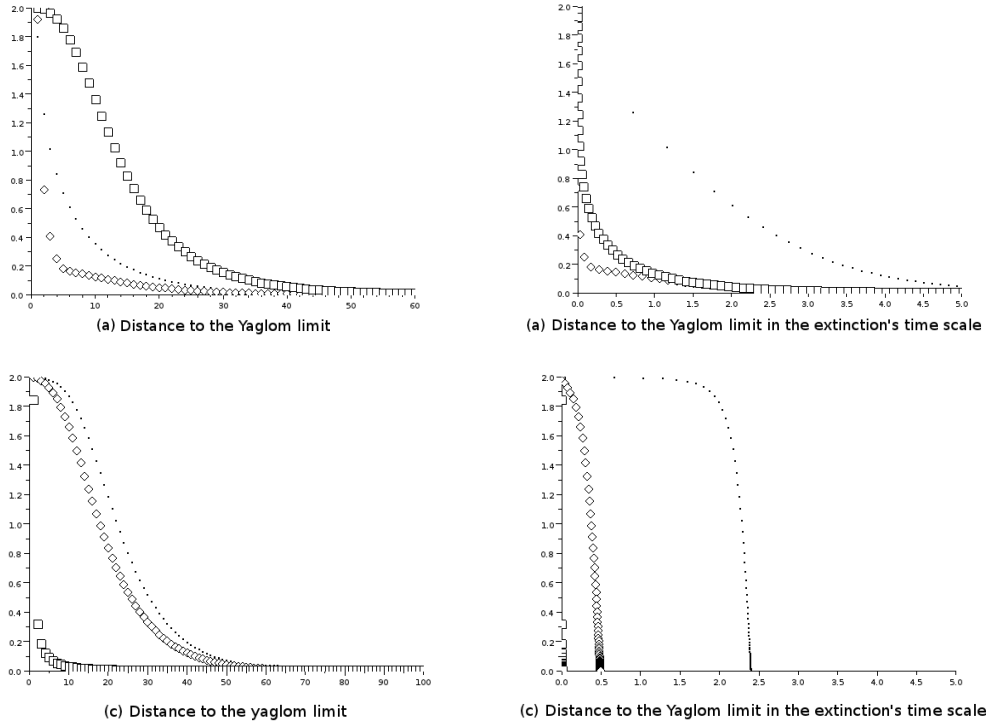


Figure 3: Example 2. Pictures (a) and (c) correspond to different values of λ (the following values of $\theta(\alpha) - \chi$ have been obtained by numerical computation): (a) $\lambda = 0.9$, $\theta(\alpha) = 0.100$, $\theta(\alpha) - \chi = 0.102$; (c) $\lambda = 1.1$, $\theta(\alpha) = 5.84 \times 10^{-5}$, $\theta(\alpha) - \chi = 0.103$; each curve corresponds to a given initial size of the population: $(\cdot)$ $Z_0 = 1$; $(\diamond)$ $Z_0 = 10$; $(\square)$ $Z_0 = 100$.

(c) In the third case ($\lambda = 1.1$), the Yaglom limit α is concentrated near the upper bound 100, while the extinction rate is $\theta(\alpha) = 5.84 \times 10^{-5}$. The comparison with the linear birth and death process is no more relevant, since the important factor in this case is the effect of the upper bound $N = 100$, which models the finiteness of the resources in the environment.

In Figure 3, we study the effect of the initial position and of the value of the parameter λ on the speed of convergence to the Yaglom limit and on the speed of extinction. We choose the positions $Z_0 = 1$, $Z_0 = 10$ and $Z_0 = 100$, and we look at the two different cases $\lambda = 0.9$ and $\lambda = 1.1$, which correspond to the subcritical case (a) and to the supercritical case (c) respectively. We represent, for each set of values of (λ,Z_0), the distance to the Yaglom limit, $\sup_{i\in\{1,\cdots,100\}} |\mathbb{P}_{Z_0}(Z_t = i | t < T_0) - \alpha_i|$ as a function of the time, and the same distance as a function of the logarithm of the survival probability $-\log \mathbb{P}_{Z_0}(t < T_0)$ (i.e. the extinction time scale). By numerical computation, we also obtain that

(a) $\lambda = 0.9$: $\theta(\alpha) = 0.100$ and $\theta(\alpha) - \chi = 0.102$.

(c) $\lambda = 1.1$: $\theta(\alpha) = 5.84 \times 10^{-5}$ and $\theta(\alpha) - \chi = 0.103$.

In the case (a), we have $\theta(\alpha) = 0.100 \simeq \chi - \theta(\alpha) = 0.102$ and we observe that the speed of convergence depends on the initial position in a non-trivial way: while the survival probability is smaller for the process starting from 10 than for the process starting from 100, the convergence to the Yaglom limit in the extinction's time scale happens faster

in the case $Z_0 = 10$. In the case (c), we have $\theta(\alpha) = 5.84 \times 10^{-5} \ll \chi - \theta(\alpha) = 0.103$. The speed of convergence to the Yaglom limit in the extinction's time scale depends on the initial position: if ($\square$) $Z_0 = 100$, then it is almost immediate; if ($\diamond$) $Z_0 = 10$, the distance between the conditional distribution and the Yaglom limit is equal to 0.05 when the survival probability is around $e^{-0.5} \simeq 0.61$; if ($\cdot$) $Z_0 = 1$, then this distance is equal to 0.05 when the survival probability is around $e^{-2.4} \simeq 0.091$.

3.2 The Q-process

Let us now study the marginals of the process conditioned to never be extinct.

Theorem 9. *Assume that we are in the conditions of Theorem 7. For any $i_0, i_1, \cdots, i_k \in E^*$, any $0 < s_1 < \cdots, s_k < t$, the limiting value $\lim_{t\to\infty} \mathbb{P}_{i_0}(Z_{s_1} = i_1, \cdots, Z_{s_k} = i_k | T_0 > t)$ exists.*

Let $(Y_t, t \geq 0)$ be the process starting from $i_0 \in E^*$ and defined by its finite dimensional distributions

$$\mathbb{P}_{i_0}(Y_{s_1} = i_1, \cdots, Y_{s_k} = i_k) = \lim_{t\to\infty} \mathbb{P}_{i_0}(Z_{s_1} = i_1, \cdots, Z_{s_k} = i_k | T_0 > t). \tag{14}$$

Then Y is a Markov process with values in E^* and transition probabilities given by

$$\mathbb{P}_i(Y_t = j) = e^{\theta(\alpha)t} \, \frac{\pi_i}{\pi_j} \, P_{ij}(t).$$

It is conservative, and has the unique stationary probability measure $(\alpha_j \pi_j)_j$.

Remark that the stationary probability is absolutely continuous with respect to the QSD, but, contrary to intuition, it is not equal to the QSD.

Proof. Let us denote $\theta(\alpha)$ by θ for simplicity. Let $i_0, i_1, \cdots, i_k \in E^*$ and $0 < s_1 < \cdots < s_k < t$. We introduce the filtration $\mathcal{F}_s = \sigma(Z_u, u \leq s)$. Then

$$\begin{aligned}
\mathbb{P}_{i_0}(Z_{s_1} = i_1, \cdots, Z_{s_k} = i_k \ ; \ T_0 > t) &= \mathbb{E}_{i_0}\left(\mathbf{1}_{Z_{s_1}=i_1,\cdots,Z_{s_k}=i_k} \, \mathbb{E}_{i_0}\left(\mathbf{1}_{T_0>t} | \mathcal{F}_{s_k}\right)\right) \\
&= \mathbb{E}_{i_0}(\mathbf{1}_{Z_{s_1}=i_1,\cdots,Z_{s_k}=i_k} \, \mathbb{E}_{i_k}(\mathbf{1}_{T_0>t-s_k})) \\
&\quad (\text{ by Markov property}) \\
&= \mathbb{P}_{i_0}\left(Z_{s_1} = i_1, \cdots, Z_{s_k} = i_k\right) \mathbb{P}_{i_k}(T_0 > t - s_k).
\end{aligned}$$

By Theorem 7,

$$\lim_{t\to\infty} \frac{\mathbb{P}_{i_k}(T_0 > t - s_k)}{\mathbb{P}_{i_0}(T_0 > t)} = \frac{\pi_{i_k}}{\pi_{i_0}} e^{\theta s_k}.$$

Thus

$$\lim_{t\to\infty} \mathbb{P}_{i_0}(Z_{s_1} = i_1, \cdots, Z_{s_k} = i_k | T_0 > t) = \mathbb{P}_{i_0}(Z_{s_1} = i_1, \cdots, Z_{s_k} = i_k) \frac{\pi_{i_k}}{\pi_{i_0}} e^{\theta s_k}. \tag{15}$$

Let us now show that Y is a Markov process. We have

$$\begin{aligned}\mathbb{P}_{i_0}(Y_{s_1}=i_1,\cdots,Y_{s_k}=i_k,Y_t=j) &= e^{\theta t}\ \frac{\pi_j}{\pi_{i_0}}\ \mathbb{P}_{i_0}(Z_{s_1}=i_1,\cdots,Z_{s_k}=i_k,Z_t=j)\\ &= e^{\theta(t-s_k)}\ e^{\theta s_k}\ \frac{\pi_j}{\pi_{i_k}}\frac{\pi_{i_k}}{\pi_{i_0}}\ \mathbb{P}_{i_0}(Z_{s_1}=i_1,\cdots,Z_{s_k}=i_k)\\ &\quad\times\mathbb{P}_{i_k}(Z_{t-s_k}=j)\quad\text{(by Markov property of } Z)\\ &= \mathbb{P}_{i_0}(Y_{s_1}=i_1,\cdots,Y_{s_k}=i_k)\ \mathbb{P}_{i_k}(Y_{t-s_k}=j),\end{aligned}$$

and thus $\mathbb{P}(Y_t=j|Y_{s_1}=i_1,\cdots,Y_{s_k}=i_k)=\mathbb{P}_{i_k}(Y_{t-s_k}=j)$.

By (15) and Theorem 7, we have

$$\mathbb{P}_i(Y_t=j)=\frac{\pi_j}{\pi_i}\ \mathbb{P}_i(Z_t=j)\ e^{\theta t}\xrightarrow[t\to+\infty]{}\frac{\pi_j}{\pi_i}\ \alpha_j\ \pi_i=\alpha_j\pi_j.$$

Moreover let us compute the infinitesimal generator $\hat{L}$ of Y from the infinitesimal generator L of Z. We have for $j\neq i$,

$$\hat{L}_{ij}=\lim_{s\to 0}\hat{P}_{ij}(s)=\frac{\pi_j}{\pi_i}\ L_{ij}.$$

For $j=i$,

$$\begin{aligned}\hat{L}_{ii} &= -\lim_{s\to 0}\frac{1-\hat{P}_{ii}(s)}{s}=-\lim_{s\to 0}\frac{1-e^{\theta s}P_{ii}(s)}{s}\\ &= -\lim_{s\to 0}\frac{1-e^{\theta s}+e^{\theta s}(1-p_{ii}(s))}{s}=\theta+L_{ii}.\end{aligned}$$

We thus check that

$$\sum_{j\in E^*}\hat{L}_{ij}=\sum_{j\in E^*}\frac{\pi_j}{\pi_i}L_{ij}+\theta.$$

Since $L\pi=-\theta\pi$, then $\sum_{j\in E^*}\pi_jL_{ij}=-\theta\pi_i$ and thus $\sum_{j\in E^*}\hat{L}_{ij}=0$. □

4 QSD for birth and death processes

We are describing here the dynamics of isolated asexual populations, as for example populations of bacteria with cell binary division, in continuous time. Individuals may reproduce or die, and there is only one child per birth. The population size dynamics will be modeled by a birth and death process in continuous time. The individuals may interact, competing (for example) for resources and therefore the individual death's rate will depend on the total size of the population. In a first part, we recall and partially prove some results on the non-explosion of continuous time birth and death processes. We will also recall conditions on the birth and death rates which ensure that the process goes to extinction in finite time almost surely. In a second part, we concentrate on the cases where the process goes almost surely to zero and we study the existence and uniqueness of quasi-stationary distributions.

4.1 Birth and death processes

We consider birth and death processes with rates $(\lambda_i)_i$ and $(\mu_i)_i$, that is $\mathbb{N}$-valued pure jump Markov processes, whose jumps are $+1$ or -1, with transitions

$$\begin{aligned} i &\to i+1 \quad \text{with rate} \quad \lambda_i\ , \\ i &\to i-1 \quad \text{with rate} \quad \mu_i, \end{aligned}$$

where λ_i and μ_i, $i \in \mathbb{N}$, are non-negative real numbers.

Knowing that the process is at state i at a certain time, the process will wait for an exponential time of parameter λ_i before jumping to $i+1$ or independently, will wait for an exponential time of parameter μ_i before jumping to $i-1$. The total jump rate from state i is thus $\lambda_i + \mu_i$. We will assume in what follows that $\lambda_0 = \mu_0 = 0$. This condition ensures that 0 is an absorbing point, modeling the extinction of the population. Since these processes have a main importance in the modeling of biological processes, we study in detail their existence and extinction properties, and then their QSDs.

The most standard examples are the following ones.

1. The Yule process. For each $i \in \mathbb{N}$, $\lambda_i = \lambda i$ for a positive real number λ, and $\mu_i = 0$. There are no deaths. It's a fission model.

2. The linear birth and death process, or binary branching process. There exist positive numbers λ and μ such that $\lambda_i = \lambda i$ and $\mu_i = \mu i$. This model holds if individuals reproduce and die independently, with birth rate equal to λ and death rate equal to μ.

3. The logistic birth and death process. We assume that every individual in the population has a constant birth rate $\lambda > 0$ and a natural death rate $\mu > 0$. Moreover the individuals compete to share fixed resources, and each individual $j \neq i$ creates a competition pressure on individual i with rate $c > 0$. Thus, given that the population's size is i, the individual death rate due to competition is given by $c(i-1)$ and the total death rate is $\mu_i = \mu i + ci(i-1)$.

In the following, we will assume that $\lambda_i > 0$ and $\mu_i > 0$ for any $i \in \mathbb{N}^*$.

We denote by $(\tau_n)_n$ the sequence of the jump times of the process, either births or deaths. Let us first see under which conditions on the birth and death rates the process is well defined for all time $t \geq 0$, i.e. $\tau = \lim_n \tau_n = +\infty$ almost surely. Indeed, if $\tau = \lim_n \tau_n < \infty$ with a positive probability, the process would only be defined for $t < \tau$ on this event. There would be an accumulation of jumps near τ and the process could increase until infinity in finite time.

Let us give a necessary and sufficient condition ensuring that a birth and death process does not explode in finite time. The result is already stated in Anderson [2], but the following proof is actually far much shorter and easier to follow.

Theorem 10. *The birth and death process does not explode in finite time, almost surely, if and only if* $\sum_n r_n = +\infty$, *where*

$$r_n = \frac{1}{\lambda_n} + \sum_{k=1}^{n-1} \frac{\mu_{k+1} \cdots \mu_n}{\lambda_k \lambda_{k+1} \cdots \lambda_n} + \frac{\mu_1 \cdots \mu_n}{\lambda_1 \cdots \lambda_n}.$$

Proof. 1) Let us more generally consider a pure jump Markov process $(X_t, t \geq 0)$ with values in $\mathbb{N}$, and generator $(L_{ij}, i, j \in \mathbb{N})$. We set $q_i = -L_{ii}$. Let $(\tau_n)_n$ be the sequence of jump times of the process and $(U_n)_n$ the sequence of inter-times defined by

$$U_n = \tau_n - \tau_{n-1}, \quad \forall n \geq 1; \quad \tau_0 = 0, \quad U_0 = 0.$$

We also set $\tau_\infty = \lim_{n\to\infty} \tau_n \in [0, +\infty]$. The process does not explode in finite time almost surely (and is well defined for all time $t \in \mathbb{R}_+$), if and only if for each $i \in \mathbb{N}$

$$\mathbb{P}_i(\tau_\infty < \infty) = 0.$$

Let us show that this property is equivalent to the fact that the unique non-negative and bounded solution $x = (x_i)_{i\in\mathbb{N}}$ of $L\,x = x$ is the null solution.

For any i, we set $h_i^{(0)} = 1$ and, for $n \in \mathbb{N}^*$, $h_i^{(n)} = \mathbb{E}_i(\exp(-\sum_{k=1}^n U_k))$. For any $n \in \mathbb{N}$, we have

$$h_i^{(n+1)} = \sum_{j\neq i} \frac{L_{ij}}{q_i}\ h_j^{(n)}\ \mathbb{E}_i(\exp(-U_1)).$$

Indeed, the property is true for $n = 0$ since $\sum_{i\neq j} \frac{L_{ij}}{q_i} = 1$. Moreover, by conditioning with respect to U_1 and using the strong Markov property, we get

$$\mathbb{E}_i\left(\exp(-\sum_{k=1}^{n+1} U_k)\middle| U_1\right) \;=\; \exp(-U_1)\ \mathbb{E}_{X_{U_1}}\left(\exp(-\sum_{k=1}^{n} U_k)\right), \tag{16}$$

since the jump times of the U_1-translated process are the $\tau_n - U_1, n \in \mathbb{N}^*$. We have

$$\begin{aligned}\mathbb{E}_i\left(\mathbb{E}_{X_{U_1}}\left(\exp(-\sum_{k=1}^{n} U_k)\right)\right) &= \sum_{j\neq i} \mathbb{P}_i(X_{U_1} = j)\ \mathbb{E}_j\left(\exp(-\sum_{k=1}^{n} U_k)\right)\\ &= \sum_{j\neq i} \frac{L_{ij}}{q_i}\ \mathbb{E}_j(\exp(-\sum_{k=1}^{n} U_k)),\end{aligned}$$

since $\mathbb{P}_i(X_{U_1} = j) = \frac{L_{ij}}{q_i}$. By independence of U_1 and X_{U_1}, we deduce from (16) that

$$\mathbb{E}_i\left(\exp(-\sum_{k=1}^{n+1} U_k)\right) = \sum_{j\neq i} \frac{L_{ij}}{q_i}\ \mathbb{E}_j\left(\exp(-\sum_{k=1}^{n} U_k)\right)\ \mathbb{E}_i\left(\exp(-U_1)\right).$$

As

$$\mathbb{E}_i(\exp(-U_1)) = \int_0^\infty q_i e^{-q_i s} e^{-s} ds = \frac{q_i}{1+q_i},$$

it turns out that

$$h_i^{(n+1)} = \sum_{j\neq i} \frac{L_{ij}}{1+q_i}\ h_j^{(n)}. \tag{17}$$

Let $(x_i)_i$ be a nonnegative solution of $Lx = x$ bounded by 1, then $x_i = \sum_j L_{ij}\ x_j = L_{ii}x_i + \sum_{j\neq i} L_{ij}x_j = -q_i x_i + \sum_{j\neq i} L_{ij}x_j$, so that

$$x_i = \sum_{j\neq i} \frac{L_{ij}}{1+q_i} x_j. \tag{18}$$

Since $h_i^{(0)} = 1 \geq x_i \geq 0$ and $\frac{L_{ij}}{1+q_i} \geq 0$ for all $i,j \in E$, we deduce by iteration from (17) and (18) that $h_i^{(n)} \geq x_i \geq 0$, for any $n \in \mathbb{N}$.

Let us in the other hand define for any j the quantity $z_j = \mathbb{E}_j(e^{-\tau_\infty})$. Using $\tau_\infty = \lim_n \tau_n$, and $\tau_n = \sum_{k=1}^n U_k$, we deduce by monotone convergence that $z_j = \lim_n h_j^{(n)}$.

If the process does not explode a.s., then $\tau_\infty = \infty$ a.s., and $\lim_n h_i^{(n)} = z_i = 0$. Since $h_i^{(n)} \geq x_i \geq 0$, we deduce that $x_i = 0$. It turns out that the unique nonnegative bounded solution of $Lx = x$ is zero.

If the process explodes with positive probability, then there exists i such that $\mathbb{P}_i(\tau_\infty < \infty) > 0$. Making n tend to infinity in (17), we get $z_i = \sum_{j \neq i} \frac{L_{ij}}{1+q_i}\ z_j$. Since $z_i > 0$, z is a positive and bounded solution of $Lz = z$.

2) Let us now apply this result to the birth and death process with $\lambda_0 = \mu_0 = 0$. Then for $i \geq 1$, $L_{i,i+1} = \lambda_i$, $L_{i,i-1} = \mu_i$, $L_{i,i} = -(\lambda_i + \mu_i)$. The equation $Lx = x$ is given by $x_0 = 0$ and for all $n \geq 1$ by

$$\lambda_n x_{n+1} - (\lambda_n + \mu_n)x_n + \mu_n x_{n-1} = x_n.$$

Thus, if we set $\Delta_n = x_n - x_{n-1}$, we have $\Delta_1 = x_1$ and for $n \geq 1$, $\Delta_{n+1} = \Delta_n\ \frac{\mu_n}{\lambda_n} + \frac{1}{\lambda_n} x_n$. Let us remark that, for any n, $\Delta_n \geq 0$ and the sequence $(x_n)_n$ is nondecreasing. If $x_1 = 0$, the solution is zero. If not, we get by induction

$$\Delta_{n+1} = \frac{1}{\lambda_n} x_n + \sum_{k=1}^{n-1} \frac{1}{\lambda_k}\ \frac{\mu_{k+1}}{\lambda_{k+1}} \cdots \frac{\mu_n}{\lambda_n}\ x_k + \frac{\mu_1}{\lambda_1} \cdots \frac{\mu_n}{\lambda_n}\ x_1.$$

Letting

$$r_n = \frac{1}{\lambda_n} + \sum_{k=1}^{n-1} \frac{\mu_{k+1} \cdots \mu_n}{\lambda_k \lambda_{k+1} \cdots \lambda_n} + \frac{\mu_1 \cdots \mu_n}{\lambda_1 \cdots \lambda_n},$$

we deduce that $r_n\ x_1 \leq \Delta_{n+1} \leq r_n\ x_n$. Then

$$x_1(1 + r_1 + \cdots r_n) \leq x_{n+1} \leq x_1\ \prod_{k=1}^{n}(1 + r_k).$$

The boundedness of the sequence $(x_n)_n$ is thus equivalent to the convergence of the series $\sum_k r_k$. □

Corollary 11. *Let us consider a BD-process with birth rates $(\lambda_i)_i$. If there exists a constant $\lambda > 0$ such that*

$$\lambda_i \leq \lambda i,\ \forall i \geq 1,$$

then the process is well defined on $\mathbb{R}_+$.

The proof is immediate. It turns out that the linear BD-processes and the logistic processes are well defined on $\mathbb{R}_+$.

Let us now recall under which assumption a BD-process goes to extinction almost surely.

Proposition 12. *The BD-process goes almost-surely to extinction if and only if*

$$\sum_{k=1}^{\infty} \frac{\mu_1 \cdots \mu_k}{\lambda_1 \cdots \lambda_k} = +\infty. \tag{19}$$

Proof. Let us introduce

$$u_i := \mathbb{P}(Extinction | Z_0 = i) = \mathbb{P}_i(T_0 < \infty),$$

which is the probability to attain 0 in finite time, starting from i. As before T_0 denotes the extinction time and T_I the hitting time of any I. The Markov property yields the induction formula

$$\lambda_i \; u_{i+1} - (\lambda_i + \mu_i) \; u_i + \mu_i \; u_{i-1} = 0, \, \forall i \geq 1.$$

To resolve this equation, we firstly assume that the rates λ_i, μ_i are nonzero until some fixed level I such that $\lambda_I = \mu_I = 0$. We set for each i, $u_i^{(I)} := \mathbb{P}_i(T_0 < T_I)$. Thus $u_i = \lim_{I \to \infty} u_i^{(I)}$. Defining $U_I := \sum_{k=1}^{I-1} \frac{\mu_1 \cdots \mu_k}{\lambda_1 \cdots \lambda_k}$, an easy computation shows that for $i \in \{1, \cdots, I-1\}$,

$$u_i^{(I)} = (1 + U_I)^{-1} \sum_{k=i}^{I-1} \frac{\mu_1 \cdots \mu_k}{\lambda_1 \cdots \lambda_k}.$$

In particular, $u_1^{(I)} = \frac{U_I}{1+U_I}$. Hence, either $(U_I)_I$ tends to infinity when $I \to \infty$ and any extinction probability u_i is equal to 1 or $(U_I)_I$ converges to a finite limit U_∞ and for $i \geq 1$,

$$u_i = (1 + U_\infty)^{-1} \sum_{k=i}^{\infty} \frac{\mu_1 \cdots \mu_k}{\lambda_1 \cdots \lambda_k} < 1.$$

□

Corollary 13. *1. The linear BD-process with rates λi and μi goes almost surely to extinction if and only if $\lambda \leq \mu$.*

2. The logistic BD-process goes almost surely to extinction.

Proof. 1) If $\lambda \leq \mu$, *i.e.* when the process is sub-critical or critical, we obtain $U_I \geq I - 1$ for any $I \geq 1$. Then $(U_I)_I$ goes to infinity when $I \to \infty$ and the process goes to extinction with probability 1. Conversely, if $\lambda > \mu$, the sequence $(U_I)_I$ converges to $\frac{\mu}{\lambda - \mu}$, and an easy computation gives $u_i = (\lambda/\mu)^i$.

2) Here we have

$$\lambda_i = \lambda i \; ; \; \mu_i = \mu i + ci(i-1). \tag{20}$$

It is easy to check that (19) is satisfied. □

4.2 Quasi-stationary distributions for birth and death processes

We consider a BD-process (Z_t) with almost sure extinction. A probability measure α on $\mathbb{N}^*$ is given by a sequence $(\alpha_j)_{j\geq 1}$ of non-negative numbers such that $\sum_{j\geq 1}\alpha_j = 1$.
Our first result is a necessary and sufficient condition for such a sequence $(\alpha_j)_{j\geq 1}$ to be a QSD for Z. Thereafter we will study the set of sequences which fulfill this condition (we refer the reader to van Doorn [59] for more details).

Theorem 14. *The sequence $(\alpha_j)_{j\geq 1}$ is a QSD if and only if*

1. $\alpha_j \geq 0,\ \forall j \geq 1$ *and* $\sum_{j\geq 1}\alpha_j = 1$.
2. $\forall\ j \geq 1$,

$$\begin{aligned} \lambda_{j-1}\alpha_{j-1} - (\lambda_j+\mu_j)\alpha_j + \mu_{j+1}\alpha_{j+1} &= -\mu_1\alpha_1\alpha_j\,; \\ -(\lambda_1+\mu_1)\alpha_1 + \mu_2\alpha_2 &= -\mu_1\alpha_1^2. \end{aligned} \tag{21}$$

The next result follows immediately.

Corollary 15. *Let us define inductively the sequence of polynomials $(H_n(x))_n$ as follows: $H_1(x) = 1$ for all $x \in \mathbb{R}$ and for $n \geq 2$,*

$$\begin{aligned} \lambda_n\ H_{n+1}(x) &= (\lambda_n+\mu_n-x)\ H_n(x) - \mu_{n-1}\ H_{n-1}(x)\,; \\ \lambda_1\ H_2(x) &= \lambda_1+\mu_1-x. \end{aligned} \tag{22}$$

Then, any quasi-stationary distribution $(\alpha_j)_j$ satisfies for all $j \geq 1$,

$$\alpha_j = \alpha_1\ \pi_j\ H_j(\mu_1\alpha_1),$$

where

$$\pi_1 = 1\ ;\ \pi_n = \frac{\lambda_1\cdots\lambda_{n-1}}{\mu_2\cdots\mu_n}. \tag{23}$$

Proof of Theorem 14. By Proposition 4 and for a QSD α, there exists $\theta > 0$ such that

$$\alpha L = -\theta\,\alpha,$$

where L is the infinitesimal generator of Z restricted to $\mathbb{N}^*$. Taking the j^{th} component of this equation, we get

$$\begin{aligned} &\lambda_{j-1}\alpha_{j-1} - (\lambda_j+\mu_j)\alpha_j + \mu_{j+1}\alpha_{j+1} = -\theta\,\alpha_j,\ \forall j \geq 2 \\ &-(\lambda_1+\mu_1)\alpha_1 + \mu_2\alpha_2 = -\theta\,\alpha_1. \end{aligned}$$

Summing over $j \geq 1$, we get after re-indexing

$$0 = \sum_{j\geq 1}\lambda_j\alpha_j - (\lambda_j+\mu_j)\alpha_j + \mu_j\alpha_j = -\theta\sum_{j\geq 1}\alpha_j + \mu_1\alpha_1.$$

We deduce that $\theta = \mu_1\alpha_1$, which concludes the proof of Theorem 14.

□

The study of the polynomials (H_n) has been detailed in Van Doorn [59]. In particular it is shown that there exists a non-negative number ξ_1 such that

$$x \leq \xi_1 \iff H_n(x) > 0, \ \forall n \geq 1.$$

By Corollary 15, $\alpha_j = \alpha_1 \ \pi_j \ H_j(\mu_1\alpha_1)$. Since for any j, $\alpha_j > 0$, we have $H_j(\mu_1\alpha_1) > 0$ for all $j \geq 1$ and then

$$0 < \mu_1\alpha_1 \leq \xi_1.$$

We can immediately deduce from this property that if $\xi_1 = 0$, then there is no quasi-stationary distribution.

To go further, one has to study more carefully the spectral properties of the semi-group (P_t) and the polynomials $(H_n)_n$, as it has been done in [39], [30] and [59]. From these papers, the polynomials $(H_n)_n$ are shown to be orthogonal with respect to the spectral measure of (P_t). In addition, it yields a tractable necessary and sufficient condition for the existence of QSD based on the birth and death rates. The series (S) with general term

$$S_n = \frac{1}{\lambda_n \pi_n} \sum_{i=n+1}^{\infty} \pi_i$$

plays a crucial role. Remark that (S) converges if and only if

$$\sum_{n=1}^{\infty} \pi_n \left(\frac{1}{\mu_1} + \sum_{i=1}^{n-1} \frac{1}{\lambda_i \pi_i} \right) < +\infty.$$

Theorem 16. *([59], Theorems 3.2 and 4.1). We have the convergence*

$$\lim_{t\to\infty} \mathbb{P}_i(Z_t = j | T_0 > t) = \frac{1}{\mu_1} \ \pi_j \ \xi_1 \ H_j(\xi_1).$$

In particular, we obtain

$$\xi_1 = \lim_{t\to\infty} \mu_1 \mathbb{P}_1(Z_t = 1 | T_0 > t) \tag{24}$$

1. *If $\xi_1 = 0$, there is no QSD.*
2. *If (S) converges, then $\xi_1 > 0$ and the Yaglom limit is the unique QSD.*
3. *If (S) diverges and $\xi_1 \neq 0$, then there is a continuum of QSD, given by the one parameter family $(\hat{\alpha}_j(x))_{0<x\leq\xi_1}$:*

$$\hat{\alpha}_j(x) = \frac{1}{\mu_1} \ \pi_j \ x \ H_j(x).$$

Remark 4. 1. Formula (24) and the approximation method described in Section 6 allow us to deduce a simulation algorithm to get ξ_1.

2. Cases with more general birth and death processes have also been studied recently. Let us mention the infinite dimensional state space setting of Collet, Martínez, Méléard and San Martín [15], where each individual has a type in a continuous state space which influences its birth and death rates, and mutation on the type can occur. The authors give sufficient and quite general conditions for the existence and uniqueness of a QSD. We also refer the reader to the recent work of van Doorn [60], where a transition to the state 0 may occur from any state. The author provides sufficient conditions for the existence of QSDs.

Let us now develop some examples.

The linear case. We assume $\lambda_i = \lambda\, i\ ;\ \mu_i = \mu\, i$ and $\lambda \leq \mu$. In that case, the BD-process is a branching process, where each individual reproduces with rate λ and dies with rate μ. A straightforward computation shows that the series (S) diverges.

Setting $f_s : k \mapsto s^k$, we get by the Kolmogorov forward equation,

$$\frac{\partial P_t f_s(1)}{\partial t} = \mu P_t f_s(0) - (\lambda+\mu) P_t f_s(1) + \lambda P_t f_s(2).$$

But the branching property of the process implies $P_t f_s(2) = (P_t f_s(1))^2$, while $f_s(0) = 1$ so that

$$\frac{\partial P_t f_s(1)}{\partial t} = \mu - (\lambda+\mu) P_t f_s(1) + \lambda\,(P_t f_s(1))^2\,.$$

Setting $m = 2\frac{\lambda}{\lambda+\mu}$, we deduce that for $s < 1$,

$$P_t f_s(1) = 1 - \frac{2(1-s)(2m-1)}{(ms+m-2)e^{-(\lambda+\mu)(m-1)t} + (1-s)m}.$$

In particular, we deduce that the generating function $F_t : s \mapsto \mathbb{E}(s^{Z_t}|Z_t > 0)$ of Z_t conditioned to $Z_t > 0$ converges when t goes to infinity:

$$F_t(s) = \frac{P_t f_s(1) - P_t f_0(1)}{1 - P_t f_0(1)} \xrightarrow[t\to\infty]{} \frac{(\lambda-\mu)s}{\lambda s - \mu}.$$

We deduce that the Yaglom limit of Z does not exist if $\lambda = \mu$ and is given by the geometric distribution with parameter $\frac{\lambda}{\mu}$ if $\lambda < \mu$:

$$\alpha_k = \left(\frac{\lambda}{\mu}\right)^{k-1}\left(1 - \frac{\lambda}{\mu}\right).$$

An easy computation yields $\xi_1 = \mu - \lambda$, since by (24), $\alpha_1 = \frac{\xi_1}{\mu}$. But the series (S) diverges so that for $\lambda < \mu$, $\xi_1 > 0$ and there is an infinite number of QSD. If $\lambda = \mu$, $\xi_1 = 0$ and there is no QSD.

The logistic case. We assume $\lambda_i = \lambda i\ ;\ \mu_i = \mu i + ci(i-1)$. Because of the quadratic term, the branching property is lost and we can not compute the Yaglom limit as above. Therefore, we have no other choice than to study the convergence of the series (S).

We have

$$\begin{aligned}\sum_{i=n+1}^{\infty}\pi_i &\leq \sum_{i=n+1}^{\infty}\left(\frac{\lambda}{c}\right)^{i-1}\frac{1}{i!}=\sum_{p=0}^{\infty}\left(\frac{\lambda}{c}\right)^{n+p}\frac{1}{(n+p+1)!}\\ &\leq \left(\frac{\lambda}{c}\right)^{n}\frac{1}{(n+1)!}\sum_{p=0}^{\infty}\left(\frac{\lambda}{c}\right)^{p}\frac{1}{p!}=\left(\frac{\lambda}{c}\right)^{n}\frac{1}{(n+1)!}\,e^{\frac{\lambda}{c}},\end{aligned}$$

since $\frac{(n+1)!}{(n+p+1)!}\leq\frac{1}{p!}$. Thus as $\frac{1}{\pi_n}\leq C\left(\frac{c}{\lambda}\right)^{n-1}\ n!$, we get

$$\frac{1}{\lambda_n\pi_n}\sum_{i=n+1}^{\infty}\pi_i\leq\frac{C}{c}\frac{1}{n(n+1)}\,e^{\frac{\lambda}{c}}.$$

Hence the series converges. Thus the Yaglom limit exists and is the unique quasi-stationary distribution.

One can obtain substantial qualitative information by looking closer to the jump rates of the process. For instance, $(\lambda-\mu)/c$ is a key value for the process. Indeed, given a population size i, the expectation of the next step is equal to $\frac{i(\lambda-\mu-c(i-1))}{\lambda i+\mu i+ci(i-1)}$. Then the sign of this expectation depends on the position of $i-1$ with respect to $(\lambda-\mu)/c$:

- If $i\leq(\lambda-\mu)/c+1$, then the expectation of the next step will be positive.
- If $i=(\lambda-\mu)/c+1$, then it will be 0.
- If $i>(\lambda-\mu)/c+1$, then it will be negative.

We deduce that the region around $(\lambda-\mu)/c$ is stable: it plays the role of a typical size for the population and we expect that the mass of the Yaglom limit is concentrated around it. The value $(\lambda-\mu)/c$ is called (by the biologists) the charge capacity of the logistic BD-process with parameters λ, μ and c. In the next section, we will consider large population processes, which means logistic BD-processes with large charge capacity.

Example 3. We develop now a numerical illustration of the logistic BD-process case. Across the whole example, the value of the charge capacity $\frac{\lambda-\mu}{c}$ is fixed, arbitrarily chosen equal to 9.

In order to illustrate the concept of charge capacity, we represent in Figure 4 a random path of a logistic birth and death process with initial size $Z_0=1$ and with parameters $\lambda=10$, $\mu=1$ and $c=1$. We observe that the process remains for long times in a region around the charge capacity. Moreover, we remark that the process remains mainly below the charge capacity; this is because the jumps rate are higher in the upper region, so that it is less stable than the region below the charge capacity.

Let us now compare the Yaglom limits (numerically computed using the approximation method presented in Section 6) of two different logistic BD processes whose charge capacities are equal to 9 (see Figure 5):

(a) $Z^{(a)}$, whose parameters are $\lambda=10$, $\mu=1$ and $c=1$,

(b) $Z^{(b)}$, whose parameters are $\lambda=10$, $\mu=7$ and $c=1/3$.

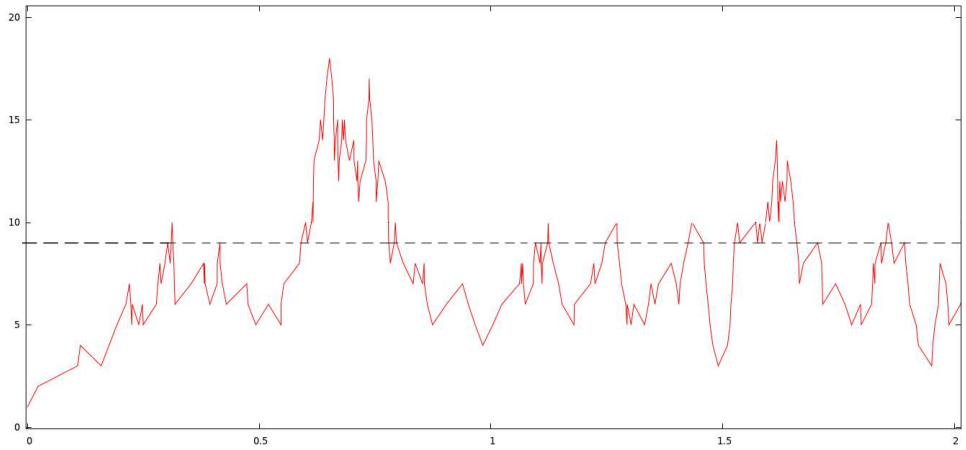


Figure 4: Example 3. A random path of a logistic birth and death process with initial size $Z_0 = 1$ and with parameters $\lambda = 10$, $\mu = 1$ and $c = 1$

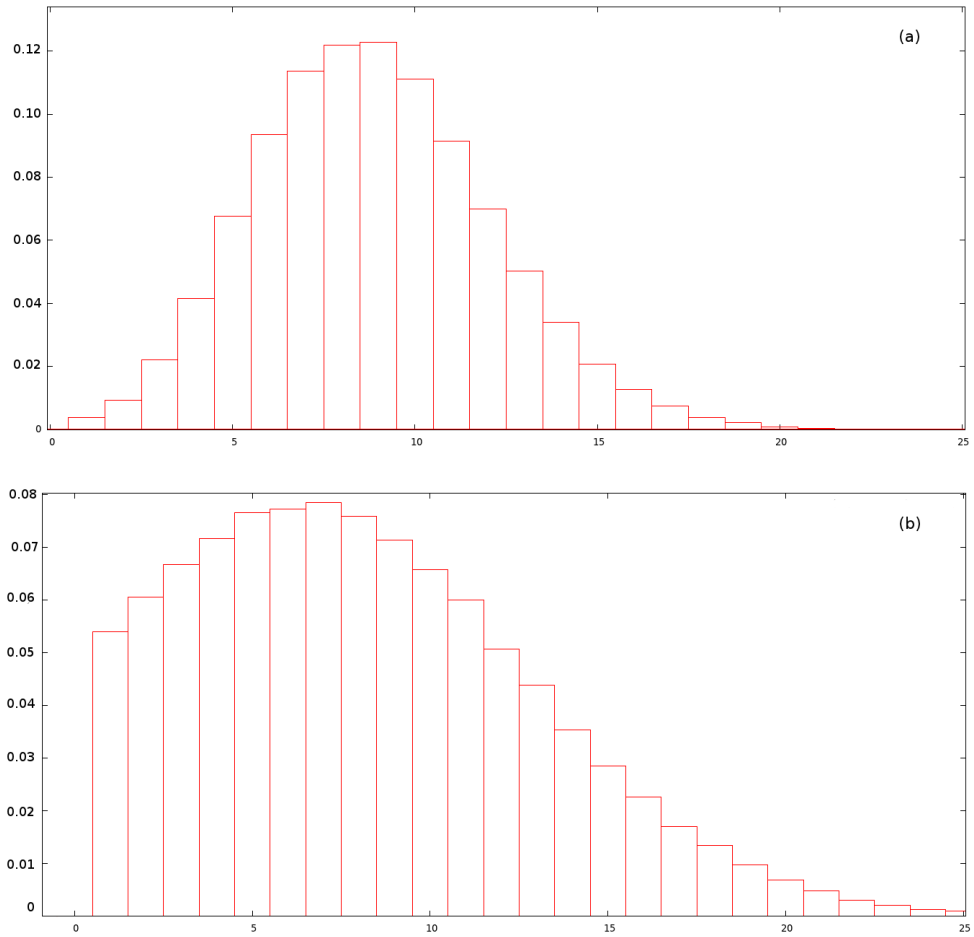


Figure 5: Example 3. The Yaglom limits of two logistic birth and death processes with the same charge capacity $\frac{\lambda-\mu}{c} = 9$: (a) $\lambda = 10$, $\mu = 1$ and $c = 1$; (b) $\lambda = 10$, $\mu = 7$ and $c = 1/3$.

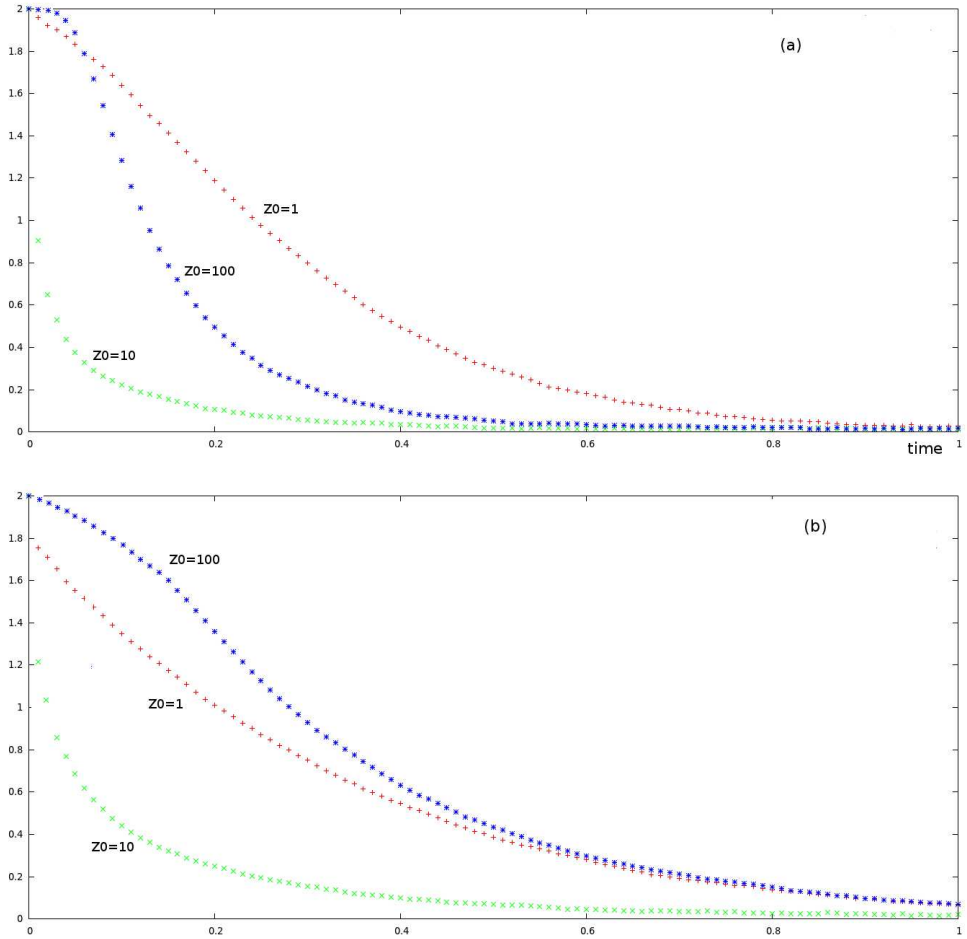


Figure 6: Example 3. Evolution of the distance between the conditioned distribution and the Yaglom limits of two logistic birth and death processes with the same charge capacity $\frac{\lambda-\mu}{c} = 9$: (a) $\lambda = 10$, $\mu = 1$ and $c = 1$; (b) $\lambda = 10$, $\mu = 7$ and $c = 1/3$.

We observe that the Yaglom limits of $Z^{(a)}$ and $Z^{(b)}$ are supported by a region which is around the charge capacity. We also remark that the Yaglom limit of the process $Z^{(b)}$ has a more flat shape than the Yaglom limit of $Z^{(a)}$. This is because the competition parameter of $Z^{(b)}$ is small in comparison with the birth and death parameters, so that the drift toward the charge capacity is small too, both above and below the charge capacity.

We compute now the distance between the conditioned distribution and the Yaglom limit for the two processes $Z^{(a)}$ and $Z^{(b)}$ for different values of the initial state, namely $Z_0 = 1$, $Z_0 = 10$ and $Z_0 = 100$. The numerical results are represented in Figure 6. We observe a strong dependence between the speed of convergence and the initial position of the processes. In the case of $Z^{(a)}$, it only takes a very short time to the process starting from 100 to reach the charge capacity, because the competition parameter is relatively high and so is the drift downward the charge capacity. On the contrary, in the case of $Z^{(b)}$, it takes a longer time for the process to come back from 100 to the charge capacity, so that the speed of convergence to the Yaglom limit is slow. In both cases, the convergence to the Yaglom limit happens very fast when starting from the value 10, because it is near the charge capacity.

5 The logistic Feller diffusion process

5.1 A large population model

We are now rescaling logistic birth and death processes with a parameter $K \in \mathbb{N}^*$ modeling a large population with small individuals, *i.e.* a population with a large initial size of order K and a large charge capacity assumption. The individual's weights (or biomasses) are assumed to be equal to $\frac{1}{K}$ and we study the limiting behavior of the total biomass process $(\frac{Z_t^K}{K}, t \geq 0)$ when K tends to infinity, Z_t^K being the population's size at time t. In what follows, λ, μ and c are fixed positive constants.
In Subsection 5.1.1, individual birth and death rates are assumed to be constant and the competition rate depends linearly on the individual biomass $\frac{1}{K}$. In Subsection 5.1.2, we investigate the qualitative differences of evolutionary dynamics across populations with allometric demographics: life-lengths and reproduction times are assumed to be proportional to the individual's weights.

In both cases, the charge capacity of (Z^K) will be $(\lambda - \mu)K/c$.

5.1.1 Convergence to the logistic equation

Given a parameter K scaling the population's size, we consider the logistic BD-process Z^K with birth, death and competition parameters λ, μ and c/K respectively. We assume that the initial value of Z^K is of order K, in the sense that there exists a non-negative real random variable X_0 such that

$$\frac{Z_0^K}{K} \xrightarrow[K\to\infty]{} X_0 \quad \text{with} \;\; \mathbb{E}(X_0^3) < +\infty.$$

We consider the total biomass process defined by $X^K = Z^K/K$ for all $K \geq 1$ and are interested in the limit of X^K when $K \to \infty$. The transitions of the process $(X_t^K, t \geq 0)$ are the following ones:

$$\begin{array}{rcll} \frac{i}{K} & \to & \frac{i+1}{K} & \text{with rate } \; \lambda i = \lambda K \frac{i}{K}; \\ \frac{i}{K} & \to & \frac{i-1}{K} & \text{with rate } \; \mu i + \frac{c}{K}\, i(i-1) = K \frac{i}{K}\left(\mu + c(\frac{i}{K} - \frac{1}{K})\right). \end{array}$$

Theorem 17. *Assume that X_0 is a positive number x_0. Then, the process $(X_t^K, t \geq 0)$ converges in law in $\mathbb{D}([0,T], \mathbb{R}_+)$ to the unique continuous (in time) deterministic function solution of*

$$x(t) = x_0 + \int_0^t (\lambda - \mu - cx(s))x(s)ds.$$

Remark 5. The function x is thus solution of the ordinary differential equation

$$\dot{x} = (\lambda - \mu)x - cx^2 \;;\; x(0) = x_0, \tag{25}$$

called logistic equation. This equation has been historically introduced as the first macroscopic model describing populations regulated by competition (cf. [64]). In Theorem 17 above, it is obtained as the limit of properly scaled stochastic jump models.

The function x solution of (25) hits 0 in finite time if $\lambda < \mu$, while it remains positive forever if $\lambda > \mu$, converging in the long term to its charge capacity $\frac{\lambda-\mu}{c}$. Thus at this scale extinction does not happen.

Proof of Theorem 17. The Markov process $(X_t^K, t \geq 0)$ is well defined and its infinitesimal generator is given, for any measurable and bounded function ϕ, by

$$L_K\phi(x) = \lambda K x \left(\phi(x + \frac{1}{K}) - \phi(x)\right) + K(\mu x + cx(x - \frac{1}{K})) \left(\phi(x - \frac{1}{K}) - \phi(x)\right). \quad (26)$$

Hence, by Dynkin's theorem ([21] Prop. IV-1.7), the process

$$\phi(X_t^K) - \phi(X_0^K) - \int_0^t L_K\phi(X_s^k)ds \quad (27)$$

defines a local martingale, and a martingale, as soon as each term in (27) is integrable. In particular, taking $\phi(x) = x$ gives that $(X_t^K, t \geq 0)$ is a semimartingale and there exists a local martingale M^K such that

$$X_t^K = X_0^K + M_t^K + \int_0^t X_s^K \left(\lambda - \mu - c\left(X_s^K - \frac{1}{K}\right)\right) ds. \quad (28)$$

Since x_0 is deterministic and using a localization argument, we deduce that $\mathbb{E}(\sup_{t\leq T}(X_t^K)^2) < \infty$. Moreover, taking $\phi(x) = x^2$ applied to (27), and comparing with Itô's formula applied to $(X^K)^2$ prove that (M^K) is a square-integrable martingale with quadratic variation process

$$\langle M^K\rangle_t = \frac{1}{K}\int_0^t \left(\lambda + \mu + c\left(X_s^K - \frac{1}{K}\right)\right) X_s^K ds. \quad (29)$$

Let us now study the convergence in law of the sequence (X^K), when K tends to infinity. For any K, the law of X^K is a probability measure on the trajectory space $\mathbb{D}_T = \mathbb{D}([0,T], \mathbb{R}_+)$, namely the Skorohod space of left-limited and right-continuous functions from [0,T] into $\mathbb{R}_+$, endowed with the Skorohod topology. This topology makes $\mathbb{D}_T$ a Polish state, that is a metrizable complete and separable space, which is not true if $\mathbb{D}_T$ is endowed with the uniform topology. See Billingsley [9] for details.

The proof of Theorem 17 is obtained by a compactness-uniqueness argument. The uniqueness of the solution of (25) is immediate.

By a natural coupling, one may bound the birth and death process X^K stochastically from above by the Yule process Y^K started from x_0, which jumps from x to $x + \frac{1}{K}$, at the same birth times than X^K. One easily shows that $\sup_K \mathbb{E}(\sup_{t\leq T}(Y_t^K)^3) < \infty$ and thus

$$\sup_K \mathbb{E}(\sup_{t\leq T}(X_t^K)^3) < \infty.$$

From this uniform estimate, we deduce the uniform tightness of the laws of X^K (as probability measures on $\mathbb{D}_T$), using the Aldous criterion (cf. Aldous [1], Joffe-Métivier [37]). Then, by Prokhorov's Theorem, the compactness of the laws of (X^K) follows. Since

$\sup_{t\leq T}|X^K_t - X^K_{t^-}| \leq \frac{1}{K}$ and the function $x \mapsto \sup_{t\leq T}|x_t - x_{t^-}|$ is continuous on $\mathbb{D}_T$, each limiting value (in law) of the sequence (X^K) will be a pathwise continuous process. In addition using (29) and (5.1.1), it can be shown that $\lim_{K\to\infty}\mathbb{E}(\langle M^K\rangle_t) = 0$. Then, the random fluctuations disappear when K tends to infinity and the limiting values are deterministic functions. Now it remains to show that these limiting values are solutions of (25), which can be done similarly to the proof of Theorem 18 (4) stated below. □

5.1.2 The logistic Feller diffusion process

In this section, we study the logistic BD-processes Z^K with birth and death rates given by $\gamma K + \lambda$ and $\gamma K + \mu$ respectively. Here γ, λ and μ are still positive constants. We assume that the competition parameter is given by c/K, so that the charge capacity of Z^K is still $(\lambda - \mu)K/c$.

Remark 6. *This BD-process $(Z^K_t)_t$ can also be interpreted as a time-rescaled BD-process Y^K_{Kt}, whose birth, death and competition parameters are given by $\gamma + \lambda/K$, $\gamma + \mu/K$ and c/K^2 respectively, that is a critical BD-process with small pertubations.*

We are considering as in Section 5.1.1 the sequence of processes X^K defined for all $t \geq 0$ by $X^K_t = \frac{Z^K_t}{K}$.

The transitions of the process (X^K) are given by

$$\begin{aligned} \frac{i}{K} &\to \frac{i+1}{K} \quad \text{with rate} \quad \gamma K i + \lambda i \\ \frac{i}{K} &\to \frac{i-1}{K} \quad \text{with rate} \quad \gamma K i + \mu i + \frac{c}{K} i(i-1). \end{aligned} \tag{30}$$

Formula (28) giving the semi-martingale decomposition of X^K will stay true in this case with another square integrable martingale part N^K such that

$$\langle N^K\rangle_t = \frac{1}{K}\int_0^t (2\gamma K + \lambda + \mu + c\left(X^K_s - \frac{1}{K}\right) X^K_s ds.$$

One immediately observes that the expectation of this quantity does not tend to zero as K tends to infinity. Hence the fluctuations will not disappear at infinity and the limit will stay random. Let us now state the convergence theorem.

Theorem 18. *Assume $\gamma, c, \lambda, \mu > 0$ and $\lambda > \mu$.*
i) Assume that the sequence $(X^K_0)_K$ converges in law to X_0 with $\mathbb{E}(X_0^3) < \infty$. Then the sequence of processes $(X^K)_K$ with transitions (30) converges in law in $\mathcal{P}(\mathbb{D}_T)$ to the continuous process X, defined as the unique solution of the stochastic differential equation

$$dX_t = \sqrt{2\gamma X_t} dB_t + \left((\lambda - \mu)X_t - cX_t^2\right) dt;\ X_0 \in]0, +\infty[, \tag{31}$$

where $(B_t)_{t\in[0,+\infty[}$ is a standard Brownian motion.

ii) Let us introduce for each $y \geq 0$ the stopping time

$$T_y = \inf\{t \in \mathbb{R}_+, X_t = y\}. \tag{32}$$

For any $x \geq 0$, we get

$$\mathbb{P}_x(T_0 < \infty) = 1.$$

When $c = 0$, Equation (31) defines the super-critical Feller diffusion process and will explode with positive probability. In the general case where $c \neq 0$, it defines the so-called logistic Feller diffusion process following the terminology introduced by Etheridge [20] and Lambert [42]. Let us remark that the quadratic term driven by c regulates the population size, which fluctuates until it attains the absorbing point 0. Theorem 18 shows that an accumulation of a large amount of birth and death events may create stochasticity, often called by biologists ecological drift or demographic stochasticity. Contrarily to the previous case (Theorem 17), the limiting process suffers extinction almost surely.

Proof. As for Theorem 17, the proof is based on a uniqueness-compactness argument.

(1) The uniqueness of the solution of (31) follows from a general existence and pathwise uniqueness result in Ikeda-Watanabe [36] Section IV-3 or Karatzas-Shreve [38]. For a stochastic differential equation

$$dX_t = \sigma(X_t)dB_t + b(X_t)dt,$$

with σ and b smooth enough, the existence and pathwise uniqueness are determined thanks to the following scale functions: for $x > 0$,

$$\begin{aligned} Q(x) &= -\int_1^x \frac{2b(y)}{\sigma^2(y)}dy \ ; \ \Lambda(x) = \int_1^x e^{Q(z)}dz \ ; \\ \kappa(x) &= \int_1^x e^{Q(y)} \left(\int_1^y e^{-Q(z)}dz \right) dy. \end{aligned} \tag{33}$$

More precisely, it is proved that (i) $\forall x > 0$, $\mathbb{P}_x(T_0 < T_\infty) = 1$ and (ii) $\Lambda(+\infty) = +\infty$; $\kappa(0^+) < +\infty$ are equivalent. In that case, pathwise uniqueness follows, and then uniqueness in law.
In our situation, the coefficients are given by

$$\sigma(x) = \sqrt{2\gamma x} \ ; \ b(x) = (\lambda - \mu)x - cx^2,$$

so that the functions Λ and κ satisfy (ii). Thus the SDE (31) has a pathwise unique solution which reaches 0 in finite time almost surely.

(2) Let us assume that $\mathbb{E}(X_0^3) < \infty$ and prove that $\sup_K \mathbb{E}(\sup_{t \leq T}(X_t^K)^2) < \infty$. The infinitesimal generator of X^K is given by

$$\begin{aligned} \tilde{L}_K\phi(x) &= (\gamma Kx + \lambda x)K \left(\phi(x + \frac{1}{K}) - \phi(x) \right) \\ &\quad + (\gamma Kx + \mu x + cx(x - \frac{1}{K}))K \left(\phi(x - \frac{1}{K}) - \phi(x) \right). \end{aligned} \tag{34}$$

With $\phi(x) = x^3$, we obtain that

$$\begin{aligned} (X_t^K)^3 &= X_0^3 + M_t^K + \int_0^t \gamma K^2 X_s^K \left[\left(X_s^K + \frac{1}{K} \right)^3 - \left(X_s^K - \frac{1}{K} \right)^3 - \left(X_s^K \right)^3 \right] ds \\ &+ \int_0^t \lambda K X_s^K \left[\left(X_s^K + \frac{1}{K} \right)^3 - \left(X_s^K \right)^3 \right] ds \\ &\quad + \int_0^t (\mu + c(X_s^K - 1))K X_s^K \left[\left(X_s^K - \frac{1}{K} \right)^3 - \left(X_s^K \right)^3 \right] ds, \end{aligned}$$

where M^K is a local martingale. Using a standard localization argument and

$$\left(X_s^K + \frac{1}{K}\right)^3 - \left(X_s^K - \frac{1}{K}\right)^3 - \left(X_s^K\right)^3 = 6\ \frac{X_s^K}{K^2},$$

we get

$$\mathbb{E}((X_t^K)^3) \leq \mathbb{E}(X_0^3) + C\int_0^t \mathbb{E}((X_s^K)^3)ds,$$

where C is independent of K. Gronwall's lemma yields

$$\sup_{t\leq T}\sup_K \mathbb{E}(|X_t^K|^3) < \infty. \tag{35}$$

Now, thanks to (35) and Doob's inequality, we deduce from the semi-martingale decomposition of $(X_t^K)^2$ (obtained using $\phi(x) = x^2$), that

$$\sup_K \mathbb{E}(\sup_{t\leq T}|X_t^K|^2) < \infty. \tag{36}$$

(3) As previously, the uniform tightness of the laws of (X^K) is obtained from (36) and the Aldous criterion [1]. Therefore, the sequence of laws is relatively compact and it remains to characterize its limit values.

(4) As in the proof of Theorem 17, we remark that the limiting values only charge the set of continuous trajectories, since $\sup_{t\leq T}|\Delta X_t^K| \leq \frac{1}{K}$. Let $\mathbb{Q} \in \mathcal{P}(C([0,T],\mathbb{R}_+))$ be a limiting value of the sequence of laws of the processes X^K. We will identify $\mathbb{Q}$ as the (unique) law of the solution of (31) and the convergence will be proved. Let us denote $C_T = C([0,T],\mathbb{R}_+)$ and define, for $\phi \in C_b^2$ and $t > 0$, the function

$$\begin{aligned}\psi_t : C_T &\to \mathbb{R}\\ X &\mapsto \phi(X_t) - \phi(X_0) - \int_0^t \left(\gamma X_s\phi''(X_s) + ((\lambda-\mu)X_s - cX_s^2)\phi'(X_s)\right)ds,\end{aligned}$$

which is continuous $\mathbb{Q}$-a.s.. Let us show first that the process $(\psi_t(X))_t$ is a $\mathbb{Q}$-martingale. For $x \in \mathbb{R}_+$, we define

$$L\phi(x) = \gamma x\phi''(x) + ((\lambda-\mu)x - cx^2)\phi'(x).$$

Using Taylor's expansion, we immediately get (with $\tilde{L}_K$ defined in (34))

$$\begin{aligned}|\tilde{L}_K\phi(x) - L\phi(x)| &= \gamma K^2\ x\ \left|\phi(x+\frac{1}{K}) + \phi(x-\frac{1}{K}) - 2\phi(x) - \frac{1}{K^2}\phi''(x)\right|\\ &\quad +\lambda\ K\ x\left|\phi(x+\frac{1}{K}) - \phi(x) - \frac{1}{K}\phi'(x)\right|\\ &\quad +K\ (\mu x + cx(x-1))\ \left|\phi(x-\frac{1}{K}) - \phi(x) + \frac{1}{K}\phi'(x)\right|\\ &\leq \frac{C}{K}\ (x^2+1),\end{aligned} \tag{37}$$

where C doesn't depend on x and K. By (36), we deduce that $\mathbb{E}\left(|\tilde{L}_K\phi(X_t^K) - L\phi(X_t^K)|\right)$ tends to 0 as K tends to infinity, uniformly for $t \in [0,T]$.

For $s_1 < \cdots < s_k < s < t$, for $g_1, \cdots, g_k \in C_b$, we introduce the function H defined on the path space by

$$H(X) = g_1(X_{s_1}) \cdots g_k(X_{s_k}) \left(\psi_t(X) - \psi_s(X)\right).$$

Let us show now that

$$\mathbb{E}_{\mathbb{Q}}(H(X)) = 0, \tag{38}$$

which will imply that $(\psi_t(X))_t$ is a $\mathbb{Q}$-martingale.

By construction, $\psi_t^K(X^K) = \phi(X_t^K) - \phi(X_0) - \int_0^t \tilde{L}_K\phi(X_s^K)ds$ defines a martingale, then

$$\mathbb{E}\left[g_1(X_{s_1}^K) \cdots g_k(X_{s_k}^K) \left(\psi_t^K(X^K) - \psi_s^K(X^K)\right)\right] = 0.$$

In another way, this quantity is equal to

$$\begin{aligned}
&\mathbb{E}\left[g_1(X_{s_1}^K) \cdots g_k(X_{s_k}^K) \left(\psi_t^K(X^K) - \psi_s^K(X^K) - \psi_t(X^K) + \psi_s(X^K)\right)\right] \\
&+\mathbb{E}\left[g_1(X_{s_1}^K) \cdots g_k(X_{s_k}^K) \left(\psi_t(X^K) - \psi_s(X^K)\right) - g_1(X_{s_1}) \cdots g_k(X_{s_k}) \left(\psi_t(X) - \psi_s(X)\right)\right] \\
&+\mathbb{E}\left[g_1(X_{s_1}) \cdots g_k(X_{s_k}) \left(\psi_t(X) - \psi_s(X)\right)\right].
\end{aligned}$$

The first term is equal to $\mathbb{E}\left[g_1(X_{s_1}^K) \cdots g_k(X_{s_k}^K) \int_0^t \left(\tilde{L}_K\phi(X_s^K) - L\phi(X_s^K)\right) ds\right]$ and tends to 0 by (36) and (37).

The second term is equal to $\mathbb{E}(H(X^K) - H(X))$. The function $X \mapsto H(X)$ is continuous and since $H(X) \leq C\left(1 + \int_s^t (1 + X_u^2)du\right)$, it is also uniformly integrable by (35). This leads the second term to tend to 0 as K tends to infinity.

Therefore, it turns out that (38) is fulfilled and the process $\psi_t(X) = \phi(X_t) - \phi(X_0) - \int_0^t L\phi(X_s)ds$ is a $\mathbb{Q}$-martingale.

By (36) and taking $\phi(x) = x$ leads to $X_t = X_0 + M_t + \int_0^t ((\lambda - \mu)X_s - cX_s^2)ds$, where M is a martingale. Taking $\phi(x) = x^2$ on the one hand and applying Itô's formula for X_t^2 on the other hand allow us to identify

$$\langle M \rangle_t = \int_0^t 2\,\gamma\; X_s\; ds.$$

By the representation theorem proved in [38] Theorem III-4.2 or in [36], there exists a Brownian motion B such that

$$M_t = \int_0^t \sqrt{2\gamma X_s}\; dB_s.$$

That concludes the proof.

□

5.2 QSD for logistic Feller diffusion processes

5.2.1 Statement of the results

We are now interested in studying the quasi-stationarity for the logistic Feller diffusion process solution of the equation

$$dZ_t = \sqrt{Z_t} dB_t + (rZ_t - cZ_t^2)dt, \quad Z_0 > 0,$$

where the Brownian motion B and the initial state Z_0 are given, and r and c are assumed to be positive. (We have assumed that $\gamma = 1/2$). The results and proofs that are presented in Section 5.2 have been obtained by Cattiaux, Collet, Lambert, Martínez, Méléard and San Martín [13].

Let us firstly state the main theorem of this part.

Theorem 19. *Assume that Z_0, r and c are positive. Then the Yaglom limit of the process Z exists and is a QLD for Z starting from any initial distribution. As a consequence, it is the unique QSD of Z.*

Remark 7. 1) The theory studying the quasi-stationary distributions for one-dimensional diffusion processes started with Mandl [45] and has been developed by many authors. See in particular [16], [47], [57], [40]. Nevertheless in most of the papers, the diffusion and drift coefficients are regular and the "Mandl's condition" $\kappa(+\infty) = \infty$ (see (33)) is assumed. This condition is not satisfied in our case because of the degeneracy of the diffusion and the unboundedness of the drift coefficient.

2) Theorem 19 differs from the results obtained in case of drifts going slower to infinity. For example, Lambert [43] proves that if $c = 0$ and $r \leq 0$, then either $r = 0$ and there is no QSD, or $r < 0$ and there is an infinite number of QSD. Lladser and San Martín [44] show that in the case of the Ornstein-Uhlenbeck process $dY_t = dB_t - Y_t dt$, killed at 0, there is also a continuum of QSD. In the logistic Feller diffusion situation as in the logistic BD-process, the uniqueness comes from the quadratic term $c\,X_t^2$ induced by the ecological constraints.

3) We have seen that the rescaled charge capacity of the logistic birth and death process converges to the charge capacity of the logistic Feller diffusion. However, whether the rescaled Yaglom limit of the logistic birth and death process converges to the Yaglom limit of the logistic Feller diffusion process remains an open problem.

In order to prove Theorem 19, we firstly make a change of variable and introduce the process $(X_t, t \geq 0)$ defined by $X_t = 2\sqrt{Z_t}$. Of course, X is still absorbed at 0 and QSDs for Z will be easily deduced from QSDs for X. From now on, we focus on the process (X_t).

An elementary computation using Itô's formula shows that (X_t) is the Kolmogorov diffusion process defined by

$$dX_t = dB_t - q(X_t)dt, \tag{39}$$

with

$$q(x) = \frac{1}{2x} - \frac{rx}{2} + \frac{cx^3}{8}.$$

Mention that the function q is continuous on $\mathbb{R}_+^*$ but explodes at 0 as $\frac{1}{2x}$ and at infinity as $\frac{c}{8}x^3$. The strong (cubic) downward drift at infinity will force the process to live essentially in compact sets. That will provide the uniqueness of the QSD, as seen below.

We introduce the measure μ, defined by

$$\mu(dy) = e^{-Q(y)}dy,$$

where Q is given by

$$Q(y) = \int_1^y 2q(z)dz = \ln y + \frac{r}{2}(1-y^2) + \frac{c}{16}(y^4-1). \tag{40}$$

In particular $-Q/2$ is a potential of the drift $-q$. The following result clearly implies Theorem 19.

Theorem 20. *[13] Assume that X_0, r and c are positive. Then the Yaglom limit α of the process X exists.*

Moreover, there exists a positive function $\eta_1 \in \mathbb{L}^2(d\mu)$ such that

1. $$\alpha(dx) = \frac{\eta_1(x)e^{-Q(x)}}{\int_{\mathbb{R}_+^*}\eta_1(y)e^{-Q(y)}dy}dx, \tag{41}$$

2. $\forall x \in \mathbb{R}_+^*$, $\lim_{t\to\infty} e^{\theta(\alpha)t}\mathbb{P}_x(T_0 > t) = \eta_1(x)$,

3. *there exists $\chi > 0$ such that, $\forall x \in \mathbb{R}_+^*$,*

$$\lim_{t\to+\infty} e^{-(\chi-\theta(\alpha))t}\left|\mathbb{P}_x\left(X_t \in A | T_0 > t\right) - \alpha(A)\right| < +\infty.$$

4. *the QSD α attracts all initial distribution, which means that α is a QLD for X starting from any initial distribution.*

The proof of Theorem 20 will be decomposed in the next subsections.

5.2.2 Spectral theory for the killed semi-group

As previously we are interested in the semi-group of the killed process, that is, for any $x > 0$, $t > 0$ and any $f \in C_b(\mathbb{R}_+^*)$,

$$P_t f(x) = \mathbb{E}_x(f(X_t)\mathbf{1}_{t<T_0}), \tag{42}$$

with the associated infinitesimal generator given for $\phi \in C_c^2((0,+\infty))$ by

$$L\phi = \frac{1}{2}\phi'' - q\phi'.$$

We are led to develop a spectral theory for this generator in $\mathbb{L}^2(\mu)$. Though the unity function 1 does not belong to $\mathbb{L}^2(\mu)$, this space is the good functional space in which to work. The key point we firstly show is that, starting from $x > 0$, the law of the killed process at time t is absolutely continuous with respect to μ with a density belonging to $\mathbb{L}^2(\mu)$. The first step of the proof is a Girsanov Theorem.

Proposition 21. *For any bounded Borel function F defined on $\Omega = C([0,t],\mathbb{R}_+^*)$ it holds*

$$\mathbb{E}_x\left[F(\omega)\mathbf{1}_{t<T_0(\omega)}\right] = \mathbb{E}^{\mathbb{W}_x}\left[F(\omega)\mathbf{1}_{t<T_0(\omega)}\exp\left(\frac{1}{2}Q(x) - \frac{1}{2}Q(\omega_t) - \frac{1}{2}\int_0^t (q^2 - q')(\omega_s)ds\right)\right]$$

where $\mathbb{E}^{\mathbb{W}_x}$ denotes the expectation with respect to the Wiener measure starting from x and ω the current point in Ω.

Proof. It is enough to show the result for non-negative and bounded functions F. Let $\varepsilon \in (0,1)$ and $\tau_\varepsilon = T_\varepsilon \wedge T_{1/\varepsilon}$. Let us choose some ψ_ε which is a non-negative C^∞ function with compact support included in $]\varepsilon/2, 2/\varepsilon[$ such that $\psi_\varepsilon(u) = 1$ if $\varepsilon \le u \le 1/\varepsilon$. For all x such that $\varepsilon \le x \le 1/\varepsilon$ the law of the diffusion (39) coincides up to τ_ε with the law of a similar diffusion process X^ε obtained by replacing q with the cutoff function $q_\varepsilon = q\psi_\varepsilon$. For the latter we may apply Novikov criterion (cf. [53] p.332), ensuring that the law of X^ε is given via Girsanov's formula. Hence

$$\begin{aligned}
\mathbb{E}_x\left[F(\omega)\mathbf{1}_{t<\tau_\varepsilon(\omega)}\right] &= \mathbb{E}^{\mathbb{W}_x}\left[F(\omega)\,\mathbf{1}_{t<\tau_\varepsilon(\omega)}\,\exp\left(\int_0^t -q_\varepsilon(\omega_s)d\omega_s - \frac{1}{2}\int_0^t (q_\varepsilon)^2(\omega_s)ds\right)\right] \\
&= \mathbb{E}^{\mathbb{W}_x}\left[F(\omega)\,\mathbf{1}_{t<\tau_\varepsilon(\omega)}\,\exp\left(\int_0^t -q(\omega_s)d\omega_s - \frac{1}{2}\int_0^t q^2(\omega_s)ds\right)\right] \\
&= \mathbb{E}^{\mathbb{W}_x}\left[F(\omega)\mathbf{1}_{t<\tau_\varepsilon(\omega)}\exp\left(\frac{1}{2}Q(x) - \frac{1}{2}Q(\omega_t) - \frac{1}{2}\int_0^t (q^2 - q')(\omega_s)ds\right)\right]
\end{aligned}$$

integrating by parts the stochastic integral. But $\mathbf{1}_{t<\tau_\varepsilon}$ is non-decreasing in ε and converges almost surely to $\mathbf{1}_{t<T_0}$ both for $\mathbb{W}_x$ and for $\mathbb{P}_x$ (since $\mathbb{P}_x(T_0 < \infty) = 1$)). Indeed, almost surely,

$$\lim_{\varepsilon\to 0} X_{\tau_\varepsilon} = \lim_{\varepsilon\to 0} X_{\tau_\varepsilon} = \lim_{\varepsilon\to 0}\varepsilon = 0$$

so that $\lim_{\varepsilon\to 0}\tau_\varepsilon \ge T_0$. But $\tau_\varepsilon \le T_0$ yielding the equality. It remains to use Lebesgue monotone convergence theorem to finish the proof. □

Theorem 22. *For all $x > 0$ and all $t > 0$ there exists a density function $r(t,x,.)$ that satisfies*

$$\mathbb{E}_x[f(X_t)\,\mathbf{1}_{t<T_0}] = \int_0^{+\infty} f(y)\,r(t,x,y)\,\mu(dy)$$

for all bounded Borel function f. In addition, for all $t > 0$ and all $x > 0$,

$$\int_0^{+\infty} r^2(t,x,y)\,\mu(dy) \;\le (1/2\pi t)^{\frac{1}{2}}\,e^{Ct}\,e^{Q(x)}\,,$$

where

$$C = -\inf_{y>0}(q^2(y) - q'(y)) < +\infty.$$

Proof. Define

$$G(\omega) = \mathbf{1}_{t<T_0(\omega)}\,\exp\left(\frac{1}{2}\,Q(\omega_0) - \frac{1}{2}\,Q(\omega_t) - \frac{1}{2}\int_0^t (q^2 - q')(\omega_s)ds\right).$$

Denote by $e^{-v(t,x,y)} = (2\pi t)^{-\frac{1}{2}} \exp\left(-\frac{(x-y)^2}{2t}\right)$ the density at time t of the Brownian motion starting from x. According to Proposition 21, we have

$$\begin{aligned}
\mathbb{E}_x\left(f(X_t)\,\mathbf{1}_{t<T_0}\right) &= \mathbb{E}^{\mathbb{W}_x}\left(f(\omega_t)\,\mathbb{E}^{\mathbb{W}_x}(G|\omega_t)\right) \\
&= \int_0^{+\infty} f(y)\,\mathbb{E}^{\mathbb{W}_x}(G|\omega_t = y)\,e^{-v(t,x,y)}\,dy \\
&= \int_0^{+\infty} f(y)\,\mathbb{E}^{\mathbb{W}_x}(G|\omega_t = y)\,e^{-v(t,x,y)+Q(y)}\,\mu(dy)\,,
\end{aligned}$$

because $\mathbb{E}^{\mathbb{W}_x}(G|\omega_t = y) = 0$ if $y \leq 0$. In other words, the law of X_t restricted to non extinction has a density with respect to μ given by

$$r(t,x,y) = \mathbb{E}^{\mathbb{W}_x}(G|\omega_t = y)\,e^{-v(t,x,y)+Q(y)}\,.$$

Hence

$$\begin{aligned}
\int_0^{+\infty} r^2(t,x,y)\,\mu(dy) &= \int_0^{+\infty} \left(\mathbb{E}^{\mathbb{W}_x}(G|\omega_t = y)\,e^{-v(t,x,y)+Q(y)}\right)^2 \\
&\qquad \times e^{-Q(y)+v(t,x,y)}\,e^{-v(t,x,y)}\,dy \\
&= \mathbb{E}^{\mathbb{W}_x}\left(e^{-v(t,x,\omega_t)+Q(\omega_t)}\,\left(\mathbb{E}^{\mathbb{W}_x}(G|\omega_t)\right)^2\right) \\
&\leq \mathbb{E}^{\mathbb{W}_x}\left(e^{-v(t,x,\omega_t)+Q(\omega_t)}\,\mathbb{E}^{\mathbb{W}_x}(G^2|\omega_t)\right) \\
&\leq e^{Q(x)}\,\mathbb{E}^{\mathbb{W}_x}\left(\mathbf{1}_{t<T_0(\omega)}\,e^{-v(t,x,\omega_t)}\,e^{-\int_0^t (q^2-q')(\omega_s)ds}\right)\,,
\end{aligned}$$

where we have used Cauchy-Schwarz's inequality. Since $e^{-v(t,x,.)} \leq (1/2\pi t)^{\frac{1}{2}}$, the proof is completed. □

Thanks to Theorem 22, we can show, using the theory of Dirichlet forms (cf. Fukushima's book [25]) that the infinitesimal generator L of X, defined by (5.2.2), can be extended to the generator of a continuous symmetric semi-group of contractions of $\mathbb{L}^2(\mu)$ denoted by $(P_t)_{t\geq 0}$. In all what follows, and for $f, g \in \mathbb{L}^2(\mu)$, we will denote $\langle f, g\rangle_\mu = \int_{\mathbb{R}_+} f(x)g(x)\mu(dx)$. The symmetry of P_t means that $\langle P_t f, g\rangle_\mu = \langle f, P_t g\rangle_\mu$.
In Cattiaux et al. [13], the following spectral theorem in $\mathbb{L}^2(\mu)$ is proved.

Theorem 23. *The operator $-L$ has a purely discrete spectrum $0 < \lambda_1 < \lambda_2 < \cdots$. Furthermore each λ_i ($i \in \mathbb{N}^*$) is associated with a unique (up to a multiplicative constant) eigenfunction η_i of class $C^2((0,\infty))$, which satisfies the ODE*

$$\frac{1}{2}\eta_i'' - q\eta_i' = -\lambda_i\eta_i. \tag{43}$$

The sequence $(\eta_i)_{i\geq 1}$ is an orthonormal basis of $\mathbb{L}^2(\mu)$ and $\eta_1(x) > 0$ for all $x > 0$.

In addition, for each i, $\eta_i \in \mathbb{L}^1(\mu)$.

The proof of this theorem is based on a relation between the Fokker-Planck operator L and a Schrödinger operator. Indeed, let us set for $g \in \mathbb{L}^2(dx)$,

$$\tilde{P}_t g = e^{-Q/2}\ P_t(g\ e^{Q/2}).$$

$\tilde{P}_t$ is a strongly semi-group on $\mathbb{L}^2(dx)$ with generator defined for $g \in C_c^\infty((0,+\infty))$ by

$$\tilde{L}g = \frac{1}{2}\triangle g - \frac{1}{2}(q^2 - q')\ g.$$

The spectral theory for such Schrödinger operator with potential $\frac{(q^2-q')}{2}$ on the line (or the half-line) is well known (see for example the book of Berezin-Shubin [7]), but the potential $\frac{(q^2-q')}{2}$ does not belong to $\mathbb{L}^\infty_{loc}$ as generally assumed. Nevertheless, in our case $\inf(q^2 - q') > -\infty$, which ensures the compactness of the operators $\tilde{L}$ and $\tilde{P}_t$.

The following corollary of Theorem 23 is a generalization of the Perron-Frobenius Theorem in this infinite-dimensional framework.

Corollary 24. *For any bounded and measurable function f, we have*

$$P_t f =_{\mathbb{L}^2(\mu)} \sum_{i\in\mathbb{N}^*} e^{-\lambda_i t}\ \langle \eta_i, f\rangle_\mu\ \eta_i. \tag{44}$$

Proof. Fix $t > 0$ and let f be a bounded measurable function on $\mathbb{R}_+^*$. Let us first prove that $P_t f$ belongs to $\mathbb{L}^2(\mu)$. On the one hand, we have

$$\int_1^{+\infty} (P_t f(x))^2 d\mu(x) \le \|f\|_\infty^2 \int_1^\infty e^{-Q(x)} dx < \infty.$$

On the other hand, by Proposition 21, we have, for all $x \in \mathbb{R}_+^*$,

$$\begin{aligned} P_t f(x) &\le \|f\|_\infty\, e^{\frac{1}{2}Q(x)+\frac{1}{2}Ct}\, \mathbb{E}^{\mathbb{W}_x}\left[\mathbf{1}_{t<T_0(\omega)} e^{-\frac{1}{2}Q(\omega_t)}\right] \\ &\le \|f\|_\infty\, e^{\frac{1}{2}Q(x)+\frac{1}{2}Ct} \int_0^\infty e^{-\frac{1}{2}Q(y)} \frac{e^{-\frac{1}{2t}(y-x)^2}}{\sqrt{2\pi t}} dy. \end{aligned}$$

But the function

$$y \mapsto e^{-\frac{1}{2}Q(y)} = \frac{1}{\sqrt{y}} e^{-\frac{r}{4}(1-y^2)-\frac{c}{32}(y^4-1)},$$

is integrable on $]0,+\infty[$. Since $e^{-\frac{1}{2t}(y-x)^2} \le 1$, we deduce that there exists a constant $K_t > 0$ independent of x and f such that

$$P_t f(x) \le K_t \|f\|_\infty e^{\frac{1}{2}Q(x)},$$

and thus

$$\int_0^1 (P_t f(x))^2 d\mu(x) \le K_t^2 \|f\|_\infty^2.$$

Finally $(P_t f)^2$ is integrable with respect to μ, so that $P_t f \in \mathbb{L}^2(\mu)$.

Now we deduce from Theorem 23 that

$$P_t f =_{\mathbb{L}^2(\mu)} \sum_{i\in\mathbb{N}^*} \langle P_t f, \eta_i\rangle_\mu \ \eta_i \tag{45}$$

If f belongs to $\mathbb{L}^2(\mu)$, then the symmetry of P_t implies that

$$\langle P_t f, \eta_i\rangle_\mu = \langle f, P_t\eta_i\rangle_\mu = e^{-\lambda_i t}\ \langle f, \eta_i\rangle_\mu.$$

Since $\eta_i \in \mathbb{L}^1(\mu)$, we deduce from the Dominated Convergence Theorem that the equality $\langle P_t f, \eta_i\rangle_\mu = e^{-\lambda_i t}\ \langle f, \eta_i\rangle_\mu$ extends to all measurable bounded functions. This and the equality (45) allow us to conclude the proof of Corollary 24. □

5.2.3 Existence of the Yaglom limit

By Corollary 24, we have for any bounded and measurable function f,

$$\begin{aligned}\|e^{\lambda_1 t}P_t f - \langle\eta_1, f\rangle\eta_1\|^2_{\mathbb{L}^2(\mu)} &= \sum_{i\geq 2} e^{-2t(\lambda_i-\lambda_1)}|\langle\eta_i,f\rangle|^2\\ &\leq e^{-2(t-1)(\lambda_2-\lambda_1)}\sum_{i\in\mathbb{N}^*} e^{-2(\lambda_i-\lambda_1)}|\langle\eta_i,f\rangle|^2\\ &\leq e^{-2(t-1)(\lambda_2-\lambda_1)}\, e^{2\lambda_1}\, \|P_1 f\|^2_{\mathbb{L}^2(\mu)}\end{aligned}$$

Using Cauchy-Schwartz inequality, we deduce that, for any function $h \in \mathbb{L}^2(\mu)$,

$$\left|e^{\lambda_1 t}\langle P_t f,h\rangle_\mu - \langle\eta_1, f\rangle\langle\eta_1, h\rangle_\mu\right| \leq e^{-2(t-1)(\lambda_2-\lambda_1)}\|P_1 f\|_{\mathbb{L}^2(\mu)}\|h\|_{\mathbb{L}^2(\mu)}. \tag{46}$$

By Theorem 22, $\delta_x P_1$ has the density $r(1,x,.) \in \mathbb{L}^2(\mu)$ with respect to μ, so that, taking $h = r(1,x,\cdot)$,

$$\left|e^{\lambda_1 t}P_{t+1} f(x) - \langle\eta_1, f\rangle\langle\eta_1, r(1,x,\cdot)\rangle_\mu\right| \leq e^{-2(t-1)(\lambda_2-\lambda_1)}\|P_1 f\|_{\mathbb{L}^2(\mu)}\|r(1,x,\cdot)\|_{\mathbb{L}^2(\mu)}.$$

By definition of η_1, $\langle\eta_1, r(1,x,\cdot)\rangle_\mu = e^{-\lambda_1}\eta_1(x)$. Thus we have

$$e^{\lambda_1 t}P_{t+1} f(x) \xrightarrow[t\to+\infty]{} \langle\eta_1, f\rangle_\mu\, e^{-\lambda_1}\eta_1(x)$$

and

$$e^{\lambda_1 t}P_{t+1}\mathbf{1}_{\mathbb{R}^*_+}(x) \xrightarrow[t\to\infty]{} \langle\eta_1, \mathbf{1}_{\mathbb{R}^*_+}\rangle_\mu\, e^{-\lambda_1}\eta_1(x)$$

Finally, $\eta_1(x)$ being positive, for any $x \in \mathbb{R}^*_+$,

$$\frac{P_t f(x)}{P_t\mathbf{1}_{\mathbb{R}^*_+}(x)} \xrightarrow[t\to+\infty]{} \frac{\langle\eta_1, f\rangle_\mu}{\langle\eta_1, \mathbf{1}_{\mathbb{R}^*_+}\rangle_\mu} = \alpha(f),$$

where α is defined in (41). We conclude that α is the Yaglom limit for Z. We also deduce parts (2) and (3) of Theorem 20.

5.2.4 Attractiveness of any initial distribution

Let us first consider a compactly supported probability measure ν on $(0, +\infty)$. By Theorem 22, $y \mapsto \int_{E^*} r(1,x,y)\nu(dx)$ is the density of νP_1 with respect to μ. By [13, Lemma 5.3], there exists a locally bounded function Θ such that

$$r(1,x,y) \leq \Theta(x)\eta_1(y),\ \forall x,y \in (0,+\infty).$$

In particular, $h : y \mapsto \int_{E^*} r(1,x,y)\nu(dx)$ belongs to $\mathbb{L}^2$. Then we deduce from (46) that

$$\mathbb{E}_\nu\left(f(X_{t+1})|\, T_0 > t+1\right) = \frac{\nu P_{t+1}(f)}{\nu P_{t+1}(\mathbf{1}_{E^*})} \xrightarrow[t\to\infty]{} \alpha(f).$$

We conclude that α attracts any compactly supported probability measure.

Let us now prove that α attracts all initial distributions ν supported in $(0,\infty)$. We want to show that, for any probability measure ν on $\mathbb{R}_+^*$, for any Borel set A, we get

$$\lim_{t\to\infty} \mathbb{P}_\nu(X_t \in A | T_0 > t) = \alpha(A). \tag{47}$$

This is part (4) of Theorem 20 and it clearly implies the uniqueness of the QSD for X (and hence for Z).

Proposition 25. *For any $a > 0$, there exists $y_a > 0$ such that $\sup_{x>y_a} \mathbb{E}_x(e^{aT_{y_a}}) < \infty$.*

Proof. Let us remark that $\int_1^\infty e^{Q(y)} \int_y^\infty e^{-Q(z)}\, dz\, dy < \infty$. Let $a > 0$, and pick x_a large enough so that $\int_{x_a}^\infty e^{Q(x)} \int_x^\infty e^{-Q(z)}\, dz\, dx \leq \frac{1}{2a}$. Let J be the nonnegative increasing function defined on $[x_a,\infty)$ by

$$J(x) = \int_{x_a}^x e^{Q(y)} \int_y^\infty e^{-Q(z)}\, dz\, dy.$$

Then we check that $J'' = 2qJ' - 1$, so that $LJ = -1/2$. Set now $y_a = 1 + x_a$, and consider a large $M > x$. Itô's formula gives

$$\mathbb{E}_x(e^{a(t\wedge T_M \wedge T_{y_a})}\, J(X_{t\wedge T_M\wedge T_{y_a}})) = J(x) + \mathbb{E}_x\left(\int_0^{t\wedge T_M\wedge T_{y_a}} e^{as}\left(aJ(X_s) + LJ(X_s)\right) ds\right).$$

But $LJ = -1/2$, and $J(X_s) < J(\infty) \leq 1/(2a)$ for any $s \leq T_{y_a}$, so that

$$\mathbb{E}_x(e^{a(t\wedge T_M \wedge T_{y_a})}\, J(X_{t\wedge T_M\wedge T_{y_a}})) \leq J(x).$$

For $x \geq y_a$, one gets $1/(2a) > J(x) \geq J(y_a) > 0$. It follows that $\mathbb{E}_x(e^{a(t\wedge T_M\wedge T_{y_a})}) \leq 1/(2aJ(y_a))$. Letting $M \to \infty$ then $t \to \infty$, we deduce $\mathbb{E}_x(e^{aT_{y_a}}) \leq 1/(2aJ(y_a))$, by the monotone convergence theorem. So Proposition 25 is proved. □

Proving that α attracts all initial distributions requires the following estimates near 0 and ∞.

Lemma 26. *For $h \in \mathbb{L}^1(\mu)$ strictly positive on $(0,\infty)$ we have*

$$\lim_{\varepsilon\downarrow 0} \limsup_{t\to\infty} \frac{\int_0^\varepsilon h(x)\mathbb{P}_x(T_0 > t)\mu(dx)}{\int_0^{+\infty} h(x)\mathbb{P}_x(T_0 > t)\mu(dx)} = 0, \tag{48}$$

$$\lim_{M\uparrow\infty} \limsup_{t\to\infty} \frac{\int_M^\infty h(x)\mathbb{P}_x(T_0 > t)\mu(dx)}{\int_0^{+\infty} h(x)\mathbb{P}_x(T_0 > t)\mu(dx)} = 0. \tag{49}$$

Proof. We start with (48). Using Harnack's inequality (see [58, Theorem 1.1]), we have for $\varepsilon < 1$ and large t

$$\frac{\int_0^\varepsilon h(x)\mathbb{P}_x(T_0 > t)\mu(dx)}{\int_0^{+\infty} h(x)\mathbb{P}_x(T_0 > t)\mu(dx)} \le \frac{\mathbb{P}_1(T_0 > t)\int_0^\varepsilon h(z)\mu(dz)}{C\int_1^{3/2} h(x)\mu(dx)\mathbb{P}_1(T_0 > t-1)},$$

then

$$\begin{aligned}\limsup_{t\to\infty} \frac{\int_0^\varepsilon h(x)\mathbb{P}_x(T_0 > t)\mu(dx)}{\int h(x)\mathbb{P}_x(T_0 > t)\mu(dx)} &\le \limsup_{t\to\infty} \frac{\mathbb{P}_1(T_0 > t)\int_0^\varepsilon h(z)\mu(dz)}{C\int_1^{3/2} h(x)\mu(dx)\mathbb{P}_1(T_0 > t-1)} \\ &= \frac{e^{-\lambda_1}\int_0^\varepsilon h(z)\mu(dz)}{C\int_1^{3/2} h(x)\mu(dx)},\end{aligned}$$

and the first assertion of the lemma is proved.

For the second limit, we set $A_0 := \sup_{x\ge y_{\lambda_1}} \mathbb{E}_x(e^{\lambda_1 T_{y_{\lambda_1}}}) < \infty$, where y_{λ_1} is taken from Proposition 25. Then for large $M > y_{\lambda_1}$, we have

$$\mathbb{P}_x(T_0 > t) = \int_0^t \mathbb{P}_{x_0}(T_0 > u)\mathbb{P}_x(T_{x_0} \in d(t-u)) + \mathbb{P}_x(T_{x_0} > t).$$

Using $\lim_{u\to\infty} e^{\lambda_1 u}\mathbb{P}_{x_0}(T_0 > u) = \eta_1(x_0)\langle\eta_1,1\rangle_\mu$, we obtain $B_0 := \sup_{u\ge 0} e^{\lambda_1 u}\mathbb{P}_{x_0}(T_0 > u) < \infty$. Then

$$\begin{aligned}\mathbb{P}_x(T_0 > t) &\le B_0\int_0^t e^{-\lambda_1 u}\mathbb{P}_x(T_{x_0} \in d(t-u)) + \mathbb{P}_x(T_{x_0} > t) \\ &\le B_0\, e^{-\lambda_1 t}\,\mathbb{E}_x(e^{\lambda_1 T_{x_0}}) + e^{-\lambda_1 t}\,\mathbb{E}_x(e^{\lambda_1 T_{x_0}}) \le e^{-\lambda_1 t}A_0(B_0+1),\end{aligned}$$

and (49) follows immediately (since $x \ge x_0 \ge y_{\lambda_1} \Rightarrow T_{x_0} \le T_{y_{\lambda_1}}$). □

Let ν be any fixed probability distribution whose support is contained in $(0,\infty)$. We must prove (47). We begin by claiming that ν can be assumed to have a strictly positive density h, with respect to μ. Indeed, let

$$\ell(y) = \int_0^{+\infty} r(1,x,y)\nu(dx).$$

Using Tonelli's theorem we have

$$\begin{aligned}\int_0^{+\infty}\int_0^{+\infty} r(1,x,y)\nu(dx)\,\mu(dy) &= \int_0^{+\infty}\int_0^{+\infty} r(1,x,y)\,\mu(dy)\,\nu(dx) \\ &= \int_0^{+\infty} \mathbb{P}_x(T_0 > 1)\nu(dx) \le 1,\end{aligned}$$

which implies that $\int r(1,x,y)\nu(dx)$ is finite $dy-$a.s.. Finally, define $h = \ell/\int \ell d\mu$. Notice that for $d\rho = hd\mu$

$$\mathbb{P}_\nu(X_{t+1} \in \cdot \mid T_0 > t+1) = \mathbb{P}_\rho(X_t \in \cdot \mid T_0 > t),$$

showing the claim.

Consider $M > \varepsilon > 0$ and any Borel set A included in $(0,\infty)$. Then

$$\left| \frac{\int \mathbb{P}_x(X_t \in A, T_0 > t)h(x)\,\mu(dx)}{\int \mathbb{P}_x(T_0 > t)h(x)\,\mu(dx)} - \frac{\int_\varepsilon^M \mathbb{P}_x(X_t \in A, T_0 > t)h(x)\,\mu(dx)}{\int_\varepsilon^M \mathbb{P}_x(T_0 > t)h(x)\,\mu(dx)} \right|$$

is bounded by the sum of the following two terms

$$\begin{aligned} I1 &= \left| \frac{\int \mathbb{P}_x(X_t \in A, T_0 > t)h(x)\,\mu(dx)}{\int \mathbb{P}_x(T_0 > t)h(x)\,\mu(dx)} - \frac{\int_\varepsilon^M \mathbb{P}_x(X_t \in A, T_0 > t)h(x)\,\mu(dx)}{\int \mathbb{P}_x(T_0 > t)h(x)\,\mu(dx)} \right| \\ I2 &= \left| \frac{\int_\varepsilon^M \mathbb{P}_x(X_t \in A, T_0 > t)h(x)\,\mu(dx)}{\int \mathbb{P}_x(T_0 > t)h(x)\,\mu(dx)} - \frac{\int_\varepsilon^M \mathbb{P}_x(X_t \in A, T_0 > t)h(x)\,\mu(dx)}{\int_\varepsilon^M \mathbb{P}_x(T_0 > t)h(x)\,\mu(dx)} \right|. \end{aligned}$$

We have the bound

$$I1 \vee I2 \leq \frac{\int_0^\varepsilon \mathbb{P}_x(T_0 > t)h(x)\,\mu(dx) + \int_M^\infty \mathbb{P}_x(T_0 > t)h(x)\,\mu(dx)}{\int \mathbb{P}_x(T_0 > t)h(x)\,\mu(dx)}.$$

Thus, from Lemma 26 we get

$$\lim_{\varepsilon\downarrow 0,\, M\uparrow\infty} \limsup_{t\to\infty} \left| \frac{\int \mathbb{P}_x(X_t \in A, T_0 > t)h(x)\mu(dx)}{\int \mathbb{P}_x(T_0 > t)h(x)\,\mu(dx)} - \frac{\int_\varepsilon^M \mathbb{P}_x(X_t \in A, T_0 > t)h(x)\mu(dx)}{\int_\varepsilon^M \mathbb{P}_x(T_0 > t)h(x)\mu(dx)} \right| = 0.$$

On the other hand we have

$$\lim_{t\to\infty} \frac{\int_\varepsilon^M \mathbb{P}_x(X_t \in A, T_0 > t)h(x)\,\mu(dx)}{\int_\varepsilon^M \mathbb{P}_x(T_0 > t)h(x)\,\mu(dx)} = \frac{\int_A \eta_1(z)\mu(dz)}{\int_{\mathbb{R}^+} \eta_1(z)\mu(dz)} = \alpha(A),$$

since α attracts any compactly supported probability measures , and the result follows.

Example 4. We develop now a numerical illustration of this logistic Feller diffusion case. As for the logistic birth and death process (see Example 3, Section 4), the value of the charge capacity $\frac{r}{c}$ will remain equal to the fixed value 9 across the whole example.

We begin by showing in Figure 7 a random path of a logistic Feller diffusion process with initial size $Z_0 = 1$ and with parameters $r = 9$ and $c = 1$ (an Euler method is used for the numerical simulation of the random path). We observe that the process quickly attains the value of the charge capacity and remains around it for a long time.

We compare now the Yaglom limits of two different logistic Feller diffusion processes whose charge capacity is equal to 9 (see Figure 8):

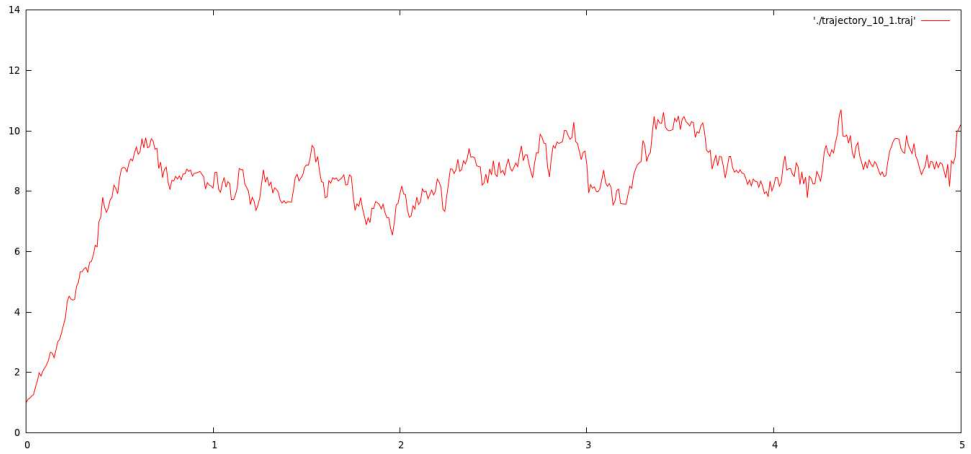


Figure 7: Example 4. A random path for logistic Feller diffusion process with initial size $Z_0 = 1$ and parameters $r = 9$ and $c = 1$

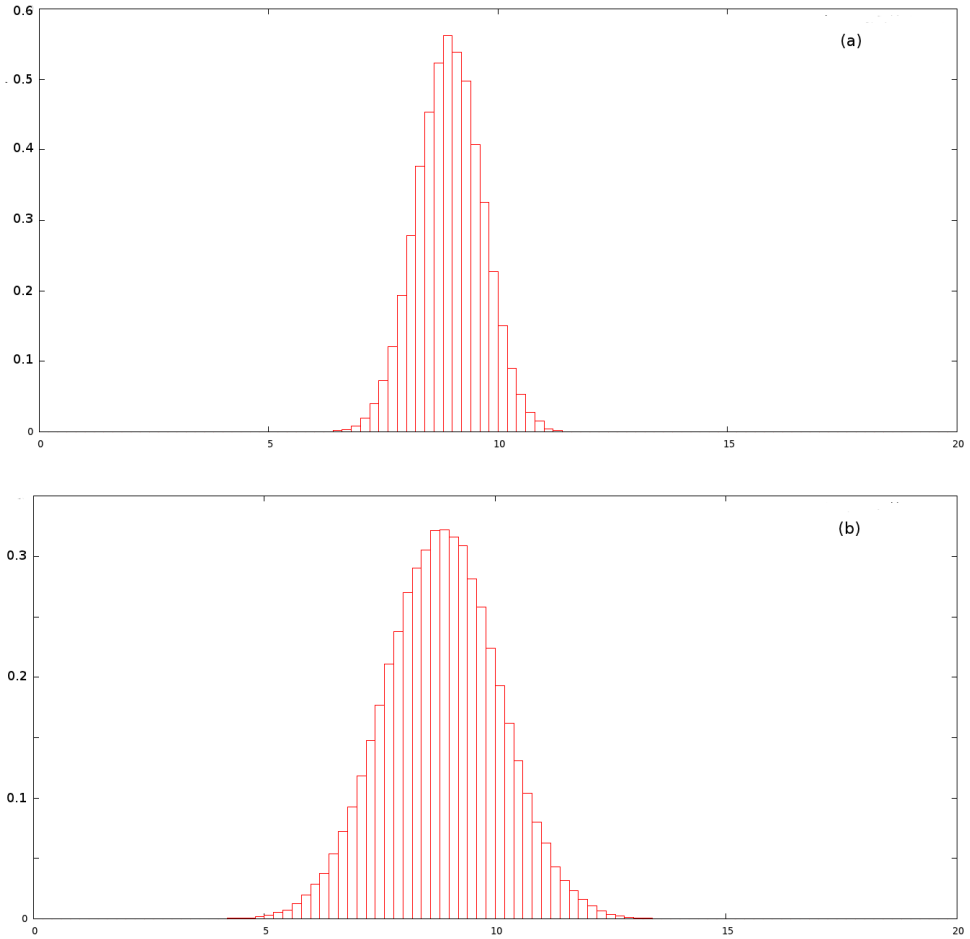


Figure 8: Example 4. The Yaglom limits of two logistic Feller diffusion processes with the same charge capacity: (a) $r = 9$ and $c = 1$; (b) $r = 3$ and $c = 1/3$.

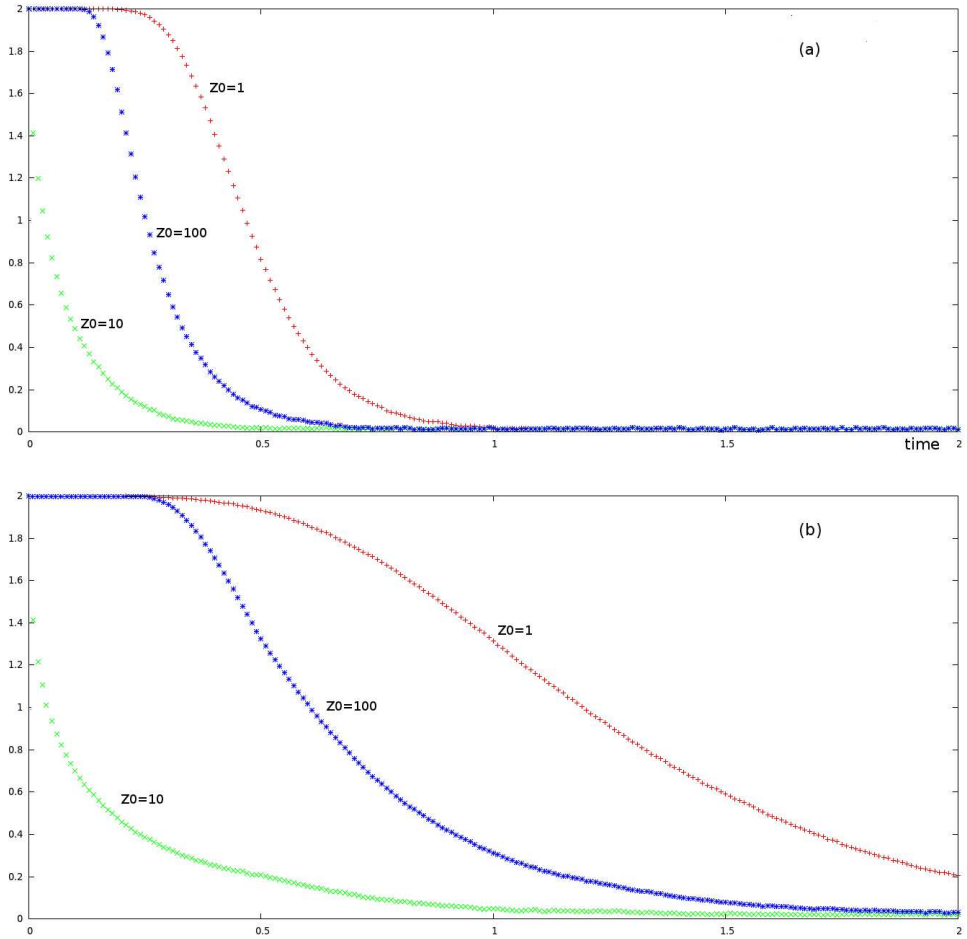


Figure 9: Example 4. Evolution of the distance between the conditioned distribution and the Yaglom limits of two logistic Feller diffusion processes with the same charge capacity $\frac{r}{c} = 9$: (a) $r = 9$ and $c = 1$; (b) $r = 3$ and $c = 1/3$.

(a) $Z^{(a)}$, whose parameters are $r = 9$ and $c = 1$,

(b) $Z^{(b)}$, whose parameters are $r = 3$ and $c = 1/3$.

As for the logistic BD-processes, we observe that the two Yaglom limits are *centered* around the charge capacity. But as a consequence of the relatively *weak* noise around the charge capacity, the Yaglom limit has clearly a smaller variation around this value in the logistic Feller diffusion case than in the logistic BD process case. We also observe that the smaller are the parameters, the flatter is the Yaglom limit and with a similar explanation as in the logistic BD-process case.

We observe now the distance between the conditional distributions of $Z^{(a)}$ and $Z^{(b)}$ and their respective Yaglom limits, for different initial states, namely $Z_0 = 1$, $Z_0 = 10$ and $Z_0 = 100$. The results, computed with the help of the approximation method studied in Section 6, are represented on figure 9. For both $Z^{(a)}$ and $Z^{(b)}$, the speed of convergence to the Yaglom limit is the highest for $Z_0 = 10$, which is quite intuitive since the value of the charge capacity is 9. We also observe that it is higher for the processes starting from 100 than for the processes starting from 1. In particular, this behavior is different than in the logistic birth and death process case.

5.3 The Q-process

Let us now describe the law of the trajectories conditioned to never attain 0.

Theorem 27. *[13] Let us fix a time s and consider B a measurable subset of $C([0,s],\mathbb{R}_+)$. Then for any $x \in \mathbb{R}_+^*$,*

$$\lim_{t\to\infty} \mathbb{P}_x(X \in B | t < T_0) = Q_x(X \in B),$$

where Q_x is the law of a continuous process with transition probabilities given by $q(s,x,y)dy$, with

$$q(s,x,y) = e^{\lambda_1 s}\ \frac{\eta_1(y)}{\eta_1(x)}\ r(s,x,y)\ e^{-Q(y)}.$$

Proof. Since $r(s,x,y)\ e^{-Q(y)}dy$ is the law of X_s started from x before extinction, we have to prove that

$$Q_x(X \in B) = e^{\lambda_1 s}\mathbb{E}_x\left(\mathbf{1}_B(X)\ \frac{\eta_1(X_s)}{\eta_1(x)}\ \mathbf{1}_{T_0>s}\right).$$

For $t > s$,

$$\frac{\mathbb{P}_x(X \in B; T_0 > t)}{\mathbb{P}_x(T_0 > t)} = \frac{\mathbb{P}_x(X \in B; T_0 > s; \mathbb{E}_{X_s}(T_0 > t-s))}{\mathbb{P}_x(T_0 > t)},$$

and we have proved that

$$\lim_{t\to\infty} \frac{\mathbb{P}_y(T_0 > t-s)}{\mathbb{P}_x(T_0 > t)} = e^{\lambda_1 s}\ \frac{\eta_1(y)}{\eta_1(x)}.$$

Then,

$$\lim_{t\to\infty} \frac{\mathbb{P}_x(X \in B; T_0 > t)}{\mathbb{P}_x(T_0 > t)} = \frac{e^{\lambda_1 s}}{\eta_1(x)}\ \mathbb{P}_x\left(\mathbf{1}_B(X)\ \eta_1(X_s)\ \mathbf{1}_{T_0>s}\right).$$

□

Corollary 28. *For any Borel set $A \subset (0,\infty)$ and any x,*

$$\lim_{s\to\infty} Q_x(X_s \in A) = \int_A \eta_1^2(y)\mu(dy) =< \eta_1, 1 >_\mu \int_A \eta_1(y)\alpha(dy).$$

Proof. Since $\mathbf{1}_A\,\eta_1 \in \mathbb{L}^2(\mu)$, thus

$$\eta_1(x)\ Q_x(X_s \in A) = \int \mathbf{1}_A(y)\ \eta_1(y)\, e^{\lambda_1 s}\ r(s,x,y)\ \mu(dy)$$

converges to $\eta_1(x)\ \int_B \eta_1^2(y)\mu(dy)$ as $s \to +\infty$, since $e^{\lambda_1 s} r(s,x,.)$ converges to $\eta_1(x)\,\eta_1(.)$ in $\mathbb{L}^2(d\mu)$. □

Remark 8. The stationary measure of the Q-process is absolutely continuous with respect to α, with Radon-Nikodym derivative $< \eta_1, 1 >_\mu\ \eta_1$.

5.4 The case of a multi-type population

Until now, we have considered a population where all individuals have the same ecological parameters. This biological assumption corresponds to the case where individuals have the same type. In this section, we generalize the previous study to a population composed of k different types. The population size process describing the dynamics of each subpopulation is given by a k-dimensional stochastic Lotka-Volterra process $Z = (Z_t^1, \cdots, Z_t^k)_{t\geq 0}$ (SLVP), which describes the size of a k-types density dependent population. This model generalizes to k types the 2-types density dependent model introduced by Cattiaux and Méléard [14].

More precisely, we consider for $i,j \in \{1, \cdots, k\}$ the coefficients

$$\gamma_i > 0 \ , \ r_i > 0 \ ; \ c_{ij} > 0, \ \forall i,j \in \{1, \cdots, k\}.$$

The process Z takes its values in $(\mathbb{R}_+)^k$ and is solution of the stochastic differential system

$$dZ_t^i = \sqrt{\gamma_i Z_t^i} dB_t^i + (r_i Z_t^i - \sum_{j=1}^{k} c_{ij} Z_t^i Z_t^j) \ dt, \tag{50}$$

where $(B^i)_{i=1,\cdots,k}$ are independent standard Brownian motions independent of the initial data Z_0. The system (50) can be obtained as (31) as approximation of renormalized k-types birth and death processes in case of large population and small life lengths and reproduction times. The coefficients r_i are the asymptotic growth rates of i-type's populations. The positive coefficients γ_i can be interpreted as demographic parameters describing the ecological timescale. The coefficient c_{ij}, for $i,j = 1, \cdots, k$, represents the pressure felt by an individual holding type i from an individual with type j. Intra-specific competition is modeled by the rates c_{ii}, while inter-specific competition is described by the coefficients $c_{ij} > 0, i \neq j$. If $c_{ij} = 0$ for all $i \neq j$, the stochastic k-dimensional process reduces to k independent Feller logistic diffusion processes. Extinction of the population is modeled by the absorbing state $(0, \cdots, 0)$ and the extinction of the subpopulation of type i is modeled by the absorbing set

$$H_i = (\mathbb{R}_+^*)^{i-1} \times \{0\} \times (\mathbb{R}_+^*)^{k-i}.$$

We denote by D the open subset of $\mathbb{R}^k$ defined by $D = (\mathbb{R}_+^*)^k$ and by ∂D its boundary. We denote by T_0 the first hitting time of $(0, \cdots, 0)$, by T_A the first hitting time of some subset A and thus by $T_{\partial D}$ the exit time of D. Of course, some of these stopping times are comparable. For example if the initial condition belongs to D,

$$T_{\partial D} \leq T_{H_i} \leq T_0, \quad \forall i = 1, \cdots, k. \tag{51}$$

On the other hand, T_{H_i} and T_{H_j} are not directly comparable for $i \neq j$.

Let us prove the existence of the SLVP .

Proposition 29. *The process $(Z_t)_t$ is well defined on $\mathbb{R}_+$. In addition, for all $x \in (\mathbb{R}_+)^k$,*

$$\mathbb{P}_x(T_0 < +\infty) = 1$$

and there exists $\lambda > 0$ such that

$$\sup_{x \in (\mathbb{R}_+)^k} \mathbb{E}_x(e^{\lambda T_0}) < +\infty.$$

Proof. The existence of the SLVP is shown by a comparison argument (cf. Ikeda-Watanabe [36] Chapter 6 Thm 1.1). Indeed, the coordinates $(Z_t^i)_t$ can be upper-bounded by the independent solutions of logistic Feller equations

$$dY_t^i = \sqrt{\gamma_i Y_t^i} dB_t^i + (r_i Y_t^i - c_{ii}(Y_t^i)^2)\ dt, \tag{52}$$

for which we have obtained in the previous section that extinction occurs a.s. in finite time and that the extinction time has some finite exponential moments. The almost sure finiteness of each T_{H_i}, hence of $T_{\partial D}$ and T_0, thus follows. □

As in the previous sections we are interested in the quasi-stationary distributions for the process (50). We firstly reduce the problem by a change of variable. Let us define $X_t = (X_t^1, \cdots, X_t^k)$ with $X_t^i = 2\sqrt{\frac{Z_t^i}{\gamma_i}}$. We obtain via Itô's formula and for any $i \in \{1, \cdots, k\}$,

$$dX_t^i \ = \ dB_t^i \ + \left(\frac{r_i X_t^i}{2} \ - \sum_{j=1}^{k} \frac{c_{ij}\gamma_j\, X_t^i (X_t^j)^2}{8} \ - \frac{1}{2X_t^i} \right). \tag{53}$$

In the following, we will focus on the symmetric case where X is a Kolmogorov diffusion, that is a Brownian motion with a drift in gradient form as

$$dX_t = dB_t \ - \ \nabla V(X_t) dt. \tag{54}$$

Let us state a necessary and sufficient condition to write the drift of (X) as in (54). The proof is obtained by computation and requires the equality of the second order cross-derivatives of V.

Proposition 30. *If the following balance conditions on the ecological parameters are satisfied,*

$$c_{ij}\gamma_j = c_{ji}\gamma_i, \quad \forall i, j, \tag{55}$$

then the process X is a Kolmogorov process with potential V given by

$$V(x^1, \cdots, x^k) = \frac{1}{2} \sum_{i=1}^{k} \left(\ln(x^i) + \frac{c_{ii}\gamma_i (x^i)^4}{16} - \frac{r_i (x^i)^2}{2} \right) + \sum_{i \neq j} c_{ij}\gamma_j\, (x^i)^2\, (x^j)^2.$$

We will establish an existence and uniqueness result for the QSD of the process (X_t). The re-statement of the results for the initial stochastic Lotka-Volterra process follows immediately, since the hitting time of $(0, \cdots, 0)$ and H_i and the exit time of D are the same for both processes (X) and (Z).

By generalizing to k-types populations the results proved in [14] for two-types populations (an easy consequence of Girsanov's theorem), we get

Proposition 31. *For all $x \in D$, for all $i \neq j$,*

$$\mathbb{P}_x(T_{\partial D} = T_{H_i \cap H_j}) = 0.$$

Let us now state the first theorem, which is concerned by conditioning on the co-existence of the k types.

Theorem 32. *Under the balance conditions* (55)*, there exists a unique quasi-stationary distribution ν for the process (X) and the absorbing set ∂D, which is the quasi-limiting distribution starting from any initial distribution: for any μ on D and any $A \subset D$,*

$$\lim_{t\to+\infty} \mathbb{P}_\mu(X_t \in A | T_{\partial D} > t) = \nu(A).$$

Furthermore, there exist $\lambda > 0$ and a positive function η such that

$$\lim_{t\to+\infty} e^{\lambda t}\, \mathbb{P}_x(X_t \in A; T_{\partial D} > t) = \eta(x)\, \nu(A).$$

Proof. The proof of the existence of a quasi-stationary distribution results from the spectral theory for the semi-group of the killed process (P_t) (related to X) established in Cattiaux-Méléard [14], Appendices A, B, C. Define the reference measure on $(\mathbb{R}_+)^k$ by

$$\mu(dx_1, \cdots, dx_k) = e^{-2V(x)}\, dx_1 \cdots dx_k.$$

As in Subsection 5.2.2, one builds a self-adjoint operator on $\mathbb{L}^2(\mu)$ which coincides with P_t for bounded functions belonging to $\mathbb{L}^2(\mu)$. Its generator L is self-adjoint on $\mathbb{L}^2(\mu)$ and

$$Lg = \frac{1}{2}\Delta g - V \cdot \nabla g, \quad \forall g \in C_0^\infty(D).$$

We check that the assumptions required in [14] Theorem A.4 are satisfied and therefore, the operator $-L$ is proved to have a purely discrete spectrum of non-negative eigenvalues and the smallest one λ is positive. The corresponding eigenfunction η is proved to be in $\mathbb{L}^1(\mu)$ and the probability measure $\nu = \frac{\eta\, d\mu}{\int_D \eta\, d\mu}$ is the Yaglom limit.
let us emphasize that the uniqueness of the quasi-stationary distribution results by [14] Proposition B.12 from the ultracontractivity of the semi-group P_t (ultracontractivity means that P_t maps continuously $\mathbb{L}^2(\mu)$ in $\mathbb{L}^\infty(\mu)$ for any $t > 0$). The proof of the latter is easily generalized from the two-types case ([14] Proposition B.14) to the k-types case.

□

Theorem 32 shows that in some cases, a stabilization of the process with co-existence of the k types will occur before one of these types disappears. Let us now come back to our initial question: the long-time behavior of the process conditioned on non-extinction. For each $i = 1, \ldots, k$, we denote by λ_i the smallest eigenvalue related to the purely discrete spectrum of the generator for the i-axis diffusion defined by the stochastic differential equation (52).

Theorem 33. *Under the balance conditions* (55)*, there exists a Yaglom limit m for the process (X) conditioned on non extinction: for any $x \neq 0$, for any $A \subset D$,*

$$\lim_{t\to+\infty} \mathbb{P}_x(X_t \in A | T_0 > t) = m(A).$$

The support of this measure is included in the k axes.
Furthermore, if there exist $i_1, \ldots, i_l \in \{1, \cdots, k\}$ such that $\lambda^{i_1} = \cdots = \lambda^{i_l} < \min_{i \neq i_1, \ldots, i_l} \lambda^i$, then this QSD is concentrated on the axes of coordinates $i_1, \ldots, i_l$.

Proof. Recall that the existence of a Yaglom limit has been proved in the case $k = 1$ (Section 5.2). In what follows, we prove by induction the existence of a Yaglom limit for any k-type system (53).

The induction assumption (A_{k-1}) is as follows: we assume that, for any $(k-1)$-type Kolmogorov process $X^{(k-1)}$ satisfying (53) with (55), there exist a constant $\lambda > 0$, a uniformly bounded function $\eta > 0$ and a probability measure ν on $(\mathbb{R}_+)^{k-1}$ such that, for any $x \in (\mathbb{R}_+)^{k-1} \setminus \{0\}$ and any bounded measurable function f on $(\mathbb{R}^+)^{k-1}$ such that $f(0) = 0$, we have

$$\begin{aligned} &\lim_{t\to\infty} e^{\lambda t}\, \mathbb{E}_x(f(X_t^{(k-1)})) = \eta(x)\nu(f); \\ &\sup_{t\geq 0,\ x\in(\mathbb{R}_+^*)^{k-1}} |e^{\lambda t}\, \mathbb{E}_x(f(X_t^{(k-1)}))| < +\infty. \end{aligned} \tag{56}$$

As mentioned above, Assumption (A_1) is already proved. Let us assume that (A_{k-1}) is true and show that (A_k) follows.

Let $X^{(k)}$ be a k-type Kolmogorov process satisfying (53) with (55). Once hitting the boundary $\partial D = \cup_{i=1}^k H_i$, the process will no more leave it. Hence, for $t \geq T_{\partial D}$, the process will stay on the union of hyperplanes H_i. Moreover, $T_{\partial D} = \inf_{i=1,\cdots,k} T_{H_i}$. Fix $i \in \{1, \cdots, k\}$ and assume that the process leaves D through H_i. The dynamics on H_i is given by the process $(U_t^{(i),j})_{j\neq i}$ defined in $(\mathbb{R}_+)^{k-1}$ by:

$$dU_t^{(i),j} \;=\; dB_t^j \;+\; \left(\frac{r_j U_t^{(i),j}}{2} \;-\; \sum_{\ell=1,\ell\neq i}^{k} \frac{c_{j\ell}\gamma_j\, U_t^{(i),j}(U_t^{(i),\ell})^2}{8} \;-\; \frac{1}{2U_t^{(i),j}} \right).$$

Remark that by Proposition 31, the process really leaves ∂D by the interior of H_i. Each system $(U^{(i),j})_{j\neq i}$ is a $(k-1)$-type kolmogorov process (53) with balance conditions. Hence, by our induction assumption (A_{k-1}), there exist for each $i \in \{1, \cdots, k\}$ a positive constant v_i, a positive function η_i and a probability measure ν_i on H_i such that (56) holds for $(U^{(i),j})_{j\neq i}$, $i \in \{1, \cdots, k\}$.

Let us define

$$v_{min} = \inf_{i\in\{1,\cdots,k\}} v_i.$$

For any bounded measurable function f on $(\mathbb{R}_+)^k$ such that $f(0) = 0$ and for all $t \geq 0$, we have

$$e^{v_{min}t}\mathbb{E}_x\left(f(X_t^{(k)})\right) = e^{v_{min}t}\mathbb{E}_x\left(f(X_t^{(k)})\mathbf{1}_{T_{\partial D}>t}\right) + \sum_{i=1}^{k} \mathbb{E}_x\left(e^{v_{min}t} f(X_t^{(k)})\mathbf{1}_{T_{\partial D}=T_{H_i}\leq t}\right), \tag{57}$$

where we used the fact that $X^{(k)}$ reaches ∂D by hitting the interior of one and only one H_i. By Theorem 32, there exist a positive constant λ', a positive function η' and a probability measure ν' on $(\mathbb{R}_+^*)^k$ such that

$$\lim_{t\to+\infty} e^{\lambda' t}\mathbb{E}_x\left(f(X_t^{(k)})\mathbf{1}_{T_{\partial D}>t}\right) = \eta'(x)\nu'(f).$$

Moreover, a similar coupling argument as in [14] yields $\lambda' > v_{min}$. We deduce that

$$\lim_{t\to\infty} e^{v_{min}t}\mathbb{E}_x\left(f(X_t^{(k)})\mathbf{1}_{T_{\partial D}>t}\right) = 0.$$

For each $i \in \{1, \cdots, k\}$, we have by the Markov property

$$\begin{aligned}\mathbb{E}_x\left(e^{v_{min}t} f(X_t^{(k)})\ \mathbf{1}_{T_{\partial D}=T_{H_i}\le t}\right) &= \mathbb{E}_x\left(e^{v_{min}t}\,\mathbf{1}_{T_{\partial D}=T_{H_i}\le t}\,\mathbb{E}_{X_{T_{\partial D}}}\left(f(U^{(i)}_{t-T_{\partial D}})\right)\right)\\ &= \mathbb{E}_x\left(e^{v_{min}T_{\partial D}}\,\mathbf{1}_{T_{\partial D}=T_{H_i}\le t}\,\mathbb{E}_{X_{T_{\partial D}}}\left(e^{v_{min}(t-T_{\partial D})}f(U^{(i)}_{t-T_{\partial D}})\right)\right). \qquad (58)\end{aligned}$$

By the induction assumption (A_{k-1}), $e^{v_{min}(t-T_{\partial D})}f(U^{(i)}_{t-T_{\partial D}})$ is uniformly bounded. Moreover the inequality $0 < v_i < \lambda'$ and Proposition 3 ensure that $\mathbb{E}_x\left(e^{v_{min}T_{\partial D}}\right) < +\infty$. Using the convergence property of the induction assumption (A_{k-1}) and the dominated convergence theorem, we deduce that

$$\lim_{t\to\infty}\mathbb{E}_x\left(e^{v_{min}t}f(X_t^{(k)})\mathbf{1}_{T_{\partial D}=T_{H_i}\le t}\right) = \begin{cases}\mathbb{E}_x\left(e^{v_{min}T_{\partial D}}\mathbf{1}_{T_{\partial D}=T_{H_i}}\eta_i(X_{T_{\partial D}})\right)\nu_i(f), \text{ if } v_i = v_{min}\\ 0, \text{ otherwise.}\end{cases}$$

We have then

$$\lim_{t\to\infty} e^{v_{min}t}\mathbb{E}_x(f(X_t^{(k)})) = \sum_{i=1}^{k}\mathbf{1}_{v_i=v_{min}}\mathbb{E}_x\left(e^{v_{min}T_{\partial D}}\mathbf{1}_{T_{\partial D}=T_{H_i}}\eta_i(X_{T_{\partial D}})\right)\nu_i(f),$$

which gives us the first part of the induction assumption (A_k).

In order to prove the second part of (A_k), let us introduce the SLVP $Y^{(k)}$ with coefficients (c'_{ij}) defined by

$$c'_{kk} = c_{kk},\ c'_{ij} = c_{ij}\ \text{ and } c'_{ki} = c'_{ik} = 0,\ \forall i,j = 1, \cdots, k-1.$$

By the same coupling argument as above, the return time to ∂D for $X^{(k)}$ is stochastically dominated by the return time to ∂D for $Y^{(k)}$, *i.e.* $\mathbb{P}_x(X_t^{(k)} \in D) \le \mathbb{P}_x(Y_t^{(k)} \in D)$ for all $t \ge 0$.

Since the $k-1$ first components of $Y^{(k)}$ are independent of the last one and since

$$\{Y_t^{(k)} \in D\} = \{(Y_t^{(k),1}, \cdots, Y_t^{(k),k-1}) \in (\mathbb{R}_+^*)^{k-1}\} \cap \{Y_t^{(k),k} \in \mathbb{R}_+^*\},$$

we have

$$\mathbb{P}_x(Y_t^{(k)} \in D) \le \mathbb{P}_x((Y_t^{(k),1}, \cdots, Y_t^{(k),k-1}) \in (\mathbb{R}_+^*)^{k-1}) \times \mathbb{P}_x(Y_t^{(k),k} \in \mathbb{R}_+^*).$$

On the one hand, the dynamic of $(Y^{(k),1}, \cdots, Y^{(k),k-1})$ is the same as $U^{(k)}$, so that, by the second part of the induction assumption (A_{k-1}) and by the definition of v_{min},

$$\sup_{t\ge 0, x\in D} e^{v_{min}t}\mathbb{P}_x((Y_t^{(k),1}, \cdots, Y_t^{(k),k-1}) \in (\mathbb{R}_+^*)^{k-1}) < +\infty.$$

On the other hand, $Y^{(k),k}$ is a one dimensional SLVP, thus we deduce from (A_1) that there exists a positive constant λ_1 such that

$$\sup_{t\ge 0, x\in D} e^{\lambda_1 t}\mathbb{P}_x(Y_t^{(k),k} \in \mathbb{R}_+^*) < +\infty.$$

As a consequence, we have

$$\sup_{t\geq 0, x\in D} e^{(v_{min}+\lambda_1)t}\mathbb{P}_x(X_t^{(k)}\in D) \leq \sup_{t\geq 0, x\in D} e^{(v_{min}+\lambda_1)t}\mathbb{P}_x(Y_t^{(k)}\in D) < +\infty$$

and we deduce that

$$\sup_{x\in D}\mathbb{E}_x(e^{v_{min}T_{\partial D}}) < +\infty.$$

For any bounded measurable function f, this immediately leads us to

$$\sup_{t\geq 0, x\in E}\mathbb{E}_x(e^{v_{min}t}\mathbf{1}_{t<T_{\partial D}}f(X_t^{(k)})) < +\infty.$$

Moreover, by Equality (58) and the second part of (A_{k-1}), we deduce that, for each $i\in\{1,\cdots,k\}$,

$$\sup_{t\geq 0, x\in E}\mathbb{E}_x\left(e^{v_{min}t}f(X_t^{(k)})\mathbf{1}_{T_{\partial D}=T_{H_i}\leq t}\right) < +\infty.$$

By Equality (57), the second part of the induction assumption (A_k) is thus proved.

By induction on $k\geq 1$, we conclude that Assumption (A_k) is true for any $k\geq 1$, thus Theorem 33 follows. □

Example 5. Let us numerically study a 3-type system and observe its long-time behavior. The 3-tuple process (Z^1, Z^2, Z^3) evolves as

$$\begin{aligned}
dZ_t^1 &= \sqrt{\gamma_1 Z_t^1}dB_t^1 + \left(r_1Z_t^1 - c_{11}(Z_t^1)^2 - c_{12}Z_t^1Z_t^2 - c_{13}Z_t^1Z_t^3\right)dt,\\
dZ_t^2 &= \sqrt{\gamma_2 Z_t^2}dB_t^2 + \left(r_2Z_t^2 - c_{21}Z_t^1Z_t^2 - c_{22}(Z_t^2)^2 - c_{23}Z_t^2Z_t^3\right)dt,\\
dZ_t^3 &= \sqrt{\gamma_3 Z_t^3}dB_t^3 + \left(r_3Z_t^3 - c_{31}Z_t^1Z_t^3 - c_{32}Z_t^2Z_t^3 - c_{33}(Z_t^3)^2\right)dt,
\end{aligned}$$

with

$$\gamma_i = 1,\ c_{ii} = 10, \quad \forall i\in\{1,2,3\} \quad \text{and} \quad c_{ij} = 0.5,\ \forall\ i\neq j\in\{1,2,3\},$$

and

$$r_1 = 1.5,\ r_2 = 1,\ r_3 = 0.5\ ;\ Z_0^1 = Z_0^2 = Z_0^3 = 1.$$

We describe the dynamics of $\mathbb{P}_{(1,1,1)}((Z_t^1,Z_t^2,Z_t^3)\in\cdot \mid T_0 > t)$. As explained above, the process conditioned on non-extinction initially behaves as a 3-type population. Then a type goes extinct, then a second one and finally it only remains one type in the population. In order to represent graphically these transitions, we compute numerically the dynamics of the probabilities of coexistence and existence of the different types as functions of time. In Figure 10, we represent

(a) the probability of coexistence of the three types $\mathbb{P}_{(1,1,1)}(Z_t^1>0, Z_t^2>0, Z_t^3>0 \mid T_0 > t)$;

(b) the probability $\mathbb{P}_{(1,1,1)}(Z_t^i>0, Z_t^j>0, Z_t^k=0 \mid T_0>t)$ of coexistence of exactly two types $i\neq j$, for each combination of types $(i,j,k) = (1,2,3)$, $(i,j,k) = (2,3,1)$ and $(i,j,k) = (1,3,2)$;

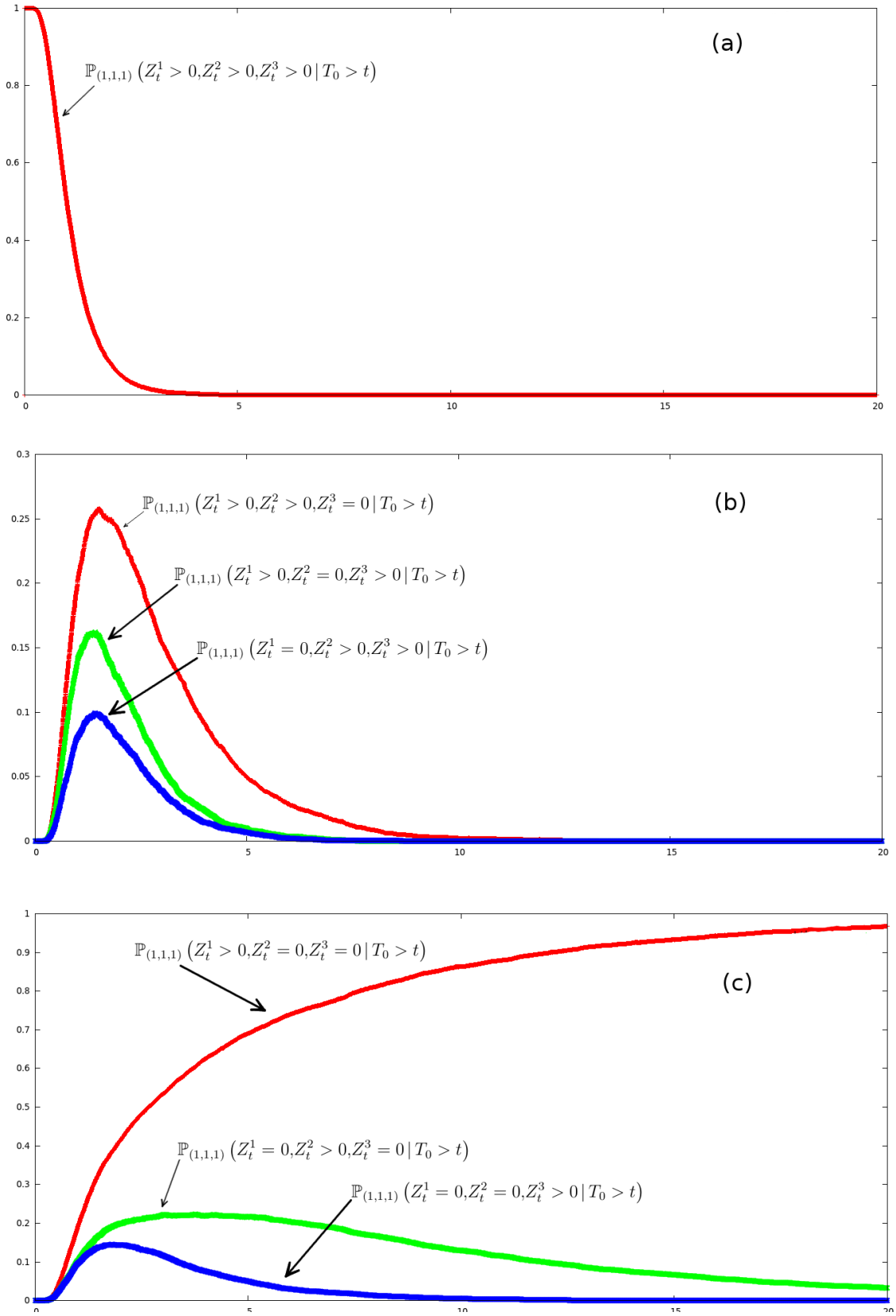


Figure 10: Dynamics of the probabilities of co-existence and existence of the different types for a 3-type stochastic Lotka-Volterra system. The horizontal axis is the time axis

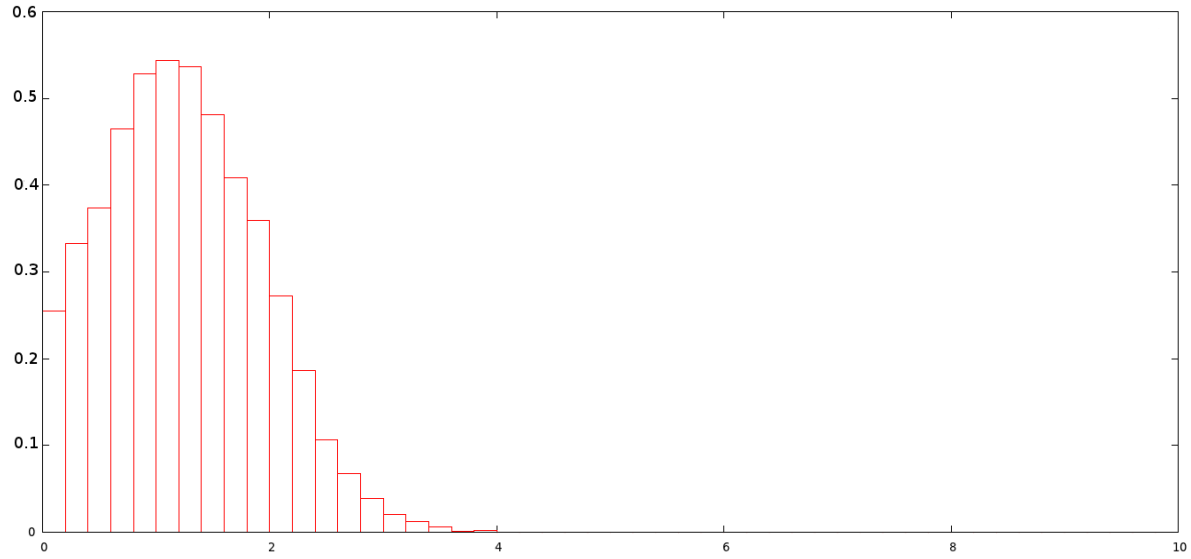


Figure 11: First marginal of the Yaglom limit of a 3-type stochastic Lotka-Volterra system. The two other marginals are equal to the null measure.

(c) the probability $\mathbb{P}_{(1,1,1)}(Z_t^i > 0, Z_t^j = 0, Z_t^k = 0 \mid T_0 > t)$ of existence of one and only one type i, for each type $i = 1, 2$ and 3.

As expected, the 3-type mode disappears quickly and the 2-type modes are transient. We also observe that the probability $\mathbb{P}_{(1,1,1)}(Z_t^1 > 0, Z_t^2 = 0, Z_t^3 = 0 \mid T_0 > t)$ converges to 1 when t increases, meaning that the last state of the population before extinction is monotype with type 1. It turns out that the support of the conditional law $\mathbb{P}_{(1,1,1)}((Z_t^1, Z_t^2, Z_t^3) \in \cdot \mid T_0 > t)$ becomes more and more concentrated on $\mathbb{R}_+^* \times \{0\} \times \{0\}$ in the long time. The Yaglom limit is thus equal to $\nu_1 \otimes \delta_{(0,0)}$, where ν_1 is the Yaglom limit of the process

$$dZ_t'^1 = \sqrt{Z_t'^1} dB_t^1 + \left(r_1 Z_t'^1 - c_{11}(Z_t'^1)^2\right) dt,$$

absorbed at 0 and is represented in Figure 11.

6 Simulation: the Fleming-Viot system

As seen in the previous sections, the spectral theory is a powerful tool to prove existence and eventually uniqueness of a QSD for a given process Z. It is based on the equivalence property of Proposition 4, stating that a probability measure α on E^* is a QSD for the killed process Z if and only if

$$\alpha L = -\theta(\alpha)\alpha, \tag{59}$$

where L denotes the infinitesimal generator of Z and $\theta(\alpha)$ a positive constant. In some cases, such as in the finite state space case, one can easily compute numerically the whole set of eigenvalues and eigenvectors of L as seen in Example 1 and Example 2. For these numerical illustrations, we used the software SCILAB and its function `spec`. We also

refer to [62] for a detailed description of some algorithms available in MATLAB for the computation of eigenfunctions and eigenvalues in large (but finite) state space cases.

In other cases, such as the logistic birth and death process of Section 4 and the logistic Feller diffusion of Section 5, solving numerically Equation (59) is too hard and we use a different approach. This approach consists in approximating the QSD and the conditioned distribution $\mathbb{P}_z(Z_t \in .|t < T_0)$ by the empirical distribution of a simulable interacting particle system. This Fleming-Viot type system, built for any number of particles $N \geq 2$, has been introduced by Burdzy, Holyst and March [10] and explored in [11] and in Grigorescu-Kang [32] for d-dimensional killed Brownian motions. It has also been studied in Villemonais [65] for multi-dimensional diffusion processes with unbounded drifts and a general result is available in [66]. Similar systems have also been considered by Ferrari-Marič [23] for continuous Markov chains in a countable state space. In this section, we explain the approximation method based on the Fleming-Viot type interacting particle systems.

Let Z be a killed Markov process which evolves in the state space E. Fix $N \geq 2$ and let $Z_0 \in E$ be its initial value. The interacting particle system with N particles $(Z^1, \cdots, Z^N)$ starts from $(Z_0, \cdots, Z_0)$ and belongs to $(E^*)^N$. The particles evolve independently from this initial position according to the law of the killed Markov process Z, until one of them hits the state 0. At that time τ_1, the killed particle jumps to the position at τ_1 of one of the $N-1$ remaining particles, chosen uniformly among them. Then the particles evolve independently according to the law of Z until one of them attains 0 (time τ_2), and so on. The sequence of jumps is denoted by $(\tau_n)_n$ and we set

$$\tau_\infty = \lim_{n\to\infty} \tau_n.$$

This procedure defines the $(E^*)^N$-valued process $(Z^1, \cdots, Z^N)$ for all time $t \in [0, \tau_\infty[$. Figure 12 shows an illustration of such a system with two particles evolving between their jumps as Markov processes absorbed in 0 and 1.

If $\tau_\infty = +\infty$ almost surely, then the Fleming-Viot particle system will be well defined at all time $t > 0$. The condition $\tau_\infty = +\infty$ is clearly fulfilled for continuous time Markov chains with bounded jump rates. In the diffusion process case, criteria have been provided in [8], [33], [65] and [66].

In that case, denote by μ_t^N the empirical distribution of $(Z^1, \cdots, Z^N)$ at time t:

$$\mu_t^N = \frac{1}{N} \sum_{i=1}^{N} \delta_{Z_t^i}, \quad \forall t \geq 0.$$

The following result is obtained in [66] by martingale method.

Theorem 34. *Assume that for all $N \geq 2$, $(Z^1, \cdots, Z^N)$ is well defined at any time $t \geq 0$. Then, for any time $t > 0$, the sequence of empirical distributions (μ_t^N) converges in law to the conditioned distribution $\mathbb{P}_{Z_0}(Z_t \in \cdot | t < T_0)$, when N goes to infinity.*

If moreover $(Z^1, \cdots, Z^N)$ is ergodic, we denote by M^N its stationary distribution and by $\mathcal{X}^N$ its empirical stationary distribution, which is defined by $\mathcal{X}^N = \frac{1}{N}\sum_{i=1}^N \delta_{z_i}$, where

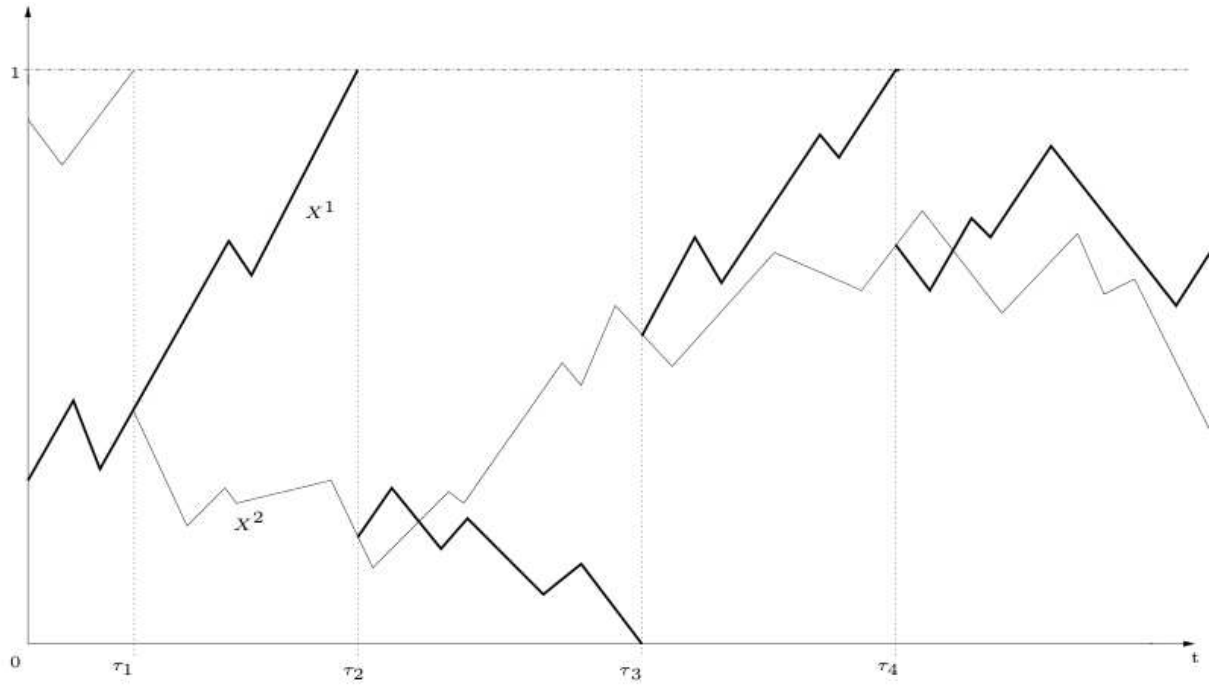


Figure 12: Fleming-Viot type system with two particles absorbed in 0 and 1.

$(z_1, \cdots, z_N) \in E^*$ is a random vector distributed with respect to M^N. In particular, μ_t^N converges in law to $\mathcal{X}^N$ when $t \to \infty$. We refer to [65] for the proof of the following theorem.

Theorem 35. *Assume that Z has a QLD α which attracts all initial distributions: for any probability measure μ on E^*,*

$$\lim_{t \to +\infty} \mathbb{P}_\mu \left(Z_t \in \cdot | t < T_0 \right) = \alpha.$$

Assume moreover that $(Z^1, \cdots, Z^N)$ is ergodic and that the family of laws of $(\mathcal{X}^N)_{N \geq 2}$ is uniformly tight. Then the sequence of random probability measures $(\mathcal{X}^N)$ converges weakly to α.

If E is a bounded subset of $\mathbb{R}^d$, $d \geq 1$, and if Z is a drifted Brownian motion with bounded drift which is killed at the boundaries of E, then the assumptions of Theorems 34 and 35 are fulfilled (see [65]). The proofs of Theorems 34 and 35 are based on a coupling argument. More general (but longer) proofs can also be found in [33] or [66]. In particular, these results provide us a numerical approximation method of the Yaglom limit for such processes.

Let us now consider the Kolmogorov diffusion process X defined in (39). In that case, the existence of the Fleming-Viot particle system remains an open problem because of the unboundedness of the drift coefficient. In order to avoid this difficulty, we introduce the law $\mathbb{P}^\varepsilon$ of the diffusion process with bounded coefficients defined by

$$dX_t^\varepsilon = dB_t - q(X_t^\varepsilon)dt \ ; \ X_0 \in (\varepsilon, 1/\varepsilon), \tag{60}$$

killed when it hits ε or $\frac{1}{\varepsilon}$. One can easily show that at any time $t \geq 0$, the conditioned distribution of X^ε converges to the one of X:

$$\mathbb{P}^\varepsilon \left(X_t^\varepsilon \in \cdot | t < T_\varepsilon \wedge T_{1/\varepsilon} \right) \xrightarrow[\varepsilon \to 0]{} \mathbb{P} \left(X_t \in \cdot | t < T_0 \right).$$

The existence of the Yaglom limit denoted by α_ϵ and the uniqueness of the QSD for $\mathbb{P}^\varepsilon$ are obtained from Pinsky [49]. The following approximation result is proved in [65] using a compactness-uniqueness argument.

Proposition 36. *The sequence $(\alpha^\varepsilon)_\varepsilon$ weakly converges to the Yaglom limit α of X as ε tends to 0.*

For all $N \geq 2$, we denote by $(X^{\varepsilon,1}, \cdots, X^{\varepsilon,N})$ the interacting particle system built as above, with the law $\mathbb{P}^\epsilon$. Since the diffusion process X^ϵ is a drifted Brownian motion with bounded drift evolving in the bounded interval $]\epsilon,1/\epsilon[$, the interacting particle system $(X^{\varepsilon,1}, \cdots, X^{\varepsilon,N})$ fulfills the assumptions of Theorems 34 and 35. Denoting by $\mu^{\epsilon,N}$ the empirical distribution of the simulable particle system $(X^{\varepsilon,1}, \cdots, X^{\varepsilon,N})$, we get

$$\lim_{\epsilon \to 0} \lim_{N \to \infty} \mu_t^{\varepsilon,N} = \mathbb{P}_{X_0} \left(X_t \in \cdot | t < T_0 \right), \forall t \geq 0,$$

and

$$\lim_{\epsilon \to 0} \lim_{N \to \infty} \lim_{t \to \infty} \mu_t^{\varepsilon,N} = \lim_{\epsilon \to 0} \alpha_\varepsilon = \alpha.$$

Then, choosing ϵ small enough and N big enough, we get a numerical approximation method for the conditioned distribution and the Yaglom limit of X.

Example 6. Let us now develop this simulation method in the case of the Wright-Fisher diffusion conditioned to be absorbed at 0, which evolves in [0,1[and is defined by

$$dZ_t = \sqrt{Z_t(1 - Z_t)} dB_t - Z_t dt,\ Z_0 = z \in]0,1[.$$

This is a model for a bi-type population in which the second type cannot disappear. In that model, Z_t is the proportion of the first type in the population at time $t \geq 0$ and $1 - Z_t$ the proportion of the other one. The existence of a Yaglom limit for this process has been proved by Huillet in [35], which also proved that it has the density $2 - 2x$ with respect to the Lebesgue measure.

Using the approximation method described above with $\epsilon = 0.001$ and $N = 10000$, we obtain numerically the density of the Yaglom limit for Z represented in Figure 13, which is very close to the function $x \mapsto 2 - 2x$ and shows the efficiency of the method.

References

[1] D. Aldous. Stopping times and tightness. *Ann. Probability*, Volume 6, Issue 2 (1978), pages 335–340.

[2] W. J. Anderson. *Continuous-time Markov chains.* Springer Series in Statistics: Probability and its Applications. Springer-Verlag, New York, 1991.

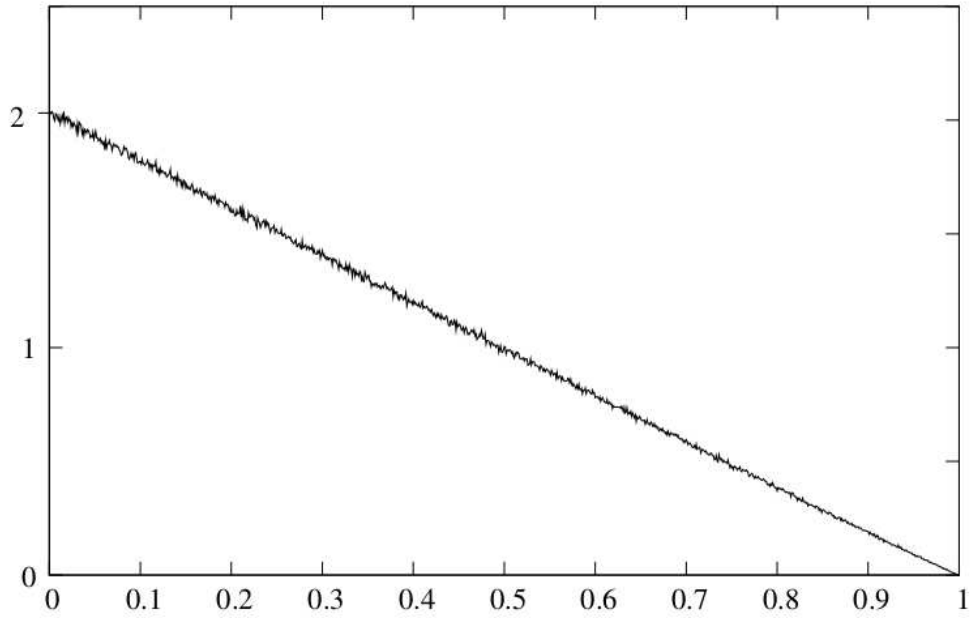


Figure 13: The Yaglom limit of the Wright-Fisher diffusion conditioned to be absorbed at 0 obtained by numerical simulation.

[3] A. Asselah, P. A. Ferrari, and P. Groisman. Quasistationary distributions and Fleming-Viot processes in finite spaces. *J. Appl. Probab.*, Volume 48, Issue 2 (2011), pages 322–332.

[4] K. B. Athreya and P. E. Ney. *Branching processes*. Springer-Verlag, New York, 1972. Die Grundlehren der mathematischen Wissenschaften, Band 196.

[5] V. Bansaye. Surviving particles for subcritical branching processes in random environment. *Stochastic Process. Appl.*, Volume 119, Issue 8 (2009), pages 2436–2464.

[6] A. D. Barbour and P. K. Pollett. Total variation approximation for quasi-stationary distributions. *J. Appl. Probab.*, Volume 47, Issue 4 (2010), pages 934–946.

[7] F. A. Berezin and M. A. Shubin. *The Schrödinger equation*, volume 66 of *Mathematics and its Applications (Soviet Series)*. Kluwer Academic Publishers Group, Dordrecht, 1991. Translated from the 1983 Russian edition by Yu. Rajabov, D. A. Leĭtes and N. A. Sakharova and revised by Shubin, With contributions by G. L. Litvinov and Leĭtes.

[8] M. Bieniek, K. Burdzy, and S. Finch. Non-extinction of a fleming-viot particle model. *Probab. Theory Related Fields* (2011), pages 1–40.

[9] P. Billingsley. *Convergence of probability measures*. John Wiley & Sons Inc., New York, 1968.

[10] K. Burdzy, R. Holyst, D. Ingerman, and P. March. Configurational transition in a fleming-viot-type model and probabilistic interpretation of laplacian eigenfunctions. *J. Phys. A*, Volume 29 (1996), pages 2633–2642.

[11] K. Burdzy, R. Hołyst, and P. March. A Fleming-Viot particle representation of the Dirichlet Laplacian. *Comm. Math. Phys.*, Volume 214, Issue 3 (2000), pages 679–703.

[12] Carey, P. Liedo, D. Orozco, and J. W. Vaupel. Slowing of mortality rates at older ages in large medfly cohorts. *Science*, Volume 258 (1992), pages 457–461.

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

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

[15] P. Collet, S. Martínez, S. Méléard, and J. San Martín. Quasi-stationary distributions for structured birth and death processes with mutations. *Probab. Theory Related Fields*, Volume 151, Issue 1 (2011), pages 191–231.

[16] P. Collet, S. Martínez, and J. San Martín. Asymptotic laws for one-dimensional diffusions conditioned to nonabsorption. *Ann. Probab.*, Volume 23, Issue 3 (1995), pages 1300–1314.

[17] P. Coolen-Schrijner and E. A. van Doorn. Quasi-stationary distributions for a class of discrete-time Markov chains. *Methodol. Comput. Appl. Probab.*, Volume 8, Issue 4 (2006), pages 449–465.

[18] J. N. Darroch and E. Seneta. On quasi-stationary distributions in absorbing discrete-time finite Markov chains. *J. Appl. Probability*, Volume 2 (1965), pages 88–100.

[19] J. N. Darroch and E. Seneta. On quasi-stationary distributions in absorbing continuous-time finite Markov chains. *J. Appl. Probability*, Volume 4 (1967), pages 192–196.

[20] A. M. Etheridge. Survival and extinction in a locally regulated population. *Ann. Appl. Probab.*, Volume 14, Issue 1 (2004), pages 188–214.

[21] S. N. Ethier and T. G. Kurtz. *Markov processes, Characterization and convergence.* Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986.

[22] P. A. Ferrari, H. Kesten, S. Martínez, and P. Picco. Existence of quasi-stationary distributions. A renewal dynamical approach. *Ann. Probab.*, Volume 23, Issue 2 (1995), pages 501–521.

[23] P. A. Ferrari and N. Marić. Quasi stationary distributions and Fleming-Viot processes in countable spaces. *Electron. J. Probab.*, Volume 12, Issue 24 (2007), pages 684–702.

[24] P. A. Ferrari, S. Martínez, and P. Picco. Some properties of quasi-stationary distributions in the birth and death chains: a dynamical approach. In *Instabilities and nonequilibrium structures, III (Valparaíso, 1989)*, volume 64 of *Math. Appl.* (1991), pages 177–187.

[25] M. Fukushima. *Dirichlet forms and Markov processes*, volume 23 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam, 1980.

[26] F. Galton and H. W. Watson. On the probability of the extinction of families. *Available at http://galton.org/essays/1870-1879/galton-1874-jaigi-family-extinction.pdf*, 1974.

[27] F. R. Gantmacher. *The theory of matrices. Vols. 1, 2.* Translated by K. A. Hirsch. Chelsea Publishing Co., New York, 1959.

[28] B. Gompertz. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Phil. Trans. R. Soc. London*, Volume 115 (1825), pages 513–583.

[29] G. L. Gong, M. P. Qian, and Z. X. Zhao. Killed diffusions and their conditioning. *Probab. Theory Related Fields*, Volume 80, Issue 1 (1988), pages 151–167.

[30] P. Good. The limiting behavior of transient birth and death processes conditioned on survival. *J. Austral. Math. Soc.*, Volume 8 (1968), pages 716–722.

[31] M. Greenwood and J. Irwin. The biostatistics of senility. *Human Biology*, Volume 11, Issue 1 (1939), pages 1–23.

[32] I. Grigorescu and M. Kang. Hydrodynamic limit for a Fleming-Viot type system. *Stochastic Process. Appl.*, Volume 110, Issue 1 (2004), pages 111–143.

[33] I. Grigorescu and M. Kang. Immortal particle for a catalytic branching process. *Probab. Theory Related Fields* (2011), pages 1–29. 10.1007/s00440-011-0347-6.

[34] A. G. Hart and P. K. Pollett. New methods for determining quasi-stationary distributions for Markov chains. *Math. Comput. Modelling*, Volume 31, Issue 10-12 (2000), pages 143–150.

[35] T. Huillet. On Wright Fisher diffusion and its relatives. *J. Stat. Mech.-Theory E.*, Volume 11 (2007), pages 6–+.

[36] N. Ikeda and S. Watanabe. *Stochastic differential equations and diffusion processes*. North-Holland Mathematical Library, 1989.

[37] A. Joffe and M. Métivier. Weak convergence of sequences of semimartingales with applications to multitype branching processes. *Adv. in Appl. Probab.*, Volume 18, Issue 1 (1986), pages 20–65.

[38] I. Karatzas and S. E. Shreve. *Brownian motion and stochastic calculus*, Volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1988.

[39] S. Karlin and J. L. McGregor. The differential equations of birth-and-death processes, and the Stieltjes moment problem. *Trans. Amer. Math. Soc.*, Volume 85 (1957), pages 489–546.

[40] R. Knobloch and L. Partzsch. Uniform conditional ergodicity and intrinsic ultracontractivity. *Potential Analysis*, Volume 33 (2010), pages 107–136.

[41] M. Kolb and D. Steinsaltz. Quasilimiting behavior for one-dimensional diffusions with killing. *To appear in Ann. Probab.*.

[42] A. Lambert. The branching process with logistic growth. *Ann. Appl. Probab.*, Volume 15, Issue 2 (2005), pages 1506–1535.

[43] A. Lambert. Quasi-stationary distributions and the continuous-state branching process conditioned to be never extinct. *Electron. J. Probab.*, Volume 12, Issue 14 (2007), pages 420–446.

[44] M. Lladser and J. San Martín. Domain of attraction of the quasi-stationary distributions for the Ornstein-Uhlenbeck process. *J. Appl. Probab.*, Volume 37, Issue 2 (2000), pages 511–520.

[45] P. Mandl. Spectral theory of semi-groups connected with diffusion processes and its application. *Czechoslovak Math. J.*, Volume 11, Issue 86 (1961), pages 558–569.

[46] S. Martínez, P. Picco, and J. San Martín. Domain of attraction of quasi-stationary distributions for the Brownian motion with drift. *Adv. in Appl. Probab.*, Volume 30, Issue 2 (1998), pages 385–408.

[47] S. Martínez and J. San Martín. Classification of killed one-dimensional diffusions. *Ann. Probab.*, Volume 32, Issue 1 (2004), pages 530–552.

[48] A. G. Pakes and P. K. Pollett. The supercritical birth, death and catastrophe process: limit theorems on the set of extinction. *Stochastic Process. Appl.*, Volume 32, Issue 1 (1989), pages 161–170.

[49] R. G. Pinsky. On the convergence of diffusion processes conditioned to remain in a bounded region for large time to limiting positive recurrent diffusion processes. *Ann. Probab.*, Volume 13, Issue 2 (1985), pages 363–378.

[50] P. Pollett. Quasi-stationary distributions : a bibliography. http://www.maths.uq .edu.au/∼pkp/papers/qsds/qsds.pdf.

[51] P. K. Pollett and D. E. Stewart. An efficient procedure for computing quasi-stationary distributions of Markov chains with sparse transition structure. *Adv. in Appl. Probab.*, Volume 26, Issue 1 (1994), pages 68–79.

[52] O. Renault, R. Ferrière, and J. Porter. The quasi-stationary route to extinction. Private communication.

[53] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.

[54] E. Seneta and D. Vere-Jones. On quasi-stationary distributions in discrete-time Markov chains with a denumerable infinity of states. *J. Appl. Probability*, Volume 3 (1966), pages 403–434.

[55] D. Serre. *Matrices*, volume 216 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2002. Theory and applications, Translated from the 2001 French original.

[56] D. Steinsaltz and S. N. Evans. Markov mortality models: Implications of quasistationarity and varying initial conditions. *Theo. Pop. Bio.*, Volume 65 (2004), 319–337.

[57] D. Steinsaltz and S. N. Evans. Quasistationary distributions for one-dimensional diffusions with killing. *Trans. Amer. Math. Soc.*, Volume 359, Issue 3 (2007), pages 1285–1324.

[58] N. S. Trudinger. Pointwise estimates and quasilinear parabolic equations. *Comm. Pure Appl. Math.*, Volume 21 (1968), pages 205–226.

[59] E. A. van Doorn. Quasi-stationary distributions and convergence to quasi-stationarity of birth-death processes. *Adv. in Appl. Probab.*, Volume 23, Issue 4 (1991), pages 683–700.

[60] E. A. van Doorn. Conditions for the existence of quasi-stationary distributions for birth-death processes with killing. *Memorandum No. 1949, Department of Applied Mathematics, University of Twente*, 2011.

[61] E. A. van Doorn and P. K. Pollett. Quasi-stationary distributions for reducible absorbing Markov chains in discrete time. *Markov Process. Related Fields*, Volume 15, Issue 2 (2009), pages 191–204.

[62] E. A. van Doorn and P. K. Pollett. Quasi-stationary distributions. *Memorandum No. 1945, Department of Applied Mathematics, University of Twente*, 2011.

[63] D. Vere-Jones. Some limit theorems for evanescent processes. *Austral. J. Statist.*, Volume 11 (1969), pages 67–78.

[64] P. F. Verhulst. Notice sur la loi que la population suit dans son accroissement. *Corr. Math. et Phys.*, 1938.

[65] D. Villemonais. Interacting particle systems and Yaglom limit approximation of diffusions with unbounded drift. *Electronic Journal of Probability*, Volume 16 (2011), pages 1663–1692.

[66] D. Villemonais. Interacting particle processes and approximation of Markov processes conditioned to not be killed. *ArXiv e-prints*, 2011.

[67] A. M. Yaglom. Certain limit theorems of the theory of branching random processes. *Doklady Akad. Nauk SSSR (N.S.)*, Volume 56 (1947), pages 795–798.